Lagrangeov i Markovljev spektar

Tomislav Pejković

PMF – Matematički odsjek, 2018/2019.

Sadržaj

Predgovor

Ova skripta nastala je za potrebe predavanja iz kolegija *Lagrangeov i Markovljev spektar* koji se održavao na doktorskom studiju pri Matematičkom odsjeku Prirodoslovno-matematičkog fakulteta u Zagrebu od studenog 2018. do lipnja 2019. godine.

Popis najvažnije korištene literature dan je na kraju skripte. U nastavku ukratko opisujemo obrađene teme. Prva dva poglavlja su uvodnog tipa. Posebno je važno prvo u kojem se navode osnovne definicije i dokazuju rezultati o verižnim razlomcima koji se koriste kroz čitav kolegij. U drugom poglavlju obrađene su binarne kvadratne forme i to s realnim koeficijentima. U trećem poglavlju se dokazuju osnovni rezultati o strukturi spektara. Najopsežnije poglavlje je četvrto u kojem se detaljno iznosi klasična Markovljeva teorija o strukturi spektara na intervalu $(0, 3)$ gdje se oni podudaraju. Peto poglavlje nadovezuje se na prethodno i donose se neki vrlo parcijalni rezultati vezani uz poznatu slutnju o jedinstvenosti Markovljevih brojeva. U šestom poglavlju sabrani su najvažniji rezultati o spektrima na intervalu $(3, +\infty)$. U sedmom poglavlju nastavlja se proučavanje tih dijelova spektara, ali s metričke točke gledišta. Na koncu su vrlo površno spomenuta dva moguća analogona klasičnih spektara.

Uz skriptu su na internetu dostupne i tri riješene zadaće koje mogu poslužiti za provjeru usvojenosti nekih od pojmova proučavanih na kolegiju.

Poglavlje 1

Verižni razlomci

1.1 Osnovne definicije i rezultati

Verižni razlomak je izraz oblika

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}} \tag{1.1}$$

koji označavamo $[a_0, a_1, a_2, \ldots]$, a može biti konačan ili beskonačan. Veličine $a_0, a_1, a_2, \ldots$ koje zovemo *parcijalni kvocijenti* mogu biti cijeli, realni ili kompleksni brojevi te funkcije u takvim varijablama. Ako je a_i cijeli broj za svaki i te $a_i > 0$ za $i \geq 1$, onda govorimo o *jednostavnom* verižnom razlomku i kod nas će gotovo uvijek biti takvi.

Verižnom razlomku (1.1) pridružujemo dva niza

$$\begin{aligned} p_{-2} &= 0, & p_{-1} &= 1, & p_n &= a_n p_{n-1} + p_{n-2}, \\ q_{-2} &= 1, & q_{-1} &= 0, & q_n &= a_n q_{n-1} + q_{n-2}, \end{aligned} \tag{1.2}$$

za $n \geq 0$. Matrično to zapisujemo

$$\begin{bmatrix} p_{-1} & p_{-2} \\ q_{-1} & q_{-2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} = \begin{bmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{bmatrix} \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix},$$

pa je

$$\begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} = \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix}. \tag{1.3}$$

Indukcijom se pokazuje da je $\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n]$ za $n \geq 0$. Baza za $n = 0, 1$ je trivijalna, a pretpostavimo li da tvrdnja vrijedi za verižne razlomke s manje od $n+1$ parcijalnih kvocijenata, imamo

$$\begin{aligned} [a_0, a_1, \ldots, a_{n-1}, a_n] &= [a_0, a_1, \ldots, a_{n-1} + \frac{1}{a_n}] \\ &= \frac{\left(a_{n-1} + \frac{1}{a_n}\right) p_{n-2} + p_{n-3}}{\left(a_{n-1} + \frac{1}{a_n}\right) q_{n-2} + q_{n-3}} = \frac{(a_n a_{n-1} + 1) p_{n-2} + a_n p_{n-3}}{(a_n a_{n-1} + 1) q_{n-2} + a_n q_{n-3}} \\ &= \frac{a_n p_{n-1} + p_{n-2}}{a_n q_{n-1} + q_{n-2}} = \frac{p_n}{q_n}. \end{aligned}$$

Transponiranjem matrice u (1.3) dobivamo

$$\begin{bmatrix} p_n & q_n \\ p_{n-1} & q_{n-1} \end{bmatrix} = \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{n-1} & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix}$$

te je stoga

$$\frac{p_n}{p_{n-1}} = [a_n, a_{n-1}, \ldots, a_0], \qquad \frac{q_n}{q_{n-1}} = [a_n, a_{n-1}, \ldots, a_1]. \tag{1.4}$$

Promotrimo (konačni ili beskonačni) jednostavni verižni razlomak $[a_0, a_1, a_2, \ldots]$. Niz nazivnika $(q_n)_{n\geq 1}$ je strogo rastući. Iz (1.2) je za $n \geq 0$

$$q_n \geq 2^{\lfloor \frac{n}{2} \rfloor}, \tag{1.5}$$

gdje smo s $\lfloor \cdot \rfloor$ označili funkciju najveće cijelo. Lako je dati i bolju ogradu

$$q_n \geq F_{n+1} > \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^{n+1} - 1. \tag{1.6}$$

Ovdje je s $F_0 = 0$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$ dan niz Fibonaccijevih brojeva, a odmah vidimo i da će rast niza $(q_n)_{n\geq 0}$ biti najsporiji za verižni razlomak $[1, 1, 1, \ldots]$ kojemu je $q_n = F_{n+1}$.

Uzmemo li u (1.3) determinantu, dobivamo

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n+1} = (-1)^{n-1} \tag{1.7}$$

iz čega je jasno da su *konvergente* $\frac{p_n}{q_n}$ $(n \geq 0)$ skraćeni razlomci, tj. najveći zajednički djelitelj $(p_n, q_n) = 1$, ali i $(p_n, p_{n-1}) = (q_n, q_{n-1}) = 1$. Također je iz (1.7)

$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^{n-1}}{q_{n-1} q_n}. \tag{1.8}$$

U slučaju beskonačnog verižnog razlomka, niz konvergenti je Cauchyjev i zato konvergira prema

$$\frac{p_0}{q_0} + \sum_{n=1}^{\infty} \left(\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} \right) = \frac{p_0}{q_0} + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{q_{n-1} q_n}.$$

Vrijedi

$$p_n q_{n-2} - q_n p_{n-2} = \begin{vmatrix} p_n & p_{n-2} \\ q_n & q_{n-2} \end{vmatrix} = \begin{vmatrix} p_n - p_{n-2} & p_{n-2} \\ q_n - q_{n-2} & q_{n-2} \end{vmatrix} = \begin{vmatrix} a_n p_{n-1} & p_{n-2} \\ a_n q_{n-1} & q_{n-2} \end{vmatrix} = (-1)^n a_n$$

iz čega je

$$\frac{p_n}{q_n} - \frac{p_{n-2}}{q_{n-2}} = \frac{(-1)^n a_n}{q_{n-2} q_n}.$$

Vidimo da je niz konvergenti s parnim indeksima rastući, a niz konvergenti s neparnim indeksima padajući i oba niza teže istom broju. Time smo i beskonačnom jednostavnom verižnom razlomku pridružili realan broj $\lim_{n\to\infty} \frac{p_n}{q_n}$ koji nazivamo vrijednošću verižnog razlomka.

Krenimo sada obratno, tj. od realnog broja prema pridruženom verižnom razlomku. Pomoću Euklidovog algoritma svaki racionalan broj možemo razviti u konačni jednostavni verižni razlomak $[a_0, a_1, \ldots, a_n]$ koji je jedinstven ako zahtijevamo da vrijedi $n = 0$ ili $a_n \geq 2$. Bez tog zahtjeva, možemo pisati i $[a_0, a_1, \ldots, a_n - 1, 1]$ za isti racionalni broj što ćemo ponekad koristiti ako želimo da nam broj parcijalnih kvocijenata bude određene parnosti.

Svaki realan broj α razvijamo u verižni razlomak $[a_0, a_1, a_2, \ldots]$ ovako:

$$\alpha_0 = \alpha, \quad a_i = \lfloor \alpha_i \rfloor, \quad \alpha_{i+1} = \frac{1}{\alpha_i - a_i}, \ i \geq 0.$$

Za iracionalan α su svi α_i iracionalni brojevi te je za sve $i \geq 1$, $\alpha_i > 1$ i zato $a_i \geq 1$ te očito vrijedi $a_i < \alpha_i < a_i + 1$.

Brojeve α_i zovemo *potpuni kvocijenti*. Iz $\alpha = [a_0, a_1, \ldots, a_n, \alpha_{n+1}]$, označimo li niz konvergenti verižnog razlomka $[a_0, a_1, \ldots]$ s $(\frac{p_k}{q_k})_k$, dobivamo pomoću (1.2) važnu jednakost

$$\alpha = \frac{\alpha_{n+1} p_n + p_{n-1}}{\alpha_{n+1} q_n + q_{n-1}}. \tag{1.9}$$

Pomoću (1.9), (1.7) i (1.4) dobivamo ključno aproksimacijsko svojstvo za verižne razlomke. Naime,

$$\alpha - \frac{p_n}{q_n} = \frac{\alpha_{n+1} p_n + p_{n-1}}{\alpha_{n+1} q_n + q_{n-1}} - \frac{p_n}{q_n} = \frac{\alpha_{n+1} p_n q_n + p_{n-1} q_n - \alpha_{n+1} p_n q_n - p_n q_{n-1}}{(\alpha_{n+1} q_n + q_{n-1}) q_n},$$

pa je

$$\begin{aligned} \alpha - \frac{p_n}{q_n} &= \frac{(-1)^n}{q_n(\alpha_{n+1} q_n + q_{n-1})} = \frac{(-1)^n}{q_n^2(\alpha_{n+1} + \beta_{n+1})} \\ q_n \alpha - p_n &= \frac{(-1)^n}{\alpha_{n+1} q_n + q_{n-1}}, \\ \alpha_{n+1} &= [a_{n+1}, a_{n+2}, \ldots], \\ \beta_{n+1} &= \frac{q_{n-1}}{q_n} = [0, a_n, a_{n-1}, \ldots, a_1]. \end{aligned} \tag{1.10}$$

Odavde zbog rasta od $(q_n)_n$ slijedi da niz $(\frac{p_n}{q_n})_n$ teži prema α. Ako su dva beskonačna verižna razlomka jednaki, lako se vidi da su im početni, pa onda induktivno i svi, parcijalni kvocijenti jednaki. Time smo ustanovili bijekciju između skupa iracionalnih brojeva i svih beskonačnih jednostavnih verižnih razlomaka.

Koristeći matrični zapis djelovanja razlomljene linearne funkcije, jednakost (1.9) možemo pisati

$$\alpha = \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} \alpha_{n+1},$$

gdje je odgovarajuća matrica prema (1.7) *unimodularna*, odnosno determinanta joj je ± 1. Sve unimodularne matrice čine grupu $\mathrm{GL}(2, \mathbb{Z})$ i lako se vidi da komponiranje razlomljenih linearnih funkcija odgovara množenju matrica u toj grupi. Tako imamo

$$\alpha_{n+1} = \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix}^{-1} \alpha = \begin{bmatrix} q_{n-1} & -p_{n-1} \\ -q_n & p_n \end{bmatrix} \alpha = \frac{q_{n-1}\alpha - p_{n-1}}{-q_n \alpha + p_n}. \tag{1.11}$$

Kažemo da su iracionalni brojevi α i β *ekvivalentni* ako postoji $T \in \mathrm{GL}(2, \mathbb{Z})$ tako da je $\beta = T\alpha$. Lako se provjeri da je ovo zaista relacija ekvivalencije te koristimo oznaku $\alpha \sim \beta$.

Teorem 1.1 (Serret). *Neka su $\alpha = [a_0, a_1, a_2, \ldots]$ i $\beta = [b_0, b_1, b_2, \ldots]$ iracionalni brojevi. Tada su α i β ekvivalentni ako i samo ako postoje cijeli brojevi k i l takvi da je $a_{k+n} = b_{l+n}$ za sve $n \geq 0$.*

Lema 1.2. *Neka je $T = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ unimodularna matrica i $\beta = T\alpha$, $\alpha > 1$, $c > d > 0$. Tada su $\frac{b}{d}$ i $\frac{a}{c}$ dvije uzastopne konvergente od β, recimo $\frac{p_{n-2}}{q_{n-2}}$ i $\frac{p_{n-1}}{q_{n-1}}$ te vrijedi $\alpha = \beta_n$.*

Dokaz leme. Prikažimo $\frac{a}{c}$ kao konačni jednostavni verižni razlomak

$$\frac{a}{c} = [a_0, a_1, \ldots, a_{n-1}] = \frac{p_{n-1}}{q_{n-1}}.$$

Budući da su a i c relativno prosti i $c > 0$, imamo $a = p_{n-1}$, $c = q_{n-1}$. Izaberemo n tako da je

$$p_{n-1}q_{n-2} - p_{n-2}q_{n-1} = \det T = ad - bc = p_{n-1}d - bq_{n-1},$$

pa dobivamo

$$p_{n-1}(q_{n-2} - d) = q_{n-1}(p_{n-2} - b).$$

Kako je $(p_{n-1}, q_{n-1}) = 1$, slijedi $q_{n-1} \mid (q_{n-2} - d)$. No $0 \leq q_{n-2} \leq q_{n-1}$ i $0 < d < q_{n-1}$, pa je $|q_{n-2} - d| < q_{n-1}$ i stoga $q_{n-2} - d = 0$. To povlači $p_{n-2} - b = 0$ te smo dobili

$$\beta = \begin{bmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{bmatrix} \alpha.$$

Dakle, $\beta = [a_0, a_1, \ldots, a_{n-1}, \alpha]$, a budući je $\alpha > 1$, vidimo da je $\alpha = \beta_n$ i $\frac{b}{d} = \frac{p_{n-2}}{q_{n-2}}$, $\frac{a}{c} = \frac{p_{n-1}}{q_{n-1}}$ su susjedne konvergente od β. $\square$

Dokaz Teorema 1.1. Pretpostavimo da postoje cijeli brojevi k i l takvi da je $a_{k+n} = b_{l+n}$ za sve $n \geq 0$. Drugim riječima $\alpha = [a_0, a_1, \ldots, a_{k-1}, \alpha_k]$, $\beta = [b_0, b_1, \ldots, b_{l-1}, \beta_l]$ i $\alpha_k = \beta_l$. Prema (1.9) i (1.7) je $\alpha \sim \alpha_k$ i $\beta \sim \beta_l$, pa je i $\alpha \sim \beta$.

Pretpostavimo sada da su α i β ekvivalentni, tj. $\beta = T\alpha$, $T = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, $\det T = \pm 1$. Bez smanjenja općenitosti možemo pretpostaviti da je $c\alpha + d > 0$ jer inače zamijenimo T s $-T$. Neka su $\frac{p_n}{q_n}$ konvergente od α i definirajmo $T_{n-1} = \left[\begin{smallmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{smallmatrix}\right]$. Tada je $\alpha = T_{n-1}\alpha_n$, pa je $\beta = TT_{n-1}\alpha_n$. Vrijedi

$$TT_{n-1} = \begin{bmatrix} ap_{n-1} + bq_{n-1} & ap_{n-2} + bq_{n-2} \\ cp_{n-1} + dq_{n-1} & cp_{n-2} + dq_{n-2} \end{bmatrix} = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}.$$

Imamo

$$cp_{n-1} + dq_{n-1} = q_{n-1}\left(c\frac{p_{n-1}}{q_{n-1}} + d\right) = c',$$
$$cp_{n-2} + dq_{n-2} = q_{n-2}\left(c\frac{p_{n-2}}{q_{n-2}} + d\right) = d'.$$

Budući je $c\alpha + d > 0$ i $\lim_{n\to\infty} \frac{p_n}{q_n} = \alpha$, zaključujemo da su za dovoljno velike n brojevi c' i d' pozitivni. Nadalje je $a'd' - b'c' = \det T \cdot \det T_{n-1} = \pm 1$. Odaberimo parnost od n tako da je $c' > d'$, tj. za $c \geq 0$ biramo n paran, a za $c < 0$ biramo n neparan. Sada su zadovoljene sve pretpostavke Leme 1.2, pa je $\alpha_n = \beta_m$ za neki m, što je i trebalo dokazati. $\square$

Još jedan dokaz težeg dijela Serretovog teorema dat ćemo kasnije kad budemo proučavali opću linearnu grupu $\mathrm{GL}(2, \mathbb{Z})$.

1.2 Periodičnost

Beskonačni verižni razlomak u kojemu se neki blok parcijalnih kvocijenata ponavlja u nedogled zovemo *periodski* verižni razlomak i pišemo u obliku

$$\alpha = [a_0, a_1, \dots, a_{k-1}, \overline{a_k, a_{k+1}, \dots, a_{k+m-1}}], \tag{1.12}$$

što znači da je $a_{n+m} = a_n$ za sve $n \geq k$. Ako je $k = 0$, tj. nema pretperioda, govorimo o *čisto periodskom* verižnom razlomku $[\overline{a_0, \dots, a_{m-1}}]$.

Budući da su svi periodski verižni razlomci beskonačni, njihove vrijednosti su nužno iracionalni brojevi. Označimo za α iz (1.12) s $\alpha_n = [a_n, a_{n+1}, \dots]$ potpune kvocijente, s $\frac{p_n}{q_n}$ konvergente te s

$$P = \begin{bmatrix} a_k & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_{k+m-1} & 1 \\ 1 & 0 \end{bmatrix} \quad \text{i} \quad T = \begin{bmatrix} p_{k-1} & p_{k-2} \\ q_{k-1} & q_{k-2} \end{bmatrix}.$$

Imamo $\alpha = T\alpha_k$ i $\alpha_k = P\alpha_{k+m}$. Zbog periodičnosti je $\alpha_k = \alpha_{k+m}$ te dobivamo $\alpha_k = P\alpha_k$, a onda i $\alpha = TPT^{-1}\alpha$. Odavde slijedi da je α iracionalni korijen kvadratnog polinoma s cjelobrojnim koeficijentima, odnosno α je *kvadratna iracionalnost*. Svaki se realan broj α koji je kvadratna iracionalnost može zapisati u obliku $\alpha = \frac{s+\sqrt{d}}{t}$, gdje su s, d, t cijeli brojevi, $t \neq 0$ i $d > 0$ nije potpun kvadrat. Drugi korijen navedenog polinoma koji se poništava u α zovemo *konjugat* od α i označavamo α' te je očito $\alpha' = \frac{s-\sqrt{d}}{t}$. Operacija konjugiranja komutira sa zbrajanjem, oduzimanjem, množenjem i dijeljenjem racionalnih brojeva i kvadratnih iracionalnosti s istom vrijednošću broja d, tj. brojeva iz istog kvadratnog proširenja od $\mathbb{Q}$.

Teorem 1.3 (Euler, Lagrange)**.** *Razvoj u jednostavni verižni razlomak realnog broja α je periodski ako i samo ako je α kvadratna iracionalnost.*

Dokaz. Lakši smjer smo upravo dokazali.

Neka je $\alpha = \frac{a+\sqrt{b}}{c}$, $a, b, c \in \mathbb{Z}$, $c \neq 0$, $b > 0$, b nije potpuni kvadrat. Proširujući razlomak kojim je zadan α s $|c|$ dobivamo $\alpha_0 = \alpha = \frac{s_0+\sqrt{d}}{t_0}$, gdje su uvjeti na s_0, d, t_0 kao na a, b, c, ali je sada i sigurno ispunjeno $t_0 \mid (d - s_0^2)$. Razvijemo α u $[a_0, a_1, a_2, \dots]$ i imamo za potpune kvocijente α_i

$$a_i = \lfloor \alpha_i \rfloor, \quad \alpha_{i+1} = \frac{1}{\alpha_i - a_i} = \frac{1}{\frac{s_i+\sqrt{d}}{t_i} - a_i} = \frac{(a_i t_i - s_i) + \sqrt{d}}{\frac{d-(a_i t_i - s_i)^2}{t_i}},$$

pa je

$$\alpha_{i+1} = \frac{s_{i+1} + \sqrt{d}}{t_{i+1}}, \quad s_{i+1} = a_i t_i - s_i, \quad t_{i+1} = \frac{d - s_{i+1}^2}{t_i}$$

za $i \geq 0$. Brojevi s_{i+1} i t_{i+1} su cijeli jer t_i dijeli $d - s_i^2$, pa zato i $d - s_{i+1}^2$, a $t_{i+1} \neq 0$ jer d nije potpun kvadrat.

Konjugiranjem (1.9) dobivamo

$$\alpha_0' = \frac{\alpha_{n+1}' p_n + p_{n-1}}{\alpha_{n+1}' q_n + q_{n-1}},$$

odnosno

$$\alpha_{n+1}' = \frac{\alpha_0' q_{n-1} - p_{n-1}}{-\alpha_0' q_n + p_n} = -\frac{q_{n-1}}{q_n} \left(\frac{\alpha_0' - \frac{p_{n-1}}{q_{n-1}}}{\alpha_0' - \frac{p_n}{q_n}} \right).$$

Kad n teži u beskonačnost, $\frac{p_{n-1}}{q_{n-1}}$ i $\frac{p_n}{q_n}$ teže prema $\alpha_0 \neq \alpha_0'$, pa izraz u zagradi teži prema 1. Stoga za dovoljno velike n imamo $-1 < \alpha_{n+1}' < 0$, a jer je $\alpha_{n+1} > 1$ vrijede sljedeće tvrdnje za sve dovoljno velike n.

$$
\begin{array}{rclcl}
\alpha_n - \alpha_n' > 0 & \Rightarrow & \dfrac{2\sqrt{d}}{t_n} > 0 & \Rightarrow & t_n > 0, \\
\alpha_n + \alpha_n' > 0 & \Rightarrow & \dfrac{2s_n}{t_n} > 0 & \Rightarrow & s_n > 0, \\
\alpha_n' < 0 & \Rightarrow & s_n - \sqrt{d} < 0 & \Rightarrow & s_n < \sqrt{d}, \\
\alpha_n' > -1 & \Rightarrow & \dfrac{s_n - \sqrt{d}}{t_n} > -1 & \Rightarrow & \sqrt{d} - s_n < t_n, \\
\alpha_n > 1 & \Rightarrow & \dfrac{s_n + \sqrt{d}}{t_n} > 1 & \Rightarrow & t_n < \sqrt{d} + s_n.
\end{array}
$$

Stoga su s_n i t_n cijeli brojevi takvi da je

$$0 < s_n < \sqrt{d} \qquad \text{i} \qquad 0 < \sqrt{d} - s_n < t_n < \sqrt{d} + s_n.$$

Budući da par (s_n, t_n) može poprimiti samo konačno mnogo različitih vrijednosti, mora doći do ponavljanja, a to znači da se ponavljaju i potpuni kvocijenti od tog mjesta nadalje, što znači da je verižni razlomak periodski. □

Ključna ideja u dokazu prethodnog teorema je bila da za dovoljno velike n vrijedi $-1 < \alpha_n' < 0$. Za kvadratnu iracionalnost α kažemo da je *reducirana* ako je $\alpha > 1$ i $-1 < \alpha' < 0$.

Teorem 1.4 (Galois)**.** *Kvadratna iracionalnost α ima čisto periodski razvoj u jednostavni verižni razlomak ako i samo ako je α reducirana.*

Dokaz. Pretpostavimo da je razvoj od α čisto periodski, $\alpha = [\overline{a_0, a_1, \ldots, a_{m-1}}]$. Imamo $\alpha > a_0 = a_m \geq 1$. Također je

$$\alpha = [a_0, a_1, \ldots, a_{m-1}, \alpha] = \frac{\alpha p_{m-1} + p_{m-2}}{\alpha q_{m-1} + q_{m-2}},$$

pa je α korijen polinoma

$$f(x) = q_{m-1}x^2 + (q_{m-2} - p_{m-1})x - p_{m-2}.$$

Drugi korijen ovog polinoma je α' te je zbog $\alpha > 1$ dovoljno provjeriti da f ima nultočku između -1 i 0, a to sigurno vrijedi jer f ima različite predznake u tim točkama

$$
\begin{aligned}
f(0) &= -p_{m-2} < 0 \\
f(-1) &= q_{m-1} - q_{m-2} + p_{m-1} - p_{m-2} > 0.
\end{aligned}
$$

Obratno, pretpostavimo da je $\alpha = \alpha_0$ reducirana, tj. $\alpha_0 > 1$ i $-1 < \alpha_0' < 0$. Budući je $\alpha_0' = a_0 + \frac{1}{\alpha_1'}$, vrijedi $\frac{1}{\alpha_1'} = -a_0 + \alpha_0' < -a_0 \leq -1$ te zato $-1 < \alpha_1' < 0$. Indukcijom slijedi da je $-1 < \alpha_i' < 0$ za sve $i \geq 0$. Pretpostavimo da ne vrijedi tvrdnja teorema, tj. da α nije čisto periodski i oblika je (1.12) s $k \geq 1$. Tada a_{k-1} nije jednako a_{k+m-1} jer bi inače period počinjao jedno mjesto prije. No iz $\alpha_k = \alpha_{k+m}$ slijedi da je

$$\alpha_{k-1} - \alpha_{k+m-1} = \left(a_{k-1} + \frac{1}{\alpha_k}\right) - \left(a_{k+m-1} + \frac{1}{\alpha_{k+m}}\right) = a_{k-1} - a_{k+m-1}$$

cijeli broj različit od nule, pa je i $\alpha_{k-1}' - \alpha_{k+m-1}'$ takav, a to nije moguće jer se taj broj nalazi u intervalu $(-1, 1)$. □

Propozicija 1.5 (Galois)**.** *Neka je* $\alpha = [\overline{a_0, a_1, \dots, a_{m-1}}]$*. Tada je* $\frac{-1}{\alpha'} = [\overline{a_{m-1}, a_{m-2}, \dots, a_0}]$ *i zato* $-\alpha' = [0, \overline{a_{m-1}, a_{m-2}, \dots, a_0}]$.

Dokaz. Imamo

$$\alpha_0 = a_0 + \frac{1}{\alpha_1}, \quad \alpha_1 = a_1 + \frac{1}{\alpha_2}, \quad \dots \quad \alpha_k = a_k + \frac{1}{\alpha_{k+1}}, \quad \dots \quad \alpha_{m-1} = a_{m-1} + \frac{1}{\alpha_0},$$

pa konjugiranjem i rješavanjem u obrnutom poretku, dobivamo

$$\frac{-1}{\alpha'_0} = a_{m-1} + \frac{1}{\frac{-1}{\alpha'_{m-1}}}, \quad \dots \quad \frac{-1}{\alpha'_{k+1}} = a_k + \frac{1}{\frac{-1}{\alpha'_k}}, \quad \dots \quad \frac{-1}{\alpha'_1} = a_0 + \frac{1}{\frac{-1}{\alpha'_0}}.$$

Kako je svaki od $\frac{-1}{\alpha'_k}$ prema Teoremu 1.4 veći od 1, možemo ih promatrati kao potpune kvocijente broja

$$\frac{-1}{\alpha'} = \frac{-1}{\alpha'_0} = [\overline{a_{m-1}, a_{m-2}, \dots, a_0}].$$

□

Propozicija 1.6. *Verižni razlomci konjugiranih kvadratnih iracionalnosti imaju međusobno inverzne (obrnute) periode.*

Dokaz. Za kvadratnu iracionalnost $\alpha = [a_0, a_1, \dots, a_{k-1}, \overline{a_k, a_{k+1}, \dots, a_{k+m-1}}]$ imamo prema (1.9) $\alpha = \left[\begin{smallmatrix} p_{k-1} & p_{k-2} \\ q_{k-1} & q_{k-2} \end{smallmatrix}\right] \alpha_k$, gdje je $\alpha_k = [\overline{a_k, a_{k+1}, \dots, a_{k+m-1}}]$ čisto periodski verižni razlomak, pa prema prethodnoj propoziciji vrijedi $\frac{-1}{\alpha'_k} = [\overline{a_{k+m-1}, a_{k+m-2}, \dots, a_k}]$. Sada za konjugat od α dobivamo

$$\alpha' = \begin{bmatrix} p_{k-1} & p_{k-2} \\ q_{k-1} & q_{k-2} \end{bmatrix} \alpha'_k = \begin{bmatrix} -p_{k-2} & p_{k-1} \\ -q_{k-2} & q_{k-1} \end{bmatrix} \frac{-1}{\alpha'_k}$$

iz čega zaključujemo da su α' i $\frac{-1}{\alpha'_k}$ ekvivalentni brojevi, a to prema Serretovom teoremu 1.1 znači da se niz parcijalnih kvocijenata od nekog mjesta u α' i od nekog mjesta u α'_k podudaraju.

Primijetimo da smo ovdje radi sažetije izreke propozicije prešutno dopustili da period u razvoju α' u verižni razlomak počinje i poslije prvog mogućeg pojavljivanja, tj. nismo nužno uzeli najkraći mogući pretperiod. □

Propozicija 1.7 (Legendre)**.** *Za racionalan broj* $d > 1$ *takav da* $\sqrt{d}$ *nije racionalan, imamo* $\sqrt{d} = [a_0, \overline{a_1, a_2, \dots, a_2, a_1, 2a_0}]$.

Dokaz. Kvadratna iracionalnost $\sqrt{d} + \lfloor \sqrt{d} \rfloor$ je reducirana, pa ima čisto periodski verižni razlomak $\sqrt{d} + \lfloor \sqrt{d} \rfloor = [\overline{2a_0, a_1, a_2, \dots, a_{m-1}}]$, gdje je $a_0 = \lfloor \sqrt{d} \rfloor$. Prema propoziciji 1.5 je

$$[0, \overline{a_1, a_2, \dots, a_{m-1}, 2a_0}] = \sqrt{d} - \lfloor \sqrt{d} \rfloor = -(\sqrt{d} + \lfloor \sqrt{d} \rfloor)' = [0, \overline{a_{m-1}, a_{m-2}, \dots, a_1, 2a_0}]$$

iz čega usporedbom parcijalnih kvocijenata slijedi tvrdnja propozicije. □

1.3 Aproksimacijska svojstva

U sljedećim teoremima sabrat ćemo neka od aproksimacijskih svojstava verižnih razlomaka.

Teorem 1.8. *Neka je* $(\frac{p_n}{q_n})_n$ *niz konvergenti te* $(\alpha_n)_n$ *niz potpunih kvocijenata realnog broja* $\alpha = [a_0, a_1, a_2, \ldots]$.

(a) Fiksiramo li u $\alpha = [a_0, a_1, \ldots, a_{n-1}, \alpha_n]$ *brojeve* $a_0, a_1, \ldots, a_{n-1}$ *te promatramo* α *kao funkciju varijable* $\alpha_n \in (1, +\infty)$*, ta će funkcija biti rastuća za paran* n*, a padajuća za neparan* n*. Vrijednosti* α *leže u intervalu između* $\frac{p_{n-1}+p_{n-2}}{q_{n-1}+q_{n-2}}$ *i* $\frac{p_{n-1}}{q_{n-1}}$*. Duljina tog intervala je* $\frac{1}{q_{n-1}(q_{n-1}+q_{n-2})}$.

(b) Ako se $\beta = [a_0, a_1, \ldots, a_{n-1}, b_n, b_{n+1}, \ldots] = [a_0, a_1, \ldots, a_{n-1}, \beta_n]$ *podudara s* α *u parcijalnim kvocijentima* $a_0, a_1, \ldots, a_{n-1}$*, ali je* $a_n \neq b_n$*, onda za paran* n *imamo* $\alpha > \beta$ *ako i samo ako je* $a_n > b_n$ *dok za neparan* n *imamo* $\alpha > \beta$ *ako i samo ako je* $a_n < b_n$*. Udaljenost brojeva* α *i* β *je*

$$|\alpha - \beta| = \frac{|\alpha_n - \beta_n|}{q_{n-1}^2(\alpha_n + r)(\beta_n + r)} < \frac{1}{2^{n-2}}, \tag{1.13}$$

gdje je $r = [0, a_{n-1}, a_{n-2}, \ldots, a_1]$.

Dokaz. Za $n = 0$ sve tvrdnje koje imaju smisla, tj. rast funkcije u (a) dijelu i usporedivost brojeva α i β u (b) dijelu teorema, očito vrijede, pa u nastavku uzimamo da je $n \geq 1$.

(a) Iz (1.10) imamo

$$\alpha = \frac{p_{n-1}}{q_{n-1}} + \frac{(-1)^{n-1}}{q_{n-1}(\alpha_n q_{n-1} + q_{n-2})} = \frac{p_{n-1} + \frac{p_{n-2}}{\alpha_n}}{q_{n-1} + \frac{q_{n-2}}{\alpha_n}},$$

pa je uz fiksirane p_{n-1}, q_{n-1}, q_{n-2} funkcija $\alpha_n \mapsto \alpha$ na intervalu $\alpha_n \in (1, +\infty)$ strogo rastuća za n paran, a strogo padajuća za n neparan. Zbog monotonosti ove neprekidne funkcije zaključujemo da α poprima sve vrijednosti između $[a_0, a_1, \ldots, a_{n-1}, 1] = \frac{p_{n-1}+p_{n-2}}{q_{n-1}+q_{n-2}}$ i $[a_0, a_1, \ldots, a_{n-1}, +\infty] = [a_0, a_1, \ldots, a_{n-1}] = \frac{p_{n-1}}{q_{n-1}}$. Duljina tog intervala je $\frac{1}{q_{n-1}(q_{n-1}+q_{n-2})}$.

(b) Direktno iz (1.10) dobivamo

$$\begin{aligned}|\alpha - \beta| &= \left|\left(\alpha - \frac{p_{n-1}}{q_{n-1}}\right) - \left(\beta - \frac{p_{n-1}}{q_{n-1}}\right)\right| \\ &= \left|\frac{(-1)^{n-1}}{q_{n-1}^2(\alpha_n + r)} - \frac{(-1)^{n-1}}{q_{n-1}^2(\beta_n + r)}\right| = \left|\frac{\alpha_n - \beta_n}{q_{n-1}^2(\alpha_n + r)(\beta_n + r)}\right|.\end{aligned}$$

Budući je $\alpha_n > 1$, $\beta_n > 1$ i $r > 0$, vrijedi

$$\left|\frac{1}{\alpha_n + r} - \frac{1}{\beta_n + r}\right| < 1,$$

pa prema (1.5) imamo

$$|\alpha - \beta| < \frac{1}{q_{n-1}^2} \leq \frac{1}{2^{2\lfloor\frac{n-1}{2}\rfloor}} \leq \frac{1}{2^{n-2}}.$$

□

Teorem 1.9. *(a) Za racionalan broj* α *postoji eksplicitna konstanta* $c(\alpha)$ *takva da je*

$$\left|\alpha - \frac{p}{q}\right| > \frac{1}{c(\alpha) \cdot q}$$

za sve racionalne brojeve $\frac{p}{q} \neq \alpha$.

U nastavku uzimamo da je $\alpha = [a_0, a_1, \ldots]$ *iracionalan broj s nizom konvergenti* $(\frac{p_n}{q_n})$.

(b) *Vrijedi*

$$\frac{1}{(a_{n+1}+2)q_n^2} \leq \frac{1}{q_n(q_n+q_{n+1})} < \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}} \leq \frac{1}{a_{n+1}q_n^2} \leq \frac{1}{q_n^2}, \tag{1.14}$$

pa postoji beskonačno mnogo racionalnih brojeva $\frac{p}{q}$ *takvih da je*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}.$$

(c) *(Fatou) Ako je* $|\alpha - \frac{p}{q}| < \frac{1}{q^2}$, *onda postoji* n *takav da je*

$$\frac{p}{q} \in \left\{\frac{p_n}{q_n}, \frac{p_n - p_{n-1}}{q_n - q_{n-1}}, \frac{p_n + p_{n-1}}{q_n + q_{n-1}}\right\}.$$

(d) *(Vahlen) Barem jedna od dvije uzastopne konvergente od* α *zadovoljava nejednakost*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{2q^2}.$$

(e) *(Legendre) Neka su* p *i* q *cijeli brojevi takvi da je* $q \geq 1$ *i*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{2q^2}.$$

Tada je $\frac{p}{q}$ *neka konvergenta od* α.

(f) *(Borel) Barem jedna od tri uzastopne konvergente od* α *zadovoljava nejednakost*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}.$$

(Hurwitz) Postoji beskonačno mnogo racionalnih brojeva $\frac{p}{q}$ *takvih da je*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}.$$

Dokaz. (a) Ako je $\alpha = \frac{m}{n}$, gdje su m i n cijeli brojevi i $\frac{p}{q} \neq \alpha$, onda je

$$\left|\alpha - \frac{p}{q}\right| = \left|\frac{mq - np}{nq}\right| \geq \frac{1}{|n|q}.$$

(b) Koristeći (1.10) i činjenicu da je $a_{n+1} < \alpha_{n+1} < a_{n+1} + 1$, dobivamo

$$\frac{1}{(a_{n+1}+2)q_n^2} \leq \frac{1}{q_n(q_n+q_{n+1})} = \frac{1}{q_n\big((a_{n+1}+1)q_n + q_{n-1}\big)} < \left|\alpha - \frac{p_n}{q_n}\right| \quad \text{i}$$

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n(a_{n+1}q_n + q_{n-1})} = \frac{1}{q_n q_{n+1}} \leq \frac{1}{a_{n+1}q_n^2} \leq \frac{1}{q_n^2}.$$

(c) Ova tvrdnja, koju u nastavku nećemo trebati, slijedi uvrštavanjem $c = 1$ u općenitiji rezultat Worleyja i Dujelle kojeg navodimo u nastavku.

Neka je α proizvoljan realan broj te c pozitivan realan broj. Ako racionalan broj $\frac{p}{q}$ zadovoljava $|\alpha - \frac{p}{q}| < \frac{c}{q^2}$, onda je

$$\frac{p}{q} = \frac{rp_n \pm sp_{n-1}}{rq_n \pm sq_{n-1}}$$

za neki $n \geq 0$ i nenegativne cijele brojeve r, s takve da je $rs < 2c$ te neki izbor predznaka.

(e) I ovaj rezultat slijedi iz upravo spomenutog teorema Worleyja i Dujelle, ali ćemo ga mi dokazati drugačije.

Neka su $q > 0$ i p cijeli brojevi takvi da je $|\alpha - \frac{p}{q}| < \frac{1}{2q^2}$, odnosno $-\frac{1}{2} < q(q\alpha - p) < \frac{1}{2}$. Ako je $q = 1$, vidimo da je $\frac{p}{q} = \lfloor \alpha \rfloor$ kad je razlomljeni dio $\{\alpha\} = \alpha - \lfloor \alpha \rfloor < \frac{1}{2}$, odnosno $\frac{p}{q} = \lfloor \alpha \rfloor + 1$ za $\{\alpha\} > \frac{1}{2}$, a to su uistinu konvergente od α. Za $q > 1$, zbog $(p, q) = 1$, što uvijek možemo pretpostaviti, postoje cijeli brojevi b i d takvi da je $dp - bq = \pm 1$, $q > d > 0$ i predznak od $dp - bq$ je isti kao predznak od $-q(q\alpha - p)$. Naime, ako nam predznak ne odgovara, zamijenimo b, d sa $p - b$, $q - d$. Neka je sada $T = \left[\begin{smallmatrix} p & b \\ q & d \end{smallmatrix}\right]$ i ω zadan sa $\alpha = T\omega$. Imamo

$$\begin{aligned} \omega = T^{-1}\alpha &= \begin{bmatrix} d & -b \\ -q & p \end{bmatrix} \alpha = \frac{dq\alpha - bq}{-q^2\alpha + pq} = \frac{dq\alpha + \det T - dp}{-q(q\alpha - p)} \\ &= \frac{\det T}{-q(q\alpha - p)} - \frac{d}{q} > 2 - 1 = 1. \end{aligned}$$

Ispunjeni su uvjeti Leme 1.2, pa zaključujemo da je $\frac{p}{q}$ konvergenta od α.

(d) Znamo iz (1.10) da je

$$\left|\alpha - \frac{p_n}{q_n}\right| = \frac{1}{q_n^2(\alpha_{n+1} + \beta_{n+1})},$$

gdje je $\alpha_{n+1} = [a_{n+1}, a_{n+2}, \ldots]$ i $\beta_{n+1} = [0, a_n, a_{n-1}, \ldots, a_1]$. Stoga je dovoljno pokazati da ne postoji prirodan broj n takav da je $\alpha_n + \beta_n \leq 2$ i $\alpha_{n+1} + \beta_{n+1} \leq 2$. Pretpostavimo suprotno, tj. da takav n postoji. Primijetimo da se jednakosti ne mogu postizati jer su $\alpha_i + \beta_i$ iracionalni brojevi kao suma iracionalnog α_i i racionalnog β_i. Iz prve nejednakosti je sada $a_n + \frac{1}{\alpha_{n+1}} + \beta_n < 2$, pa je $\frac{1}{\alpha_{n+1}} < 2 - (a_n + \beta_n) = 2 - \frac{1}{\beta_{n+1}}$. Iz druge pak nejednakosti imamo $\alpha_{n+1} < 2 - \beta_{n+1}$ te množenjem s dobivenom nejednakošću slijedi

$$1 < \left(2 - \frac{1}{\beta_{n+1}}\right)(2 - \beta_{n+1}),$$

odnosno $0 < 2 - \beta_{n+1} - \frac{1}{\beta_{n+1}}$ ili $(\beta_{n+1} - 1)^2 < 0$ što nije moguće.

(f) Druga tvrdnja slijedi direktno iz prve. Slično kao u prethodnom dokazu, pretpostavimo da prva tvrdnja ne vrijedi, tj. za neki n imamo $\alpha_i + \beta_i \leq \sqrt{5}$ za $i \in \{n, n+1, n+2\}$. Iz prve dvije nejednakosti dobijemo kao prije

$$1 \leq \left(\sqrt{5} - \frac{1}{\beta_{n+1}}\right)\left(\sqrt{5} - \beta_{n+1}\right),$$

odnosno $\beta_{n+1}^2 - \sqrt{5}\beta_{n+1} + 1 \le 0$. Ponovivši isto zaključivanje s drugom i trećom nejednakošću, dobivamo $\beta_{n+2}^2 - \sqrt{5}\beta_{n+2} + 1 \le 0$. Stoga su β_{n+1} i β_{n+2} sadržani u intervalu $\left[\frac{\sqrt{5}-1}{2}, \frac{\sqrt{5}+1}{2}\right]$, odnosno, budući da su β_i uvijek racionalni, u intervalu $\left(\frac{\sqrt{5}-1}{2}, \frac{\sqrt{5}+1}{2}\right)$. No, zbog $a_{n+1} \ge 1$ imamo

$$\beta_{n+2} = \frac{1}{a_{n+1} + \beta_{n+1}} < \frac{1}{1 + \frac{\sqrt{5}-1}{2}} = \frac{\sqrt{5}-1}{2}$$

čime smo došli do kontradikcije. $\square$

Teorem 1.10. *Neka je α iracionalan broj s nizom konvergenti $(\frac{p_n}{q_n})$. Aproksimacije konvergentama su sve bolje, tj.*

$$\left|\alpha - \frac{p_n}{q_n}\right| < \left|\alpha - \frac{p_{n-1}}{q_{n-1}}\right| \quad i \quad |q_n\alpha - p_n| < |q_{n-1}\alpha - p_{n-1}| \qquad za\ sve\ n \ge 1.$$

Neka je $n \ge 1$ i p, q su cijeli brojevi takvi da je $q > 0$.

Ako je $\left|\alpha - \frac{p}{q}\right| < \left|\alpha - \frac{p_n}{q_n}\right|$, onda je $q > q_n$.

Ako je $|q\alpha - p| < |q_n\alpha - p_n|$, onda je $q \ge q_{n+1}$.

Dokaz. Iz Teorema 1.9.b imamo

$$|q_n\alpha - p_n| < \frac{1}{q_{n+1}} \le \frac{1}{q_{n-1} + q_n} < |q_{n-1}\alpha - p_{n-1}|,$$

pa je $|q_n\alpha - p_n| < |q_{n-1}\alpha - p_{n-1}|$, a množenjem ove nejednakosti s $\frac{1}{q_n} \le \frac{1}{q_{n-1}}$ proizlazi $|\alpha - \frac{p_n}{q_n}| < |\alpha - \frac{p_{n-1}}{q_{n-1}}|$ za sve $n \ge 1$.

Dovoljno je još dokazati da iz $|q\alpha - p| < |q_n\alpha - p_n|$ slijedi $q \ge q_{n+1}$. Naime, pretpostavimo li da ta tvrdnja vrijedi, iz $|\alpha - \frac{p}{q}| < |\alpha - \frac{p_n}{q_n}|$ nužno mora slijediti da je $q > q_n$ jer bi u suprotnom iz $|\alpha - \frac{p}{q}| < |\alpha - \frac{p_n}{q_n}|$ i $q \le q_n$ množenjem dobili $|q\alpha - p| < |q_n\alpha - p_n|$ za neki $q < q_{n+1}$.

Neka je, dakle, $|q\alpha - p| < |q_n\alpha - p_n|$ i pretpostavimo da je $q < q_{n+1}$. Možemo uzeti da je $\frac{p}{q}$ skraćeni razlomak. Za sustav jednadžbi

$$\begin{aligned} p_n x + p_{n+1} y &= p \\ q_n x + q_{n+1} y &= q \end{aligned}$$

je matrica sustava unimodularna, tj. determinante ± 1, pa postoji jedinstveno rješenje koje je nužno cjelobrojno $x = r$, $y = s$. Ako je $r = 0$, onda je $s = \frac{q}{q_{n+1}} > 0$ i stoga $s \ge 1$ te zato $q \ge q_{n+1}$ što ne može biti. Ako je $s = 0$, onda je $r = \frac{p}{p_n} = \frac{q}{q_n}$ te zbog $(p_n, q_n) = (p, q) = 1$ imamo $p = p_n$, $q = q_n$ što je u kontradikciji s $|q\alpha - p| < |q_n\alpha - p_n|$.

Pokazali smo da je $rs \ne 0$ i tvrdimo da su r i s različitih predznaka. Za $s < 0$ je $q_n r = q - q_{n+1}s > 0$, pa je $r > 0$. Ako je $s \ge 1$, onda je $q_n r = q - q_{n+1}s \le q - q_{n+1} < 0$, pa je $r < 0$.

Kako su $\alpha - \frac{p_n}{q_n}$ i $\alpha - \frac{p_{n+1}}{q_{n+1}}$ različitih predznaka, to su i $q_n\alpha - p_n$ i $q_{n+1}\alpha - p_{n+1}$ različitih predznaka, a onda $r(q_n\alpha - p_n)$ i $s(q_{n+1}\alpha - p_{n+1})$ imaju isti predznak iz čega zbog

$$q\alpha - p = (q_n r + q_{n+1}s)\alpha - (p_n r + p_{n+1}s) = r(q_n\alpha - p_n) + s(q_{n+1}\alpha - p_{n+1})$$

zaključujemo

$$|q\alpha - p| = |r| \cdot |q_n\alpha - p_n| + |s| \cdot |q_{n+1}\alpha - p_{n+1}| > |q_n\alpha - p_n|,$$

što je u kontradikciji s početnom pretpostavkom, pa smo završili dokaz. $\square$

Primijetimo da nam Teorem 1.10 omogućuje drugi dokaz Teorema 1.9.e. Pretpostavimo da je $\frac{p}{q}$ takav da vrijedi $|\alpha - \frac{p}{q}| < \frac{1}{2q^2}$, ali $\frac{p}{q}$ nije konvergenta od α. Postoji $n \geq 1$ takav da je $q_n \leq q < q_{n+1}$. Tada je prema prethodnom teoremu $|q_n\alpha - p_n| \leq |q\alpha - p| < \frac{1}{2q}$ iz čega je $|\alpha - \frac{p_n}{q_n}| < \frac{1}{2qq_n}$. Zbog $\frac{p}{q} \neq \frac{p_n}{q_n}$ je $|pq_n - qp_n| \geq 1$, pa imamo

$$\frac{1}{qq_n} \leq \frac{|pq_n - qp_n|}{qq_n} = \left|\frac{p}{q} - \frac{p_n}{q_n}\right| \leq \left|\alpha - \frac{p}{q}\right| + \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{2q^2} + \frac{1}{2qq_n}.$$

To znači da je $\frac{1}{qq_n} < \frac{1}{q^2}$, odnosno $q < q_n$ te smo dobili kontradikciju i Teorem 1.9.e je dokazan.

Prisjetimo se da za svaki algebarski broj α, tj. korijen nekog polinoma s racionalnim koeficijentima, postoji polinom (različit od nulpolinoma) minimalnog stupnja s cjelobrojnim koeficijentima koji se poništava u α. Ako još tražimo da je najveći zajednički djelitelj koeficijenata tog polinoma jednak 1 i da je vodeći koeficijent pozitivan, takav polinom je jedinstven i zovemo ga *minimalni polinom* od α nad $\mathbb{Z}$. Minimalni polinom je ireducibilan nad $\mathbb{Q}$, a njegov stupanj nazivamo *stupnjem* algebarskog broja α.

Teorem 1.11 (Liouville). *Neka je α realan algebarski broj stupnja $d \geq 1$. Tada postoji konstanta $c(\alpha) > 0$ tako da vrijedi*

$$\left|\alpha - \frac{p}{q}\right| > \frac{1}{c(\alpha) \cdot q^d} \tag{1.15}$$

za sve racionalne brojeve $\frac{p}{q}$, gdje je $q > 0$ i $\frac{p}{q} \neq \alpha$.

Za $c(\alpha)$ možemo uzeti

$$1 + \max_{|t-\alpha|\leq 1} |P'(t)| \qquad \textit{ili} \qquad 1 + \sum_{i=1}^{d} \frac{1}{i!}|P^{(i)}(\alpha)|,$$

gdje je $P(x)$ minimalni polinom od α nad $\mathbb{Z}$.

Dokaz. Neka je $\frac{p}{q}$ racionalan broj takav da je $|\alpha - \frac{p}{q}| \leq 1$. Budući je $P(x)$ ireducibilan i ima cjelobrojne koeficijente, imamo $P(\frac{p}{q}) \neq 0$, te zato $|q^d P(\frac{p}{q})| \geq 1$. Iz Lagrangeovog teorema srednje vrijednosti slijedi da postoji t između α i $\frac{p}{q}$ tako da je

$$\left|P\left(\frac{p}{q}\right)\right| = \left|P(\alpha) - P\left(\frac{p}{q}\right)\right| = |P'(t)| \cdot \left|\alpha - \frac{p}{q}\right|.$$

Iz razvoja $P(x)$ u Taylorov red oko α je

$$\left|P\left(\frac{p}{q}\right)\right| = \left|\sum_{i=1}^{d} \frac{P^{(i)}(\alpha)}{i!}\left(\frac{p}{q} - \alpha\right)^i\right| \leq \left|\alpha - \frac{p}{q}\right| \cdot \sum_{i=1}^{d} \frac{|P^{(i)}(\alpha)|}{i!}.$$

Kombinirajući prethodne činjenice dobivamo tvrdnju teorema. □

Thue, Siegel, Dyson i Geljfond su smanjivali eksponent d u (1.15) da bi konačno Roth 1955. pokazao kako za svaki $\varepsilon > 0$ nejednakost (1.15) možemo zamijeniti s

$$\left|\alpha - \frac{p}{q}\right| > \frac{1}{c(\alpha, \varepsilon) \cdot q^{2+\varepsilon}}.$$

Nažalost, za razliku od Liouvilleovog rezultata, sva navedena poboljšanja bila su neefektivna, tj. ne možemo dati nikakvu gornju ogradu za konstantu u nazivniku razlomka s desne strane nejednakosti. Prema trenutno poznatim rezultatima, za općeniti α je uz efektivnu konstantu, eksponent d moguće smanjiti samo vrlo malo, na $d - \tau(\alpha)$, gdje je

$$\tau(\alpha) < \frac{1}{10^{26d-3}d^{14d}}.$$

1.4 Lagrangeov spektar

Neka je $\varepsilon > 0$ proizvoljan. Nije teško pokazati da, s obzirom na Lebesgueovu mjeru, za skoro sve realne brojeve α, nejednadžba

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\varepsilon}}$$

ima najviše konačno rješenja u racionalnim brojevima $\frac{p}{q}$. Rothov teorem stoga kaže da se algebarski brojevi, kojih je prebrojivo mnogo, ponašaju kao skoro svi realni brojevi.

Vidjeli smo da

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}$$

ima beskonačno mnogo racionalnih rješenja $\frac{p}{q}$ za sve iracionalne brojeve α. Štoviše, prema Teoremu 1.9.f ta će tvrdnja vrijediti i ako podijelimo desnu stranu te nejednadžbe s $\sqrt{5}$. Ako, dakle, općenito ne možemo povećati eksponent 2 u q^2, od interesa je pitanje možemo li povećati konstantu $\sqrt{5}$ s kojom množimo q^2.

Definicija 1.1. Neka je α iracionalan realan broj. *Lagrangeov broj* od α definiramo kao

$$L(\alpha) = \sup\{L \in \mathbb{R} \,:\, \left|\alpha - \frac{p}{q}\right| < \frac{1}{L \cdot q^2} \text{ za beskonačno mnogo } \frac{p}{q} \in \mathbb{Q}\}.$$

Teorem 1.9.f kaže da je uvijek $L(\alpha) \geq \sqrt{5}$.

Lako se dobiva ekvivalentna definicija Lagrangeovog broja.

Definicija 1.2. Za iracionalan broj α, Lagrangeov broj $L(\alpha)$ definiramo s

$$L(\alpha) = \limsup_{q \to +\infty} \frac{1}{q\|q\alpha\|} = \frac{1}{\liminf_{q \to +\infty} q\|q\alpha\|},$$

gdje smo sa $\|x\|$ označili udaljenost realnog broja x od najbližeg cijelog broja, a za niz realnih brojeva $(x_n)_{n\geq 1}$ je

$$\limsup_{n \to +\infty} x_n = \inf_{k\geq 1} \sup_{n \geq k} x_n$$

najveće gomilište i

$$\liminf_{n \to +\infty} x_n = \sup_{k\geq 1} \inf_{n \geq k} x_n$$

je najmanje gomilište.

Definicija 1.3. *Lagrangeov spektar* je

$$\mathcal{L} = \{L(\alpha) \,:\, \alpha \in \mathbb{R} \setminus \mathbb{Q}\}.$$

Napomenimo da se ponegdje, posebno u starijoj literaturi, u definiciji Lagrangeovog spektra gleda $\frac{1}{L(\alpha)}$ umjesto $L(\alpha)$. Mi ćemo uvijek koristiti gornju definiciju.

Propozicija 1.12. *Označimo za iracionalan broj* $\alpha = [a_0, a_1, a_2, \dots]$ *s*

$$\lambda_n(\alpha) = [a_n, a_{n+1}, \dots] + [0, a_{n-1}, a_{n-2}, \dots, a_1], \quad n \geq 2.$$

Tada je $L(\alpha) = \limsup_{n\to\infty} \lambda_n(\alpha)$.

Dokaz. Iz teorema 1.9.e je jasno da je kod određivanja $L(\alpha)$ dovoljno promatrati konvergente $\frac{p_n}{q_n}$ od α. U (1.10) smo pokazali da je

$$\alpha - \frac{p_n}{q_n} = \frac{(-1)^n}{q_n^2(\alpha_{n+1} + \beta_{n+1})},$$

gdje je $\alpha_{n+1} = [a_{n+1}, a_{n+2}, \dots]$, $\beta_{n+1} = \frac{q_{n-1}}{q_n} = [0, a_n, a_{n-1}, \dots, a_1]$. Sada je

$$q_n|q_n\alpha - p_n| = \frac{1}{\alpha_{n+1} + \beta_{n+1}} = \frac{1}{\lambda_{n+1}(\alpha)}, \tag{1.16}$$

pa tvrdnja slijedi direktno iz maloprije dane ekvivalentne definicije $L(\alpha)$. □

Propozicija 1.13. *Ekvivalentni brojevi* α *i* $\widehat{\alpha}$ *imaju isti Lagrangeov broj,* $L(\alpha) = L(\widehat{\alpha})$.

Dokaz. Neka je $\alpha \sim \widehat{\alpha}$. Tada postoje k i l takvi da je

$$\alpha = [a_0, a_1, \dots, a_k, \gamma], \qquad \widehat{\alpha} = [\widehat{a}_0, \widehat{a}_1, \dots, \widehat{a}_l, \gamma],$$

gdje je $\gamma = [c_1, c_2, c_3, \dots]$. Za $n \geq 2$ je

$$\begin{aligned} \lambda_{k+n}(\alpha) &= [c_n, c_{n+1}, \dots] + [0, c_{n-1}, \dots, c_1, a_k, \dots, a_1], \\ \lambda_{l+n}(\widehat{\alpha}) &= [c_n, c_{n+1}, \dots] + [0, c_{n-1}, \dots, c_1, \widehat{a}_l, \dots, \widehat{a}_1]. \end{aligned}$$

Prema Teoremu 1.8.b je

$$|\lambda_{k+n}(\alpha) - \lambda_{l+n}(\widehat{\alpha})| \leq \frac{1}{2^{n-2}}.$$

Neka je $(n_i)_i$ rastući niz prirodnih brojeva takav da $(\lambda_{k+n_i}(\alpha))_i$ konvergira prema $L(\alpha) = \limsup_{n\to+\infty} \lambda_{k+n}(\alpha)$. Tada je očito da niz $(\lambda_{l+n_i}(\widehat{\alpha}))_i$ također konvergira prema $L(\alpha)$, pa je $L(\widehat{\alpha}) \geq L(\alpha)$. Potpuno analogno pokažemo da je $L(\alpha) \geq L(\widehat{\alpha})$, pa slijedi tvrdnja propozicije. □

Pokazat ćemo da je na pitanje vrijedi li obrat, tj. povlači li $L(\alpha) = L(\widehat{\alpha})$ nužno da je $\alpha \sim \widehat{\alpha}$, općenito odgovor ne. No, ograničimo li se na $L(\alpha) \in \mathcal{L}_{<3} = \mathcal{L} \cap (0, 3)$, poznata neriješena slutnja o kojoj ćemo kasnije puno više reći, davala bi potvrdan odgovor.

Korolar 1.14. *Neka je* α *iracionalan broj.*

(a) *Ako je* $L(\alpha) < 3$*, onda je* α *ekvivalentan nekom* $\widehat{\alpha} = [\widehat{a}_0, \widehat{a}_1, \widehat{a}_2, \dots]$*, gdje su svi* $\widehat{a}_i \in \{1, 2\}$.

(b) *Za* $\alpha \sim \frac{1+\sqrt{5}}{2}$ *je* $L(\alpha) = \sqrt{5}$. *Za* $\alpha \sim 1 + \sqrt{2} \sim \sqrt{2}$ *je* $L(\alpha) = \sqrt{8}$.

(c) *Ako je* $L(\alpha) < 3$ *i* α *nije ekvivalentan ni* $\frac{1+\sqrt{5}}{2}$ *ni* $1 + \sqrt{2}$*, onda* α *ima beskonačno mnogo jedinica i beskonačno mnogo dvojki u razvoju u verižni razlomak.*

Dokaz. (a) Neka je $\alpha = [a_0, a_1, a_2, \ldots]$. Ako je $a_i \geq 3$, onda je $\lambda_i(\alpha) = [a_i, a_{i+1}, \ldots] + [0, a_{i-1}, \ldots, a_1] > 3$, pa zbog $L(\alpha) = \limsup_n \lambda_n(\alpha) < 3$, može postojati samo konačno mnogo takvih a_i. Zamijenimo li a_0 i te a_i s 1 dobivamo broj $\widehat{\alpha}$ ekvivalentan broju α kojemu su svi parcijalni kvocijenti 1 i 2.

(b) Budući je

$$\frac{1+\sqrt{5}}{2} = 1 + \frac{1}{\frac{1+\sqrt{5}}{2}} = [1, 1, 1, \ldots] = [\overline{1}] \qquad \text{i}$$

$$1+\sqrt{2} = 2 + \frac{1}{1+\sqrt{2}} = [2, 2, 2, \ldots] = [\overline{2}],$$

imamo

$$\lambda_n\left(\frac{1+\sqrt{5}}{2}\right) = [1, 1, 1, \ldots] + \frac{1}{\underbrace{[1, 1, \ldots, 1]}_{n-1}},$$

$$\lambda_n\left(1+\sqrt{2}\right) = [2, 2, 2, \ldots] + \frac{1}{\underbrace{[2, 2, \ldots, 2]}_{n-1}},$$

pa je

$$L\left(\frac{1+\sqrt{5}}{2}\right) = \limsup_{n\to+\infty} \lambda_n\left(\frac{1+\sqrt{5}}{2}\right) = \frac{1+\sqrt{5}}{2} + \left(\frac{1+\sqrt{5}}{2}\right)^{-1} = \sqrt{5},$$

$$L\left(1+\sqrt{2}\right) = \limsup_{n\to+\infty} \lambda_n\left(1+\sqrt{2}\right) = 1+\sqrt{2} + \left(1+\sqrt{2}\right)^{-1} = 2\sqrt{2} = \sqrt{8}.$$

(c) Iz (a) dijela ovog korolara znamo da je α za kojeg je $L(\alpha) < 3$ ekvivalentan $\widehat{\alpha}$ koji ima samo 1 i 2 u razvoju u verižni razlomak. Budući je $\widehat{\alpha} \sim \alpha \not\sim \frac{1+\sqrt{5}}{2}$, vidimo da $\widehat{\alpha}$ ne može sadržavati samo konačno mnogo parcijalnih kvocijenata jednakih 2, a zbog $\widehat{\alpha} \sim \alpha \not\sim 1+\sqrt{2}$ zaključujemo da $\widehat{\alpha}$ ne može sadržavati samo konačno mnogo parcijalnih kvocijenata jednakih 1. □

Iz Korolara 1.14.b vidimo da je $\min \mathcal{L} = \sqrt{5}$ i da je rezultat u Teoremu 1.9.f općenito najbolji mogući.

Propozicija 1.15. *Neka je α kvadratna iracionalnost s periodom $a_0, a_1, \ldots, a_{m-1}$ u razvoju α u verižni razlomak. Označimo s $\rho_i = [\overline{a_i, a_{i+1}, \ldots, a_{i-1}}]$ za $0 \leq i \leq m-1$. Tada je*

$$L(\alpha) = \max_{0\leq i\leq m-1} (\rho_i - \rho_i').$$

Dokaz. Zbog Teorema 1.1 i Propozicije 1.13 možemo bez smanjenja općenitosti uzeti da je $\alpha = [\overline{a_0, a_1, \ldots, a_{m-1}}]$. Sada je

$$L(\alpha) = \max_{0\leq i\leq m-1} \limsup_{k\to+\infty} \lambda_{i+km}(\alpha),$$

gdje je $\lambda_{i+km}(\alpha)$ jednak sumi $[\overline{a_i, a_{i+1}, \ldots, a_{i-1}}] = \rho_i$ i drugog pribrojnika koji se podudara s $[0, \overline{a_{i-1}, a_{i-2}, \ldots, a_0, a_{m-1}, \ldots, a_i}]$ u prvih $i+km$ parcijalnih kvocijenata. Dakle, $\lambda_{i+km}(\alpha)$ uz fiksirani $i \in \{0, 1, \ldots, m-1\}$ teži prema $\rho_i - \rho_i'$ kada k teži u beskonačnost zbog Teorema 1.8 i Propozicije 1.5. Time je dokazana propozicija. □

Iz prethodne propozicije vidimo da je za određivanje Lagrangeovog broja kvadratne iracionalnosti dovoljno usporediti konačno mnogo brojeva te da je Lagrangeov broj kvadratne iracionalnosti uvijek kvadratna iracionalnost.

Primjer. Promotrimo $\alpha = [\overline{1\,2\,1}]$. Ne moramo pisati zareze između parcijalnih kvocijenata ako su svi parcijalni kvocijenti jednoznamenkasti, a vidjet ćemo da će nam samo takvi verižni razlomci biti od interesa. Rješavajući kvadratnu jednadžbu dobivamo

$$\alpha = \rho_0 = \frac{1+\sqrt{10}}{3}, \quad \rho_1 = [\overline{2\,1\,1}] = \frac{2+\sqrt{10}}{2}, \quad \rho_2 = [\overline{1\,1\,2}] = \frac{2+\sqrt{10}}{3}$$

i zato je

$$L(\alpha) = \max_i(\rho_i - \rho_i') = \max\left\{\frac{2\sqrt{10}}{3}, \sqrt{10}, \frac{2\sqrt{10}}{3}\right\} = \sqrt{10}.$$

Ovaj nam primjer pokazuje i da činjenica kako su svi parcijalni kvocijenti nekog broja 1 ili 2 ne jamči da je njegov Lagrangeov broj manji od 3.

Pokazat ćemo kasnije da se $\mathcal{L}_{<3}$, Lagrangeov spektar ispod 3, sastoji od niza brojeva koji konvergiraju prema 3 i svaki od njih je Lagrangeov broj za prebrojivo mnogo realnih brojeva, točnije, postiže se na skupu nekih kvadratnih iracionalnosti. Za $L(\alpha) = 3$ je situacija bitno drugačija kako pokazuje sljedeći Markovljev rezultat.

Teorem 1.16. *Postoji neprebrojivo mnogo realnih brojeva α takvih da je $L(\alpha) = 3$, te da nikoja dva među njima nisu ekvivalentna.*

Dokaz. Neka je $(r_n)_{n\geq 1}$ strogo rastući niz prirodnih brojeva i neka je

$$\alpha = [\underbrace{1\,1\,\ldots\,1}_{r_1}\,2\,2\,\underbrace{1\,1\,\ldots\,1}_{r_2}\,2\,2\,\underbrace{1\,1\,\ldots\,1}_{r_3}\,2\,2\,\ldots]. \tag{1.17}$$

Ako odaberemo $n \geq 2$ tako da je $a_n = 1$, onda je $\lambda_n(\alpha) = \alpha_n + \beta_n < 2 + 1 = 3$. Ako n prolazi nizom indeksa takvih da je $a_n = a_{n+1} = 2$, onda zbog $\lim_i r_i = +\infty$, imamo

$$\lim_n \lambda_n(\alpha) = [2\,2\,1\,1\,\ldots] + [0\,1\,1\,\ldots] = 2 + \frac{1}{2+\frac{\sqrt{5}-1}{2}} + \frac{\sqrt{5}-1}{2} = 3.$$

Konačno, ako n prolazi skupom indeksa za koje je $a_{n-1} = a_n = 2$, onda je

$$\lim_n \lambda_n(\alpha) = [2\,1\,1\,\ldots] + [0\,2\,1\,1\,\ldots] = 2 + \frac{\sqrt{5}-1}{2} + \frac{1}{2+\frac{\sqrt{5}-1}{2}} = 3.$$

Dakle, $L(\alpha) = \limsup_n \lambda_n(\alpha) = 3$.

Trebamo još pokazati da je skup međusobno neekvivalentnih brojeva α definiranih s (1.17) neprebrojiv. Brojevi α i $\widehat{\alpha}$ su ekvivalentni ako i samo ako im se pripadni nizovi $(r_n)_{n\geq 1}$ i $(\widehat{r}_n)_{n\geq 1}$ podudaraju od nekog mjesta nadalje. Za takve nizove ćemo reći da su ekvivalentni. Pretpostavimo da neekvivalentnih nizova ima prebrojivo mnogo, recimo da su to $R_1, R_2, \ldots$, gdje R_i predstavlja niz $r_{i_1} < r_{i_2} < \ldots$. Možemo uzeti da je $R_1 = (n)_{n\geq 1}$ niz svih prirodnih brojeva. Za $i > 1$, R_i nije ekvivalentan R_1, pa postoji beskonačno mnogo prirodnih brojeva koji nisu sadržani u R_i.

Za $i > 1$ sa $S_i = (s_{i_n})_{n\geq 1}$ označimo niz prirodnih brojeva koji je komplementaran nizu R_i. Znamo da je svaki S_i beskonačan. Sada ćemo definirati niz T na sljedeći način. Najprije izaberemo $t_1 \in S_2$, a zatim izaberemo $t_2, t_3, \ldots$ tako da vrijedi

$$\begin{array}{cccc} t_1 \in S_2 & t_1 < t_2 \in S_3, & & \\ 1 + t_2 < t_3 \in S_2 & t_3 < t_4 \in S_3, & t_4 < t_5 \in S_4, & \\ 1 + t_5 < t_6 \in S_2 & t_6 < t_7 \in S_3, & t_7 < t_8 \in S_4, & t_8 < t_9 \in S_5, \\ & \vdots & & \end{array}$$

Očito je T rastući niz prirodnih brojeva. Za svaki $i \geq 2$, beskonačno mnogo članova niza T je sadržano u S_i, pa stoga nije sadržano u R_i. Prema tome, T nije ekvivalentan niti jednom od $R_2, R_3, \ldots$. Budući da svaki od elemenata $t_3, t_6, t_{10}, \ldots$ od T koji pripada S_2 premašuje svog prethodnika za barem 2, zaključujemo da T nije ekvivalentan niti s R_1. Dakle, T nije ekvivalentan niti jednom R_i, što je u suprotnosti s pretpostavkom da niz $(R_i)_{i\geq 1}$ sadrži po jedan element iz svake klase ekvivalencije. $\square$

Podsjetimo se da realne brojeve koji nisu algebarski zovemo *transcendentni* brojevi. Budući da algebarskih brojeva ima samo prebrojivo mnogo, iz prethodnog teorema direktno dobivamo sljedeći rezultat.

Korolar 1.17. *Postoje transcendentni brojevi α za koje je $L(\alpha) = 3$.*

Rothov teorem povlači da je svaki realan broj koji se može dovoljno dobro aproksimirati racionalnim brojevima nužno transcendentan. Međutim, Korolar 1.17 pokazuje da postoje transcendentni brojevi čije su racionalne aproksimacije skoro najgore moguće.

Definicija 1.4. Za realan broj α kažemo da je *loše aproksimabilan* ako postoji konstanta $c > 0$ takva da je $\left|\alpha - \frac{p}{q}\right| > \frac{1}{c\cdot q^2}$ za sve $\frac{p}{q} \in \mathbb{Q} \setminus \{\alpha\}$.

Prema Liouvilleovom teoremu 1.11, svaka je kvadratna iracionalnost loše aproksimabilna. Budući da kvadratna iracionalnost zbog periodičnosti verižnog razlomka ima samo konačno mnogo različitih parcijalnih kvocijenata, idući rezultat daje novi dokaz spomenute činjenice.

Propozicija 1.18. *Realan broj α je loše aproksimabilan ako i samo ako je niz parcijalnih kvocijenata u njegovu razvoju u verižni razlomak omeđen. Stoga je skup svih loše aproksimabilnih brojeva neprebrojiv.*

Dokaz. Za racionalne brojeve je tvrdnja očita, pa pretpostavimo da je $\alpha = [a_0, a_1, a_2, \ldots]$ iracionalan s nizom konvergenti $(\frac{p_n}{q_n})_n$. Neka je α loše aproksimabilan i neka je $c > 0$ konstanta takva da je $\left|\alpha - \frac{p}{q}\right| > \frac{1}{c\cdot q^2}$ za svaki racionalan broj $\frac{p}{q}$. Iz Teorema 1.9.b je za sve $n \geq 0$

$$\frac{1}{c\cdot q_n^2} < \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{a_{n+1}q_n^2},$$

pa je $a_{n+1} < c$ i niz parcijalnih kvocijenata je omeđen.

Obratno, pretpostavimo da parcijalni kvocijenti od α nisu veći od konstante M. Tada za sve $n \geq 0$ imamo $q_{n+1} \leq (M+1)q_n$. Ako za racionalan broj $\frac{p}{q}$ vrijedi $\left|\alpha - \frac{p}{q}\right| < \frac{1}{2q^2}$, prema Teoremu 1.9.e je $\frac{p}{q} = \frac{p_n}{q_n}$ za neki $n \geq 0$ te je iz b dijela istog teorema

$$\left|\alpha - \frac{p}{q}\right| > \frac{1}{q(q+q_{n+1})} \geq \frac{1}{(M+2)q^2}.$$

To znači da je α loše aproksimabilan. $\square$

Vidimo da je $L(\alpha)$ konačan broj ako i samo ako je α loše aproksimabilan broj, pa su nam za proučavanje Lagrangeovog spektra od interesa samo vrijednosti $L(\alpha)$ na tom skupu. Borel je 1909. pokazao da je skup loše aproksimabilnih brojeva Lebesgueove mjere 0.

Slutnja je da nijedan algebarski broj stupnja barem 3 nije loše aproksimabilan, ali ne postoji nijedan rezultat tog tipa.

Poglavlje 2

Binarne kvadratne forme

2.1 Uvod

U ovom poglavlju promatrat ćemo *binarne kvadratne forme*

$$f(x,y) = ax^2 + bxy + cy^2, \qquad a,b,c \in \mathbb{R},$$

tj. homogene polinome u dvije varijable, drugog stupnja, s **realnim** koeficijentima. *Diskriminanta* od f je broj $d = d(f) = b^2 - 4ac$. Nadopunjavanjem do kvadrata na dva načina, dobivamo

$$4af(x,y) = (2ax+by)^2 - dy^2, \tag{2.1}$$
$$4cf(x,y) = (2cy+bx)^2 - dx^2. \tag{2.2}$$

Kvadratnu formu ćemo radi jednostavnosti često zapisivati matrično. Formi f pridružena je simetrična matrica $F = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$ te onda djelovanje te forme možemo zapisati

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = X^\tau F X,$$

gdje je $X = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$, a s $^\tau$ smo označili transponiranje. Vidimo da je diskriminanta forme povezana s determinantom pripadne matrice $d = -4 \det F$.

Za formu f kažemo da je

pozitivno (negativno) *definitna* ako na $\mathbb{R}^2 \setminus \{(0,0)\}$ poprima samo pozitivne (negativne) vrijednosti,

pozitivno (negativno) *semidefinitna* ako poprima samo nenegativne (nepozitivne) vrijednosti,

indefinitna ako poprima i pozitivne i negativne vrijednosti.

Iz (2.1) i (2.2) vidimo da je forma definitna ako i samo ako je $d < 0$, semidefinitna ako i samo ako je $d \leq 0$, te indefinitna ako i samo ako je $d > 0$. Ako je f negativno definitna (semidefinitna) forma, onda je $-f$ pozitivno definitna (semidefinitna) i obratno, pa ćemo u nastavku promatrati samo pozitivno definitne (semidefinitne) forme.

Zanimat će nas vrijednosti koje forma poprima na $\mathbb{Z}^2 \setminus \{(0,0)\}$. Isključili smo $x = y = 0$ jer je u toj točki vrijednost forme uvijek 0.

Definicija 2.1. Kažemo da kvadratna forma f *reprezentira* realan broj λ ako postoje cijeli brojevi x_0, y_0 takvi da je $f(x_0, y_0) = \lambda$. Ako je pritom $(x_0, y_0) = 1$, onda kažemo da f *primitivno reprezentira* λ.

Definicija 2.2. Kvadratne forme f i f' s matricama F i F' su *ekvivalentne* ako postoji $U = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right] \in \mathrm{SL}(2, \mathbb{Z})$ tako da je

$$F' = U^\tau F U.$$

Drugim riječima, supstitucijom oblika $x = px' + qy'$, $y = rx' + sy'$, gdje je $ps - qr = 1$, dobivamo

$$f'(x', y') = f(x, y).$$

Ako je $f'(x', y') = a'x'^2 + b'x'y' + cy'^2$, imamo

$$\begin{aligned} a' &= f(p, r) = ap^2 + bpr + cr^2, \\ b' &= 2apq + b(ps + qr) + 2crs, \\ c' &= f(q, s) = aq^2 + bqs + cs^2. \end{aligned} \tag{2.3}$$

Pišemo $f \sim f'$.

Lako je provjeriti da je $\sim$ zaista relacija ekvivalencije. Napomenimo da se ponegdje u literaturi koriste supstitucije čije su matrice $U \in \mathrm{GL}(2, \mathbb{Z})$, tj. dopušta se $\det U = \pm 1$, pa se onda slučaj kada je $\det U = 1$ zove prava ekvivalencija, a za $\det U = -1$ govori se o nepravoj ekvivalenciji formi. Mi nećemo trebati tu širu definiciju ekvivalencije.

Propozicija 2.1. *Neka su f i f' ekvivalentne forme te λ realan broj. Tada vrijedi:*

1) f reprezentira λ ako i samo ako f' reprezentira λ;

2) f primitivno reprezentira λ ako i samo ako f' primitivno reprezentira λ;

3) diskriminante od f i f' su jednake;

4) forme f i f' su istog tipa (definitnosti).

Dokaz. Sve tvrdnje lako se pokazuju pomoću matričnog zapisa formi i njihove ekvivalentnosti. Koristimo oznake kao u definiciji ekvivalentnosti formi.

Ako imamo $f(x_0, y_0) = \lambda$, tj. za $X_0 = \left[\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right]$ je $X_0^\tau F X_0 = \lambda$, onda je $X_0^\tau (U^{-1})^\tau F' U^{-1} X_0 = \lambda$, pa je $f'(x'_0, y'_0) = \lambda$, gdje je $\begin{bmatrix} x'_0 \\ y'_0 \end{bmatrix} = U^{-1} \left[\begin{smallmatrix} x_0 \\ y_0 \end{smallmatrix}\right]$. Budući da i U i U^{-1} imaju cjelobrojne koeficijente, vidimo da je $(x_0, y_0) > 1$ ako i samo ako je $(x'_0, y'_0) > 1$. Također,

$$d(f') = -4 \det F' = -4 \det U^\tau \det F \det U = -4 \det F = d(f).$$

Konačno, ekvivalentne forme očito poprimaju iste vrijednosti ne samo na $\mathbb{Z}^2 \setminus \{(0, 0)\}$, nego i na $\mathbb{R}^2 \setminus \{(0, 0)\}$, pa moraju biti iste definitnosti. $\square$

Lema 2.2. *Ako f primitivno reprezentira realan broj λ, onda je f ekvivalentna nekoj formi f' za koju je $f'(1, 0) = a' = \lambda$.*

Dokaz. Postoje relativno prosti cijeli brojevi x_0, y_0 takvi da je $f(x_0, y_0) = \lambda$. Možemo uzeti da je y_0 nenegativan, pa razvijemo $\frac{x_0}{y_0}$ u verižni razlomak $[a_0, a_1, a_2, \ldots, a_n]$, gdje uzimamo da je n neparan. Stavimo li

$$\frac{P_n}{Q_n} = [a_0, a_1, \ldots, a_n] = \frac{x_0}{y_0}, \qquad \frac{P_{n-1}}{Q_{n-1}} = [a_0, a_1, \ldots, a_{n-1}],$$

imamo

$$\det \begin{bmatrix} P_n & P_{n-1} \\ Q_n & Q_{n-1} \end{bmatrix} = (-1)^{n-1} = 1$$

te je $f \sim f'$, gdje je $f'(x, y) = f(P_n x + P_{n-1} y, Q_n x + Q_{n-1} y) = a'x^2 + b'xy + c'y^2$. Budući je $a' = f(P_n, Q_n) = f(x_0, y_0) = \lambda$, dobili smo traženu formu. $\square$

Za kvadratnu formu f u nastavku ćemo proučavati

$$m(f) = \inf \left\{ |f(x, y)| \,:\, (x, y) \in \mathbb{Z}^2 \setminus \{(0, 0)\} \right\}.$$

Budući da je $m(\lambda f) = |\lambda| \cdot m(f)$ za svaki realan broj λ, ovu ćemo veličinu učiniti neovisnom o skalaru kojim je forma pomnožena gledajući

$$\frac{\sqrt{|d(f)|}}{m(f)}.$$

Uzeli smo ovu, a ne njoj recipročnu vrijednost da bismo uskladili izbor s našom definicijom Lagrangeovog spektra kako ćemo vidjeti kasnije. Analogno onome što je kod te definicije spomenuto, neki autori ovdje kreću od $m(f)/\sqrt{|d(f)|}$.

2.2 Pozitivno definitne forme

Prije prelaska na definitne forme, promatramo slučaj pozitivno semidefinitne forme koja nije definitna, a to znači da je $d(f) = 0$. Spomenimo da neki autori, kao primjerice Cassels, nazivaju semidefinitnima samo takve forme. Barem jedan od brojeva a i c mora biti različit od nule jer bi inače f bila konstanta 0. Stoga možemo pretpostaviti $a \neq 0$, pa iz (2.1) vidimo da je $f(x, y) = a\left(x + \frac{b}{2a} y\right)^2$. Ako je $\frac{b}{2a}$ racionalan broj, očito je $m(f) = 0$ i postiže se minimum.

Ako je $\frac{b}{2a}$ iracionalan, iz Teorema 1.9.b znamo da postoji beskonačno mnogo $(x, y) \in \mathbb{Z}^2$, $y \neq 0$ takvih da je

$$\left| x + \frac{b}{2a} y \right| = \left| -\frac{b}{2a} - \frac{x}{y} \right| \cdot |y| < \frac{1}{|y|^2} \cdot |y| = \frac{1}{|y|}.$$

Budući je u skupu uređenih parova koji to zadovoljavaju y neograničen, zaključujemo da je $m(f) = 0$ i infimum se ne postiže.

U nastavku promatramo samo nedegenerirani slučaj $d(f) \neq 0$, tj. pozitivno definitne forme. Dakle, $d < 0$, a iz (2.1) i (2.2) je $a > 0$ i $c > 0$.

Očito je od interesa znati kada su dvije forme ekvivalentne. U tu svrhu izabiremo iz svake klase ekvivalencije neke predstavnike.

Definicija 2.3. Reći ćemo da je pozitivno definitna kvadratna forma $f(x, y) = ax^2 + bxy + cy^2$ *reducirana* ako je

$$-a < b \leq a < c \qquad \text{ili} \qquad 0 \leq b \leq a = c. \tag{2.4}$$

Teorem 2.3. *Svaka pozitivno definitna kvadratna forma je ekvivalentna nekoj reduciranoj formi.*

Lema 2.4. *Neka je $f(x,y) = ax^2+bxy+cy^2$ pozitivno definitna forma i t pozitivan realan broj. Ako za realne brojeve x_0, y_0 vrijedi $f(x_0,y_0) \leq t$, onda je*

$$x_0^2 \leq \frac{4ct}{|d|}, \qquad y_0^2 \leq \frac{4at}{|d|}.$$

Zato f reprezentira samo konačno mnogo brojeva manjih od t.

Dokaz. Iz (2.2) je $-dx_0^2 \leq 4cf(x_0,y_0) \leq 4ct$, a iz (2.1) je $-dy_0^2 \leq 4af(x_0,y_0) \leq 4at$, pa slijedi tvrdnja leme. □

Dokaz Teorema 2.3. Koristimo supstitucije čije su matrice

$$U_s = \begin{bmatrix} 0 & 1 \\ -1 & s \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^s, \tag{2.5}$$

gdje je s cijeli broj. Primjenom takve supstitucije iz forme $f(x,y) = ax^2 + bxy + cy^2$ dobivamo

$$f'(x,y) = a'x^2 + b'xy + c'y^2 = cx^2 - (b+2cs)xy + (a+bs+cs^2)y^2.$$

Izborom cijelog broja s takvog da je $-c < -(b+2cs) \leq c$ dobivamo $-a' < b' \leq a'$. Ako je $a' < c'$, stajemo jer smo dobili reduciranu formu. Ako je $a' = c'$ i $b' < 0$ primijenimo supstituciju s matricom U_0 te dobivamo reduciranu formu. Ako je $a' > c'$, ponavljamo gornji postupak s formom f'.

Pretpostavimo da za neku formu f redukcijski algoritam ne staje, te označimo s $f_i(x,y) = a_ix^2 + b_ixy + c_iy^2$ formu dobivenu nakon i-tog redukcijskog koraka. Budući da novi korak provodimo samo ako je $a_i > c_i$ i vrijedi $a_{i+1} = c_i$, dobivamo da je $(a_i)_{i\geq 0}$ strogo padajući niz pozitivnih realnih brojeva koje reprezentira forma f. Prema Lemi 2.4 to nije moguće jer postoji samo konačno mnogo brojeva manjih od $a = a_0$ koje f reprezentira. Dakle, redukcijski algoritam uvijek staje. □

Korijeni polinoma $f(x,1) = ax^2+bx+c$ su $\frac{-b\pm\sqrt{d}}{2a}$, pa iz (2.4) vidimo da je f reducirana pozitivno definitna forma ako i samo ako je $\frac{-b+\sqrt{d}}{2a}$ u *fundamentalnoj domeni* klasične modularne grupe $\mathrm{SL}(2,\mathbb{Z})$

$$\begin{aligned} &\{z \in \mathbb{H} : |z| > 1, -\frac{1}{2} < \mathrm{Re}(z) < \frac{1}{2}\} \\ &\qquad \cup \{z \in \mathbb{H} : |z| \geq 1, \mathrm{Re}(z) = -\frac{1}{2}\} \\ &\qquad\qquad \cup \{z \in \mathbb{H} : |z| = 1, -\frac{1}{2} < \mathrm{Re}(z) \leq 0\}, \end{aligned}$$

gdje smo s $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) \geq 0\}$ označili gornju poluravninu kompleksne ravnine.

Teorem 2.5. *Ako su f i f' dvije ekvivalentne reducirane pozitivno definitne forme, onda je $f = f'$.*

Dakle, svaka klasa ekvivalencije pozitivno definitnih kvadratnih formi sadrži točno jednu reduciranu formu.

Dokaz. Neka su $f(x,y) = ax^2 + bxy + cy^2$ i $f'(x,y) = a'x^2 + b'xy + c'y^2$ dvije ekvivalentne reducirane forme. Ako su $x, y \in \mathbb{Z} \setminus \{0\}$ i $|x| \geq |y|$, uzimajući u obzir $a \geq |b|$, imamo

$$f(x,y) \geq |x|(a|x| - |by|) + c|y|^2 \geq |x|^2(a - |b|) + c|y|^2 \geq a - |b| + c.$$

Analogno se dobiva $f(x,y) \geq a - |b| + c$ za $|y| \geq |x|$. Dakle, tri najmanje vrijednosti koje može poprimiti $f(x,y)$ za $x, y \in \mathbb{Z}$, $(x,y) = 1$ su a, c i $a - |b| + c$ i to upravo u tom redoslijedu, a poprimaju se za $(x,y) = (1,0)$, $(0,1)$, te $(1,1)$ ili $(1,-1)$. Budući da po Propoziciji 2.1.2 forma f' poprima iste vrijednosti za $(x,y) = 1$ kao i f, te budući da je f' također reducirana, zaključujemo da je $a = a'$.

Pretpostavimo da je $a < c$. Ako bi bilo $a = c'$, onda bi broj a imao više primitivnih reprezentacija pomoću forme f', nego pomoću forme f. Naime, $f(x,y) = a$ samo za $(x,y) = (\pm 1, 0)$, dok je $f'(\pm 1, 0) = f'(0, \pm 1) = a$. Stoga je $a < c'$, pa je $c = c'$. Iz $b^2 = d + 4ac = b'^2$ dobivamo $|b| = |b'|$. Dakle, treba još samo pokazati da $b' = -b$ povlači $b = 0$.

Kako je f' reducirana, imamo $-a = -a' < b' = -b$, pa je $-a < b < a < c$. Prema tome je $f(x,y) \geq a - |b| + c > c > a$ za sve $x, y \in \mathbb{Z} \setminus \{0\}$. Neka je $\left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right]$ matrica prijelaza iz f u f'. Tada je prema (2.3) $a' = f(p,r)$ i $c' = f(q,s)$. Budući je $a = a' = f(p,r)$ i $(p,r) = 1$, imamo $p = \pm 1$ i $r = 0$. Iz $ps - qr = 1$ slijedi $s = \pm 1$, te iz $c = c' = f(q,s)$ uz $(q,s) = 1$, imamo $q = 0$. To znači da je $b' = 2apq + b(ps + qr) + 2crs = b$ te je stoga $b = 0$.

Preostaje razmotriti slučaj $a = c$. Tada broj a ima barem 4 primitivne reprezentacije pomoću f, pa mora imati i barem 4 primitivne reprezentacije pomoću f' što povlači $c' = a' = a = c$. Ponovno dobivamo da je $|b| = |b'|$, ali iz definicije reduciranosti u ovom slučaju imamo da je $b \geq 0$ i $b' \geq 0$, pa je $b = b'$. $\square$

Ako je $f(x,y) = ax^2 + bxy + cy^2$ reducirana forma, onda je $-d = 4ac - b^2 \geq 3ac$, pa za $a, b, c \in \mathbb{Z}$ vrijedi $|b| \leq a \leq c \leq \frac{1}{3}|d|$, što znači da reduciranih cjelobrojnih formi s fiksiranom diskriminantom d ima najviše konačno mnogo, taj broj naziva se broj klasa od d.

U dokazu Teorema 2.5 vidjeli smo da je najmanja nenul vrijednost koju poprima reducirana pozitivno definitna forma f upravo $f(1,0)$. Dakle, za reduciranu pozitivno definitnu formu $f(x,y) = ax^2 + bxy + cy^2$ imamo

$$\frac{\sqrt{|d(f)|}}{m(f)} = \frac{\sqrt{|d|}}{a}.$$

No, zbog $-d = 4ac - b^2 \geq 3ac \geq 3a^2$ je $\sqrt{|d|}/a \geq \sqrt{3}$.

Za svaki $t \geq \sqrt{3}$ forma $f(x,y) = x^2 + xy + \frac{t^2+1}{4}y^2$ je reducirana pozitivno definitna kvadratna forma s diskriminantom $-t^2$, pa je za nju $\sqrt{|d(f)|}/m(f) = t$. Budući da je svaka forma ekvivalentna nekoj reduciranoj formi i ekvivalentne forme imaju iste diskriminante i reprezentiraju iste brojeve, zaključujemo da je

$$\left\{ \frac{\sqrt{|d(f)|}}{m(f)} \,:\, f \text{ pozitivno definitna forma} \right\} = [\sqrt{3}, +\infty).$$

Kao što ćemo vidjeti, za indefinitne forme je situacija znatno drugačija.

2.3 Indefinitne forme i Markovljev spektar

Neka je dana indefinitna forma $f(x,y) = ax^2 + bxy + cy^2$ s realnim koeficijentima a, b, c i diskriminantom $d(f) = d > 0$. Označimo li korijene od $f(x,1)$ s

$$\vartheta = \frac{-b+\sqrt{d}}{2a}, \qquad \varphi = \frac{-b-\sqrt{d}}{2a}, \tag{2.6}$$

vidimo da se f faktorizira $f(x,y) = a(x-\vartheta y)(x-\varphi y)$. Budući da će nas, kao i prije, zanimati $m(f)$, pretpostavljat ćemo da su ϑ i φ iracionalni brojevi. To je i razlog zašto smo gore prešutno pretpostavili $a \neq 0$. Naime, za $a = c = 0$ dobivamo neinteresantnu formu $f(x,y) = bxy$, a za $a = 0$, $c \neq 0$, polinom $f(1,y)$ ima racionalni korijen $y = 0$.

Pogledajmo na primjeru jednog svojstva kakva je veza između racionalnih aproksimacija realnih brojeva i veličine $\sqrt{|d(f)|}/m(f)$ koju promatramo za kvadratne forme. Iz Teorema 1.9.f znamo da postoje nizovi $(x_n)_n$ i $(y_n)_n$ brojeva iz $\mathbb{Z} \setminus \{0\}$ tako da je

$$\left|\vartheta - \frac{x_n}{y_n}\right| < \frac{1}{\sqrt{5}\cdot y_n^2}, \qquad |y_n| \to +\infty.$$

Budući je $\sqrt{d} = a(\vartheta - \varphi)$, imamo

$$\frac{\sqrt{d(f)}}{m(f)} \geq \frac{\sqrt{d}}{|f(x_n,y_n)|} = \frac{|a(\vartheta-\varphi)|}{|a(x_n-\vartheta y_n)(x_n-\varphi y_n)|} = \frac{|\vartheta-\varphi|}{|\frac{x_n}{y_n}-\varphi|}\cdot\frac{1}{|\vartheta-\frac{x_n}{y_n}|y_n^2}.$$

Zbog $d \neq 0$ je $\vartheta \neq \varphi$. Očito $\frac{x_n}{y_n} \to \vartheta$, pa za svaki $\varepsilon > 0$, imamo za dovoljno velike n, $|\vartheta-\varphi|/|\frac{x_n}{y_n}-\varphi| > 1-\varepsilon$ iz čega je gornji izraz veći od $(1-\varepsilon)\sqrt{5}$. Zaključujemo da za indefinitne forme f takve da $f(x,1)$ i $f(1,y)$ nemaju racionalnih korijena, vrijedi

$$\frac{\sqrt{d(f)}}{m(f)} \geq \sqrt{5}. \tag{2.7}$$

Za $f(x,y) = x^2 + xy - y^2$ se očito postiže donja ograda u ovoj nejednakosti.

Definicija 2.4. *Markovljev spektar* je

$$\mathcal{M} = \left\{\frac{\sqrt{d(f)}}{m(f)} : f \text{ indefinitna binarna kvadratna forma s realnim koeficijentima}\right\}.$$

Iz (2.7) je $\min\mathcal{M} = \sqrt{5}$.

Uz matricu prijelaza $U = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right]$ s determinantom 1, dobivamo iz forme f ekvivalentnu formu f', tj. $f'(x,y) = f(px+qy, rx+sy)$, pa je

$$\begin{aligned} f'(x,y) &= a\big((px+qy)-\vartheta(rx+sy)\big)\big((px+qy)-\varphi(rx+sy)\big) \\ &= a(p-r\vartheta)(p-r\varphi)\left(x - \frac{s\vartheta-q}{-r\vartheta+p}y\right)\left(x-\frac{s\varphi-q}{-r\varphi+p}y\right). \end{aligned}$$

Također imamo

$$f'(x,y) = a'x^2 + b'xy + c'y^2 = a'(x-\vartheta^* y)(x-\varphi^* y),$$

gdje je

$$\vartheta^* = \frac{-b' + \sqrt{d}}{2a'}, \qquad \varphi^* = \frac{-b' - \sqrt{d}}{2a'}.$$

Koristimo * umjesto $'$ u oznaci korijena kako ne bi bilo nikakve mogućnosti zabune s konjugacijom kvadratnih iracionalnosti.

Očito je $f'(1,0) = a' = a(p - r\vartheta)(p - r\varphi)$ kako i slijedi iz (2.3). Pomoću tih formula i

$$a(\vartheta\varphi) = c, \quad 2a\vartheta = -b + \sqrt{d}, \quad 2a\varphi = -b - \sqrt{d}, \quad \det U = ps - qr = 1,$$

precizirajmo kako se preslikavaju korijeni od $f(x,1)$ pri prijelazu na ekvivalentnu formu. Vrijedi

$$\begin{aligned}
\frac{s\vartheta - q}{-r\vartheta + p} &= \frac{2a(s\vartheta - q)(-r\varphi + p)}{2a(-r\vartheta + p)(-r\varphi + p)} \\
&= \frac{-(2apq - 2a\vartheta ps - 2a\varphi qr + 2a\vartheta\varphi rs)}{2a'} \\
&= \frac{-\big(2apq + (b - \sqrt{d})ps + (b + \sqrt{d})qr + 2crs\big)}{2a'} \\
&= \frac{-b' + (ps - qr)\sqrt{d}}{2a'} = \frac{-b' + \sqrt{d}}{2a'} = \vartheta^*.
\end{aligned}$$

Dakle, ako je $f'(x,y) = f(px + qy, rx + sy)$, onda je

$$\begin{gathered}
\vartheta^* = U^{-1}\vartheta, \quad \vartheta = U\vartheta^*, \qquad \varphi^* = U^{-1}\varphi, \quad \varphi = U\varphi^*, \\
U = \begin{bmatrix} p & q \\ r & s \end{bmatrix}, \qquad U^{-1} = \begin{bmatrix} s & -q \\ -r & p \end{bmatrix}.
\end{gathered} \tag{2.8}$$

Vidimo da je $\vartheta \sim \vartheta^*$ i $\varphi \sim \varphi^*$, tj. možemo reći da su ekvivalentnim formama pridruženi ekvivalentni korijeni.

Definicija 2.5. Reći ćemo da je indefinitna forma $f(x,y) = ax^2 + bxy + cy^2$ s diskriminantom d *reducirana* ako je

$$0 < \sqrt{d} - b < 2|a| < \sqrt{d} + b. \tag{2.9}$$

Uvjet reduciranosti (2.9) je ekvivalentan

$$\big|\sqrt{d} - 2|a|\big| < b < \sqrt{d}. \tag{2.10}$$

Zbog $(\sqrt{d} - b)(\sqrt{d} + b) = -4ac$ je uvjet reduciranosti ekvivalentan i

$$0 < \sqrt{d} - b < 2|c| < \sqrt{d} + b. \tag{2.11}$$

Iz (2.9) je jasno da za fiksiranu diskriminantu, reduciranih indefinitnih formi s cjelobrojnim koeficijentima ima najviše konačno mnogo.

Propozicija 2.6. *Indefinitna forma* $f(x,y) = ax^2 + bxy + cy^2 = a(x - \vartheta y)(x - \varphi y)$ *je reducirana ako i samo ako je*

$$|\vartheta| < 1, \qquad |\varphi| > 1 \qquad i \qquad \vartheta\varphi < 0. \tag{2.12}$$

Ovdje, kao i uvijek, uzimamo da su korijeni ϑ i φ od $f(x,1)$ određeni s (2.6).

Dokaz. Pretpostavimo da je f reducirana. Dijeljenjem (2.9) s $2|a|$ dobivamo $|\vartheta| < 1$ i $|\varphi| > 1$. Također je $0 < b < \sqrt{d}$, pa zbog $d = b^2 - 4ac$ imamo $ac < 0$. Odatle je $\vartheta\varphi = \frac{c}{a} < 0$.

Pretpostavimo sada da vrijedi (2.12). Iz $\frac{c}{a} = \vartheta\varphi < 0$ su a i c suprotnih predznaka te je $|b| < \sqrt{d}$. Nadalje, $(-b+\sqrt{d})(-b-\sqrt{d}) = 4a^2\vartheta\varphi < 0$ povlači da su $-b+\sqrt{d}$ i $b+\sqrt{d}$ istog predznaka, odnosno zbog $b < \sqrt{d}$, moraju biti pozitivni. Iz $|\vartheta| < 1$ je $|-b+\sqrt{d}| < 2|a|$, dok iz $|\varphi| > 1$ slijedi $2|a| < |b+\sqrt{d}|$. Dakle, vrijedi (2.9), pa je f reducirana. $\square$

Primijetimo da se u slučaju kada forma ima cjelobrojne koeficijente, reduciranost forme f podudara s reduciranošću korijena od $f(x,1)$ ili $f(x,-1)$ kako smo je definirali za kvadratne iracionalnosti kod proučavanja verižnih razlomaka.

Teorem 2.7. *Svaka indefinitna forma je ekvivalentna nekoj reduciranoj formi.*

Dokaz. Neka je $f(x,y) = ax^2 + bxy + cy^2$ indefinitna forma. Primjenom supstitucije s matricom $U_s = \left[\begin{smallmatrix} 0 & 1 \\ -1 & s \end{smallmatrix}\right]$, dobivamo

$$f'(x,y) = a'x^2 + b'xy + c'y^2 = cx^2 - (b+2cs)xy + (a+bs+cs^2)y^2,$$

pri čemu $s \in \mathbb{Z}$ odabiremo na sljedeći način.

(1) Ako je $|c| \geq \sqrt{d}$, onda uzimamo s tako da je $-|c| < -b-2cs \leq |c|$, tj. $-|a'| < b' \leq |a'|$. Sada je

$$|c'| = \frac{|d-b'^2|}{4|a'|} \leq \frac{d+|a'|^2}{4|a'|} \leq \frac{|a'|^2+|a'|^2}{4|a'|} = \frac{|a'|}{2} = \frac{|c|}{2},$$

pa ponavljanjem ovakvih supstitucija dobivamo nakon konačno koraka formu kojoj je koeficijent uz y^2 po apsolutnoj vrijednosti manji od $\sqrt{d}$ te nastupa sljedeći slučaj.

(2) Ako je $|c| < \sqrt{d}$, onda uzimamo s tako da je $\sqrt{d} - 2|c| < -b-2cs \leq \sqrt{d}$, tj. $\sqrt{d} - 2|a'| < b' \leq \sqrt{d}$.

Primijetimo da bi $b' = \sqrt{d}$ značilo $a' = 0$ ili $c' = 0$. U prvom slučaju bi bilo $c = 0$, pa bi $f(x,1)$ imao korijen $x = 0$. U drugom slučaju bi imali $a + bs + cs^2 = 0$, pa bi $f(1,y)$ imao cjelobrojni korijen $y = s$. Te smo mogućnosti još na početku isključili. Spomenutu jednakost koeficijenta uz xy i diskriminante možemo isključiti i u nastavku jer prema (2.8), $f(x,1)$ ili $f(1,y)$ ima racionalni korijen ako i samo ako isto vrijedi i za svaku formu ekvivalentnu f.

(2.a) Pretpostavimo da smo dobili $|a'| \leq |c'|$. Množenjem $0 < \sqrt{d} - b' < 2|a'|$ sa $4|a'|\cdot|c'| = \left|\sqrt{d}-b'\right|\cdot\left|\sqrt{d}+b'\right|$ dobivamo $|\sqrt{d}+b'| > 2|c'| \geq 2|a'| > \sqrt{d}-b' > 0$, pa b' mora biti pozitivan. Zato vrijedi (2.9) za f' te je to reducirana forma.

(2.b) Pretpostavimo da smo dobili $|a'| > |c'|$. Prisjetimo se da je $|a'| < \sqrt{d}$. Promotrimo sada c'.

(2.b.i) Ako je $|c'| \leq \frac{\sqrt{d}}{2}$, onda primjenom odgovarajućeg $U_{s'}$ na f' dobivamo $f''(x,y) = a''x^2 + b''xy + c''y^2$ takvu da je $\sqrt{d} - 2|a''| < b'' < \sqrt{d}$ i $|a''| = |c'| \leq \frac{\sqrt{d}}{2}$. Stoga je

$$\left|\sqrt{d} - 2|a''|\right| = \sqrt{d} - 2|a''| < b'' < \sqrt{d},$$

pa je prema (2.10) f'' reducirana forma.

(2.b.ii) Neka je $|c'| > \frac{\sqrt{d}}{2}$. Zbog $|a'| < \sqrt{d}$ imamo $-\sqrt{d} < \sqrt{d} - 2|a'| < b' < \sqrt{d}$ što povlači $0 < \sqrt{d} - b' < 2|a'|$ i $\sqrt{d} + b' > 0$. Stoga je

$$\frac{\sqrt{d} + b'}{2|c'|} = \frac{2|a'|}{\sqrt{d} - b'} > 1$$

što daje $-b' + 2|c'| < \sqrt{d}$. S druge strane, $|b'| < \sqrt{d}$ i $|c'| > \frac{\sqrt{d}}{2}$ povlači

$$-b' + 2|c'| > 2|c'| - \sqrt{d} = \left|\sqrt{d} - 2|c'|\right|.$$

Nakon primjene supstitucije s matricom U_s za $s = \frac{-|c'|}{c'}$, dobivamo formu

$$\begin{aligned} f''(x,y) &= c'x^2 - \left(b' + 2c'\frac{-|c'|}{c'}\right)xy + \left(a + b\frac{-|c'|}{c'} + c\Big(\frac{-|c'|}{c'}\Big)^2\right)y^2 \\ &= c'x^2 + (-b' + 2|c'|)xy + \left(a - b\frac{|c'|}{c'} + c\right)y^2 \end{aligned}$$

koja je prema gore dokazanom i (2.10) reducirana. □

Dok za definitne forme u svakoj klasi ekvivalencije postoji jedinstvena reducirana forma, za indefinitne forme može biti više, pa i beskonačno mnogo reduciranih formi u istoj klasi, ali one posjeduju elegantnu strukturu, čine takozvane lance.

Definicija 2.6. *Desna susjedna forma* od $f(x,y) = ax^2 + bxy + cy^2$ je njoj ekvivalentna forma f_d koja se dobiva supstitucijom s matricom $U_s = \left[\begin{smallmatrix} 0 & 1 \\ -1 & s \end{smallmatrix}\right]$, tj.

$$f_d(x,y) = f(y, -x + sy) = cx^2 - (b + 2cs)xy + c'y^2$$

za neki cijeli broj s. Analogno, *lijeva susjedna forma* od f dobiva se supstitucijom s matricom $\left[\begin{smallmatrix} s & -1 \\ 1 & 0 \end{smallmatrix}\right] = U_s^{-1}$, odnosno

$$f_l(x,y) = f(sx - y, x) = a'x^2 - (b + 2as)xy + ay^2.$$

Lako se vidi da su forme $ax^2 + bxy + cy^2$ i $cx^2 + b'xy + c'y^2$ s istom diskriminantom susjedne ako i samo ako je $\frac{b+b'}{2c} \in \mathbb{Z}$, tj. $b + b' \equiv 0 \pmod{2c}$.

Pomoću susjednih formi ćemo organizirati ekvivalentne forme u pogodnu strukturu. Primijetimo da prema dokazu Teorema 2.7 možemo iz svake indefinitne forme konačnim nizom formi takvim da je svaka od njih desna susjedna forma svoga prethodnika, doći do reducirane forme.

Propozicija 2.8. *Svaka reducirana forma ima jedinstvenu reduciranu desnu susjednu formu i jedinstvenu reduciranu lijevu susjednu formu.*

Dokaz. Pokažimo najprije tvrdnju za desnu susjednu formu. Neka su ϑ i φ korijeni pridruženi reduciranoj formi $f(x,y) = ax^2 + bxy + cy^2$ kao u (2.6). Tada ϑ i a imaju isti predznak, dok c ima njima suprotan predznak. Neka je s cijeli broj istog predznaka kao ϑ takav da je $|s| = \lfloor 1/|\vartheta| \rfloor$. Budući da je $|\vartheta| < 1$, vidimo da s nije nula i formi $f_d(x,y) = f(y, -x + sy)$ je pridružen prvi korijen

$$\vartheta_d = \begin{bmatrix} 0 & 1 \\ -1 & s \end{bmatrix}^{-1} \vartheta = \begin{bmatrix} s & -1 \\ 1 & 0 \end{bmatrix} \vartheta = \frac{s\vartheta - 1}{\vartheta} = s - \frac{1}{\vartheta} \tag{2.13}$$

prema (2.8). Stoga je $|\vartheta_d| < 1$ i predznak od ϑ_d je suprotan onome od s i ϑ. Kako je φ suprotnog predznaka od ϑ i s, drugi korijen pridružen f_d , tj. $\varphi_d = s - \frac{1}{\varphi}$ ima isti predznak kao s. Stoga je $|\varphi_d| > 1$. Prema Propoziciji 2.6 zaključujemo da je f_d reducirana desna susjedna forma od f, pa moramo još samo pokazati da je jedinstvena.

Budući da je f reducirana, imamo $a\vartheta > 0$ i $ac < 0$, dok iz reduciranosti od f_d slijedi $c\vartheta_d > 0$. Zato je $(ac)^2\vartheta\vartheta_d < 0$, pa su ϑ i ϑ_d suprotnog predznaka. Sada $|\vartheta| < 1$, $|\vartheta_d| < 1$ i (2.13) povlače da je s istog predznaka kao ϑ te da je $|s| = \lfloor 1/|\vartheta| \rfloor$. Dakle, s i stoga f_d su jedinstveno određeni.

Iz ekvivalencije uvjeta (2.9) i (2.11) vidimo da je $f(x, y)$ reducirana forma ako i samo ako je $\widehat{f}(x, y) = f(y, x)$ reducirana forma. Budući da $\widehat{f}$ ima jedinstvenu reduciranu desnu susjednu formu $\widehat{f}_d$, dobivamo da f ima jedinstvenu reduciranu lijevu susjednu formu $f_l(x, y) = \widehat{f}_d(y, x)$. Matrično to uz samorazumljive oznake i $J = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, možemo zapisati

$$\widehat{F} = J^\tau F J, \quad \widehat{F}_d = U_s^\tau \widehat{F} U_s, \quad J^\tau \widehat{F}_d J = (JU_sJ)^\tau F (JU_sJ) = (U_s^{-1})^\tau F U_s^{-1} = F_l. \qquad \square$$

Propozicija 2.8 nam govori da je svaka reducirana forma f element *lanca* ekvivalentnih reduciranih formi

$$\ldots, f_{-2}, f_{-1}, f_0, f_1, f_2, \ldots$$

pri čemu su f_{k-1} i f_{k+1} jedinstvena lijeva i desna reducirana susjedna forma od f_k. Budući da predznak prvog koeficijenta formi u lancu alternira, možemo uzeti da $f_0(x, y)$ ima pozitivan prvi koeficijent i tada su forme u danom lancu

$$f_k(x, y) = (-1)^k a_k x^2 + b_k xy + (-1)^{k+1} a_{k+1} y^2,$$

gdje je

$$f_{k+1}(x, y) = f_k(y, -x + s_k y)$$

tako da su a_k, b_k i $(-1)^k s_k$ pozitivni brojevi za svaki k (a_k zbog izbora f_0, b_k zbog reduciranosti forme, $(-1)^k s_k$ jer je s_k istog predznaka kao $(-1)^k a_k$ što smo vidjeli u dokazu Propozicije 2.8).

Neka je $\ell_k = (-1)^k s_k$ prirodan broj i neka su

$$\vartheta_k = \frac{-b_k + \sqrt{d}}{2(-1)^k a_k}, \qquad \varphi_k = \frac{-b_k - \sqrt{d}}{2(-1)^k a_k}$$

kao u (2.6) korijeni od $f_k(x, 1)$. Iz (2.9) vidimo da je $0 < (-1)^k \vartheta_k < 1$ i $(-1)^{k+1}\varphi_k > 1$. Zbog (2.8) je

$$\vartheta_k = U_{s_k}\vartheta_{k+1} = \frac{1}{-\vartheta_{k+1} + s_k}, \qquad \varphi_k = U_{s_k}\varphi_{k+1} = \frac{1}{-\varphi_{k+1} + s_k},$$

pa imamo

$$(-1)^k\vartheta_k = \frac{1}{\ell_k + (-1)^{k+1}\vartheta_{k+1}}, \qquad (-1)^{k+2}\varphi_{k+1} = \ell_k + \frac{1}{(-1)^{k+1}\varphi_k}.$$

Zato $(-1)^k\vartheta_k$ i $(-1)^{k+1}\varphi_k$ imaju razvoje u verižne razlomke

$$(-1)^k\vartheta_k = [0, \ell_k, \ell_{k+1}, \ldots], \qquad (-1)^{k+1}\varphi_k = [\ell_{k-1}, \ell_{k-2}, \ell_{k-3}, \ldots]. \tag{2.14}$$

Dakle, korijeni ϑ_0 i φ_0 od $f_0(x,1)$ su preko verižnih razlomaka povezani s korijenima $f_k(x,1)$ za svaku formu u lancu

$$\begin{aligned}
&\vartheta_0 = [0, \ell_0, \ell_1, \ldots, \ell_{k-1} + (-1)^k \vartheta_k], \quad (-1)^{k+1}\varphi_k = [\ell_{k-1}, \ell_{k-2}, \ldots, \ell_1, \ell_0, -\varphi_0] \qquad \text{za } k > 0,\\
&(-1)^k \vartheta_k = [0, \ell_k, \ell_{k+1}, \ldots, \ell_{-1} + \vartheta_0], \quad -\varphi_0 = [\ell_{-1}, \ell_{-2}, \ldots, \ell_k, (-1)^{k+1}\varphi_k] \qquad \text{za } k < 0.
\end{aligned} \tag{2.15}$$

Primijetimo da vrijedi

$$(-1)^k \vartheta_k + (-1)^{k+1}\varphi_k = \frac{\sqrt{d}}{a_k} = \frac{\sqrt{d}}{|f_k(1,0)|}. \tag{2.16}$$

Forme u istom lancu su očito ekvivalentne, ali vrijedi i obrat.

Teorem 2.9. *Svake dvije ekvivalentne reducirane indefinitne forme pripadaju istom lancu.*

Dokaz. Budući da svaku formu možemo, ako treba, zamijeniti s njoj susjednom reduciranom formom, dovoljno je pokazati tvrdnju teorema za dvije različite reducirane forme $f(x,y) = ax^2+bxy+cy^2$ i $f'(x,y) = a'x^2+b'xy+c'y^2$ koje su ekvivalentne i imaju pozitivan prvi koeficijent, tj. $a, a' > 0$. Stoga za korijene ϑ i φ od $f(x,1)$ te ϑ^* i φ^* od $f'(x,1)$ prema (2.6) i (2.9) imamo $\vartheta, \vartheta^* \in (0,1)$ i $\varphi, \varphi^* \in (-\infty, -1)$. Neka je $f'(x,y) = f(px+qy, rx+sy)$, gdje je $U = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right] \in \mathrm{SL}(2,\mathbb{Z})$. Uzevši eventualno $-U$ umjesto U, možemo pretpostaviti da je ili $p = 0$ i $r \geq 1$ ili $p > 0$.

Ako je $p = 0$ i $r \geq 1$, onda $-qr = 1$ povlači $q = -1$ i $r = 1$. Prema (2.8) imamo

$$\begin{aligned}
\vartheta = U\vartheta^* = \frac{-1}{\vartheta^* + s} \qquad &\text{pa je} \qquad s = -\frac{1}{\vartheta} - \vartheta^* < -1,\\
\varphi = U\varphi^* = \frac{-1}{\varphi^* + s} \qquad &\text{pa je} \qquad s = -\frac{1}{\varphi} - \varphi^* > 1,
\end{aligned}$$

što nije moguće. Stoga je $p \neq 0$ i imamo $p > 0$. kako je

$$\begin{bmatrix} p & q \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix}^{-1} = \begin{bmatrix} p & q \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s & -q \\ -r & p \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -r & p \end{bmatrix},$$

ponovno iz (2.8) imamo

$$\begin{bmatrix} p & q \\ 0 & 1 \end{bmatrix} \vartheta^* = \begin{bmatrix} 1 & 0 \\ -r & p \end{bmatrix} \vartheta, \qquad \begin{bmatrix} p & q \\ 0 & 1 \end{bmatrix} \varphi^* = \begin{bmatrix} 1 & 0 \\ -r & p \end{bmatrix} \varphi,$$

odnosno

$$\left(\frac{p}{\vartheta} - r\right)(p\vartheta^* + q) = 1, \tag{2.17}$$

$$\left(\frac{p}{\varphi} - r\right)(p\varphi^* + q) = 1. \tag{2.18}$$

Ako je $r = 0$, onda je $p = s = 1$ te iz (2.17) slijedi $\frac{\vartheta^*+q}{\vartheta} = 1$, odnosno $q = \vartheta - \vartheta^* \in (-1,1)$, pa je $q = 0$. Ako je $q = 0$, onda je opet $p = s = 1$ te iz (2.18) slijedi $(\frac{1}{\varphi} - r)\varphi^* = 1$, odnosno $r = \frac{1}{\varphi} - \frac{1}{\varphi^*} \in (-1,1)$, pa je $r = 0$. U oba slučaja dobivamo $f'(x,y) = f(x,y)$ što je u suprotnosti s početnom pretpostavkom da su f i f' različite. Dakle, q i r su različiti od nule te smo do sada pokazali da je $p > 0$ i $qr \neq 0$.

Ako je $q > 0$, onda je $p\vartheta^* + q > 1$, pa je $0 < \frac{p}{\vartheta} - r < 1$ prema (2.17) što daje $0 < p < \frac{p}{\vartheta} < r + 1$ iz čega je $r > 0$. Ako je $r > 0$, onda je $\frac{p}{\varphi} - r < -1$, pa je $-1 < p\varphi^* + q < 0$ prema (2.18) što daje $0 < p < p(-\varphi^*) < q + 1$ iz čega je $q > 0$. Zato je $qr > 0$. Kako je $ps = qr + 1$ i $p > 0$, vidimo da je $s > 0$. Dakle, smijemo uzeti da su svi p, q, r, s prirodni brojevi jer bi za q i r negativne mogli promatrati "inverznu" ekvivalenciju $f(x, y) = f'(sx - qy, -rx + py)$.

Sada imamo prirodne brojeve p, q, r, s takve da je $ps - qr = 1$, a maloprije smo vidjeli da to povlači $p < q + 1$ i $p < r + 1$, odnosno $p \leq q$ i $p \leq r$. kako je $qr < ps$, imamo također $r < s$ i $q < s$. Iz $\vartheta = U\vartheta^*$ je $\vartheta = \left[\begin{smallmatrix} q & p \\ s & r \end{smallmatrix}\right] \frac{1}{\vartheta^*}$, pa su ispunjeni svi uvjeti za primjenu Leme 1.2. Ova lema povlači da je $\vartheta = [0, a_0, a_1, \ldots, a_{k-1}, \frac{1}{\vartheta^*}] = [0, a_0, a_1, \ldots, a_{k-1} + \vartheta^*]$, gdje je $\frac{p}{r} = [0, a_0, \ldots, a_{k-2}]$, $\frac{q}{s} = [0, a_0, \ldots, a_{k-1}]$. Usporedimo li ovaj verižni razlomak s (2.14), tj. (2.15), iz jedinstvenosti razvoja u verižni razlomak slijedi da je k paran i ϑ^* je prvi korijen od $f''(x, 1)$, gdje je f'' forma koja se nalazi k mjesta desno od f u njezinu lancu.

Iz (1.4) je $\frac{q}{p} = [a_{k-1}, a_{k-2}, \ldots, a_0, 0] = [a_{k-1}, a_{k-2}, \ldots, a_1]$ i $\frac{s}{r} = [a_{k-1}, a_{k-2}, \ldots, a_0]$, pa su to uzastopne konvergente. Budući je $\varphi^* = U^{-1}\varphi = \left[\begin{smallmatrix} s & -q \\ -r & p \end{smallmatrix}\right] \varphi = \left[\begin{smallmatrix} -s & -q \\ r & p \end{smallmatrix}\right] (-\varphi)$, vrijedi $-\varphi^* = \left[\begin{smallmatrix} s & q \\ r & p \end{smallmatrix}\right] (-\varphi)$, iz čega slijedi $(-1)^{k+1}\varphi^* = [a_{k-1}, a_{k-2}, \ldots, a_0, -\varphi]$, pa usporedbom s (2.15) dobivamo da je drugi korijen od $f''(x, 1)$ upravo φ^*.

To znači da f' i f'' imaju istu diskriminantu, a $f'(x, 1)$ i $f''(x, 1)$ imaju iste korijene, pa zaključujemo da je $f'' = f'$. Dakle, f' se nalazi u lancu od f. □

Teorem 2.10 (Lagrange)**.** *Neka je $\ldots, f_{-1}, f_0, f_1, \ldots$ jedan lanac reduciranih formi diskriminante d. Neka je λ primitivno reprezentiran nekom, pa i svakom formom u tom lancu. Ako je $|\lambda| < \frac{1}{2}\sqrt{d}$, onda je $f_k(1, 0) = \lambda$ za neki cijeli broj k.*

Dokaz. Neka je λ kao u izreci teorema. Prema Lemi 2.2 postoji forma $f(x, y) = \lambda x^2 + bxy + cy^2$ ekvivalentna formama u danom lancu. Primjenom supstitucije s matricom $\left[\begin{smallmatrix} 1 & s \\ 0 & 1 \end{smallmatrix}\right]$ dobivamo ekvivalentnu formu f' s matricom

$$\begin{bmatrix} 1 & 0 \\ s & 1 \end{bmatrix} \begin{bmatrix} \lambda & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \begin{bmatrix} 1 & s \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda & \lambda s + \frac{b}{2} \\ \lambda s + \frac{b}{2} & \lambda s^2 + bs + c \end{bmatrix},$$

tj. $f'(x, y) = a'x^2 + b'xy + c'y^2 = \lambda x^2 + (2\lambda s + b)xy + (\lambda s^2 + bs + c)y^2$. Izaberemo cijeli broj s tako da je $\sqrt{d} - 2|\lambda| < 2\lambda s + b < \sqrt{d}$, odnosno $\sqrt{d} - 2|\lambda| < b' < \sqrt{d}$. Zbog $|\lambda| < \frac{1}{2}\sqrt{d}$ je $\sqrt{d} - 2|\lambda| > 0$ te imamo

$$\left|\sqrt{d} - 2|\lambda|\right| = \sqrt{d} - 2|\lambda| < b' < \sqrt{d},$$

što prema (2.10) znači da je forma f' reducirana, pa je prema Teoremu 2.9 u danom lancu formi. □

Propozicija 2.11. *Neka je $\ldots, f_{-1}, f_0, f_1, \ldots$ lanac reduciranih formi i f forma ekvivalentna f_0. Tada je $m(f) = \inf_{k \in \mathbb{Z}} |f_k(1, 0)|$.*

Dokaz. Prema Teoremu 2.7 je forma f ekvivalentna reduciranoj formi $ax^2 + bxy + cy^2$ za koju je $-4ac = d - b^2 \in (0, d)$, pa je $|a| < \frac{1}{2}\sqrt{d}$ ili $|c| < \frac{1}{2}\sqrt{d}$. Kako su i a i c primitivno reprezentirani, zaključujemo da f primitivno reprezentira barem jedan broj λ za koji je $|\lambda| < \frac{1}{2}\sqrt{d}$. No, to znači da je

$$m(f) = \inf \left\{ |f(x, y)| \, : \, (x, y) \in \mathbb{Z}^2, \, (x, y) = 1, \, |f(x, y)| < \frac{1}{2}\sqrt{d} \right\},$$

pa iz Teorema 2.10 slijedi tvrdnja propozicije. □

Teorem 2.12. *Neka je f indefinitna forma i $(f_i)_{i\in\mathbb{Z}}$ lanac reduciranih formi koje su ekvivalentne f. Pretpostavimo da je $f_0(1,0)>0$ i neka su $\vartheta_0\in(0,1)$, $\varphi_0\in(-\infty,-1)$ korijeni od $f_0(x,1)$. Neka je $\vartheta_0=[0,\ell_0,\ell_1,\ell_2,\ldots]$ i $-\varphi_0=[\ell_{-1},\ell_{-2},\ell_{-3},\ldots]$. Tada je*

$$\frac{\sqrt{d(f)}}{m(f)}=\sup_{k\in\mathbb{Z}}\big(\ell_k+[0,\ell_{k+1},\ell_{k+2},\ldots]+[0,\ell_{k-1},\ell_{k-2},\ldots]\big).$$

Dokaz. Tvrdnja slijedi direktno iz Propozicije 2.11 te jednakosti (2.16) i (2.15). □

Ponekad ćemo trebati potpuniju karakterizaciju prvih koeficijenata formi u lancu.

Propozicija 2.13. *Uz oznake kao u (2.14) i (2.16), za $k\geq 0$ vrijedi*

$$a_k=f_k(1,0)=f_0(P_{k-1},Q_{k-1}),$$

gdje su $\frac{P_k}{Q_k}=[0,\ell_0,\ell_1,\ldots,\ell_{k-1}]$ konvergente u razvoju korijena $\vartheta_0=[0,\ell_0,\ell_1,\ldots]\in(0,1)$ od $f_0(x,1)$ u verižni razlomak.

Dokaz. Označimo li s F_i matricu pridruženu formi f_i te s

$$U_{s_i}=\begin{bmatrix}0&1\\-1&s_i\end{bmatrix}=\begin{bmatrix}0&1\\-1&(-1)^i\ell_i\end{bmatrix}$$

matricu prijelaza iz f_i u f_{i+1}, imamo $F_k=U^\tau F_0U$, gdje je

$$\begin{aligned}
U&=\prod_{i=0}^{k-1}U_{s_i}=\prod_{i=0}^{k-1}\begin{bmatrix}0&1\\-1&(-1)^i\ell_i\end{bmatrix}\\
&=\kappa_k\begin{bmatrix}0&1\\1&0\end{bmatrix}\left(\prod_{i=0}^{k-1}\begin{bmatrix}0&(-1)^i\\1&0\end{bmatrix}\begin{bmatrix}0&1\\-1&(-1)^i\ell_i\end{bmatrix}\begin{bmatrix}0&(-1)^{i+1}\\1&0\end{bmatrix}\right)\begin{bmatrix}0&(-1)^k\\1&0\end{bmatrix}\\
&=\kappa_k\begin{bmatrix}0&1\\1&0\end{bmatrix}\left(\prod_{i=0}^{k-1}\begin{bmatrix}\ell_i&1\\1&0\end{bmatrix}\right)\begin{bmatrix}0&(-1)^k\\1&0\end{bmatrix}\\
&=\kappa_k\begin{bmatrix}P_k&P_{k-1}\\Q_k&Q_{k-1}\end{bmatrix}\begin{bmatrix}0&(-1)^k\\1&0\end{bmatrix}=\kappa_k\begin{bmatrix}P_{k-1}&(-1)^kP_k\\Q_{k-1}&(-1)^kQ_k\end{bmatrix}.
\end{aligned}$$

Ovdje smo dobili $\frac{P_k}{Q_k}=[0,\ell_0,\ell_1,\ldots,\ell_{k-1}]$ korištenjem matričnog zapisa verižnog razlomka (1.3). Također, iz $\left[\begin{smallmatrix}0&(-1)^{i+1}\\1&0\end{smallmatrix}\right]^2=(-1)^{i+1}\left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]$ je

$$\kappa_k=\prod_{i=0}^{k-2}(-1)^{i+1}=(-1)^{\sum_{i=1}^{k-1}i}=(-1)^{\frac{(k-1)k}{2}}=(-1)^{\binom{k}{2}}.$$

Iz (2.3) je sada $a_k=f_k(1,0)=f_0(\kappa_kP_{k-1},\kappa_kQ_{k-1})=f_0(P_{k-1},Q_{k-1})$. □

Propozicija 2.14. *Ako lanac formi $(f_i)_{i\in\mathbb{Z}}$ sadrži barem jednu formu s cjelobrojnim koeficijentima, onda sve forme u tom lancu imaju cjelobrojne koeficijente i takav lanac je periodski, tj. postoji $k\in\mathbb{N}$ takav da je $f_i=f_{i+k}$ za svaki $i\in\mathbb{Z}$.*

Lanac formi je periodski ako i samo ako postoji realan broj $\lambda>0$ takav da su za neku (pa onda i za svaku) formu f u lancu, koeficijenti od λf cjelobrojni. Posebno, ako u lancu neka forma ima racionalne koeficijente, onda sve forme u lancu imaju racionalne koeficijente i lanac je periodski.

Dokaz. Ako je u lancu $(f_i)_{i\in\mathbb{Z}}$ neka forma cjelobrojna, onda su i sve druge forme cjelobrojne budući da su međusobno ekvivalentne i koeficijenti su im vezani jednakostima tipa (2.3). Sve forme u danom lancu imaju istu diskriminantu. Iz (2.9) se vidi da postoji samo konačno reduciranih cjelobrojnih formi sa zadanom diskriminantom, pa i u danom lancu može biti samo konačno mnogo različitih formi. Jedinstvenost lijeve i desne reducirane susjedne forme povlači periodičnost formi u lancu.

Lako je provjeriti da ako reduciranu formu pomnožimo nekim pozitivnim brojem, forma i dalje ostaje reducirana. Dakle, ako sve forme u lancu pomnožimo istim pozitivnim realnim brojem, ponovno dobivamo lanac jer susjedne forme ostaju susjedne. Ako postoji forma f u lancu takva da je λf cjelobrojna forma za neki $\lambda > 0$, onda je $(\lambda f_i)_{i\in\mathbb{Z}}$ prema maloprije pokazanom, lanac s cjelobrojnim formama, pa stoga i periodski. Zaključujemo da je i $(f_i)_{i\in\mathbb{Z}}$ periodski lanac.

S druge strane, ako je lanac periodski, onda se iz forme f u tom lancu supstitucijom s matricom $U = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right] \in \mathrm{SL}(2,\mathbb{Z})$ ponovno dobiva forma f. Stoga je za korijene ϑ, φ od $f(x,1)$ prema (2.8) nužno $\vartheta = U\vartheta$ i $\varphi = U\varphi$, odnosno vrijedi

$$r\vartheta^2 + (s-p)\vartheta - q = 0, \qquad r\varphi^2 + (s-p)\varphi - q = 0.$$

Primijetimo da je $U \neq \pm\left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ što se može pokazati slično onome što smo radili u dokazu Propozicije 2.13. Zato su ϑ i φ kao iracionalni brojevi nužno kvadratne iracionalnosti koje su korijeni istog cjelobrojnog kvadratnog polinoma, pa su zato konjugirane. Stoga forma $\frac{f(x,y)}{f(1,0)} = (x-\vartheta y)(x-\varphi y)$ ima racionalne koeficijente te λf ima cjelobrojne koeficijente za neki $\lambda > 0$. □

Prethodna propozicija daje nam nove dokaze Teorema 1.3 i 1.4. Završimo ovo poglavlje jednim kriterijem za periodičnost lanaca. Obrat sljedećeg teorema slijedi direktno iz Propozicije 2.14.

Teorem 2.15. *Ako za neki lanac formi $(f_i)_{i\in\mathbb{Z}}$ postoje tri različita indeksa $k,m,n \in \mathbb{Z}$ takva da je*

$$\frac{f_m(1,0)}{f_k(1,0)} \in \mathbb{Q} \qquad i \qquad \frac{f_n(1,0)}{f_k(1,0)} \in \mathbb{Q},$$

onda je dani lanac periodski.

Dokaz. Bez smanjenja općenitosti možemo pretpostaviti da je $0 < k < m < n$ jer uvijek možemo uzeti da je f_0 dovoljno daleko lijevo u lancu. Također, podijelimo sve forme u lancu s $f_0(1,0) > 0$ tako da možemo uzeti $f_0(x,y) = x^2 + bxy + cy^2$. Sada iz Propozicije 2.13 imamo

$$\begin{aligned} f_0(P_{k-1}, Q_{k-1}) &= f_k(1,0), \\ f_0(P_{m-1}, Q_{m-1}) &= f_m(1,0), \\ f_0(P_{n-1}, Q_{n-1}) &= f_n(1,0), \end{aligned}$$

gdje su $\frac{P_i}{Q_i}$, $i \in \{k-1, m-1, n-1\}$ različite konvergente prvog korijena od $f_0(x,1)$. Matrično možemo pisati

$$\begin{bmatrix} P_{k-1}^2 & P_{k-1}Q_{k-1} & Q_{k-1}^2 \\ P_{m-1}^2 & P_{m-1}Q_{m-1} & Q_{m-1}^2 \\ P_{n-1}^2 & P_{n-1}Q_{n-1} & Q_{n-1}^2 \end{bmatrix} \begin{bmatrix} 1 \\ b \\ c \end{bmatrix} = \begin{bmatrix} f_k(1,0) \\ f_m(1,0) \\ f_n(1,0) \end{bmatrix} = f_k(1,0) \begin{bmatrix} 1 \\ \frac{f_m(1,0)}{f_k(1,0)} \\ \frac{f_n(1,0)}{f_k(1,0)} \end{bmatrix}.$$

Promatramo ovo kao sustav u $1, b, c$. Determinanta matrice sustava se računa korištenjem Vandermondeove determinante.

$$\begin{vmatrix} P_{k-1}^2 & P_{k-1}Q_{k-1} & Q_{k-1}^2 \\ P_{m-1}^2 & P_{m-1}Q_{m-1} & Q_{m-1}^2 \\ P_{n-1}^2 & P_{n-1}Q_{n-1} & Q_{n-1}^2 \end{vmatrix} =$$
$$= (Q_{k-1}Q_{m-1}Q_{n-1})^2 \left(\frac{P_{k-1}}{Q_{k-1}} - \frac{P_{m-1}}{Q_{m-1}}\right)\left(\frac{P_{k-1}}{Q_{k-1}} - \frac{P_{n-1}}{Q_{n-1}}\right)\left(\frac{P_{m-1}}{Q_{m-1}} - \frac{P_{n-1}}{Q_{n-1}}\right) \neq 0.$$

Primjenom Cramerovog pravila najprije dobivamo da je $f_k(1,0) \in \mathbb{Q}$, pa je zbog uvjeta teorema i $f_m(1,0), f_n(1,0) \in \mathbb{Q}$. Daljnjom primjenom Cramerovog pravila slijedi da su b i c racionalni brojevi, pa je f_0 forma s racionalnim koeficijentima iz čega je po Propoziciji 2.14 lanac periodski. $\square$

Poglavlje 3

Osnovni rezultati o spektrima

Podsjetimo se da sa $\mathbb{Z}$, $\mathbb{N}$, $\mathbb{N}_0$ označavamo redom skup cijelih, pozitivnih cijelih (prirodnih) i nenegativnih cijelih brojeva.

Motivirani rezultatima iz teorije verižnih razlomaka i indefinitnih kvadratnih formi, uvodimo sljedeće definicije. Neka je dan obostrano beskonačan niz $A = (a_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$ s elementima iz skupa prirodnih brojeva. Definiramo za svaki $i \in \mathbb{Z}$

$$\begin{aligned}\lambda_i(A) &= [a_{i-1}, a_{i-2}, a_{i-3}, \ldots]^{-1} + a_i + [a_{i+1}, a_{i+2}, a_{i+3}, \ldots]^{-1} \\ &= [a_i, a_{i-1}, a_{i-2}, \ldots] + [a_{i+1}, a_{i+2}, a_{i+3}, \ldots]^{-1} \\ &= [a_{i-1}, a_{i-2}, a_{i-3}, \ldots]^{-1} + [a_i, a_{i+1}, a_{i+2}, \ldots].\end{aligned} \tag{3.1}$$

Lagrangeov broj od A je

$$L(A) = \limsup_{i\in\mathbb{Z}} \lambda_i(A) = \max\{\limsup_{i\to+\infty} \lambda_i(A),\ \limsup_{i\to-\infty} \lambda_i(A)\}.$$

Markovljev supremum od A je

$$M(A) = \sup_{i\in\mathbb{Z}} \lambda_i(A).$$

Uvijek je $L(A) \leq M(A)$. Za čisto periodski niz $A = (a_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$, tj. takav da za neki $m \in \mathbb{N}$ i svaki $n \in \mathbb{Z}$ vrijedi $a_{n+m} = a_n$, očito je i $(\lambda_i(A))_{i\in\mathbb{Z}}$ čisto periodski, pa poprima samo konačno mnogo vrijednosti i svaku od njih beskonačno mnogo puta te je zato $L(A) = M(A)$.

Propozicija 3.1. *Za Lagrangeov i Markovljev spektar vrijedi*

$$\mathcal{L} = \{L(A) : A \in \mathbb{N}^{\mathbb{Z}}\}, \qquad \mathcal{M} = \{M(A) : A \in \mathbb{N}^{\mathbb{Z}}\}.$$

Dokaz. Neka je dan iracionalan broj $\alpha = [a_0, a_1, a_2, \ldots]$, $a_0 \in \mathbb{Z}$, $a_i \in \mathbb{N}$ za $i \geq 1$. Tada se korištenjem Teorema 1.8.b lako dobiva da je $L(\alpha) = L(A)$, gdje je $A = (b_i)_{i\in\mathbb{Z}}$ definiran sa $b_i = a_{i+1}$ za $i \geq 0$ i $b_i = a_{-i}$ za $i < 0$, tj. $A = (\ldots, a_3, a_2, a_1, a_1, a_2, a_3, \ldots)$.

Obratno, za $A = (a_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$ se najprije definiraju brojevi $\alpha_d = [a_0, a_1, a_2, \ldots]$ i $\alpha_l = [a_{-1}, a_{-2}, a_{-3}, \ldots]$ koji su očito iracionalni jer su im verižni razlomci beskonačni. Zatim je ponovno lako vidjeti da je $L(A) = \max\{L(\alpha_d), L(\alpha_l)\}$. Iz ovih činjenica slijedi tvrdnja propozicije za Lagrangeov spektar.

Za Markovljev spektar tvrdnja slijedi direktno iz Teorema 2.12. Za potvrdu jednog smjera u dokazu treba samo primijetiti da je obostrano beskonačnom nizu $A = (a_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$ pridružena reducirana indefinitna forma $f_0(x, y) = (x - \vartheta_0 y)(x - \varphi_0 y)$, gdje je $\vartheta_0 = [0, a_0, a_1, a_2, \ldots]$ i $\varphi_0 = -[a_{-1}, a_{-2}, a_{-3}, \ldots]$. $\square$

U dokazima će nam često biti od koristi sljedeća lema.

Lema 3.2. *Neka je zadan obostrano beskonačan niz $A = (a_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$. Ako je $M(A)$ konačan, onda postoji obostrano beskonačan niz $B \in \mathbb{N}^{\mathbb{Z}}$ takav da je $M(A) = M(B) = \lambda_0(B)$. Ako je $L(A)$ konačan, onda postoji $C \in \mathbb{N}^{\mathbb{Z}}$ takav da je $L(A) = M(C) = \lambda_0(C)$.*

Dokaz. Brojevi $M(A)$ i $L(A)$ su konačni ako i samo ako je skup $\{a_i : i \in \mathbb{Z}\}$ ograničen, tj. postoji samo konačno mnogo vrijednosti koje elementi a_i mogu poprimiti.

Ako je već $\lambda_n(A) = M(A)$ za neki n, možemo uzeti da je $B = (b_i)_{i\in\mathbb{Z}}$ definiran sa $b_i = a_{i+n}$ za sve i, tj. B je dobiven pomakom niza A. U suprotnom imamo niz $(n_i)_{i\in\mathbb{N}}$ različitih brojeva u $\mathbb{Z}$ takav da je $\lambda_{n_i}(A) < M(A)$ te $\lambda_{n_i}(A) \to M(A)$ kad $i \to +\infty$.

Iz ograničenosti a_i-ova slijedi da postoji b_0 koji je jednak beskonačno mnogo a_{n_i}-ova. Nastavivši slično, za svaki prirodan broj k postoje brojevi b_{-k} i b_k takvi da je blok

$$b_{-k}, \ldots, b_{-1}, b_0, b_1, \ldots, b_k$$

jednak bloku

$$a_{n_i-k}, \ldots, a_{n_i-1}, a_{n_i}, a_{n_i+1}, \ldots, a_{n_i+k}$$

za beskonačno mnogo $i \in \mathbb{Z}$. Niz $B = (b_i)_{i\in\mathbb{Z}}$ konstruiran na ovaj način očito zadovoljava $M(A) = M(B) = \lambda_0(B)$. Ovdje se, kao i u nastavku, koristi Teorem 1.8.b.

Konstrukcija za Lagrangeov broj provodi se analogno te dobivamo $C = (c_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$ takav da je $L(A) = M(C) = \lambda_0(C)$. Treba primijetiti da bi postojanje nekog $k \in \mathbb{Z}$ takvog da je $\lambda_k(C) > \lambda_0(C)$ povlačilo postojanje beskonačno mnogo po volji dugačkih blokova u A koji se podudaraju sa $\ldots, c_{k-1}, c_k, c_{k+1}, \ldots$ što bi davalo beskonačno različitih i za koje je $\lambda_i(A) > \frac{\lambda_k(C)+\lambda_0(C)}{2} > \lambda_0(C)$, a to je u suprotnosti s činjenicom da je $\lambda_0(C) = L(A)$, tj. najveće gomilište od $(\lambda_i(A))_{i\in\mathbb{Z}}$. □

Korolar 3.3. *Neka je za svaki $j \in \mathbb{N}$ s $A_j = (a_i^{(j)})_{i\in\mathbb{Z}}$ označen obostrano beskonačan niz prirodnih brojeva. Ako je skup $\{a_i^{(j)} : i \in \mathbb{Z}, j \in \mathbb{N}\}$ omeđen, onda postoji obostrano beskonačan niz $B = (b_i)_{i\in\mathbb{Z}}$ takav da za svaki $m \in \mathbb{N}$ postoji beskonačan podniz $(j_k)_{k\in\mathbb{N}}$ skupa prirodnih brojeva takav da za svaki $k \in \mathbb{N}$ vrijedi $a_i^{(j_k)} = b_i$ za sve $-m \le i \le m$.*

Dokaz. Konstrukcija niza B već je dana u dokazu Leme 3.2. Naime, budući da $a_i^{(j)}$ imaju zajedničku ogradu, postoji beskonačno $j \in \mathbb{N}$ za koje je $a_0^{(j)}$ isti, nazovimo tu vrijednost b_0. Uzevši iz $(A_j)_{j\in\mathbb{N}}$ samo članove s tim j, a takvih je beskonačno mnogo, nastavljamo postupak i dobivamo $b_{-1}, b_1, b_{-2}, b_2, \ldots$ te tako niz $B = (b_i)_{i\in\mathbb{Z}}$.

Napomenimo da ako želimo dobiti podniz $A' = (A'_j)_{j\in\mathbb{N}}$ od $(A_j)_{j\in\mathbb{N}}$ koji konvergira po koordinatama prema B, treba u svakom koraku kod određivanja novog b_i dodati u A' i jedan element iz odgovarajućeg podniza od (A_j) te taj element u nastavku isključiti iz promatranja sve profinjenijih podnizova od (A_j). □

Lema 3.2 ponekad se zove Lema o kompaktnosti. Naime, označimo li s $\mathcal{A}$ konačan skup vrijednosti koje a_i-ovi u dokazu te leme poprimaju, možemo na $\mathcal{A}$ uvesti diskretnu topologiju (svaki podskup je otvoren), pa na $\mathcal{A}^{\mathbb{Z}}$ dobivamo odgovarajuću produktnu topologiju. Tvrdnja se onda svodi na činjenicu da je $\mathcal{A}^{\mathbb{Z}}$ kao prebrojivi produkt nizovno kompaktnih skupova ponovno nizovno kompaktan.

Lema o kompaktnosti može se iskazati i preko indefinitnih binarnih kvadratnih formi.

Lema 3.4. *Za svaku indefinitnu formu f, postoji forma g iste diskriminante takva da je $m(f) = m(g)$ i g primitivno reprezentira $m(g)$.*

Dokaz. Neka je formi f ekvivalentna forma f_0 koja je reducirana i $f_0(1,0) > 0$. Neka su, slično kao u (2.14), $[0, a_0, a_1, \ldots]$ i $-[a_{-1}, a_{-2}, a_{-3}, \ldots]$ korijeni od $f_0(x, 1)$. Time je definiran obostrano beskonačan niz $A = (a_i)_{i\in\mathbb{Z}}$. Prema Teoremu 2.12 je $\sqrt{d(f)}/m(f) = M(A)$. Lema 3.2 kaže da postoji $B = (b_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$ takav da je $M(A) = M(B) = \lambda_{-1}(B)$. Neka je $\kappa \in (0, +\infty)$ takav da $g(x, y) = \kappa(x - [0, b_0, b_1, \ldots]y)(x + [b_{-1}, b_{-2}, b_{-3}, \ldots]y)$ ima diskriminantu $d(f)$. Kao u (2.16), dobivamo sada

$$\frac{\sqrt{d(f)}}{m(f)} = M(A) = M(B) = \lambda_{-1}(B) = \frac{\sqrt{d(g)}}{g(1,0)} = \frac{\sqrt{d(g)}}{m(g)}.$$

te je $m(f) = m(g) = g(1, 0)$, pa je tražena forma g. □

Spomenimo da u nastavku radimo sa standardnom topologijom na skupu realnih brojeva.

Teorem 3.5. *Markovljev spektar je zatvoren skup.*

Dokaz. Neka je $(M(A_j))_{j\in\mathbb{N}}$ konvergentan niz u $\mathcal{M}$. Zbog leme o kompaktnosti možemo bez smanjenja općenitosti pretpostaviti da je $M(A_j) = \lambda_0(A_j)$ za svaki j. Iz konvergencije slijedi da svi $M(A_j)$ imaju zajedničku gornju ogradu, pa je ograničen i skup svih brojeva koji se pojavljuju u barem jednom A_j. Stoga iz Korolara 3.3 dobivamo obostrano beskonačan niz B. Reindeksirajući (A_j), možemo uzeti da je niz (A'_j) iz dokaza navedenog korolara upravo $(A_j)_{j\in\mathbb{N}}$. Iz Teorema 1.8.b i $M(A_j) = \lambda_0(A_j)$ dobivamo da za svaki $n \in \mathbb{Z}$ vrijedi

$$\lambda_n(B) = \lim_{j\to\infty} \lambda_n(A_j) \leq \lim_{j\to\infty} \lambda_0(A_j) = \lambda_0(B),$$

tj. $M(B) = \lambda_0(B) = \lim_{j\to\infty} M(A_j)$. □

Idući teorem slijedi direktno iz druge tvrdnje u Lemi 3.2.

Teorem 3.6. *Lagrangeov spektar je podskup Markovljevog spektra, tj. $\mathcal{L} \subset \mathcal{M}$.*

Pokazat ćemo kasnije da je inkluzija prava, tj. postoje brojevi koji jesu u Markovljevom, ali nisu u Lagrangeovom spektru.

Sada ćemo dati još preciznije opise spektara.

Teorem 3.7. *Neka je dan skup realnih brojeva*

$$\mathcal{P} = \{M(A) : A \text{ je čisto periodski obostrano beskonačan niz}\}.$$

Tada je $\mathcal{L} = \mathrm{cl}(\mathcal{P})$, gdje smo sa $\mathrm{cl}(\mathcal{P})$ označili (topološki) zatvarač od $\mathcal{P}$.

Primijetimo da u definiciji skupa $\mathcal{P}$ zbog periodičnosti nizova A imamo $M(A) = L(A)$.

Teorem 3.8. *Neka je dan skup realnih brojeva*

$$\begin{aligned}\mathcal{B} = \{M(A) : A &\text{ je obostrano beskonačan niz koji je} \\ &\text{od nekog mjesta nalijevo i od nekog mjesta nadesno periodski}\}.\end{aligned}$$

Tada je $\mathcal{M} = \mathrm{cl}(\mathcal{B})$.

Naglasimo da za obostrano beskonačni niz A iz definicije od $\mathcal{B}$ period koji se pojavljuje u lijevom dijelu i period u desnom dijelu niza A ne moraju biti isti.

Dokaz Teorema 3.7. Najprije dokazujemo $\mathcal{L} \subset \mathrm{cl}(\mathcal{P})$. Neka je A proizvoljan obostrano beskonačan niz za koji je $L(A)$ konačno. Za svaki $\varepsilon > 0$, konstruirat ćemo čisto periodski niz C takav da je $|L(C) - L(A)| < \varepsilon$.

Zbog Teorema 1.8.b postoji $m \in \mathbb{N}$ tako da je

$$\begin{aligned} \big|[0, a_{i+1}, \ldots, a_{i+m}, x] - [0, a_{i+1}, a_{i+2}, \ldots]\big| < \frac{\varepsilon}{4} \\ \big|[0, a_{i-1}, \ldots, a_{i-m}, x] - [0, a_{i-1}, a_{i-2}, \ldots]\big| < \frac{\varepsilon}{4} \end{aligned} \tag{3.2}$$

za proizvoljne $i \in \mathbb{Z}$ i $x \geq 1$. Po definiciji je $L(A) = \limsup_i \lambda_i(A)$, pa postoji $(k(j))_{j\in\mathbb{N}}$ takav da je $L(A) = \lim_{j\to\infty} \lambda_{k(j)}(A)$. Bez smanjenja općenitosti možemo pretpostaviti da $k(j+1) - k(j) \to +\infty$ kad $j \to \infty$ jer inače umjesto A promatramo obrnuti niz $(\ldots, a_2, a_1, a_0, a_{-1}, a_{-2}, \ldots)$. Definirajmo nizove $A_j = (a_i^{(j)})_{i\in\mathbb{Z}}$ sa $a_i^{(j)} = a_{k(j)+i}$. Uvrstivši $(A_j)_{j\in\mathbb{N}}$ i $2m$ u Korolar 3.3, dobivamo niz B i podniz od $(A_j)_{j\in\mathbb{N}}$ koji ćemo radi jednostavnosti zapisa jednako bilježiti. Neka je $J \in \mathbb{N}$ dovoljno velik tako da je

$$|\lambda_{k(j)}(A) - L(A)| < \frac{\varepsilon}{2} \quad \text{za sve } j \geq J \tag{3.3}$$

te vrijedi

(i) $k(J+1) - k(J) \geq 2m$;

(ii) za svaki $j \geq J$ je $a_{k(j)+i} = b_i$ za sve i takve da je $-2m \leq i \leq 2m$;

(iii) $\lambda_i(A) < L(A) + \frac{\varepsilon}{2}$ za sve $i \geq k(J)$.

Za $v = k(J+1) - k(J)$, definiramo čisto periodski niz $C \in \mathbb{N}^{\mathbb{Z}}$ duljine perioda v sa $c_i = a_{k(J)+i}$ za $0 \leq i < v$ te proširivanjem po periodičnosti za ostale $i \in \mathbb{Z}$.

Neka je $1 \leq i \leq m$. Imamo $0 < v - i < v$ jer je $v \geq 2m$, pa vrijedi

$$\begin{aligned} c_{-i} = c_{v-i} = a_{k(J)+v-i} = a_{k(J+1)-i} \overset{\text{(ii)}}{=} b_{-i} \overset{\text{(ii)}}{=} a_{k(J)-i}, \\ c_{v-1+i} = c_{i-1} = a_{k(J)+i-1} \overset{\text{(ii)}}{=} b_{i-1} \overset{\text{(ii)}}{=} a_{k(J+1)+i-1} = a_{k(J)+v-1+i}. \end{aligned}$$

Dakle, $c_i = a_{k(J)+i}$ za sve cijele brojeve i takve da je $-m \leq i \leq v + m - 1$. Neka je i proizvoljan cijeli broj te i_0 ostatak pri dijeljenju i sa v. Vrijedi $0 \leq i_0 < v$, pa iz upravo pokazanog i periodičnosti od C slijedi

$$\begin{aligned} \lambda_i(C) = \lambda_{i_0}(C) &= [c_{i_0}, c_{i_0+1}, \ldots, c_{i_0+m}, x] + [0, c_{i_0-1}, \ldots, c_{i_0-m}, y] \\ &= [a_{k(J)+i_0}, a_{k(J)+i_0+1}, \ldots, a_{k(J)+i_0+m}, x] + [0, a_{k(J)+i_0-1}, \ldots, a_{k(J)+i_0-m}, y], \end{aligned} \tag{3.4}$$

gdje su $x, y \in (1, +\infty)$. Iz (3.2) i (3.4) je $|\lambda_i(C) - \lambda_{k(J)+i_0}(A)| < \frac{\varepsilon}{2}$.

Sada iz (iii) slijedi

$$\lambda_i(C) < \lambda_{k(J)+i_0}(A) + \frac{\varepsilon}{2} < L(A) + \varepsilon$$

za sve $i \in \mathbb{Z}$. S druge stane, za $i = nv$, $n \in \mathbb{Z}$, je $i_0 = 0$, pa imamo iz $|\lambda_i(C) - \lambda_{k(J)}(A)| < \frac{\varepsilon}{2}$ i (3.3) da je $|\lambda_i(C) - L(A)| < \varepsilon$. Time smo pokazali $|L(C) - L(A)| < \varepsilon$.

Kako bismo dokazali obratnu inkluziju $\mathrm{cl}(\mathcal{P}) \subset \mathcal{L}$, pretpostavimo da je $(A_j)_{j\in\mathbb{N}}$ niz čisto periodskih obostrano beskonačnih nizova tako da je $\lim_{j\to\infty} L(A_j) = L$. Zbog periodičnosti od A_j možemo pretpostaviti da je $L(A_j) = M(A_j) = \lambda_0(A_j)$ za svaki $j \in \mathbb{N}$.

Neka je $B = (b_i)_{i\in\mathbb{Z}}$ niz koji dobijemo iz Korolara 3.3 te, ako je potrebno, zamijenimo $(A_j)_{j\in\mathbb{N}}$ s nekim podnizom tako da za svaki $m \in \mathbb{N}$ postoji $J \in \mathbb{N}$ pri čemu $j \geq J$ povlači

$$a_i^{(j)} = b_i \quad \text{za sve } i \in \mathbb{Z} \text{ takve da je } -m \leq i \leq m. \tag{3.5}$$

Ako je p_j duljina perioda od A_j, možemo dodatno pretpostaviti da je $p_{j+1} \geq p_j$ za svaki $j \in \mathbb{N}$. Definiramo funkciju $f : \mathbb{N}_0 \to \mathbb{N}_0$ sa $f(0) = 0$ i $f(j) = p_1 + p_2 + \cdots + p_j$ za $j \in \mathbb{N}$. Definiramo niz $C = (c_k)_{k\in\mathbb{Z}}$ sa $c_{f(j)+i} = a_i^{(j+1)}$ za $j \geq 0$, $0 \leq i < p_{j+1}$ te $c_k = c_{-k}$ za $k < 0$. Pokazat ćemo da je $L(C) = L$.

Uzmimo $\varepsilon > 0$ po volji te neka je $m \in \mathbb{N}$ takav da je $\frac{\varepsilon}{4} > \frac{1}{2^{m-1}}$. Neka je $J \in \mathbb{N}$ takav da vrijedi (3.5) i $p_j > m$ za $j \geq J$.

Uzmimo $j \geq J$, $0 \leq i < p_{j+1}$, $-m \leq r \leq m$. Tada je $-m \leq i + r \leq p_{j+1} + m - 1$. Ako je $0 \leq i + r < p_{j+1}$, iz definicije od C je $c_{f(j)+i+r} = a_{i+r}^{(j+1)}$.

Za $-m \leq i + r \leq -1$ je $0 < p_j + i + r < p_j$, pa imamo

$$c_{f(j)+i+r} = c_{f(j-1)+p_j+i+r} = a_{p_j+i+r}^{(j)} = a_{i+r}^{(j)} = b_{i+r} = a_{i+r}^{(j+1)}.$$

Za $p_{j+1} \leq i + r \leq p_{j+1} + m - 1$ je $0 \leq i + r - p_{j+1} \leq m - 1 < p_{j+2}$, pa imamo

$$c_{f(j)+i+r} = c_{f(j+1)+i+r-p_{j+1}} = a_{i+r-p_{j+1}}^{(j+2)} = b_{i+r-p_{j+1}} = a_{i+r-p_{j+1}}^{(j+1)} = a_{i+r}^{(j+1)}.$$

Iz Teorema 1.8.b slijedi da je za $j \geq J$

$$-\frac{\varepsilon}{2} < \lambda_{f(j)+i}(C) - \lambda_i(A_{j+1}) < \frac{\varepsilon}{2} \quad \text{za sve } i,\ 0 \leq i < p_{j+1}. \tag{3.6}$$

Budući je $L = \lim_{j\to\infty} L(A_j) = \lim_{j\to\infty} \lambda_0(A_j)$, možemo, ako treba, J još povećati tako da je za sve $j \geq J$

$$-\frac{\varepsilon}{2} < \lambda_0(A_j) - L < \frac{\varepsilon}{2}. \tag{3.7}$$

Iz (3.6) za $j = k \geq J$, $i = 0$ te (3.7) za $j = k + 1$, dobivamo

$$\lambda_{f(k)}(C) > \lambda_0(A_{k+1}) - \frac{\varepsilon}{2} > L - \varepsilon. \tag{3.8}$$

S druge strane, $M(A_j) = \lambda_0(A_j)$ za sve j povlači da za sve $j \geq J$ i sve $i \in \mathbb{Z}$ vrijedi $\lambda_i(A_j) \leq \lambda_0(A_j)$, pa je po (3.7), $\lambda_i(A_j) < L + \frac{\varepsilon}{2}$ iz čega zbog (3.6) slijedi $\lambda_n(C) < L + \varepsilon$ za $n \geq f(J)$, odnosno zbog simetričnosti C, za $|n| \geq f(J)$. Ovo nam zajedno sa (3.8) daje $|L(C) - L| < \varepsilon$, a kako je $\varepsilon > 0$ bio proizvoljan, dobivamo konačno $L = L(C)$ čime je dokaz dovršen. □

Dokaz Teorema 3.8. Inkluzija $\mathrm{cl}(\mathcal{B}) \subset \mathcal{M}$ slijedi iz $\mathcal{B} \subset \mathcal{M}$ te činjenice da je po Teoremu 3.5 skup $\mathcal{M}$ zatvoren.

Da bismo dokazali obratnu inkluziju, promatramo proizvoljni $A \in \mathbb{N}^{\mathbb{Z}}$ takav da je $M(A) = \lambda_0(A)$ te ćemo za $\varepsilon > 0$ po volji konstruirati eventualno periodski niz $C \in \mathbb{N}^{\mathbb{Z}}$ (tj. periodski od nekog indeksa nalijevo i od nekog indeksa nadesno) za koji je $|M(C) - M(A)| < \varepsilon$. Konstrukcija je slična kao u prvom dijelu dokaza prethodnog teorema.

Prema Teoremu 1.8.b postoji $m \in \mathbb{N}$ takav da je za svaki $i \in \mathbb{Z}$ te proizvoljne $x, y \in (1, +\infty)$

$$\big|[a_i, a_{i+1}, \ldots, a_{i+m}, x] + [0, a_{i-1}, \ldots, a_{i-m}, y] - \lambda_i(A)\big| < \varepsilon. \tag{3.9}$$

Koristeći Korolar 3.3 dobivamo niz $(k(i,+))_{i\in\mathbb{N}}$ prirodnih brojeva takav da je

(i) $k(1,+) \geq m$ i $k(i+1,+) - k(i,+) \geq 2m$

te obostrano beskonačan niz $B^+ = (b_i^+)_{i\in\mathbb{Z}} \in \mathbb{N}^\mathbb{Z}$ i prirodan broj J_+ takve da

(ii) za svaki $j \geq J_+$ je $a_{k(j,+)+i} = b_i^+$ za sve $|i| \leq m$.

Na isti način dobivamo niz $(k(i,-))_{i\in\mathbb{N}}$ negativnih brojeva, niz $B^- = (b_i^-)_{i\in\mathbb{Z}} \in \mathbb{N}^\mathbb{Z}$ i prirodan broj J_- za koje vrijede odgovarajuće modifikacije od (i) i (ii).

Za $J = \max\{J_+, J_-\}$ definiramo (jednostrano beskonačne) čisto periodske nizove $C^+, C^- \in \mathbb{N}^\mathbb{N}$ s periodima redom

$$a_{k(J,+)+1}, a_{k(J,+)+2}, \dots, a_{k(J+1,+)} \qquad \text{i} \qquad a_{k(J,-)-1}, a_{k(J,-)-2}, \dots, a_{k(J+1,-)}$$

s duljinama perioda $v^+ = k(J+1,+) - k(J,+)$ i $v^- = k(J,-) - k(J+1,-)$, respektivno. Primijetimo da je zbog (i), $v^+, v^- \geq 2m$. Sada definiramo traženi niz

$$C = (c_i)_{i\in\mathbb{Z}} = \widetilde{C^-}, a_{k(J,-)}, \dots, a_{-1}, a_0, a_1, \dots, a_{k(J,+)}, C^+,$$

gdje smo sa $\widetilde{\cdot}$ označili obrnuti niz. Dokazat ćemo da je $|M(C) - M(A)| < \varepsilon$.

Budući je $k(J, \pm) \geq m$, (3.9) povlači da je

$$\lambda_0(C) > \lambda_0(A) - \varepsilon = M(A) - \varepsilon.$$

Pokazat ćemo da je

$$\lambda_i(C) < M(A) + \varepsilon \quad \text{za svaki } i \geq 0, \tag{3.10}$$

a analognim zaključivanjem dobiva se da (3.10) vrijedi za $i < 0$.

Iz definicije od C je $c_i = a_i$ za $k(J,-) \leq i \leq k(J,+)$.

Iz definicije od C i C^+ je $c_{k(J,+)+i} = c_i^+ = a_{k(J,+)+i}$ za $1 \leq i \leq v^+$.

Svojstvo (ii) i $v^+ > m$ povlači da je za $1 \leq i \leq m$

$$c_{k(J+1,+)+i} = c_{k(J+1,+)+i-v^+} = c_{k(J,+)+i} = a_{k(J,+)+i} = b_i^+ = a_{k(J+1,+)+i}.$$

Stoga je $c_i = a_i$ za sve i takve da je $k(J,-) \leq i \leq k(J+1,+) + m$, pa (3.9) daje

$$\lambda_i(C) < \lambda_i(A) + \varepsilon \quad \text{za} \quad 0 \leq i \leq k(J+1,+) = k(J,+) + v^+.$$

Za $i > k(J,+) + v^+$ zbog $v^+ > m$ dobivamo

$$\lambda_i(C) = [c_i^+, \dots, c_{i+m}^+, x] + [0, c_{i-1}^+, \dots, c_{i-m}^+, y]$$

za neke $x, y \in (1, +\infty)$. Slično kao što smo pokazali (3.4) u dokazu Teorema 3.7, dobivamo i ovdje

$$\lambda_i(C) = [a_n, a_{n+1}, \dots, a_{n+m}, x] + [0, a_{n-1}, \dots, a_{n-m}, y]$$

za neki $n \in \mathbb{Z}$ i neke $x, y \in (1, +\infty)$. Zato iz (3.9) slijedi $\lambda_i(C) < \lambda_n(A) + \varepsilon \leq M(A) + \varepsilon$. Time je završen dokaz (3.10), pa stoga i dokaz teorema. □

Teorem 3.9. *Obostrano beskonačni niz $A \in \mathbb{N}^\mathbb{Z}$ je čisto periodski ako i samo ako postoje tri različita indeksa $k, m, n \in \mathbb{Z}$ takva da su $\dfrac{\lambda_m(A)}{\lambda_k(A)}$ i $\dfrac{\lambda_n(A)}{\lambda_k(A)}$ racionalni brojevi.*

Dokaz. Ako je A čisto periodski očito postoje $k, m, n \in \mathbb{Z}$ takvi da je $\lambda_k(A) = \lambda_m(A) = \lambda_n(A)$, pa vrijedi tvrdnja teorema.

Pretpostavimo da je dan $A = (\ell_i)_{i\in\mathbb{Z}}$ takav da su $\lambda_m(A)/\lambda_k(A)$ i $\lambda_n(A)/\lambda_k(A)$ racionalni brojevi za neke različite $k, m, n \in \mathbb{Z}$. Definiramo kvadratnu formu $f_0(x, y) = (x - \vartheta_0 y)(x - \varphi_0 y)$, gdje je $\vartheta_0 = [0, \ell_0, \ell_1, \ldots]$ i $\varphi_0 = -[\ell_{-1}, \ell_{-2}, \ell_{-3}, \ldots]$. Zbog $\vartheta_0 \in (0, 1)$ i $\varphi_0 = (-\infty, -1)$, to je reducirana indefinitna forma s diskriminantom $d > 0$ te stoga pripada nekom lancu formi $(f_i)_{i\in\mathbb{Z}}$. Usporedbom s jednakostima (2.14) i (2.16) vidimo da je

$$\lambda_i(A) = |\vartheta_{i+1} - \varphi_{i+1}| = \frac{\sqrt{d}}{|f_{i+1}(1, 0)|}$$

za sve cijele brojeve i. Prema pretpostavci su sada

$$\frac{f_{m+1}(1, 0)}{f_{k+1}(1, 0)} = \pm\frac{\lambda_k(A)}{\lambda_m(A)} \quad \text{i} \quad \frac{f_{n+1}(1, 0)}{f_{k+1}(1, 0)} = \pm\frac{\lambda_k(A)}{\lambda_n(A)}$$

racionalni brojevi, pa iz Teorema 2.15 slijedi da je lanac formi $(f_i)_{i\in\mathbb{Z}}$ periodski iz čega zaključujemo da je i A čisto periodski. $\square$

Korolar 3.10. *Obostrano beskonačni niz $A \in \mathbb{N}^{\mathbb{Z}}$ je čisto periodski ako i samo ako postoje tri različita indeksa $k, m, n \in \mathbb{Z}$ takva da je $\lambda_k(A) = \lambda_m(A) = \lambda_n(A)$.*

Broj tri u prethodnom korolaru se ne može smanjiti na dva što vidimo npr. iz niza $A \in \mathbb{N}^{\mathbb{Z}}$ definiranog sa $a_{-1} = a_1 = 3$, $a_i = 1$ inače, tj. $A = \overline{1}\,3\,1\,3\,\overline{1}$. Očito je $M(A) = \lambda_{-1}(A) = \lambda_1(A)$, ali A nije čisto periodski. Ako želimo primjer niza koji nije periodski ni ulijevo ni udesno, možemo primjerice redefinirati $a_{9^k} = a_{9^k+1} = a_{-9^k} = a_{-9^k-1} = 2$ za $k \in \mathbb{N}$. I za takav niz nije teško pokazati da je $M(A) = \lambda_{-1}(A) = \lambda_1(A)$.

Poglavlje 4

Elementi spektara manji od 3

Slijedeći izlaganje prema Reutenauerovoj novoj knjizi, dokazat ćemo u ovom opširnijem poglavlju Markovljev teorem iz kojega proizlazi da se Lagrangeov i Markovljev spektar u intervalu $(0, 3)$ podudaraju i jednaki su

$$\Big\{\sqrt{9 - \frac{4}{m^2}} \ : \ m \in \mathbb{N} \text{ za koji postoje } m_1, m_2 \in \mathbb{Z} \text{ tako da je } m^2 + m_1^2 + m_2^2 = 3mm_1m_2\Big\}.$$

Za $m = 1$ i $m = 2$ dobivamo najmanje elemente ovog skupa, $\sqrt{5} = 2.236\ldots$ i $\sqrt{8} = 2.828\ldots$ koje smo već susreli, primjerice u Korolaru 1.14. Idući element dobivamo za $m = 5$ i jednak je $\sqrt{221/25} = 2.973\ldots$.

4.1 Definicije o riječima, monoidima i grupama

Budući da ćemo u nastavku dosta koristiti kombinatoriku na riječima, uvodimo ovdje neke osnovne definicije. Skup A nazivamo *alfabet*. *Slovo* je element alfabeta. *Riječ* je konačan niz slova, uključujući i prazan niz koji zovemo *prazna riječ*. *Duljina* riječi w je duljina niza, označavamo je $|w|$. Za dano slovo $a \in A$, sa $|w|_a$ označavamo broj pojavljivanja slova a u riječi w.

Slobodni monoid na A je skup svih riječi na A i označavamo ga A^*. Skup A^* je monoid pri čemu je množenje dano *spajanjem* riječi: iz riječi u i v dobivamo uv tako da najprije napišemo u, a odmah zatim v. Neutralni element ovog monoida je prazna riječ, pa ju stoga označavamo 1, a ako postoji opasnost od zabune (primjerice kada je $A = \mathbb{N}$ te je 1 slovo), označavamo ju ε. Ovaj je monoid *slobodan* na A što znači da svaka funkcija s A u neki monoid ima jedinstveno proširenje do homomorfizma monoida s A^* u dani monoid.

Za riječi w, u, v takve da je $w = uv$ govorimo o *faktorizaciji* od w, ta je faktorizacija netrivijalna ako su i u i v neprazne riječi.

Ako su riječi m, u, v, w takve da je $w = umv$, kažemo da je u *prefiks* od w, v *sufiks* od w, a m *faktor* od w; kažemo da je *pravi* ako je različit od od w.

Dvije riječi w, w' su *konjugirane* ako postoje riječi u, v takve da je $w = uv$ i $w' = vu$. *Konjugacijska klasa* neke riječi je skup svih njoj konjugiranih riječi, tj. skup svih njezinih *konjugata*.

Obratna riječ riječi $w = a_1 \ldots a_n$, $a_i \in A$ je $\widetilde{w} = a_n \ldots a_1$. Riječ w je *palindrom* ako je $w = \widetilde{w}$. Ako su dvije riječi konjugirane, onda su i njima obratne riječi konjugirane, to dobivamo iz jednakosti $\widetilde{uv} = \widetilde{v}\,\widetilde{u}$ i $\widetilde{vu} = \widetilde{u}\,\widetilde{v}$.

Komutativna slika riječi w je A-torka $(|w|_a)_{a \in A}$, tj. slika od w s obzirom na kanonski homomorfizam monoida sa slobodnog monoida A^* na slobodni komutativni monoid ${\mathbb{N}_0}^A$.

Niz s elementima iz A još zovemo *beskonačna riječ*. Za takav niz $(a_n)_{n\in\mathbb{N}_0}$ kažemo da je *periodski* ako za neki $p \in \mathbb{N}$ i neki $i \in \mathbb{N}$ vrijedi $a_n = a_{n+p}$ čim je $n \geq i$. Takav niz je *čisto periodski* ako zahtijevamo da je $i = 0$. Ako je s periodski kao u definiciji, svaki konačan faktor od s oblika $a_n \dots a_{n+p-1}$ za $i \geq i$ zovemo *periodski uzorak* ili periodski blok od s.

Možemo pomnožiti konačnu i beskonačnu riječ, ali samo u tom poretku, spajanjem riječi kao prije. Prefiksi i faktori beskonačnih riječi su konačne riječi i definiraju se na očit način.

Ako je $w = a_0 \dots a_{k-1}$ neprazna riječi, gdje su $a_i \in A$, onda s w^∞ označavamo beskonačnu riječ $(b_n)_{n\in\mathbb{N}_0}$ takvu da za svaki $n \in \mathbb{N}_0$ imamo $b_n = a_r$, gdje je r ostatak pri cjelobrojnom dijeljenju n sa k. Riječ w^∞ ponekad pišemo $www \dots w \dots$, a prije smo je pisali $\overline{w}$.

Obostrano beskonačna riječ je niz $(a_n)_{n\in\mathbb{Z}}$ indeksiran skupom cijelih brojeva. Kažemo da je *(čisto) periodski* ako za neki $p \in \mathbb{N}$ i svaki $n \in \mathbb{Z}$ vrijedi $a_n = a_{n+p}$. Za nepraznu konačnu riječ $w = a_0 \dots a_{k-1}$, sa ${}^\infty w^\infty$ označavamo periodsku obostrano beskonačnu riječ $(b_n)_{n\in\mathbb{Z}}$ definiranu za svaki $n \in \mathbb{Z}$ sa $b_n = a_r$, gdje je r ostatak pri cjelobrojnom dijeljenju n sa k.

Sa $F(a, b)$ označavamo *slobodnu grupu* s dva generatora a, b. Svaki element slobodne grupe može se prikazati kao *reducirana riječ*, tj. produkt generatora i njihovih inverza bez faktora xx^{-1} i $x^{-1}x$, gdje je $x \in \{a, b\}$. Slobodni monoid $\{a, b\}^*$ je podmonoid od $F(a, b)$ budući da je svaka riječ u $\{a, b\}^*$ nužno reducirana. Postoji kanonski epimorfizam grupa sa $F(a, b)$ na $\mathbb{Z}^2 = \mathbb{Z}^{\{a,b\}}$, sliku elementa od $F(a, b)$ zovemo njegovom *komutativnom slikom* te se ta terminologija slaže s onom koju smo uveli za slobodni monoid.

Korisno je prikazati endomorfizam slobodnog monoida $\{a, b\}^*$ ili slobodne grupe $F(a, b)$ pomoću uređenog para (u, v), gdje se $a \mapsto u$, $b \mapsto v$ s obzirom na dani endomofizam. Ovaj je prikaz kompatibilan s kanonskim ulaganjem monoida endomorfizama slobodnog monoida $\{a, b\}^*$ u monoid endomorfizama slobodne grupe $F(a, b)$. *Abelizacija* takvog endomorfizma φ je 2×2 matrica $\mathrm{Ab}(\varphi) = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right]$, gdje je (p, r) komutativna slika od u, a (q, s) komutativna slika od v. Ako je $h = \varphi(g)$, komutativna slike (i, j) od g je vezana s komutativnom slikom (k, l) od h preko

$$\begin{bmatrix} k \\ l \end{bmatrix} = \begin{bmatrix} p & q \\ r & s \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix}.$$

Abelizacijsko preslikavanje Ab je homomorfizam monoida s monoida $\mathrm{End}(F(a, b))$ endomorfizama slobodne grupe $F(a, b)$ u multiplikativni monoid matrica $\mathbb{Z}^{2\times 2}$. Iz toga slijedi da ako je takav endomorfizam invertibilan (tj. ako je automorfizam od $F(a, b)$), onda je odgovarajuća matrica invertibilna, pa pripada $\mathrm{GL}(2, \mathbb{Z})$.

Sada će nam trebati jedan standardni rezultat.

Propozicija 4.1. *Matrice* $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ *i* $\left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$ *generiraju grupu* $\mathrm{SL}(2, \mathbb{Z})$. *Matrice* $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ *i* $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ *generiraju grupu* $\mathrm{GL}(2, \mathbb{Z})$.

Dokaz. Neka je $U = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$, $V = \left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]$, $J = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$.

Pokažimo da U i V generiraju $\mathrm{SL}(2, \mathbb{Z})$, tj. svaku matricu $M \in \mathrm{SL}(2, \mathbb{Z})$ može se izraziti kao riječ u $U^{\pm 1}$ i $V^{\pm 1}$. U nastavku ćemo elemente q i s od $M = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right] \in \mathrm{SL}(2, \mathbb{Z})$ označiti redom sa $q(M)$ i $s(M)$. Neka je $T = UV^{-1}U = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right] \in \mathrm{SL}(2, \mathbb{Z})$.

Ako je $q = 0$, onda je $p = s = \pm 1$, pa je ili $M = \left[\begin{smallmatrix} 1 & 0 \\ r & 1 \end{smallmatrix}\right] = V^r$ ili $M = \left[\begin{smallmatrix} -1 & 0 \\ r & -1 \end{smallmatrix}\right] = -I \cdot V^{-r} = T^2 V^{-r}$. Stoga se M može prikazati kao riječ u $U^{\pm 1}$ i $V^{\pm 1}$.

Ako je $s = 0$, onda je $qr = -1$. Ili je $q = -r = 1$ i onda $M = \left[\begin{smallmatrix} p & 1 \\ -1 & 0 \end{smallmatrix}\right] = U^{-p}T$ ili je $q = -r = -1$ i onda $M = \left[\begin{smallmatrix} p & -1 \\ 1 & 0 \end{smallmatrix}\right] = U^pT^3$. U oba slučaja se M može prikazati kao riječ u $U^{\pm 1}$ i $V^{\pm 1}$.

Pretpostavimo sada da ni $q = q(M)$ ni $s = s(M)$ nisu nula. Vrijedi

$$q(UM) = q(\begin{bmatrix} p+r & q+s \\ r & s \end{bmatrix}) = q(M) + s(M) \quad \text{i} \quad s(UM) = s(M), \tag{4.1}$$

$$q(TM) = q(\begin{bmatrix} r & s \\ -p & -q \end{bmatrix}) = s(M) \quad \text{i} \quad s(TM) = -q(M). \tag{4.2}$$

Iz (4.1) slijedi da množenjem M slijeva odgovarajućom pozitivnom ili negativnom potencijom od U dobivamo matricu U^nM takvu da je $0 \leq |q(U^nM)| < |s(U^nM)|$. Zbog (4.2) vidimo da možemo zamijeniti uloge $\pm q$ i $\pm s$ tako da matricu pomnožimo slijeva sa T. Na taj način možemo smanjivati apsolutne vrijednosti od q i s dok jedna od njih ne iščezne. Dakle, množeći M slijeva potencijama od U i T, možemo svesti dokaz na slučaj $q = 0$ ili $s = 0$ koji smo već obradili. Zaključujemo da U i V generiraju grupu $\mathrm{SL}(2,\mathbb{Z})$.

Homomorfizam $\det : \mathrm{GL}(2,\mathbb{Z}) \to \{1, -1\}$ nam govori da je jezgra $\mathrm{SL}(2,\mathbb{Z})$ normalna podgrupa od $\mathrm{GL}(2,\mathbb{Z})$ indeksa 2. Stoga je $GL(2,\mathbb{Z}) = \mathrm{SL}(2,\mathbb{Z}) \cup J \cdot \mathrm{SL}(2,\mathbb{Z})$ te U, V i J generiraju $\mathrm{GL}(2,\mathbb{Z})$. No, $V = JUJ$, pa U i J generiraju $\mathrm{GL}(2,\mathbb{Z})$. □

Propozicija 4.2. *Slika restrikcije abelizacije* Ab *na grupu* $\mathrm{Aut}(F(a,b))$ *automorfizama od* $F(a,b)$ *je grupa* $\mathrm{GL}(2,\mathbb{Z})$.

Dokaz. Kao što smo vidjeli, slika navedene restrikcije je sadržana u $\mathrm{GL}(2,\mathbb{Z})$. No, $\left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ i $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ su redom abelizacije automorfizama (a, ab) i (b, a) od $F(a,b)$, pa su sadržane u slici. Zato iz Propozicije 4.1 zaključujemo da je abelizacijsko preslikavanje s $\mathrm{Aut}(F(a,b))$ u $\mathrm{GL}(2,\mathbb{Z})$ surjektivno. □

Napravimo kratku digresiju za alternativni dokaz težeg dijela Serretovog teorema koji je dao Bombieri.

Dokaz nužnosti u Teoremu 1.1. Neka su α i β ekvivalentni iracionalni brojevi, tj. $\beta = M\alpha$, $M \in \mathrm{GL}(2,\mathbb{Z})$. Treba pokazati da se verižni razlomak od β i verižni razlomak od α podudaraju od nekog mjesta u prvom i nekog mjesta u drugom. Iz Propozicije 4.1 vidimo kako je dovoljno provjeriti da navedena tvrdnja vrijedi ako je $M \in \{U^{\pm 1}, J\}$, gdje su $U = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ i $J = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. Primijetimo da je $J^{-1} = J$.

Neka je $\alpha = [a_0, a_1, a_2, \ldots]$. Imamo $U\alpha = \alpha + 1 = [a_0 + 1, a_1, a_2, \ldots]$ i $U^{-1}\alpha = \alpha - 1 = [a_0 - 1, a_1, a_2, \ldots]$, pa tvrdnja vrijedi za U i U^{-1}. Za matricu J tvrdnju pokazujemo kroz niz slučajeva koji pokrivaju sve moguće α. Imamo $J\alpha = \frac{1}{\alpha}$ te vrijedi

$$\begin{aligned}
&J[a_0, \gamma] = [0, a_0, \gamma] && (a_0 \geq 1) \\
&J[0, \gamma] = [\gamma] && \\
&J[-1, a_1, \gamma] = [-2, 1, a_1 - 2, \gamma] && (a_1 > 2) \\
&J[-1, 2, a_2, \gamma] = [-2, a_2 + 1, \gamma] && \\
&J[-1, 1, a_2, a_3, \gamma] = [-a_2 - 2, 1, a_3 - 1, \gamma] && (a_3 > 1) \\
&J[-1, 1, a_2, 1, a_4, \gamma] = [-a_2 - 2, a_4 + 1, \gamma] && \\
&J[-2, a_1, \gamma] = [-1, 2, a_1 - 1, \gamma] && (a_1 > 1) \\
&J[-2, 1, a_2, \gamma] = [-1, a_2 + 2, \gamma] && \\
&J[a_0, a_1, \gamma] = [-1, 1, -a_0 - 2, 1, a_1 - 1, \gamma] && (a_0 \leq -3,\ a_1 > 1) \\
&J[a_0, 1, a_2, \gamma] = [-1, 1, -a_0 - 2, a_2 + 1, \gamma] && (a_0 \leq -3)
\end{aligned}$$

Lako je provjeriti ove jednakosti korištenjem (1.3). Pokažimo to, primjerice, za zadnji identitet.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a_0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a_2 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \gamma & 1 \\ 1 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} 1+\gamma+a_2\gamma & 1+a_2 \\ 1+a_0+a_0\gamma+a_2\gamma+a_0a_2\gamma & a_0+a_2+a_0a_2 \end{bmatrix},$$

$$\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} -a_0-2 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a_2+1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \gamma & 1 \\ 1 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} -1-\gamma-a_2\gamma & -1-a_2 \\ -1-a_0-a_0\gamma-a_2\gamma-a_0a_2\gamma & -a_0-a_2-a_0a_2 \end{bmatrix}.$$

Kvocijent elementa u prvom stupcu je isti za obje matrice, pa traženi identitet vrijedi. Analogno se pokazuju i ostali slučajevi. □

4.2 Riječi

Kod definiranja i proučavanja svojstava Christoffelovih riječi koristit ćemo geometrijski pristup koji omogućuje elegantnije dokaze potrebnih tvrdnji.

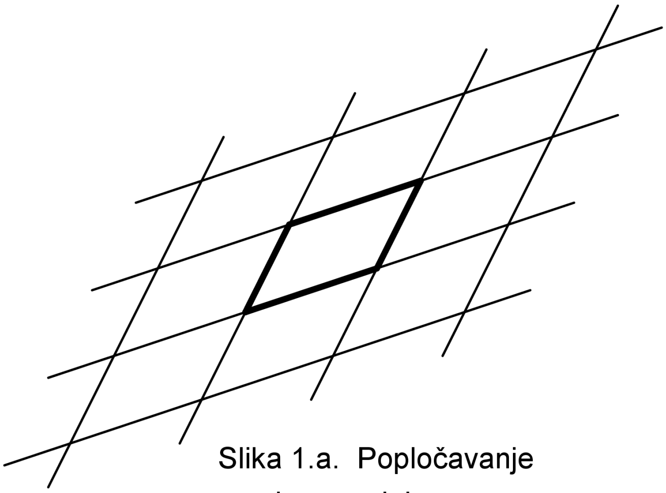
4.2.1 Popločavanje ravnine paralelogramom

Ravnina je Kartezijeva ravnina $\mathbb{R}^2$. *Cjelobrojna rešetka* ili *diskretna ravnina* je skup točaka s cjelobrojnim koordinatama, tj. skup $\mathbb{Z}^2$. Element cjelobrojne rešetke zove se *cjelobrojna točka* ili *cjelobrojni vektor.*

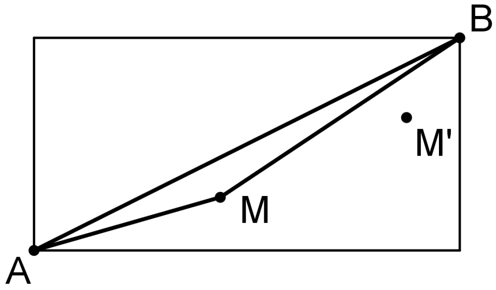
Baza od $\mathbb{Z}^2$ je par cjelobrojnih vektora takvih da je svaki vektor iz $\mathbb{Z}^2$ jedinstvena linearna kombinacija s koeficijentima iz $\mathbb{Z}$ tih dvaju vektora. Lako se vidi da dva cjelobrojna vektora $u = (a, b)$ i $v = (c, d)$ čine bazu od $\mathbb{Z}^2$ ako i samo ako je matrica $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ iz $\mathrm{GL}(2, \mathbb{Z})$, tj. determinanta joj je ± 1. Tu ćemo determinantu označavati $\det(u, v)$.

Promotrimo paralelogram u ravnini, pretpostavljamo da je nedegeneriran, tj. nije sadržan u jednom pravcu. *Vektor stranice* paralelograma je bilo koji od četiri moguća vektora koje dobivamo izabirući stranicu i jednu od orijentacija. Svaka dva vektora stranica koji odgovaraju susjednim stranicama čine bazu od $\mathbb{R}^2$ promatranog kao realnog vektorskog prostora. Paralelogram određuje popločavanje ravnine koje dobivamo tako da ga uzastopce translatiramo za vektore stranica, vidi Sliku 1.a. Pretpostavimo li da je jedan od vrhova paralelograma u ishodištu, vidimo da je aditivna podgrupa od $\mathbb{R}^2$ generirana vektorima stranica tog paralelograma upravo skup svih vrhova svih paralelograma u popločavanju. *Vektor stranice* trokuta definira se analogno.

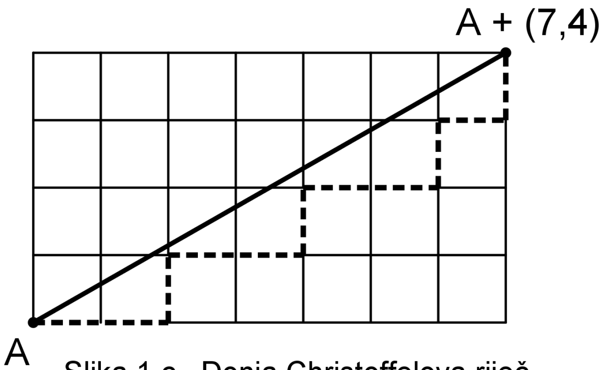
Primitivni trokut (odnosno *primitivni paralelogram*) je proizvoljan trokut (odnosno paralelogram) kojemu su vrhovi cjelobrojne točke i ne sadrži druge cjelobrojne točke (gledajući i na stranicama). Ako uzmemo neki paralelogram s cjelobrojnim vrhovima, njegovo središte je $v \in \mathbb{R}^2$ tako da je $2v \in \mathbb{Z}^2$. Centralna simetrija s obzirom na to središte preslikava vektor $u \in \mathbb{Z}^2$ u $2v-u$, pa zaključujemo da simetrija s obzirom na središte preslikava cjelobrojnu rešetku u nju samu. Ako, dakle, presiječemo paralelogram s cjelobrojnim vrhovima jednom od dijagonala, dobiveni trokuti su istovremeno primitivni ili istovremeno nisu primitivni, a primitivni su ako i samo ako je sam paralelogram primitivan.



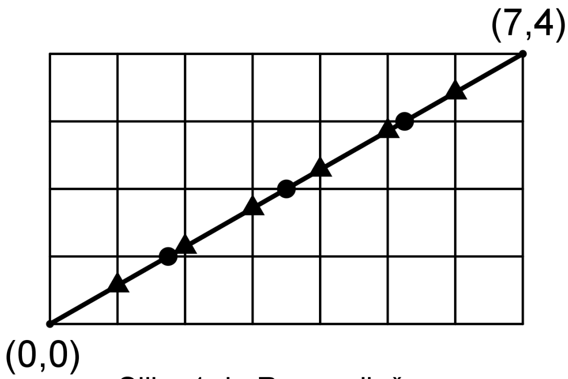
Slika 1.a. Popločavanje ravnine paralelogramom



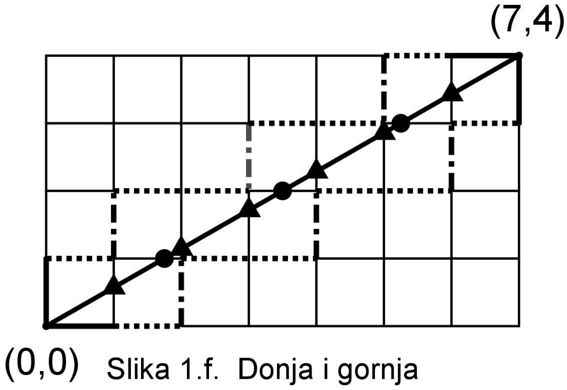
Slika 1.b. Jedinstvena točka M takva da je AMB primitivni trokut



Slika 1.c. Donja Christoffelova riječ aabaabaabab nagiba 4/7

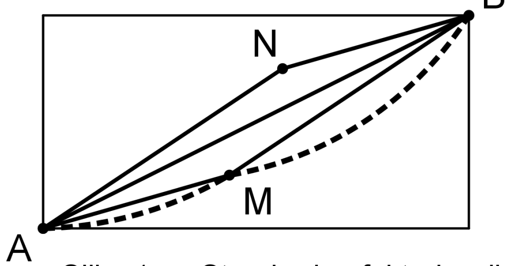


Slika 1.d. Rezna riječ abaabaaba pridružena 4/7

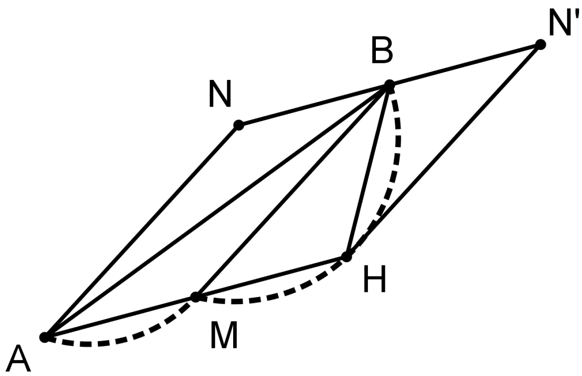


Slika 1.f. Donja i gornja Christoffelova riječ te rezna riječ

Slika 1.e. Bijekcija između koraka i sjecišta



Slika 1.g. Standardna faktorizacija



Slika 1.h. u je prefiks od v

Paralelogram s cjelobrojnim vrhovima je primitivan ako i samo ako bilo koja dva vektora koji odgovaraju njegovim susjednim stranicama čine bazu od $\mathbb{Z}^2$. Geometrijski se ta tvrdnja dokazuje tako da promatramo popločavanje ravnine danim paralelogramom kao što smo prije opisali. Svi su paralelogrami u tom popločavanju primitivni ako i samo ako je jedan od njih primitivan, a u tom slučaju nema cjelobrojnih točaka izvan skupa vrhova svih tih paralelograma. Stoga je podgrupa generirana vektorima susjednih stranica danog paralelograma jednaka $\mathbb{Z}^2$ ako i samo ako je taj paralelogram primitivan. U tom slučaju ta dva vektora razapinju čitavu cjelobrojnu rešetku te čine bazu jer su linearno nezavisni.

Iz ovoga slijedi da je trokut s cjelobrojnim vrhovima primitivan ako i samo ako bilo koja dva vektora stranica čine bazu od $\mathbb{Z}^2$.

Također možemo zaključiti da je za cjelobrojne vektore u i v, paralelogram s vrhovima 0, u, v, $u+v$ primitivan ako i samo ako je $\det(u,v) = \pm 1$. Možemo biti i precizniji ako pretpostavimo da su vektori u i v u prvom kvadrantu. Tada je $\det(u,v) > 0$ ako i samo ako je nagib α od u manji od nagiba β od v. Zaista, u je pozitivno proporcionalan vektoru $(1,\alpha)$, dok je v pozitivno proporcionalan $(1,\beta)$. Stoga je $\det(u,v)$ pozitivno proporcionalno $\det\left[\begin{smallmatrix} 1 & \alpha \\ 1 & \beta \end{smallmatrix}\right] = \beta - \alpha$. Slučaj kada su nagibi od u ili v beskonačni lako je zasebno obraditi. Zaključujemo da je za nenul cjelobrojne vektore u i v u prvom kvadrantu takve da je nagib od u manji od nagiba od v, paralelogram kojeg oni određuju primitivan ako i samo ako je $\det(u,v) = 1$.

Kod dokaza da se svaka Christoffelova riječ može na jedinstven način prikazati kao produkt dviju takvih riječi, trebat će nam sljedeća tvrdnja.

Korolar 4.3. *Neka su A i B dvije cjelobrojne točke takve da je $B = A+(p,q)$, gdje su p,q relativno prosti prirodni brojevi. Postoji jedinstvena točka M koja leži unutar pravokutnika s dijagonalom $\overline{AB}$, ali ispod te dijagonale takva da je AMB primitivni trokut.*

Dokaz. Rezultat i dokaz su ilustrirani na Slici 1.b.

Neka je M cjelobrojna točka u pravokutniku ispod dijagonale i najbliža toj dijagonali. Budući da su p i q relativno prosti, ne postoji cjelobrojna točka na otvorenoj dužini AB. Zaključujemo da je trokut AMB primitivan.

Kako bismo dokazali jedinstvenost, pretpostavimo da postoji još jedna točka M' takva da je trokut $AM'B$ primitivan. Budući da je AMB primitivan trokut, vektori $\overrightarrow{AM}, \overrightarrow{AB}$ čine bazu od $\mathbb{Z}^2$. Njihova determinanta je stoga ± 1, odnosno prema maloprije izvedenom rezultatu, ta je determinanta jednaka 1 jer je nagib od $\overrightarrow{AM}$ manji od nagiba vektora $\overrightarrow{AB}$. Slično izvodimo i da je $\det(\overrightarrow{AM'}, \overrightarrow{AB}) = 1$. Stoga je $\det(\overrightarrow{AM}, \overrightarrow{AB}) = 1 = \det(\overrightarrow{AM'}, \overrightarrow{AB}$, pa uzevši razliku, dobivamo $\det(\overrightarrow{MM'}, \overrightarrow{AB}) = 0$, odnosno MM' i AB su paralelni. Kako je dužina $\overline{MM'}$ strogo ispod dijagonale, kraća je od $\overline{AB}$, pa točka $A + \overrightarrow{MM'}$ leži na otvorenoj dužini AB što je kontradikcija. □

4.2.2 Christoffelove riječi

Put u rešetki je niz uzastopnih koraka u ravnini, gdje je svaki *korak* dužina $[(x,y),(x+1,y)]$ ili $[(x,y),(x,y+1)]$ pri čemu su $x,y \in \mathbb{Z}$. U ovom potpoglavlju će putovi biti konačni, ali ćemo kasnije proučavati i beskonačne putove.

Neka su p,q nenegativni relativno prosti cijeli brojevi. Promotrimo dužinu od neke cjelobrojne točke A do $B = A + (p,q)$ i put u rešetki od A do B koji se nalazi ispod te dužine takav da poligon omeđen dužinom i putom nema cjelobrojnih točaka u nutrini. Ako je dan potpuno uređen alfabet $\{a < b\}$, *donja Christoffelova riječ nagiba* $\frac{q}{p}$ je riječ u slobodnom monoidu $\{a, b\}^*$ koja kodira navedeni put pri čemu a predstavlja horizontalni i b vertikalni korak. Na Slici 1.c prikazan je navedeni put za $(p, q) = (7, 4)$ kojemu odgovara donja Christoffelova riječ *aabaabaabab* nagiba $\frac{4}{7}$. Primijetimo da je nagib donje Christoffelove riječi jednak nagibu dužine koja spaja krajnje točke odgovarajućeg puta u rešetki. Također, za definiciju Christoffelove riječi je izbor točke A nebitan, pa često uzimamo $A = (0, 0)$. Kažemo da navedeni put i donja Christoffelova riječ *diskretiziraju odozdo* dužinu $\overline{AB}$.

Analogno se definira *gornja Christoffelova riječ* nagiba $\frac{q}{p}$ promatrajući put u rešetki koji se nalazi iznad dane dužine. Budući da je pravokutnik s nasuprotnim vrhovima A i B i stranicama paralelnim koordinatnim osima simetričan s obzirom na svoje središte, vidimo da je gornja Christoffelova riječ danog nagiba obratna riječ donje Christoffelove riječi istog nagiba.

Christoffelova riječ je riječ koja je donja ili gornja Christoffelova riječ.

Označimo s E automorfizam slobodnog monoida $\{a, b\}^*$ koji zamjenjuje a i b. Tada E zamjenjuje donje i gornje Christoffelove riječi. Taj automorfizam šalje donje Christoffelove riječi nagiba $\frac{q}{p}$ u gornje Christoffelove riječi nagiba $\frac{p}{q}$. Kako bismo to provjerili, dovoljno je primjerice na Slici 1.c uzeti $A = (0, 0)$ i zatim primijeniti simetriju s obzirom na pravac $y = x$.

Slova a i b su Christoffelove riječi, istovremeno donje i gornje, nagiba redom 0 i ∞ te odgovaraju slučajevima $(p, q) = (1, 0)$ i $(p, q) = (0, 1)$. Sve ostale Christoffelove riječi nazivamo *prave*. Riječi $a^n b$ i ab^n su za svaki $n \in \mathbb{N}_0$ donje Christoffelove riječi.

Sljedeća karakterizacija donjih Christoffelovih riječi može poslužiti za algebarsko dokazivanje tvrdnji koje ćemo mi pokazati geometrijski.

Teorem 4.4. *Neka je $w = a_1 a_2 \dots a_{p+q}$ donja Christoffelova riječ nagiba $\frac{q}{p}$. Neka je za $i \in \mathbb{N}_0$, $k(i) = \left\lfloor \frac{q}{p+q} i \right\rfloor$. Tada je*

$$a_i = \begin{cases} a & \text{ako je } k(i) - k(i-1) = 0, \\ b & \text{ako je } k(i) - k(i-1) = 1, \end{cases} \tag{4.3}$$

za sve $i \in \{1, \dots, p+q\}$.

Dokaz. Tvrdnju je trivijalno provjeriti za riječi a i b, pa pretpostavimo da je w prava donja Christoffelova riječ, tj. p i q su relativno prosti prirodni brojevi.

Neka je v riječ definirana s (4.3). Dokazat ćemo da je $v = w$ tako da pokažemo da za svaki $n \in \{1, \dots, p+q\}$ među prvih n slova u riječi v i u riječi w ima jednak broj slova b.

U prvih n slova riječi v ima

$$\sum_{i=1}^{n} \big(k(i) - k(i-1)\big) = k(n) - k(0) = \left\lfloor \frac{q}{p+q} n \right\rfloor$$

slova b. Ako označimo s x broj slova a i s y broj slova b u prvih n slova riječi w, iz definicije Christoffelove riječi slijedi

$$x + y = n, \qquad \left\lfloor \frac{q}{p}(x-1) \right\rfloor \leq y \leq \left\lfloor \frac{q}{p} x \right\rfloor.$$

Odavde koristeći činjenicu da su x i y nenegativni cijeli brojevi, a p i q relativno prosti prirodni brojevi, dobivamo nakon malo računa da je $y = \left\lfloor \frac{q}{p+q} n \right\rfloor$ što je i trebalo dokazati. $\square$

Označimo sa C *konjugator*, tj. preslikavanje koje svaku riječ $a_1 \dots a_n$ duljine n preslikava u riječ $a_2 \dots a_n a_1$. Time je definirano djelovanje cikličke grupe $\mathbb{Z}/n\mathbb{Z}$ na skup riječi duljine n. Za riječ w duljine n, sve njoj konjugirane riječi su $C^i(w)$, $i = 0, \dots, n-1$.

Propozicija 4.5. *Neka je w Christoffelova riječ. Tada w nije potencija nijedne druge riječi. Ako je w duljine n, sve riječi $C^i(w)$, $i = 0, \dots, n-1$ su različite.*

Dokaz. Očito je u donjoj ili gornjoj Christoffelovoj riječi nagiba $\frac{q}{p}$ broj slova a jednak p dok je broj slova b jednak q. To posebno povlači da su za Christoffelovu riječ w, brojevi $|w|_a$ i $|w|_b$ relativno prosti te w ne može biti potencija neke druge riječi.

Neka je w duljine n i pretpostavimo da nisu sve navedene riječi različite. Tada je $C^d(w) = w$ za neki pravi djelitelj d od n, pa je w potencija riječi duljine d što je u kontradikciji s upravo dokazanom tvrdnjom. □

4.2.3 Palindromi

Neka su p i q relativno prosti prirodni brojevi. *Reznu riječ m pridruženu $\frac{q}{p}$* dobivamo na sljedeći način. Neka su A i B cjelobrojne točke takve da je $\overrightarrow{AB} = (p, q)$. Promatramo sjecišta otvorene dužine AB s pravcima cjelobrojne rešetke. Ako je pravac rešetke vertikalan, sjecište označimo s a, a ako je pravac horizontalan, sjecište označimo s b. Tada je m riječ dobivena čitanjem tih slova tako da se krećemo s lijeva nadesno i zabilježimo slova pri prelasku navedenih sjecišta. Na Slici 1.d prikazana je rezna riječ pridružena $\frac{4}{7}$. Primijetimo da za reznu riječ m pridruženu $\frac{q}{p}$ imamo $|m|_a = p - 1$, $|m|_b = q - 1$.

Jasno je da je svaka rezna riječ palindrom. Zaista, pri centralnoj simetriji s obzirom na središte definirajućeg pravokutnika s dijagonalom $\overline{AB}$, ta se dijagonala preslikava u samu sebe, a vertikalni (horizontalni) pravci rešetke se ponovno preslikavaju u vertikalne (horizontalne) pravce rešetke.

Teorem 4.6. *Prava donja Christoffelova riječ nagiba r je oblika amb za neki palindrom m koji je rezna riječ pridružena r. Pripadna gornja Christoffelova riječ je bma.*

Dokaz. Dokazujemo prvu tvrdnju teorema, pa će druga slijediti iz onoga što smo vidjeli u prethodnom odjeljku. Primijetimo da prava donja Christoffelova riječ uvijek počinje s a i završava s b. Na Slici 1.e prikazana je bijekcija između koraka puta koji definira donju Christoffelovu riječ (izbacivši prvi i zadnji korak) i sjecišta koja definiraju reznu riječ. □

Prethodni teorem je ilustriran na Slici 1.f. gdje su prikazane donja i gornja Christoffelova riječ nagiba $\frac{4}{7}$. Njihovi putovi zajedno čine zmijoliki poligon što ilustrira činjenicu da su oblika amb i bma za istu riječ m koja je rezna riječ pridružena $\frac{4}{7}$.

Teorem 4.7. *Neka su p i q relativno prosti prirodni brojevi. Neka su $P = \{ip : 0 < i < q\}$ i $Q = \{jq : 0 < j < p\}$. Tada su skupovi P i Q disjunktni. Neka je $P \cup Q = \{r_1 < r_2 < \dots < r_{p+q-2}\}$. Definirajmo riječ $m = a_1 a_2 \dots a_{p+q-2}$ na alfabetu $\{a < b\}$ uvjetima $a_i = a$ ako je $r_i \in Q$ i $a_i = b$ ako je $r_i \in P$. Tada je m rezna riječ pridružena $\frac{q}{p}$. Svaka rezna riječ može se dobiti na ovaj način.*

Dokaz. Neka je w rezna riječ pridružena $\frac{q}{p}$. Pogledajmo na kojim mjestima u riječi w se pojavljuje slovo a. Proizvoljnom slovu a u riječi w odgovara sjecište otvorene dužine AB s vertikalnim pravcem cjelobrojne rešetke. Neka je to sjecište točka (x, y). Tada je $x \in \{1, \dots, p-1\}$, $y = \frac{q}{p}x$. Primijetimo da zbog $(p, q) = 1$ ne može biti $y \in \mathbb{N}$. Prije odgovarajućeg slova a, ima u w točno $x - 1$ slova a i $\lfloor y \rfloor$ slova b, tj. promatrano slovo a je

$x + \lfloor y \rfloor$ po redu slovo u riječi w i vrijedi $\lfloor y \rfloor < \frac{q}{p}x < \lfloor y \rfloor + 1$, tj. $\lfloor y \rfloor p < xq < (\lfloor y \rfloor + 1)p$. To znači da su od $xq \in Q$ manji točno elementi $p, 2p, \ldots, \lfloor y \rfloor p$ iz P i $q, 2q, \ldots, (x-1)q$ iz Q. Dakle, na mjestu $x + \lfloor y \rfloor$ u riječi m je slovo a.

Time smo pokazali da je na svakom mjestu gdje se u w pojavljuje slovo a ujedno i slovo a u riječi m. Budući da obje riječi imaju jednako slova a i jednako slova b, zaključujemo da su jednake. □

Kažemo da je riječ m iz prethodnog teorema dobivena superpozicijom dvaju periodskih pojava.

4.2.4 Standardna faktorizacija

Christoffelove riječi mogu se konstruirati rekurzivno. Prvi korak u toj konstrukciji slijedi iz idućeg rezultata.

Teorem 4.8. *Svaka prava Christoffelova riječ može se na jedinstven način prikazati kao produkt dviju Christoffelovih riječi.*

Ova faktorizacija prave Christoffelove riječi naziva se *standardna faktorizacija.* U dokazu ćemo pokazati da su dva faktora u standardnoj faktorizaciji donje Christoffelove riječi opet donje Christoffelove riječi. Kažemo da je (u, v) *Christoffelov par* ako su sve tri riječi u, v, uv *donje* Christoffelove riječi.

Dokaz.

1. Dovoljno je tvrdnju teorema dokazati za donje Christoffelove riječi. Naime, gornje su dobivene obrtanjem donjih, a ako je donja Christoffelova riječ produkt dviju Christoffelovih riječi, one moraju biti donje. Zaista, uz $a < b$, prava Christoffelova riječ je donja ako i samo ako počinje s a te također ako i samo ako završava s b.

2. Neka je $\frac{q}{p}$ nagib prave donje Christoffelove riječi w. Neka su A i B cjelobrojne točke tako da je $B = A + (p, q)$. U pravokutniku kojemu je dijagonala dužina $\overline{AB}$, izaberimo M kao u Korolaru 4.3. Tada je trokut AMB primitivan. Ne postoji cjelobrojna točka na otvorenoj dužini AM, pa je $\overrightarrow{AM} = (i, j)$, gdje su i, j relativno prosti. Slično, $\overrightarrow{MB} = (k, l)$, gdje su k, l relativno prosti. Tada je Christoffelova riječ nagiba $\frac{q}{p}$ nastala spajanjem Christoffelovih riječi nagiba $\frac{j}{i}$ i $\frac{l}{k}$. Na Slici 1.g je riječ w prikazana iscrtkanom krivuljom dobivena kao produkt dviju riječi predstavljenih iscrtkanim krivuljama od A do M i od M do B.

3. Obratno, ako je donja Christoffelova riječ w produkt dviju takvih riječi $w = uv$, onda u definirajućem pravokutniku za w koji ima dijagonalu $\overline{AB}$, na putu u rešetki kodiranom riječju w postoji točka M koja odgovara faktorizaciji $w = uv$. Sada je trokut AMB primitivan. Zaista, putovi koji odgovaraju riječima u i v su redom ispod dužina AM i MB. Stoga je trokut sadržan u poligonu kojeg omeđuju put koji odgovara riječi w i dužina AB, a znamo da taj poligon nema cjelobrojnih točaka u nutrini. Točka M je prema Korolaru 4.3 jedinstvena, pa su riječi u, v jedinstvene. □

Na Slici 1.g je točka N definirana preko $\overrightarrow{AN} = \overrightarrow{MB}$. Tada je pravokutnik $AMBN$ primitivan.

Za primjer standardne faktorizacije pogledajmo Sliku 1.c. Najbliža točka je $A + (2, 1)$, pa je standardna faktorizacija odgovarajuće Christoffelove riječi $(aab)(aabaabab)$.

Korolar 4.9. *Za donje Christoffelove riječi u, v s komutativnim slikama (i, j) i (k, l), tim redom, sljedeći uvjeti su ekvivalentni:*

(i) uv je donja Christoffelova riječ;

(ii) $\det \left[\begin{smallmatrix} i & j \\ k & l \end{smallmatrix}\right] = 1$*;*

(iii) Paralelogram kojemu su stranice vektori (i, j) i (k, l) je primitivan i nagib od (i, j) je manji nego nagib od (k, l).

Ako vrijede ovi uvjeti, $w = uv$ je standardna faktorizacija od w.

Dokaz. (i) povlači (ii): Uz oznake iz dokaza Korolara 4.3 i trećeg dijela dokaza Teorema 4.8 imamo $(i, j) = \overrightarrow{AM}$, $(k, l) = \overrightarrow{MB}$. Tada, kao u dokazu navedenog korolara, vrijedi $1 = \det(\overrightarrow{AM}, \overrightarrow{AB}) = \det(\overrightarrow{AM}, \overrightarrow{AM} + \overrightarrow{MB}) = \det(\overrightarrow{AM}, \overrightarrow{MB})$.

(ii) povlači (iii): Ovo slijedi iz diskusije prije Korolara 4.3.

(iii) povlači (i): Neka je točka A ishodište te $\overrightarrow{AM} = (i, j)$, $\overrightarrow{MB} = (k, l)$, $\overrightarrow{AN} = (k, l)$. Paralelogram koji se spominje u uvjetu (iii) je $AMBN$. Iz uvjeta na nagibe slijedi da je točka M ispod dužine AM. Na Slici 1.g je u diskretizacija odozdo dužine $\overline{AM}$, a v je diskretizacija odozdo dužine $\overline{MB}$. Kako je paralelogram primitivan, diskretizacija od $\overline{AB}$ je uv što stoga mora biti donja Christoffelova riječ.

Zadnja tvrdnja slijedi iz Teorema 4.8. □

Korolar 4.10. *Neka je $w = uv$ donja Christoffelova riječ sa svojom standardnom faktorizacijom.*

(i) Tada su $u(uv)$ i $(uv)v$ donje Christoffelove riječi s naznačenom standardnom faktorizacijom.

(ii) Ako je w duljine barem 3, onda je ili u pravi prefiks od v te je $v = uv'$ standardna faktorizacija od v ili v pravi sufiks od u te je $u = u'v$ standardna faktorizacija od u.

Dokaz. Označimo komutativne slike od u i v redom s (i, j) i (k, l). Prema (ii) u prethodnom korolaru, vrijedi $il - jk = 1$.

(i) Komutativna slika od uv je $(i + k, j + l)$. Imamo $(i + k)l - (j + l)k = 1$, pa je prema (i) u prethodnom korolaru, $(uv)v$ donja Christoffelova riječ te ta riječ ima naznačenu standardnu faktorizaciju. Za $u(uv)$ je dokaz sličan.

(ii) Pretpostavljamo da je duljina od w barem 3, tj. $i+j+k+l \geq 3$. Tada je $j+k > 0$ jer bi u protivnom bilo $j = k = 0$ i $1 = il - jk = il$ te stoga $i = l = 1$, odnosno $i + j + k + l = 2$ što daje kontradikciju.

Prema Teoremu 4.8, dovoljno je dokazati da je ili $v = uv'$ gdje je v' donja Christoffelova riječ ili $u = u'v$ gdje je u' donja Christoffelova riječ.

(ii.a) Pretpostavimo najprije da je $i \leq k$. Promotrimo cjelobrojne točke A, $M = A+(i, j)$, $N = A+(k, l)$, $B = A+(i+k, j+l) = M+(k, l) = N+(i, j)$ kao na Slici 1.g. Točka M je ispod dijagonale AB jer je nagib $\frac{j}{i}$ od $\overrightarrow{AM}$ manji nego nagib $\frac{l}{k}$ od $\overrightarrow{MB}$ budući da je $il - jk = 1$. Paralelogram $AMBN$ je primitivan. Translacijom tog paralelograma za vektor $\overrightarrow{AM}$ dobivamo paralelogram koji označimo $MHN'B$. Uvjet $i \leq k$ povlači da prva koordinata od $H = A+(2i, 2j)$

nije veća nego prva koordinata od B kao što je prikazano na Slici 1.h. Oba su paralelograma primitivni, pa su jedine cjelobrojne točke u njihovoj uniji njihovi vrhovi A, M, H, N', B, N. Posebno, nema cjelobrojnih točaka u nutrini trokuta AHB. Zaključujemo da je diskretizacija odozdo dužine $\overline{AB}$ jednaka diskretizaciji od $\overline{AM}$ za kojom slijedi diskretizacija od $\overline{MH}$ i onda diskretizacija od $\overline{HB}$. Kako je $\overrightarrow{AM} = \overrightarrow{MH}$, prve su dvije diskretizacije jednake što znači da je u prefiks od v. Štoviše $v = uv'$ gdje je v' diskretizacija od $\overline{HB}$ iz čega slijedi da je v' donja Christoffelova riječ.

(ii.b) Pretpostavimo sada da je $i > k$. Imamo $il - jk = 1$, pa mora biti $l \leq j$. Zaista, kad bi bilo $l > j$, vrijedilo bi $il \geq (k+1)(j+1) = jk + j + k + 1 > jk + 1$, jer je $j + k > 0$ kako smo pokazali na početku (ii), te bi slijedilo $il - jk > 1$ što nije moguće.

Involutorni automorfizam E slobodnog monoida $\{a, b\}^*$ koji zamjenjuje a i b zamjenjuje donje i gornje Christoffelove riječi. I obrtanje također zamjenjuje donje i gornje Christoffelove riječi. Budući da je $w = uv$, riječ $E(\widetilde{w})$ ima, prema jedinstvenosti u Teoremu 4.8, standardnu faktorizaciju $E(\widetilde{v})E(\widetilde{u})$. Komutativne slike od $E(\widetilde{v})$ i $E(\widetilde{u})$ su (l, k) i (j, i) tim redom. Kako je $l \leq j$, argument iz dijela (ii.a) pokazuje da je $E(\widetilde{u}) = E(\widetilde{v})m$ za neku donju Christoffelovu riječ m. Stoga je $\widetilde{u} = \widetilde{v}E(m)$ i $u = u'v$ gdje je $u' = E(\widetilde{m})$ donja Christoffelova riječ. $\square$

4.2.5 Stablo Christoffelovih parova

Ako je (u, v) Christoffelov par, onda iz Korolara 4.10 slijedi da su (u, uv) i (uv, v) također Christoffelovi parovi. Očito je (a, b) Christoffelov par. Ako, dakle, konstruiramo beskonačno binarno stablo s korijenom (a, b) i pravilom danim idućom slikom, onda su svi čvorovi toga stabla Christoffelovi parovi.

$$\begin{array}{ccc} & (u,v) & \\ \swarrow & & \searrow \\ (u,uv) & & (uv,v) \end{array}$$

Pravilo konstrukcije stabla Christoffelovih parova

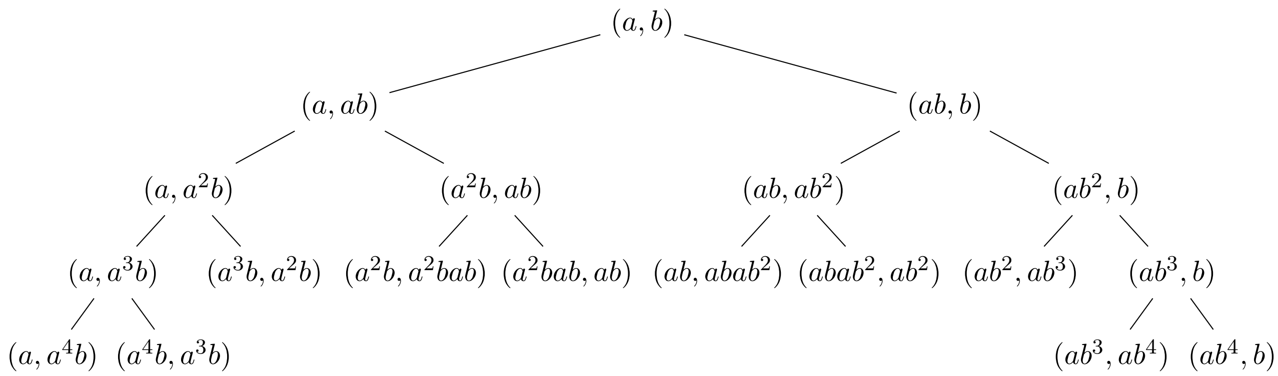
Obratno, svaki Christoffelov par se pojavljuje na ovom stablu, što slijedi iz drugog dijela Korolara 4.10. Isti korolar također povlači da su svi Christoffelovi parovi na stablu različiti. Zaista, čvor (u, v) različit od korijena je ili lijevo ili desno dijete, ali ne može biti oboje jer su uvjeti "u je pravi prefiks od v" i "v je pravi sufiks od u" međusobno isključivi. Zaključivanje dalje ide rekurzivno.

Primijetimo da ako se (u, v) pojavljuje na stablu, onda uv mora biti donja Christoffelova riječ. Svaka prava donja Christoffelova riječ pojavljuje se točno jednom kao produkt elemenata nekog uređenog para u stablu zbog jedinstvenosti standardne faktorizacije.

Idući rezultat je lagana posljedica pravila konstrukcije svih Christoffelovih parova.

Korolar 4.11. *Parovi $(|u|, |v|)$ za sve Christoffelove parove (u, v) čine upravo skup svih parova relativno prostih prirodnih brojeva.*

Ako je prava donja Christoffelova riječ $w = uv$ dana sa svojom standardnom faktorizacijom, tada barem jedna od u ili v nije prava ako i samo ako je $w = a^n b$ ili $w = ab^n$, gdje je n prirodan broj. Standardna faktorizacija je tada $a(a^{n-1}b)$ ili $(ab^{n-1})b$, respektivno.



Početni dio stabla Christoffelovih parova

4.2.6 Sturmovski morfizmi

Promatramo potpuno uređeni alfabet $\{a < b\}$. Sjetimo se da sa (u, v) označavamo endomorfizam f slobodnog monoida $\{a < b\}^*$ definiran s $f(a) = u$, $f(b) = v$. Sada definiramo sljedeća četiri endomorfizma

$$G = (a, ab), \qquad D = (ba, b), \qquad \widetilde{G} = (a, ba), \qquad \widetilde{D} = (ab, b).$$

Propozicija 4.12. *Endomorfizmi G, D, $\widetilde{G}$, $\widetilde{D}$ slobodnog monoida $\{a < b\}^*$ su injekcije.*

Dokaz. Neka je $w = G(u) = G(v)$. Ako je $|w| = 0$, onda su u i v prazne riječi. Ako w završava s a, onda je zadnje slovo u riječima u i v upravo a, a ako w završava s b, onda je zadnje slovo od u i v slovo b. Sada argument možemo nastaviti indukcijom po $|w|$ te dobivamo $u = v$. Analogno se pokaže injektivnost ostalih navedenih preslikavanja. □

Nije teško pokazati da su preslikavanja iz prethodne propozicije zapravo automorfizmi slobodne grupe $F(a, b)$, ali ta nam činjenica neće trebati.

U nastavku ćemo promatrati nekoliko beskonačnih potpunih binarnih stabala s korijenom. Za svaki čvor takvog stabla, postoji jedinstveni put od korijena do tog čvora. Taj je put kodiran riječju $w \in \{a < b\}^*$, gdje a znači "lijevo", a b znači "desno". Tako dobivamo bijekciju između $\{a < b\}^*$ i skupa čvorova stabla te kažemo da *riječ w kodira čvor*. Primjerice, u gornjem stablu čvor (a^2bab, ab) odgovara riječi abb.

Označimo s $\mathrm{End}(\{a, b\}^*)$ monoid endomorfizama slobodnog monoida $\{a, b\}^*$.

Teorem 4.13. *Neka je $\rho : \{a, b\}^* \to \mathrm{End}(\{a, b\}^*)$ homomorfizam monoida takav da je $\rho(a) = G$ i $\rho(b) = \widetilde{D}$. Neka je (u, v) proizvoljan čvor u stablu Christoffelovih parova i neka ga kodira riječ $x \in \{a < b\}^*$. Tada u monoidu endomorfizama slobodnog monoida $\{a, b\}^*$ vrijedi jednakost $(u, v) = \rho(x)$.*

Pogledamo li opet primjer s prethodnog stabla, imamo $(a^2bab, ab) = \rho(abb) = G \circ \widetilde{D} \circ \widetilde{D}$. Zaista, $G \circ \widetilde{D} \circ \widetilde{D}(a) = G \circ \widetilde{D}(ab) = G(abb) = aabab$ i $G \circ \widetilde{D} \circ \widetilde{D}(b) = G \circ \widetilde{D}(b) = G(b) = ab$.

Dokaz. Za korijen, tj. $x = 1$, tvrdnja vrijedi. Opći slučaj dokazujemo indukcijom. Pretpostavimo da je $(u, v) = \rho(x)$. Tada je $\rho(x)(a) = u$, $\rho(x)(ab) = uv$, $\rho(x)(b) = v$. Stoga imamo

$$\rho(xa)(a) = \rho(x) \circ G(a) = \rho(x)(a) = u, \quad \text{i} \quad \rho(xa)(b) = \rho(x) \circ G(b) = \rho(x)(ab) = uv,$$
$$\rho(xb)(a) = \rho(x) \circ \widetilde{D}(a) = \rho(x)(ab) = uv \quad \text{i} \quad \rho(xb)(b) = \rho(x) \circ \widetilde{D}(b) = \rho(x)(b) = v.$$

U prvom slučaju je $\rho(xa) = (u, uv)$, dok je u drugom $\rho(xb) = (uv, v)$ što je u skladu s pravilom konstrukcije stabla. □

Korolar 4.14. *Slika svake donje Christoffelove riječi po G ili $\widetilde{D}$ je donja Christoffelova riječ.*

Dokaz. Tvrdnja je jasna za donje Christoffelove riječi duljine 1, tj. a i b jer je ab donja Christoffelova riječ. Promotrimo sada pravu donju Christoffelovu riječ $w = uv$ sa svojom standardnom faktorizacijom. Tada se (u, v) pojavljuje na stablu Christoffelovih parova i, promatran kao endomorfizam slobodnog monoida $\{a, b\}^*$, jednak je $\rho(x)$ prema Teoremu 4.13 i ondje uvedenoj notaciji. Dakle, $\rho(x)(a) = u$ i $\rho(x)(b) = v$. Riječ ax kodira neki čvor (u', v') u stablu, pa je $u'v'$ donja Christoffelova riječ. Prema teoremu je

$$
\begin{aligned}
G(u) &= G(\rho(x)(a)) = G \circ \rho(x)(a) = \rho(a) \circ \rho(x)(a) = \rho(ax)(a) = u', \\
G(v) &= G(\rho(x)(b)) = G \circ \rho(x)(b) = \rho(a) \circ \rho(x)(b) = \rho(ax)(b) = v'.
\end{aligned}
$$

Stoga je $G(w) = G(uv) = G(u)G(v) = u'v'$ donja Christoffelova riječ.

Dokaz da je $\widetilde{D}(w)$ donja Christoffelova riječ je sličan. □

Korolar 4.15. *Slika svake donje Christoffelove riječi po $\widetilde{G}$ ili D je konjugat donje Christoffelove riječ.*

Dokaz. Zbog Korolara 4.14 je dovoljno dokazati da su za svaku riječ w, riječi $D(w)$ i $\widetilde{D}(w)$ konjugati te slično za G i $\widetilde{G}$. Mi ćemo to učiniti samo za D. Podsjetimo se da je $D = (ba, b)$ i $\widetilde{D} = (ab, b)$. Promatrajući ih kao endomorfizme slobodne grupe $F(a, b)$, imamo $D = \iota \circ \widetilde{D}$, gdje je ι unutrašnji automorfizam $g \mapsto bgb^{-1}$. To se lako provjeri na generatorima a i b. Primijetimo da slika po $\widetilde{D}$ svake neprazne riječi iz slobodnog monoida $\{a, b\}^*$ završava s b, pa za svaku riječ w iz $\{a, b\}^*$ imamo $\widetilde{D}(w) = ub$ i $D(w) = bu$. Stoga su $D(w)$ i $\widetilde{D}(w)$ konjugati. □

Sturmovski morfizam je endomorfizam slobodnog monoida $\{a, b\}^*$ koji preslikava svaku donju Christoffelovu riječ u konjugat donje Christoffelove riječi. Ekvivalentno, preslikava svaki konjugat donje Christoffelove riječi u konjugat donje Christoffelove riječi. Tako smo pokazali da su endomorfizmi $G, D, \widetilde{G}, \widetilde{D}$ sturmovski morfizmi.

Primjer. Aigner u svojoj knjizi o Markovljevom teoremu definira Christoffelove riječi slično kao i mi, ali su kod njega koraci koji čine put u rešetki dani s $[(x, y), (x + 1, y)]$ i $[(x, y), (x + 1, y + 1)]$ pri čemu prvom tipu koraka odgovara slovo a, dok drugom tipu odgovara slovo b. Pokažimo da je na taj način definiran isti skup Christoffelovih riječi.

Neka je w Aignerova donja Christoffelova riječ koja kodira odgovarajući put od $(0, 0)$ do (p, q). Primijetimo da je nužno $p \geq q$ jer se u svakom koraku povećava prva koordinata. Za istu početnu i krajnju točku, naša donja Christoffelova riječ će biti upravo $G(w)$ i to je riječ nagiba $\frac{q}{p}$. Neka je u donja Christoffelova riječ nagiba $\frac{q}{p-q}$. Tada je prema Korolaru 4.14, $G(u)$ donja Christoffelova riječ i lako vidimo da je nagiba $\frac{q}{p}$ jer se $p - q$ slova a u riječi u pretvara u $p - q$ slova a dok se q slova b pretvara u q slova a i q slova b. Zbog jedinstvenosti donje Christoffelove riječi zadanog nagiba, vidimo da je $G(u) = G(w)$, pa iz injektivnosti od G dobivamo da je $u = w$.

Aignerova definicija Christoffelovih riječi je nešto pogodnija kod dokazivanja tvrdnji o njima algebarskim putem. Kod njega je funkcija u Teoremu 4.4 dana s $k(i) = \lfloor \frac{q}{p} i \rfloor$.

4.3 Markovljevi brojevi

U ovom potpoglavlju proučavat ćemo diofantsku jednadžbu koju nazivamo Markovljeva jednadžba. Njezina rješenja nazivaju se Markovljeve trojke. Osnovni rezultat ovog dijela bit će Teorem 4.16 koji daje bijekciju između donjih Christoffelovih riječi i Markovljevih trojki.

4.3.1 Markovljeve trojke i brojevi

Markovljeva trojka je multiskup $\{x, y, z\}$ prirodnih brojeva koji zadovoljava *Markovljevu jednadžbu*

$$x^2 + y^2 + z^2 = 3xyz.$$

Primjeri su $\{1, 1, 1\}$, $\{1, 1, 2\}$ i $\{1, 2, 5\}$. Za Markovljevu trojku kažemo da je *prava* ako su tri broja koji ju čine različiti, inače ju zovemo *nepravom*. *Markovljev broj* je prirodan broj koji je element neke Markovljeve trojke.

Promotrimo homomorfizam monoida $\mu : \{a, b\}^* \to \mathrm{SL}(2, \mathbb{Z})$ definiran s

$$\mu(a) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mu(b) = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}.$$

Smatramo da je alfabet $\{a, b\}$ potpuno uređen, $a < b$.

Matričnu konstrukciju Markovljevih trojki kako je navodimo u idućem teoremu dao je Cohn koji je zamijetio sličnost Frickeovog identiteta (zadnja jednakost u Lemi 4.19) i Markovljeve jednadžbe.

Teorem 4.16. *Funkcija koja pridružuje svakoj donjoj Christoffelovoj riječi w sa standardnom faktorizacijom uv multiskup*

$$\left\{ \frac{1}{3}\,\mathrm{Tr}(\mu(u)),\ \frac{1}{3}\,\mathrm{Tr}(\mu(v)),\ \frac{1}{3}\,\mathrm{Tr}(\mu(w)) \right\}$$

je bijekcija sa skupa svih pravih donjih Christoffelovih riječi na skup svih pravih Markovljevih trojki.

Teorem ćemo dokazati nakon nekoliko pripremnih lema. U idućoj lemi ćemo pokazati da za svaku donju Christoffelovu riječ w vrijedi $\frac{1}{3}\,\mathrm{Tr}(\mu(w)) = \mu(w)_{12}$. Stoga je trojka u teoremu jednaka i $\{\mu(u)_{12}, \mu(v)_{12}, \mu(w)_{12}\}$.

Lema 4.17. *(i) Ako je M simetrična 2×2 matrica i $N = \mu(a)M\mu(b)$, onda vrijedi $\frac{1}{3}\,\mathrm{Tr}(N) = N_{12}$.*

(ii) Ako je $N = \mu(w)$ za neku donju Christoffelovu riječ w, onda je $\frac{1}{3}\,\mathrm{Tr}(N) = N_{12}$.

Dokaz. Imamo

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} p & q \\ q & s \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 2p+q & 2q+s \\ p+q & q+s \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 10p+9q+2s & 4p+4q+s \\ 5p+7q+2s & 2p+3q+s \end{bmatrix}.$$

Trag zadnje matrice je $12p + 12q + 3s$ što je tri puta veće od elementa na mjestu $(1, 2)$. Time je prva tvrdnja dokazana.

Ako je w prava donja Christoffelova riječ, onda je po Teoremu 4.6 $w = amb$ za neki palindrom m. Stoga je $N = \mu(w) = \mu(a)\mu(m)\mu(b)$ i $\mu(m)$ je simetrična budući da je m palindrom, a matrice $\mu(a)$ i $\mu(b)$ su simetrične. Time su ispunjeni uvjeti za (i) pa je (ii) dokazano kada je w prava. Ako je $w = a$ ili b, onda odmah vidimo da (ii) vrijedi. □

Lema 4.18. *Neprave Markovljeve trojke su jedino* $\{1,1,1\}$ *i* $\{1,1,2\}$.

Dokaz. Neka je $\{x,x,y\}$ Markovljeva trojka. Tada je $2x^2+y^2=3x^2y$, pa x^2 dijeli y^2, odnosno x dijeli y, možemo staviti $y=xu$. Sada je $2x^2+x^2u^2=3x^3u$ i zato $2+u^2=3xu$. Stoga u dijeli 2 te je $u=1$ ili $u=2$. Kako je $\frac{2}{u}+u=3x$, imamo u oba slučaja $3x=3$, pa je $x=1$ i zato $y=u$. Dobili smo trojke $\{1,1,1\}$ i $\{1,1,2\}$. □

Lema 4.19 (Fricke)**.** *Neka su* A *i* B *matrice iz* $\mathrm{SL}(2,\mathbb{Z})$*. Tada vrijedi*

$$\begin{aligned}\mathrm{Tr}(A^2B)+\mathrm{Tr}(B)&=\mathrm{Tr}(A)\,\mathrm{Tr}(AB),\\ \mathrm{Tr}(AB^2)+\mathrm{Tr}(A)&=\mathrm{Tr}(AB)\,\mathrm{Tr}(B),\\ \mathrm{Tr}(A)^2+\mathrm{Tr}(B)^2+\mathrm{Tr}(AB)^2&=\mathrm{Tr}(A)\,\mathrm{Tr}(B)\,\mathrm{Tr}(AB)+\mathrm{Tr}(ABA^{-1}B^{-1})+2.\end{aligned}$$

Dokaz. Skalar α identificiramo sa skalarnom matricom αI dimenzija 2×2. Prema Hamilton-Cayleyjevom teoremu je $A^2-\mathrm{Tr}(A)A+1=0$ jer A ima determinantu 1. Stoga je $A^2+1=\mathrm{Tr}(A)A$, pa zato $A^2B+B=\mathrm{Tr}(A)AB$ te uzimajući trag dobivamo $\mathrm{Tr}(A^2B)+\mathrm{Tr}(B)=\mathrm{Tr}(A)\,\mathrm{Tr}(AB)$ što je upravo prva jednakost. Drugu dobivamo zamijenom A i B.

Hamilton-Cayleyjev teorem također povlači da je $A^{-1}=\mathrm{Tr}(A)-A$, tako da je $ABA^{-1}B^{-1}=AB(\mathrm{Tr}(A)-A)(\mathrm{Tr}(B)-B)$, odnosno

$$ABA^{-1}B^{-1}=\mathrm{Tr}(A)\,\mathrm{Tr}(B)AB-\mathrm{Tr}(A)AB^2-\mathrm{Tr}(B)ABA+ABAB. \tag{4.4}$$

Iz prvih dvaju identiteta ove leme slijedi $\mathrm{Tr}(AB^2)=\mathrm{Tr}(AB)\,\mathrm{Tr}(B)-\mathrm{Tr}(A)$ i $\mathrm{Tr}(ABA)=\mathrm{Tr}(A^2B)=\mathrm{Tr}(A)\,\mathrm{Tr}(AB)-\mathrm{Tr}(B)$. Stoga, uzevši trag u (4.4), dobivamo

$$\begin{aligned}\mathrm{Tr}(ABA^{-1}B^{-1})=\mathrm{Tr}(A)\,\mathrm{Tr}(B)\,\mathrm{Tr}(AB)-\mathrm{Tr}(A)\,\mathrm{Tr}(AB)\,\mathrm{Tr}(B)+\mathrm{Tr}(A)^2\\ -\mathrm{Tr}(B)\,\mathrm{Tr}(A)\,\mathrm{Tr}(AB)+\mathrm{Tr}(B)^2+\mathrm{Tr}(ABAB).\end{aligned}$$

Zadnji Frickeov identitet dobivamo primijenimo li još jednom Hamilton-Cayleyjev teorem, za matricu $M=AB$ u obliku $\mathrm{Tr}(M^2)=\mathrm{Tr}(M)^2-2$. □

Korolar 4.20. *Neka su* A *i* B *matrice iz* $\mathrm{SL}(2,\mathbb{Z})$ *te* $x,y,z\in\mathbb{Z}$*. Tada su sljedeća dva uvjeta ekvivalentni.*

(i) $x=\frac{1}{3}\mathrm{Tr}(A)$, $y=\frac{1}{3}\mathrm{Tr}(AB)$, $z=\frac{1}{3}\mathrm{Tr}(A^2B)$;

(ii) $x=\frac{1}{3}\mathrm{Tr}(A)$, $y=\frac{1}{3}\mathrm{Tr}(AB)$, $3xy-z=\frac{1}{3}\mathrm{Tr}(B)$.

Dokaz. Ako vrijedi (i), onda je $3xy-z=\frac{1}{3}\mathrm{Tr}(A)\,\mathrm{Tr}(AB)-\frac{1}{3}\mathrm{Tr}(A^2B)$. Prema prvom Frickeovom identitetu u prethodnoj lemi, to je jednako $\frac{1}{3}\mathrm{Tr}(B)$ te zato (ii) vrijedi. Obrat se slično pokazuje. □

Ovaj korolar ima i sljedeći simetrični iskaz koji dobivamo zamjenom A i B. Uvjet $x=\frac{1}{3}\mathrm{Tr}(B)$, $y=\frac{1}{3}\mathrm{Tr}(AB)$, $z=\frac{1}{3}\mathrm{Tr}(AB^2)$ je ekvivalentan uvjetu $x=\frac{1}{3}\mathrm{Tr}(B)$, $y=\frac{1}{3}\mathrm{Tr}(AB)$, $3xy-z=\frac{1}{3}\mathrm{Tr}(A)$.

Iduća nam lema omogućuje rekurzivnu konstrukciju Markovljevih trojki u dokazu teorema.

Lema 4.21. *Neka je* $\{x,y,z\}$ *prava Markovljeva trojka takva da je* $x<y<z$*. Tada je* $\{x,y,3xy-z\}$ *Markovljeva trojka i* $3xy-z<y$.

Dokaz. Definirajmo polinom u varijabli t sa $P(t) = t^2 - 3xyt + x^2 + y^2$. Tada je z korijen od $P(t)$. Drugi korijen je $3xy - z = \frac{x^2+y^2}{z}$. Stoga je taj korijen prirodan broj i $\{x, y, 3xy - z\}$ je Markovljeva trojka. Zbog $x \geq 1$ i $y > x$ imamo

$$P(y) = y^2 - 3xy^2 + x^2 + y^2 < y^2 - 3y^2 + y^2 + y^2 = 0,$$

pa y leži strogo između dvaju korijena od P. Kako je korijen z veći od y, drugi korijen mora biti manji, odnosno $3xy - z < y$. □

Lema 4.22. *Ako su A, B i C kvadratne matrice reda $d \geq 2$ kojima su elementi prirodni brojevi, onda za svake $i, j \in \{1, \ldots, d\}$, imamo $(AB)_{ij} > A_{ij}, B_{ij}$ te $(ABC)_{ij} > (AC)_{ij}, B_{ij}$.*

Dokaz. Imamo $(AB)_{ij} = \sum_k A_{ik}B_{kj}$. Zbog pretpostavke da je $d \geq 2$ i da su svi elementi ≥ 1, ova je suma strogo veća od A_{ij} i B_{ij}. Sada je $(ABC)_{ij} = \sum_{k,l} A_{ik}B_{kl}C_{lj} > \sum_k A_{ik}B_{kk}C_{kj} \geq \sum_k A_{ik}C_{kj} = (AC)_{ij}$. Slično se vidi $(ABC)_{ij} > B_{ij}$. □

Korolar 4.23. *Ako je $w = uv$ standardna faktorizacija donje Christoffelove riječi w, onda je $\mu(w)_{12} > \mu(u)_{12}, \mu(v)_{12}$. Nadalje, ako je u prefiks of v, onda je $\mu(u)_{12} < \mu(v)_{12}$ i $\mu(uuv)_{12} < \mu(uvv)_{12}$, a ako je v sufiks od u, onda je $\mu(v)_{12} < \mu(u)_{12}$ i $\mu(uvv)_{12} < \mu(uuv)_{12}$. Posebno, $\mu(uuv)_{12} \neq \mu(uvv)_{12}$.*

Dokaz. Za nepraznu riječ m je $\mu(m)$ kvadratna matrica reda 2 s elementima iz skupa prirodnih brojeva, pa možemo primijeniti prethodnu lemu.

Budući da je $\mu(w) = \mu(u)\mu(v)$, prva tvrdnja slijedi direktno iz navedene leme.

Ako je u prefiks od v, onda je to pravi prefiks jer w nije kvadrat te možemo pisati $v = uv'$, gdje je v' neprazna. Zato je iz leme, $\mu(u)_{12} < \mu(v)_{12}$, a iz iste leme zaključujemo da je element na mjestu $(1, 2)$ u matrici $\mu(u)\mu(u)\mu(v)$ manji od odgovarajućeg elementa u matrici $\mu(u)\mu(u)\mu(v')\mu(v) = \mu(u)\mu(v)\mu(v)$.

Drugi se slučaj dokazuje slično. Zadnja tvrdnja se dobiva direktno ako je $w = ab$ te koristeći Korolar 4.10.ii ako je $w \neq ab$. □

Iduća lema nam omogućuje da iz zadnjeg Frickeovog identiteta dobijemo Markovljevu jednadžbu.

Lema 4.24. *Ako je (u, v) Christoffelov par i $A = \mu(u)$, $B = \mu(v)$, onda vrijedi $\mathrm{Tr}(ABA^{-1}B^{-1}) = -2$.*

Dokaz. Tvrdimo da je za svaki Christoffelov par (u, v), element $uvu^{-1}v^{-1}$ slobodne grupe $F(a, b)$ konjugiran elementu $aba^{-1}b^{-1}$. Promotrimo konstrukciju Christoffelovih parova preko njihova stabla. Lako se vidi da ako tvrdnja vrijedi za (u, v), onda vrijedi i za (u, uv) i (uv, v). Zaista, $u(uv)u^{-1}(uv)^{-1} = u(uvu^{-1}v^{-1})u^{-1}$ i $(uv)v(uv)^{-1}v^{-1} = uvu^{-1}v^{-1}$. Tvrdnja očito vrijedi za korijen stabla, pa zaključujemo da je istinita za sve Christoffelove parove.

Iz pokazane tvrdnje slijedi da su $ABA^{-1}B^{-1}$ i $\mu(a)\mu(b)\mu(a)^{-1}\mu(b)^{-1}$ slične matrice, pa je $\mathrm{Tr}(ABA^{-1}B^{-1}) = \mathrm{Tr}(\mu(a)\mu(b)\mu(a)^{-1}\mu(b)^{-1})$. Imamo

$$\begin{aligned}\mu(a)\mu(b)\mu(a)^{-1}\mu(b)^{-1} &= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 12 & 5 \\ 7 & 3 \end{bmatrix} \begin{bmatrix} 3 & -7 \\ -5 & 12 \end{bmatrix} = \begin{bmatrix} 11 & -24 \\ 6 & -13 \end{bmatrix},\end{aligned}$$

čime je dokaz dovršen. □

Dokaz Teorema 4.16. Neka je w donja Christoffelova riječ sa standardnom faktorizacijom uv i neka je $(x, y, z) = \left(\frac{1}{3}\operatorname{Tr}(\mu(u)), \frac{1}{3}\operatorname{Tr}(\mu(v)), \frac{1}{3}\operatorname{Tr}(\mu(w))\right)$. Po Frickeovom identitetu u Lemi 4.19 i po Lemi 4.24, imamo $(3x)^2 + (3y)^2 + (3z)^2 = (3x)(3y)(3z)$ iz čega slijedi da je $\{x, y, z\}$ Markovljeva trojka jer su prema Lemi 4.17.ii, brojevi x, y, z prirodni. Ako je $w \neq ab$, onda su prema Korolaru 4.10.ii i Korolaru 4.23 brojevi x, y, z različiti, pa je $\{x, y, z\}$ prava Markovljeva trojka. Ako je $w = ab$, onda iz $\mu(ab) = \left[\begin{smallmatrix} 12 & 5 \\ 7 & 3 \end{smallmatrix}\right]$, dobivamo $\{x, y, z\} = \{1, 2, 5\}$ što je također prava Markovljeva trojka.

Stoga je preslikavanje iz teorema dobro definirano te ćemo sada dokazati da je surjektivno. Neka je $\{x, y, z\}$ prava Markovljeva trojka. Možemo uzeti da je $x < y < z$. Tada je prema Lemi 4.21, $\{x, y, 3xy - z\}$ Markovljeva trojka i $3xy - z < y$. Suma $x + y + (3xy - z)$ je manja od $x + y + z$ te koristimo indukciju po toj sumi.

Ako je $\{x, y, 3xy - z\}$ neprava Markovljeva trojka, onda zbog $x < y$ i $3xy - z < y$, to mora biti $\{1, 1, 2\}$ prema Lemi 4.18. Tada imamo $x = 1$, $3xy - z = 1$, $y = 2$, pa je $z = 3 \cdot 2 - 1 = 5$ te je trojka $\{x, y, z\}$ pridružena donjoj Christoffelovoj riječi ab.

Ako je $\{x, y, 3xy - z\}$ prava Markovljeva trojka, onda je po indukciji jednaka $\left\{\frac{1}{3}\operatorname{Tr}(\mu(u)), \frac{1}{3}\operatorname{Tr}(\mu(v)), \frac{1}{3}\operatorname{Tr}(\mu(w))\right\}$ za neku donju Christoffelovu riječ w sa standardnom faktorizacijom $w = uv$. Kako je maksimum y, imamo prema Korolaru 4.23, $y = \frac{1}{3}\operatorname{Tr}(\mu(w)) = \frac{1}{3}\operatorname{Tr}(\mu(uv))$. Ako je $x = \frac{1}{3}\operatorname{Tr}(\mu(u))$ i $3xy - z = \frac{1}{3}\operatorname{Tr}(\mu(v))$, onda je po Korolaru 4.20 $\{x, y, z\}$ pridružena donjoj Christoffelovoj riječi uuv. Inače je $x = \frac{1}{3}\operatorname{Tr}(\mu(v))$ i $3xy - z = \frac{1}{3}\operatorname{Tr}(\mu(u))$ te prema komentaru nakon Korolara 4.20 vidimo da je $\{x, y, z\}$ pridružena donjoj Christoffelovoj riječi uvv.

Time je dokazana surjektivnost preslikavanja kojemu ćemo sada pokazati injektivnost. Pretpostavimo da je prava Markovljeva trojka $\{x < y < z\}$ pridružena dvjema pravim donjim Christoffelovim riječim $w_1 = u_1 v_1$ i $w_2 = u_2 v_2$ s naznačenim standardnim faktorizacijama. Ako je $w_1 = ab$, onda je također $w_2 = ab$ jer bi u protivnom $w_1 = ab$ bio pravi faktor od w_2 i stoga bi po Lemi 4.22 $\mu(w_2)_{12}$ bilo strogo veće od $\mu(w_1)_{12}$ te trojke pridružene w_1 i w_2 ne bi bile jednake jer nemaju isti maksimum.

Zato možemo pretpostaviti da je $w_1, w_2 \neq ab$. Prema Korolaru 4.10 možemo uzeti da je ili u_1 pravi prefiks od v_1 i u_2 pravi prefiks od v_2 ili je pak u_1 pravi prefiks od v_1 i v_2 pravi sufiks od u_2. Preostala dva slučaja su slični, pa ih preskačemo.

U prvom slučaju po Korolaru 4.10 imamo standardne faktorizacije $v_i = u_i v_i'$, $i = 1, 2$. Zbog toga prema Korolaru 4.23 vrijedi $\mu(u_i)_{12} < \mu(v_i)_{12} < \mu(w_i)_{12}$. Budući da je Markovljeva trojka $\{x < y < z\}$ pridružena $w_i = u_i v_i$, imamo po Lemi 4.17.ii

$$\begin{aligned}
x &= \frac{1}{3}\operatorname{Tr}(\mu(u_i)) = \frac{1}{3}\operatorname{Tr}(A_i), \\
y &= \frac{1}{3}\operatorname{Tr}(\mu(v_i)) = \frac{1}{3}\operatorname{Tr}(\mu(u_i v_i')) = \frac{1}{3}\operatorname{Tr}(A_i B_i), \\
z &= \frac{1}{3}\operatorname{Tr}(\mu(w_i)) = \frac{1}{3}\operatorname{Tr}(\mu(u_i^2 v_i')) = \frac{1}{3}\operatorname{Tr}(A_i^2 B_i),
\end{aligned}$$

gdje smo stavili $A_i = \mu(u_i)$, $B_i = \mu(v_i')$. Vidimo da vrijedi (i) u Korolaru 4.20, pa mora vrijediti (ii), tj. $\{x, y, 3xy - z\} = \left\{\frac{1}{3}\operatorname{Tr}(A_i), \frac{1}{3}\operatorname{Tr}(A_i B_i), \frac{1}{3}\operatorname{Tr}(B_i)\right\}$. Markovljeva trojka pridružena $v_i = u_i v_i'$ je $t_i = \left\{\frac{1}{3}\operatorname{Tr}(\mu(u_i)), \frac{1}{3}\operatorname{Tr}(\mu(v_i')), \frac{1}{3}\operatorname{Tr}(\mu(v_i))\right\}$. Zato je $t_i = \{x, y, 3xy - z\}$ te slijedi $t_1 = t_2$. Indukcijom po duljini riječi zaključujemo da je $v_1 = v_2$, tj. $u_1 v_1' = u_2 v_2'$. Iz jedinstvenosti standardne faktorizacije slijedi $u_1 = u_2$, što povlači $w_1 = w_2$.

U drugom slučaju imamo $v_1 = u_1 v_1'$ i $u_2 = u_2' v_2$. Slično kao prije, koristeći Korolar 4.20 i simetrični iskaz istog korolara, dobivamo da je trojka $\{x, y, 3xy - z\}$ pridružena objema donjim Christoffelovim riječima v_1 i u_2. Stoga je po pretpostavci indukcije $u_1 v_1' = u_2' v_2$

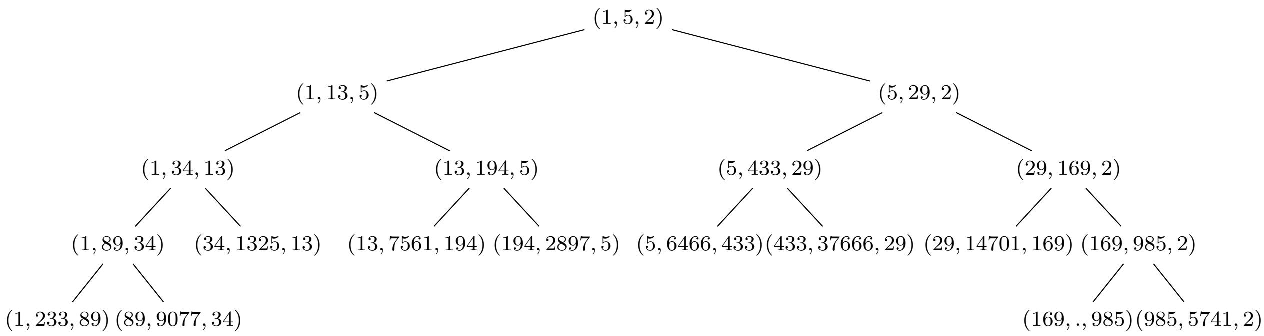
te iz jedinstvenosti faktorizacije u Christoffelove riječi, tj. Teorema 4.8, slijedi $u_1 = u_2'$, $v_1' = v_2$. Sada vidimo da je trojka $\{x, y, z\}$ pridružena riječima $w_1 = u_1 v_1 = u_1^2 v_2' = u_1^2 v_2$ i $w_2 = u_2 v_2 = u_2' v_2^2 = u_1 v_2^2$. Zato je maksimum ove trojke istovremeno jednak $\mu(u_1^2 v_2)_{12}$ i $\mu(u_1 v_2^2)_{12}$. Ovo je u kontradikciji sa zadnjom tvrdnjom u Korolaru 4.23 primijenjenom na Christoffelovu riječ $v_1 = u_1 v_2$. □

4.3.2 Stablo Markovljevih trojki

Budući da se svaki Christoffelov par pojavljuje točno jednom u stablu Christoffelovih parova, iz Teorema 4.16 slijedi da kad u tom stablu svaki čvor (u, v) zamijenimo uređenom trojkom

$$\left(\frac{1}{3}\operatorname{Tr}(\mu(u)), \frac{1}{3}\operatorname{Tr}(\mu(uv)), \frac{1}{3}\operatorname{Tr}(\mu(v))\right) \tag{4.5}$$

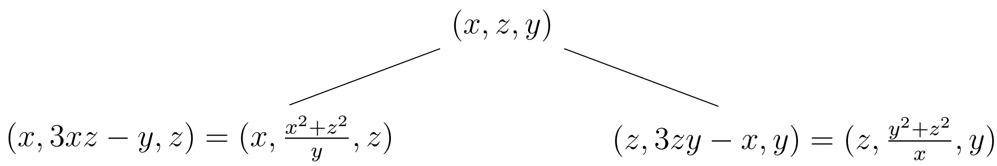
koja je prema Lemi 4.17 jednaka $(\mu(u)_{12}, \mu(uv)_{12}, \mu(v)_{12})$, dobivamo stablo u kojemu se svaka prava Markovljeva trojka pojavljuje točno jednom kao čvor, u obliku uređene trojke kojoj kao skup odgovara baš ta trojka. To se stablo naziva *stablo Markovljevih trojki*.



Početni dio stabla Markovljevih trojki

Iz izraza (4.5) i Korolara 4.23 slijedi da je u svakoj trojci (x, z, y) koja se pojavljuje u ovom stablu maksimum brojeva upravo z, no minimum može biti x ili y.

Ovo stablo može se i izravno konstruirati počevši od korijena $(1, 5, 2)$ i primjenjujući pravilo prikazano na idućoj slici.



Pravilo konstrukcije stabla Markovljevih trojki

Ispravnost ove konstrukcije slijedi iz Frickeovih identiteta. Neka je, naime, $(x, z, y) = \left(\frac{1}{3}\operatorname{Tr}(\mu(u)), \frac{1}{3}\operatorname{Tr}(\mu(uv)), \frac{1}{3}\operatorname{Tr}(\mu(v))\right)$. Tada je po Lemi 4.19

$$3xz - y = 3 \cdot \frac{1}{3}\operatorname{Tr}(\mu(u)) \cdot \frac{1}{3}\operatorname{Tr}(\mu(uv)) - \frac{1}{3}\operatorname{Tr}(\mu(v)) = \frac{1}{3}\operatorname{Tr}(\mu(u^2 v)),$$

pa je prema pravilu konstrukcije stabla Christoffelovih parova lijevo dijete na prethodnoj slici ispravno. Dokaz se analogno provodi za desno dijete.

Korolar 4.25. *Tri broja u Markovljevoj trojci su uvijek u parovima relativno prosti.*

Dokaz. Dovoljno je pokazati da su tri broja u Markovljevoj trojci relativno prosti jer iz Markovljeve jednadžbe slijedi da prost broj koji dijeli neka dva od njih, mora dijeliti i trećeg.

Tvrdnja je jasna za neprave trojke i trojku $(1, 5, 2)$, tj. korijen stabla Markovljevih trojki. Ako je (x, z, y) neki čvor u tome stablu pri čemu su x, y, z relativno prosti, onda su očito $x, 3xz - y, z$ relativno prosti, a isto tako i $z, 3zy - x, y$. Sada iz konstrukcije stabla slijedi traženi rezultat. □

4.3.3 Markovljeva slutnja o injektivnosti

Iz Teorema 4.16 slijedi da za svaki Markovljev broj m, postoji donja Christoffelova riječ w takva da je $m = \frac{1}{3}\operatorname{Tr}(\mu(w)) = \mu(w)_{12}$. Kažemo da je m Markovljev broj *pridružen* Christoffelovoj riječi w. U tom slučaju par brojeva $|w|_a$, $|w|_b$ se naziva *Frobeniusovim koordinatama* od m.

Označimo li m sa m_w, teorem nam govori da je preslikavanje $w \mapsto m_w$ surjektivno. *Markovljeva slutnja o injektivnosti* ili *slutnja o jedinstvenosti* je hipoteza da je navedeno preslikavanje injektivno. Ova slutnja prvi je puta iskazana u Frobeniusovom članku iz 1913. godine. Nekoliko ekvivalentnih formulacija slutnje kao i uz nju vezane parcijalne rezultate promotrit ćemo kasnije.

4.4 Markovljevo svojstvo

Markov je uveo jedno kombinatorno svojstvo obostrano beskonačnih riječi na binarnom alfabetu $\{a, b\}$. Riječ ima Markovljevo svojstvo ako se svaki faktor xy u toj riječi, gdje je $\{x, y\} = \{a, b\}$, može proširiti do faktora oblika $y\widetilde{m}xymx$ pri čemu je duljina od m ograničena. Markov je pokazao da takva riječ mora biti periodska, a u Teoremu 4.33 ćemo pokazati da je periodski uzorak upravo neka Christoffelova riječ.

U ovom potpoglavlju ćemo slično svojstvu uvesti i za (jednostrano) beskonačne riječi i dokazati da riječi s tim svojstvom moraju biti periodske sa sličnim periodskim uzorkom. Duljinu početnog neperiodskog dijela riječi ćemo ograničiti tako da navedeni teorem za obostrano beskonačne riječi nije teško dobiti korištenjem odgovarajućeg rezultata za jednostrano beskonačne riječi.

4.4.1 Markovljevo svojstvo beskonačnih riječi

Kažemo da beskonačna riječ s na alfabetu $\{a, b\}$ zadovoljava *Markovljevo svojstvo* ako postoje cijeli brojevi K i N takvi da za svaku faktorizaciju $s = \widetilde{u}xyv$ za koju je $\{x, y\} = \{a, b\}$ i $|u| \geq K$, vrijedi $u = myu'$, $v = mxv'$ i $|m| \leq N$. Primijetimo da je u konačna, a v beskonačna riječ.

Primjerice, lako je provjeriti da $s = a(ab)^\infty = aabababab\ldots$ zadovoljava Markovljevo svojstvo uz $K = 2$ i $N = 0$, a ne zadovoljava Markovljevo svojstvo uz $K = 1$. Drugi je primjer $(aab)^\infty = aabaabaab\ldots$ uz parametre $K = 2$ i $N = 1$.

Brojevi N i K u definiciji su nenegativni, a proširit ćemo definiciju i na slučaj kada je $N = -1$. Kažemo da s zadovoljava Markovljevo svojstvo za K i $N = -1$ ako ne postoji faktorizacija oblika $s = \widetilde{u}xyv$ pri čemu je $\{x, y\} = \{a, b\}$ i $|u| \geq K$. U tom slučaju je s od nekog mjesta konstantna, tj. $s = ux^\infty$, gdje je $x \in \{a, b\}$ i $|u| = K$.

U čitavom ovom odjeljku će nam s biti beskonačna riječ na alfabetu $\{a, b\}$ koja zadovoljava Markovljevo svojstvo uz parametre K i N.

Lema 4.26. *Neka je $s = ws_1$ za neku konačnu riječ w takvu da je $|w| = K - 1$. Tada riječ s_1 nema faktor oblika $x\widetilde{m}xymy$, gdje je $\{x, y\} = \{a, b\}$.*

Dokaz. Pretpostavimo da je $s_1 = hx\widetilde{m}xymy\ldots$. Tada je $whx\widetilde{m}xymy\ldots = ws_1 = s = \widetilde{u}xyv$, gdje je $\widetilde{u} = whx\widetilde{m}$ i $v = my\ldots$. Kako je $|u| \geq |w| + 1 = K$, iz Markovljeva svojstva dobivamo $u = nyu'$, $v = nxv'$. Stoga je $mx\widetilde{h}\widetilde{w} = u = nyu'$ i $my\ldots = v = nxv'$. Ne može biti $|m| = |n|$ jer je $x \neq y$. Ako je m kraća od n, onda su mx i my obje prefiksi od n što nije moguće, a slično imamo i ako je n kraća od m. Budući da u svakom slučaju dobivamo kontradikciju, lema je dokazana. □

Lema 4.27. *Ako je t beskonačna riječ na alfabetu $\{a, b\}$ takva da nijedna od riječi $x(xy)^ixy(yx)^iy$, $xy(xy)^jxy(yx)^jy^2$, gdje je $\{x, y\} = \{a, b\}$, $i, j \in \mathbb{N}_0$ nije faktor od t, onda je najviše jedna od riječi aa i bb faktor od t.*

Dokaz. Pretpostavimo suprotno, neka je $t = hx^2ky^2t_1$ za $\{x, y\} = \{a, b\}$. Možemo pretpostaviti da je k najkraća moguća. Tada je k različita od prazne riječi jer bi inače x^2y^2 bio faktor od t što je u suprotnosti s prvim uvjetom leme za $i = 0$. Iz minimalnosti duljine od k, imamo $k = (yx)^i$, $i \geq 1$. Riječ t_1 mora početi s x jer bi inače t sadržavao faktor $x^2(yx)^iy^3$, pa stoga i faktor $xyxy^3$ što je u kontradikciji s drugim uvjetom leme za $j = 0$. Zato možemo pisati $yt_1 = (yx)^jt_2$ za neki maksimalni prirodni broj j ili j beskonačan. Imamo $t = hx^2(yx)^iy(yx)^jt_2 = hx(xy)^ixy(yx)^jt_2$.

Ako je $j > i$, onda t sadrži kao faktor riječ $x(xy)^ixy(yx)^iy$ što je u kontradikciji s prvim uvjetom leme. Dakle, $j \leq i$, pa posebno j nije beskonačan.

Ako t_2 počinje s x, onda t sadrži $y(yx)^jx = y^2(xy)^{j-1}x^2$, pa iz minimalnosti od $|k|$ moramo imati $j - 1 \geq i$, odnosno $j > i$ što daje kontradikciju.

Ako t_2 počinje s yx, to je u kontradikciji s maksimalnošću od j.

Pretpostavimo da t_2 počinje s y^2. Tada t sadrži $x(xy)^ixy(yx)^jy^2$. Ako je $j = i$, onda t sadrži $x(xy)^ixy(yx)^iy$ što je u suprotnosti s prvim uvjetom leme. Ako je $j \in \{1, \ldots, i-1\}$, onda t sadrži $xy(xy)^jxy(yx)^jy^2$ što je u suprotnosti s drugim uvjetom leme.

Budući da u svakom slučaju dobivamo kontradikciju, lema je dokazana. □

Prisjetimo se da je G endomorfizam slobodnog monoida $\{a, b\}^*$ definiran s $G(a) = a$, $G(b) = ab$.

Lema 4.28. *Za svaki $v \in \{a, b\}^*$ je $\widetilde{G(va)} = \widetilde{G(v)a} = a\widetilde{G(v)} = G(\widetilde{v})a$. Posebno, $G(va)$ je palindrom ako i samo ako je v palindrom.*

Dokaz. Prve dvije jednakosti su očite. Treću jednakost dokazujemo indukcijom po duljini riječi v. Ako je $v = 1$, očito vrijedi. Pretpostavimo da je $v = aw$. Tada je

$$\begin{aligned} a\widetilde{G(v)} &= a\widetilde{G(a)G(w)} = a\widetilde{G(w)}a = G(\widetilde{w})aa \quad \text{[po pretpostavci indukcije]} \\ &= G(\widetilde{w})G(a)a = G(\widetilde{w}a)a = G(\widetilde{aw})a = G(\widetilde{v})a. \end{aligned}$$

Pretpostavimo sada da je $v = bw$. Vrijedi

$$\begin{aligned} a\widetilde{G(v)} &= a\widetilde{G(b)G(w)} = a\widetilde{G(w)}\widetilde{G(b)} = a\widetilde{G(w)}ba = G(\widetilde{w})aba \quad \text{[po pretpostavci indukcije]} \\ &= G(\widetilde{w})G(b)a = G(\widetilde{w}b)a = G(\widetilde{bw})a = G(\widetilde{v})a. \end{aligned}$$

Dokaz zadnje tvrdnje ove leme slijedi iz idućeg niza ekvivalencija, pri čemu u zadnjem koraku koristimo injektivnost od G koju smo pokazali u Propoziciji 4.12

$$G(va) \text{ palindrom } \Leftrightarrow \widetilde{G(va)} = G(va) \Leftrightarrow G(\widetilde{v})a = G(va) \Leftrightarrow G(\widetilde{v}) = G(v) \Leftrightarrow \widetilde{v} = v. \ \square$$

Lema 4.29. *Ako je $s = G(t)$ i $N \geq 0$, onda t zadovoljava Markovljevo svojstvo uz parametre $K-1$ i $N-1$.*

Dokaz. Ako t nema faktorizaciju oblika $t = \widetilde{u}abv$ ili $t = \widetilde{u}bav$ sa $|u| \geq K-1$, onda očito t zadovoljava Markovljevo svojstvo uz parametre $K-1$ i $N-1$.

U protivnom, t ima faktorizaciju ovog oblika. Neka je $t = \widetilde{u}abv$ sa $|u| \geq K-1$. Tada je prema Lemi 4.28, $s = G(t) = G(\widetilde{u})aabG(v) = (\widetilde{G(u)a})abG(v)$ i $|\widetilde{G(u)a}| \geq K$. Iz Markovljevog svojstva za s slijedi $G(u)a = nbu'$, $G(v) = nav'$, gdje je $|n| \leq N$. Sada u' ne može biti prazna, pa je $u' = u''a$ i $nbu'' = G(u)$ je u slici od G. Stoga n mora završiti sa a, tj. $n = n'a$ i n' je u slici od G, odnosno $n' = G(m)$, $n = G(m)a$. Tako je $G(u) = nbu'' = G(m)abu''$ i zato $u = mbu_1$. Nadalje, $G(v) = G(m)aav'$ i stoga $v = mav_1$. Konačno je $|m| \leq |G(m)| = |n'| = |n| - 1 < |n| \leq N$.

Ako je $t = \widetilde{u}bav$, dokaz se provodi slično. Treba promatrati jednakost $G(v) = nbv'$ koja povlači $n = n'a$ te daljnje detalje preskačemo. $\square$

Prisjetimo se da je D endomorfizam slobodnog monoida $\{a,b\}^*$ definiran s $D(a) = ba$, $D(b) = b$.

Korolar 4.30. *Ako je $s = G(t)$ ili $s = D(t)$ i $N \geq 0$, onda t zadovoljava Markovljevo svojstvo uz parametre $K-1$ i $N-1$.*

Dokaz. Ako je $s = G(t)$, to je upravo Lema 4.29. Ako je $s = D(t)$, rezultat se dobiva tako da u Lemama 4.28 i 4.29 pišemo D umjesto G te zamijenimo a i b čime sve tvrdnje ostaju vrijediti. $\square$

Teorem 4.31. *Ako beskonačna riječ s na alfabetu $\{a,b\}$ zadovoljava Markovljevo svojstvo uz parametre K i N, onda je $s = uw^\infty$ za neku konačnu riječ u i neku donju Christoffelovu riječ w. Duljina od u je omeđena funkcijom koja ovisi samo o K i N.*

Dokaz. Pokažimo najprije da postoji funkcija $f(K,N)$ defnirana za cijele brojeve $K \geq 0$ i $N \geq -1$ takva da je $f(K,-1) \geq K$ i za $N \geq 0$ vrijedi $K + 2f(N, N-1) + 1 = f(K,N)$. Možemo uzeti $f(K,N) = K + 2^{N+3} - 2N - 5$. Zaista, tada je $f(K,-1) = K + 2^2 + 2 - 5 = K + 1 \geq K$ i za $N \geq 0$ je

$$K+2f(N,N-1)+1 = K+2(N+2^{N+2}-2N+2-5)+1 = K+2^{N+3}-2N-6+1 = f(K,N).$$

Dokazat ćemo teorem indukcijom po N i pokazati da je $|u|$ omeđen odozgo s $f(K,N)$. Ako je $N = -1$, onda je $s = ux^\infty$ uz $|u| = K \leq f(K,-1)$, $x \in \{a,b\}$ i tvrdnja očito vrijedi.

Pretpostavimo sada da je $N \geq 0$. Neka je $s = ws_1$, gdje je $|w| = K-1$. Prema Lemi 4.26 riječ s_1 nema faktor oblika $x\widetilde{m}xymy$, pa zaključujemo da s_1 ispunjava uvjete Leme 4.27 iz čega slijedi da aa i bb nisu istovremeno faktori od s_1.

Tvrdimo da s_1 zadovoljava Markovljevo svojstvo uz parametre $K_1 = N+1$ i N. Zaista, ako je $s_1 = \widetilde{u}xyv$, $|u| \geq N+1$, onda je $s = w\widetilde{u}xyv$, $|w\widetilde{u}| \geq K+N \geq K$ i stoga $u\widetilde{w} = myu'$, $v = mxv'$, $|m| \leq N$. Iz uvjeta duljine na u slijedi da je $u = myu''$ čime je tvrdnja dokazana. Uočimo da ako uzmemo dulju riječ w te $s = ws_1$, prijašnji argumenti također vrijede.

Pretpostavimo da bb nije faktor od s_1. Tada s_1 počinje s a ili ba te možemo pretpostaviti, uzevši ako je potrebno da je riječ w za jedno slovo dulja, da s_1 počinje s a. Kako s_1 ne sadrži bb, imamo da je s_1 beskonačni produkt riječi a i ab, tj. $s_1 = G(t)$ za neku beskonačnu riječ t. Prema Korolaru 4.30, t zadovoljava Markovljevo svojstvo uz parametre N i $N-1$. Prema pretpostavci indukcije slijedi da je $t = uv^\infty$ za konačnu riječ u takvu da je $|u| \le f(N, N-1)$ i neku donju Christoffelovu riječ v. Tada je $s = ws_1 = wG(u)G(v)^\infty$. Štoviše, po Korolaru 4.14 je $G(v)$ donja Christoffelova riječ. Imamo $|wG(u)| \le K + 2|u| \le K + 2f(N, N-1) \le f(K, N)$ čime je dokaz završen u ovom slučaju.

Ako aa nije faktor od s_1, onda, simetrično, možemo pretpostaviti da je $s_1 = D(t)$. Slično kao prije, imamo $t = uv^\infty$ za neku konačnu riječ u takvu da je $|u| \le f(N, N-1)$ i neku donju Christoffelovu riječ v. Stoga je $s = ws_1 = wD(u)D(v)^\infty$. Po Korolaru 4.14 znamo da je $\widetilde{D}(v)$ donja Christoffelova riječ. Štoviše, vidjeli smo u dokazu Korolara 4.15 da je $\widetilde{D}(v) = v'b$ za neku riječ v' i da je onda $D(v) = bv'$. Zato je $s = wD(u)(bv')^\infty = wD(u)b(v'b)^\infty$. Konačno imamo $|wD(u)b| \le K + 2f(N, N-1) + 1 = f(K, N)$. □

Ovaj teorem ima i obrat: ako je $s = uw^\infty$ za neku donju Christoffelovu riječ w, tada s zadovoljava Markovljevo svojstvo. Ovo se može dokazati izravno, a može se dobiti i kao posljedica tvrdnji koje ćemo kasnije pokazati na putu prema dokazu Markovljevog teorema za verižne razlomke.

4.4.2 Markovljevo svojstvo obostrano beskonačnih riječi

Lema 4.32. *Ako su u i u' riječi te w i w' donje Christoffelove riječi na alfabetu $\{a < b\}$ takve da je $uw^\infty = u'w'^\infty$, onda je $w = w'$ i vrijedi $u = u'w^i$ ili $u' = uw^i$ za neki nenegativni cijeli broj i.*

Za proizvoljnu riječ $w \in \{a < b\}^*$, njezin *nagib* definiramo kao kvocijent $\frac{|w|_b}{|w|_a}$. Nagib riječi je očito nenegativan racionalan broj ili ∞. Ova definicija je u skladu s prijašnjom definicijom nagiba Christoffelove riječi.

Dokaz. Neka je p_n prefiks duljine n od uw^∞. Nagib od w jednak je limesu nagiba od p_n kad n teži u beskonačnost. Stoga iz pretpostavke leme slijedi da w i w' imaju isti nagib, a jer su to donje Christoffelove riječi zaključujemo da je $w = w'$.

Iz uvjeta leme slijedi da za neku faktorizaciju $w = fg$ imamo $w = gf$. No to je moguće samo ako je f ili g prazna riječ jer su po Propoziciji 4.5 sve riječi $C^k(w)$, $k \in \{0, \ldots, |w|-1\}$ različite. Stoga vrijedi i druga tvrdnja Leme. □

Kažemo da obostrano beskonačna riječ s na alfabetu $\{a, b\}$ zadovoljava *Markovljevo svojstvo* ako postoji cijeli broj N takav da za svaku faktorizaciju $s = \widetilde{u}xyv$ za koju je $\{x, y\} = \{a, b\}$, vrijedi $u = myu'$, $v = mxv'$ i $|m| \le N$. Primijetimo da su ovdje u, v, u', v' udesno beskonačne riječi dok je ulijevo beskonačna riječ $\widetilde{u}$ obratna riječi u.

Teorem 4.33. *Ako obostrano beskonačna riječ na alfabetu $\{a, b\}$ zadovoljava Markovljevo svojstvo, onda je, do na pomak, jednaka ${}^\infty w^\infty$, gdje je w neka donja Christoffelova riječ.*

Dokaz. Neka je s obostrano beskonačna riječ koja ima Markovljevo svojstvo. Za cijeli broj n, neka je s_n beskonačni niz dobiven tako da u s gledamo samo članove čiji je indeks barem n, tj. s_n je desni dio od s počevši od mjesta s indeksom n. Tada je lako provjeriti da je s_n beskonačna riječ koja zadovoljava Markovljevo svojstvo uz parametre $K = N+1$ i N. Iz Teorema 4.31 slijedi da je $s_n = u_n w_n^\infty$ za neku donju Christoffelovu riječ w_n i neku riječ u_n čija duljina je ograničena konstantom koja ne ovisi o n. Iz prethodne leme zaključujemo da je $w_n = w$ riječ koja ne ovisi o n i da je $s = {}^\infty w^\infty$ do na pomak. □

Slično kao prije, i ovaj teorem ima obrat.

4.5 Verižni razlomci i riječi

Najprije ćemo rezultate koje već znamo o verižnim razlomcima prereći u terminima njima pridruženih riječi i prirodnog uređaja koji se pri tome pojavljuje. Zatim ćemo dokazati ključnu kombinatornu činjenicu o donjim Christoffelovim riječima i njima pridruženim obratnim riječima.

4.5.1 Uređaj na verižnim razlomcima

Ako realne brojeve prikazujemo u bazi k, onda ih uspoređujemo koristeći *leksikografski uređaj*, $<_{lex}$. Zaista, realni brojevi u intervalu $[0, 1]$ predstavljeni su konačnim ili beskonačnim riječima na $\{0, 1, \dots, k-1\}$ te treba promatrati sljedeći uređaj na skupu takvih riječi. Imamo $u <_{lex} v$ ako je u pravi prefiks od v ili je $u = mau'$, $v = mbv'$ za riječi m, u', v' i znamenke a, b takve da je $a < b$. Primjerice, u bazi $k = 10$ je $0.1 < 0.1\dots$ i $0.91\dots < 0.92\dots$.

Za realne brojeve prikazane pomoću verižnih razlomaka, koristimo *alternirajući leksikografski uređaj* koji označavamo $<_{alt}$. Budući da želimo uspoređivati sve realne brojeve, promatramo skup svih nepraznih konačnih i beskonačnih riječi $a_0a_1a_2\dots$, gdje je $a_0 \in \mathbb{Z}$, $a_1, a_2, \dots \in \mathbb{N}$. Za dvije takve riječi je $u <_{alt} v$ ako vrijedi jedan od iduća dva međusobno isključiva uvjeta

(i) jedna od njih je pravi prefiks druge: ili je u pravi prefiks neparne duljine od v ili je v pravi prefiks parne duljine od u,

(ii) $u = mau'$, $v = mbv'$ za neke riječi m, u', v' i cijele brojeve a, b takve da je ili $a < b$ i m je parne duljine ili je $a > b$ i m je neparne duljine.

Primijetimo da su m u slučaju (ii) i prefiks u slučaju (i) nužno konačne riječi.

Ako su $u = au'$ i $v = av'$, gdje je $a \in \mathbb{Z}$ te su u', v' neprazne, onda je $u <_{alt} v$ ako i samo ako je $v' <_{alt} u'$. To je i objašnjenje naziva alternirajući za ovaj uređaj: najprije usporedimo prve znamenke (tj. slova) obiju riječi prema uobičajenom uređaju, onda, ako vrijedi jednakost, usporedimo druge znamenke prema obrnutom uređaju, zatim, ako vrijedi jednakost, usporedimo treće znamenke prema uobičajenom uređaju i nastavimo tako. Na taj način se lako provjeri da je alternirajući leksikografski uređaj potpun uređaj. Također vidimo da ako je $u \leq_{alt} v$, onda za riječ w parne duljine imamo $wu \leq_{alt} wv$ dok za w neparne duljine imamo $wu \geq_{alt} wv$.

Za $u = a_0a_1a_2\dots$, označimo s $[u]$ verižni razlomak $[a_0, a_1, a_2, \dots]$.

Teorem 4.34. *Preslikavanje $u \mapsto [u]$ sa skupa riječi kao gore u skup realnih brojeva zadovoljava: ako je $u <_{alt} v$, onda je $[u] \leq [v]$.*

Primijetimo da se jednakost može postići samo ako su u i v konačne riječi takve da je $u = ma$ i $v = m(a-1)1$ ili obratno. Primjeri za ovaj teorem su nejednakosti $[2] < [2, 1, \ldots]$, $[2, 1, 3, \ldots] < [2, 1]$, $[2, \ldots] < [3, \ldots]$, $[2, 3, \ldots] < [2, 1, \ldots]$.

Dokaz. Dokaz slijedi iz Teorema 1.8, ali ćemo ga ipak zapisati.

Ako je $u = au'$ za $a \in \mathbb{Z}$, onda je najveće cijelo od $[u]$ jednako a te je ili u' prazna riječi i $[u] = a$ ili je u' neprazna i $[u] = a + [u']^{-1}$.

Pretpostavimo da je $u <_{alt} v$. U nastavku će biti $a, b \in \mathbb{Z}$ dok će u', v' biti neprazne riječi.

Ako je u duljine 1, tj. $u = a$, onda je $[u] = a$ i mora biti ili $v = b$ takav da je $a < b$ ili $v = bv'$ takav da je $a \leq b$. U prvom slučaju je $[u] < [v]$, a u drugom slučaju vrijedi ista nejednakost jer je $[v] = b + [v']^{-1}$ te $[v'] > 0$ budući da v' ima samo pozitivna slova.

Uzmimo sada $u = au'$. Ako je v duljine 1, onda je $v = b$ i ne možemo imati $b = a$ jer bi inače v bio prefiks neparne duljine od u i stoga $v <_{alt} u$. Ne može biti ni $b < a$, pa je $a < b$. Stoga je $a + 1 \leq b$ i zato $[u] = a + [u']^{-1} \leq b = [v]$ jer je $[u'] \geq 1$.

Ako je $v = bv'$, onda ne možemo imati $a > b$. Ako je $a = b$, vrijedi $u' >_{alt} v'$, pa indukcijom po duljini najvećeg zajedničko prefiksa od u i v dobivamo $[u'] \geq [v']$ i stoga $[u] \leq [v]$. Ako je $a < b$, izravno dobivamo $[u] < [v]$ jer je najveće cijelo od $[u]$ jednako a, dok je najveće cijelo od $[v]$ jednako b. $\square$

Korolar 4.35. *Neka su su u i v riječi kao gore takve da je $[u] \neq [v]$. Ova je pretpostavka sigurno točna ako je jedna od riječi konačna, a druga beskonačna ili ako je jedna od riječi pravi prefiks druge. Tada je $[u] < [v]$ ekvivalentno s $u <_{alt} v$.*

Iduća lema je vrlo jednostavna, pa ispuštamo dokaz. Kažemo da je u riječi u svako slovo *ponovljeno* ako je $u = a_0a_0a_1a_1a_2a_2\ldots$.

Lema 4.36. *Neka su u i v konačne ili beskonačne riječi u kojima je svako slovo ponovljeno i nijedna od tih riječi nije prefiks druge. Tada je $u <_{lex} v$ ako i samo ako je $u <_{alt} v$.*

Ograničenje u uvjetu leme je nužno jer, primjerice, imamo $11 <_{lex} 1122$ i $11 >_{alt} 1122$.

4.5.2 Verižni razlomci pridruženi Christoffelovim riječima

Neka je w donja Christoffelova riječ i $\widetilde{w}$ njoj pridružena gornja Christoffelova riječ kao u Teoremu 4.6. Označimo kao prije s $m = m_w$ Markovljev broj pridružen w. Označimo s χ homomorfizam sa slobodnog monoida $\{a, b\}^*$ u slobodni monoid $\{1, 2\}^*$ takav da je $\chi(a) = 11$ i $\chi(b) = 22$. Prisjetimo se da je homomorfizam monoida $\mu : \{a, b\}^* \to \mathrm{SL}(2, \mathbb{Z})$ definiran s $\mu(a) = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right]$ i $\mu(b) = \left[\begin{smallmatrix} 5 & 2 \\ 2 & 1 \end{smallmatrix}\right]$. Neka je $\mu(w) = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right]$. Tada po Lemi 4.17 vrijedi $q = m = \frac{p+s}{3}$. Primijetimo također da su brojevi p, q, r, s pozitivni i cijeli.

Teorem 4.37. *Neka su realni brojevi $x = [\chi(w)^\infty]$ i $y = [\chi(\widetilde{w})^\infty]$ dani svojim razvojem u verižne razlomke. Tada je*

$$x = \frac{px + q}{rx + s}, \quad y = \frac{py + r}{qy + s} = \frac{p - s + \sqrt{(s-p)^2 + 4qr}}{2q} = \frac{p - s + \sqrt{9m^2 - 4}}{2m}, \quad x = -\frac{1}{y'}.$$

Podsjetimo se da je y' konjugat kvadratne iracionalnosti y.

Dokaz. Neka je $\chi(w) = a_1 \dots a_n$. Tada je $\chi(\widetilde{w}) = a_n \dots a_1$. Kako je

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \mu(a), \qquad \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} = \mu(b),$$

imamo

$$\begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix} = \mu(w) = \begin{bmatrix} p & q \\ r & s \end{bmatrix}, \qquad \begin{bmatrix} a_n & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} a_1 & 1 \\ 1 & 0 \end{bmatrix} = \mu(w)^\tau = \begin{bmatrix} p & r \\ q & s \end{bmatrix}.$$

Sada korištenjem (1.9) i (1.3) dobivamo $x = \frac{px+q}{rx+s}$ i $y = \frac{py+r}{qy+s}$ dok $x = -\frac{1}{y'}$ slijedi iz Propozicije 1.5.

Budući da je y pozitivan broj koji zadovoljava jednadžbu $qy^2 + (s-p)y - r = 0$, a q i r su pozitivni brojevi, vidimo da vrijedi druga jednakost za y iz tvrdnje teorema. Treba još samo primijetiti da je $q = m$ te $(s-p)^2 + 4qr = (s+p)^2 - 4ps + 4qr = (s+p)^2 - 4 = 9m^2 - 4$ jer je determinanta od $\mu(w)$ jednaka 1. □

Korolar 4.38. *Neka su x, y kao u prethodnom teoremu. Tada je* $y + \dfrac{1}{x} = \sqrt{9 - \dfrac{4}{m^2}}$.

Dokaz. Vrijedi $y + \dfrac{1}{x} = y - y' = \dfrac{p - s + \sqrt{9m^2 - 4}}{2m} - \dfrac{p - s - \sqrt{9m^2 - 4}}{2m} = \dfrac{\sqrt{9m^2 - 4}}{m}$. □

4.5.3 Markovljev supremum obostrano beskonačnog niza

Prisjetimo se da je za obostrano beskonačni niz $s = (a_n)_{n \in \mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$ i cijeli broj i,

$$\lambda_i(s) = [a_{i-1}, a_{i-2}, a_{i-3}, \dots]^{-1} + a_i + [a_{i+1}, a_{i+2}, a_{i+3}, \dots]^{-1}$$

te je Markovljev supremum od s jednak $M(s) = \sup_{i \in \mathbb{Z}} \lambda_i(s)$.

Teorem 4.39. *Neka je w neka donja Christoffelova riječ kojoj je pridružen Markovljev broj m. Neka je s obostrano beskonačni niz ${}^\infty\chi(w)^\infty$. Tada je $M(s) = \sqrt{9 - \frac{4}{m^2}}$.*

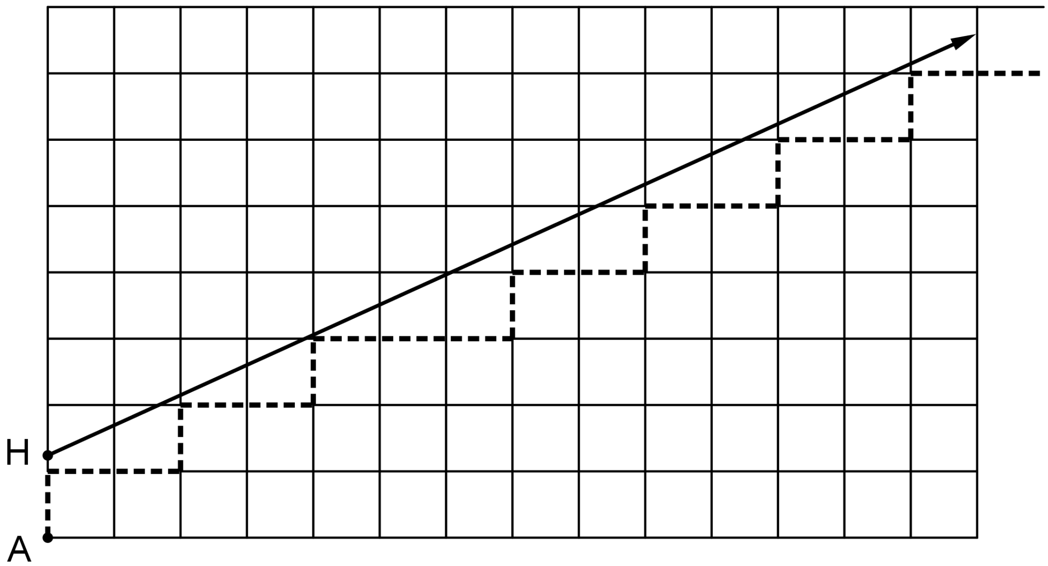
Ovaj ćemo teorem dokazati korištenjem Teorema 4.37 i sljedećeg kombinatornog svojstva Christoffelovih riječi.

Teorem 4.40. *Neka je w donja Christoffelova riječ i $\widetilde{w}$ odgovarajuća gornja Christoffelova riječ. Tada su w i $\widetilde{w}$ konjugirane riječi te je u njihovoj konjugacijskoj klasi w najmanja, dok je $\widetilde{w}$ najveća po leksikografskom uređaju.*

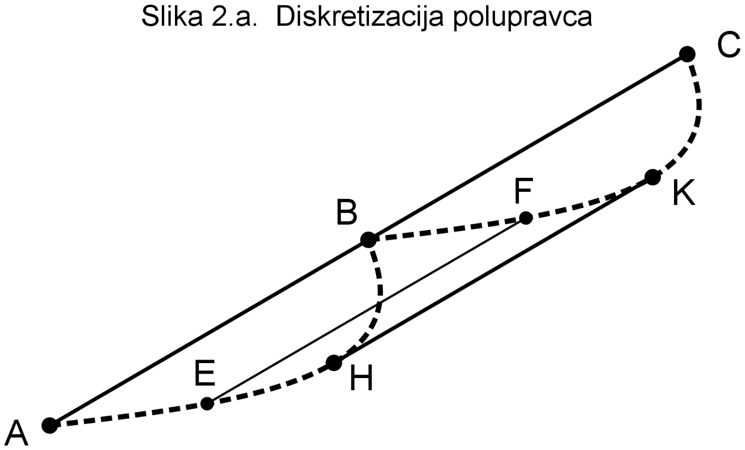
Kako bismo dokazali ovaj teorem, proširit ćemo definiciju diskretizacije s dužina i konačnih riječi na polupravce i beskonačne riječi.

Promatramo parove (A, ℓ) gdje je A cjelobrojna točka u ravnini, a ℓ polupravac nenegativnog nagiba. Kažemo da je takav par u *vertikalnom položaju* ako točka A i točka H u kojoj polupravac počinje imaju istu x-koordinatu te se A nalazi ispod H (vidi Sliku 2.a). *Vertikalna udaljenost* između A i ℓ je duljina $|AH|$.

Slično kao kod definicije donje Christoffelove riječi, *diskretizacija odozdo* para (A, ℓ) je beskonačna riječ na alfabetu $\{a < b\}$ koja kodira beskonačni put u rešetki koji počinje u A te između polupravca ℓ i tog puta nema cjelobrojnih točaka osim onih na samom putu. Diskretizacija na Slici 2.a je beskonačna riječ *baabaabaaabaabaabaaba* ....



Slika 2.a. Diskretizacija polupravca



Slika 2.b. Donja i gornja Christoffelova riječ su konjugati

Lema 4.41. *Neka su (A, ℓ) i (A, ℓ') dva para u vertikalnom položaju, gdje su ℓ i ℓ' paralelni polupravci takvi da je vertikalna udaljenost između A i ℓ manja od vertikalne udaljenosti A i ℓ'. Tada diskretizacija odozdo para (A, ℓ) po leksikografskom uređaju nije veća od diskretizacije odozdo para (A, ℓ').*

Ako je nagib polupravaca iz prethodne leme racionalan, onda diskretizacije mogu biti jednake. Trivijalan primjer za to su polupravci u prvom kvadrantu koji leže na pravcima s jednadžbama $y = \frac{1}{3}$ i $y = \frac{2}{3}$ te točka $A = (0, 0)$. No, ako je nagib iracionalan, nije teško pokazati da mora postojati cjelobrojna točka strogo između dvaju polupravaca, pa su diskretizacije različite i jedna od njih je zato strogo manja od druge. Za dokaz ove činjenice dovoljno je vidjeti da je niz razlomljenih dijelova $(\{\kappa n\})_{n \in \mathbb{N}}$ gust u $[0, 1]$ kad je κ iracionalan broj. To se može dobiti iz Teorema 1.9.b.

Dokaz. Po pretpostavci je polupravac ℓ ispod polupravca ℓ'. Promotrimo put u rešetki koji diskretizira (A, ℓ) i put koji diskretizira (A, ℓ'). Oni imaju istu početnu točku. Uočimo, ako postoji, prvi korak gdje se razlikuju. Taj korak je nužno horizontalan za prvi put, a vertikalan za drugi put, pa je prva riječ manja od druge. $\square$

Korolar 4.42. *Neka su (A, ℓ) i (A', ℓ') dva para u vertikalnom položaju, gdje su ℓ i ℓ' paralelni polupravci takvi da je vertikalna udaljenost između A i ℓ manja od vertikalne udaljenosti A' i ℓ'. Tada diskretizacija odozdo para (A, ℓ) po leksikografskom uređaju nije veća od diskretizacije odozdo para (A', ℓ').*

Dokaz. Translatiranjem za cjelobrojni vektor svodimo tvrdnju na slučaj $A = A'$ što pokriva upravo prethodna lema. □

Dokaz Teorema 4.40. Promotrimo Sliku 2.b zanemarujući na trenutak točke E i F i dužinu koja ih povezuje. Cjelobrojne točke A i B odabrane su kao prije tako da $B - A$ ima nenegativne koordinate koje su relativno prosti brojevi. Točka C je definirana s $\overrightarrow{AB} = \overrightarrow{BC}$. Iscrtkana krivulja AHB predstavlja put u rešetki koji je kodiran donjom Christoffelovom riječju w. Budući je $\overrightarrow{AB} = \overrightarrow{BC}$, ista riječ kodira i put prikazan iscrtkanom krivuljom BKC. Točka H je cjelobrojna točka na putu AHB koja je najudaljenija od pravca AB. Slično vrijedi za točku K na putu BKC te imamo $\overrightarrow{AB} = \overrightarrow{HK}$.

Riječ predstavljena krivuljom HBK je očito konjugat od w, a mi ćemo pokazati da je to upravo $\widetilde{w}$, gornja Christoffelova riječ pridružena w. Za to je dovoljno pokazati da je put HBK diskretizacija odozgo dužine $\overline{HK}$, odnosno, da nema cjelobrojnih točaka u nutrini poligona $\mathcal{P}$ omeđenog s $\overline{HK}$ i putom HBK.

Dijagonalnom udaljenošću dvaju paralelnih pravaca nazovimo duljinu dužine koju na bilo kojem pravcu nagiba -1 odsijeca taj par pravaca. Dijagonalna udaljenost između pravaca AB i HK je strogo manja od $\sqrt{2}$. Naime, nema cjelobrojnih točaka ni na otvorenoj dužini AB ni u nutrini poligona omeđenog putom AHB i dužinom $\overline{AB}$, pa točka $H + (-1, 1)$ nije u tom poligonu.

Pretpostavimo sada da postoji cjelobrojna točka M u nutrini poligona $\mathcal{P}$. Ta je točka strogo ispod puta HBK, pa je za neki prirodan broj i, točka $M + (-i, i)$ na tom putu. Stoga je $M + (-1, 1)$ u nutrini poligona ili na njegovu rubu HBK. Budući da $\mathcal{P}$ leži između pravaca ABC i HK, slijedi da je dijagonalna udaljenost ovih pravaca barem $\sqrt{2}$ što je u kontradikciji s prethodnom donjom ogradom na tu dijagonalnu udaljenost.

Time smo dokazali da su w i $\widetilde{w}$ konjugati te da $\widetilde{w}$ kodira put HBK.

Neka je sada u proizvoljni konjugat od w različit od w i $\widetilde{w}$. Referirajući se na Sliku 2.b, postoji točka E na putu AHB i odgovarajuća točka F na putu BKC, tj. $\overrightarrow{AB} = \overrightarrow{EF}$, tako da u kodira put EBF. Promatramo udaljenosti točaka A, E i H od pravca AB. Udaljenost točke E je strogo veća od udaljenosti točke A koja je 0, a strogo manja od udaljenosti točke H. Naime, kada bi udaljenosti A i E bile jednake, točka E bi ležala na otvorenoj dužini AB, a kad bi udaljenosti E i H bile jednake, na istoj bi otvorenoj dužini ležala točka $A + \overrightarrow{EH}$. Obje su mogućnosti u kontradikciji s izborom točaka A i B. Budući da omjer udaljenosti točke od pravca i vertikalne udaljenosti točke od pravca ovisi samo o nagibu pravca, zaključujemo da je vertikalna udaljenost E od AB strogo veća od vertikalne udaljenosti A od AB, a strogo manja od vertikalne udaljenosti H od AB. Korolar 4.42 povlači da je $w^\infty <_{lex} u^\infty <_{lex} \widetilde{w}^\infty$. Budući da su tri riječi w, u, $\widetilde{w}$ iste duljine, zaključujemo da je $w <_{lex} u <_{lex} \widetilde{w}$. □

Korolar 4.43. *Neka je w donja Christoffelova riječ, $\widetilde{w}$ njoj pridružena gornja Christoffelova riječ te u proizvoljni konjugat od w. Tada je*

$$[\chi(\widetilde{u})^\infty]^{-1} + [\chi(u)^\infty] \leq [\chi(w)^\infty]^{-1} + [\chi(\widetilde{w})^\infty] = \sqrt{9 - \frac{4}{m^2}},$$

gdje je m Markovljev broj pridružen w.

Dokaz. Ako su dvije riječi konjugirane, onda su i njima obratne riječi konjugirane. Iz Teorema 4.40 proizlazi da je w konjugirana obratima svih svojih konjugata.

Prema Teoremu 4.40 je $w \leq_{lex} \widetilde{u}$ jer je $\widetilde{u}$ konjugat od w, a također je $u \leq_{lex} \widetilde{w}$. Budući da je χ rastuća funkcija s obzirom na leksikografski uređaj, dobivamo $\chi(w) \leq_{lex} \chi(\widetilde{u})$ i $\chi(u) \leq_{lex} \chi(\widetilde{w})$. Kako ove riječi imaju istu duljinu, slijedi da je $\chi(w)^\infty \leq_{lex} \chi(\widetilde{u})^\infty$ i $\chi(u)^\infty \leq_{lex} \chi(\widetilde{w})^\infty$. Primijetimo da je u riječima iz slike od χ svako slovo ponovljeno, pa Lema 4.36 povlači da je $\chi(w)^\infty \leq_{alt} \chi(\widetilde{u})^\infty$ i $\chi(u)^\infty \leq_{alt} \chi(\widetilde{w})^\infty$. Zato je prema Teoremu 4.34, $[\chi(\widetilde{u})^\infty]^{-1} + [\chi(u)^\infty] \leq [\chi(w)^\infty]^{-1} + [\chi(\widetilde{w})^\infty]$. Zadnja jednakost u tvrdnji korolara slijedi iz Korolara 4.38. □

Dokaz Teorema 4.39. Neka je $s = (a_n)_{n\in\mathbb{Z}} = {}^\infty\chi(w)^\infty$. Podsjetimo se da je prema dogovoru o oznakama, a_0 prvo slovo u $\chi(w)$.

Promotrimo li preslikavanje χ i obostrano beskonačnu riječ s, vidimo da za svaki paran cijeli broj i, postoji neki konjugat u od w takav da je $\ldots a_{i-3}a_{i-2}a_{i-1} = {}^\infty\chi(u)$ i $a_i a_{i+1} a_{i+2} \ldots = \chi(u)^\infty$. Tada kažemo da je u *određen* sa i. Prema Teoremu 4.40 je $\widetilde{w}$ konjugat od w, pa postoji neki paran i_0 takav da je $\widetilde{w}$ određen sa i_0.

Pretpostavimo da je i paran cijeli broj te neka je u određen sa i. Tada je

$$\lambda_i(s) = [a_{i-1}, a_{i-2}, a_{i-3}, \ldots]^{-1} + [a_i, a_{i+1}, a_{i+2}, \ldots] = [\chi(\widetilde{u})^\infty]^{-1} + [\chi(u)^\infty].$$

Slično je $\lambda_{i_0}(s) = [\chi(w)^\infty]^{-1} + [\chi(\widetilde{w})^\infty]$, pa dobivamo $\lambda_i(s) \leq \lambda_{i_0}(s) = \sqrt{9 - \frac{4}{m^2}}$ iz Korolara 4.43.

Pretpostavimo sada da je i neparan broj te neka je v određen sa $i+1$. Tada je

$$\lambda_i(s) = [a_i, a_{i-1}, a_{i-2}, \ldots] + [a_{i+1}, a_{i+2}, a_{i+3}, \ldots]^{-1} = [\chi(\widetilde{v})^\infty] + [\chi(v)^\infty]^{-1}.$$

Kako je $\widetilde{v}$ konjugat od w, opet iz Korolara 4.43 dobivamo $\lambda_i(s) \leq \lambda_{i_0}(s)$.

Zaključujemo da je $M(s) = \lambda_{i_0}(s)$. □

4.5.4 Lagrangeov broj niza

Neka je dan beskonačni verižni razlomak $[a_0, a_1, a_2, \ldots]$. Motivirani karakterizacijom iz Propozicije 1.12, definiramo *Lagrangeov broj* beskonačne riječi $u = a_0a_1a_2\ldots$ kao

$$L(u) = \limsup_{n\to\infty} \left([a_{n-1}, a_{n-2}, \ldots, a_1]^{-1} + [a_n, a_{n+1}, a_{n+2}, \ldots]\right).$$

Teorem 4.44. *Neka je w neka donja Christoffelova riječ kojoj je pridružen Markovljev broj m. Tada je Lagrangeov broj beskonačne riječi $\chi(w)^\infty$ jednak $\sqrt{9 - \frac{4}{m^2}}$.*

Dokaz je vrlo sličan dokazu Teorema 4.39 uz neke modifikacije uzrokovane time što se ovdje bavimo jednostrano beskonačnim riječima.

Dokaz. Ako za $n \geq 2$ označimo kao prije $\lambda_n(s) = [a_{n-1}, a_{n-2}, \ldots, a_1]^{-1} + [a_n, a_{n+1}, a_{n+2}, \ldots]$, onda je Lagrangeov broj od $s = \chi(w)^\infty$ jednak $\limsup_{n\to\infty} \lambda_n(s)$.

Neka je n paran broj veći od duljine riječi $\chi(w)$. Tada je za neki konjugat u od w, $[a_n, a_{n+1}, a_{n+2}, \ldots]$ jednako $[\chi(u)^\infty]$ i $[a_{n-1}, a_{n-2}, \ldots, a_1]$ je jednako $[\chi(\widetilde{u})^k p]$, gdje je $k \geq 1$ i p je prefiks od $\chi(\widetilde{u})$. Prema Teoremu 4.40 je $u \leq_{lex} \widetilde{w}$ i $\widetilde{u} \geq_{lex} w$, pa je $\chi(u) \leq_{lex} \chi(\widetilde{w})$ i $\chi(\widetilde{u}) \geq_{lex} \chi(w)$, a to po Lemi 4.36 povlači da je $\chi(u) \leq_{alt} \chi(\widetilde{w})$ i $\chi(\widetilde{u}) \geq_{alt} \chi(w)$. Kako su $\chi(u)$ i $\chi(w)$ iste duljine, slijedi $\chi(u)^\infty \leq_{alt} \chi(\widetilde{w})^\infty$ i, zbog $k \geq 1$, $\chi(\widetilde{u})^k p \geq_{alt} \chi(w)^\infty$.

Po Teoremu 4.34 dobivamo $[\chi(u)^\infty] \leq [\chi(\widetilde{w})^\infty]$ i $[\chi(\widetilde{u})^k p] \geq [\chi(w)^\infty]$. Stoga je $\lambda_n(s) = [\chi(\widetilde{u})^k p]^{-1} + [\chi(u)^\infty] \leq [\chi(w)^\infty]^{-1} + [\chi(\widetilde{w})^\infty]$.

Primijetimo da možemo uzeti paran n takav da je njime određena riječ u jednaka $\widetilde{w}$. Tada je $[\chi(\widetilde{u})^k p] = [\chi(w)^k p]$ konvergenta od $[\chi(w)^\infty]$ budući da je p prefiks od $\chi(w)$. Uzmemo li limes po takvim brojevima n, dobivamo da je Lagrangeov broj veći ili jednak $[\chi(w)^\infty]^{-1} + [\chi(\widetilde{w})^\infty]$.

Neka je sada n neparan broj veći od duljine riječi $\chi(w)$. Budući da je $n+1$ paran, imamo $\lambda_n(s) = [a_n, a_{n-1}, \ldots, a_1] + [a_{n+1}, a_{n+2}, \ldots]^{-1} = [\chi(\widetilde{u})^k p] + [\chi(u)^\infty]^{-1}$ za neki konjugat u od w, prirodan broj k i prefiks p od $\chi(\widetilde{u})$. Kao i maloprije, ali ovaj puta koristeći $w \leq_{lex} u$ i $\widetilde{u} \leq_{lex} \widetilde{w}$, dobivamo da je $\lambda_n(s) \leq [\chi(\widetilde{w})^\infty] + [\chi(w)^\infty]^{-1}$.

Stoga je Lagrangeov broj upravo $[\chi(\widetilde{w})^\infty] + [\chi(w)^\infty]^{-1}$ čime je dokaz po Korolaru 4.38 završen. $\square$

4.6 Lagrangeovi brojevi manji od tri

U ovom, više tehničkom, potpoglavlju dokazujemo ključne činjenice Markovljeve teorije. Ako niz ima Lagrangeov broj manji od 3, onda mora ispunjavati mnoge kombinatorne uvjete. Osim nekog početnog dijela niza svaki član mora biti 1 ili 2, faktori 121 i 212 se ne mogu pojavljivati, oko faktora 1122 i 2211 pojavljuju se riječi s određenim leksikografskim svojstvima, razni specijalni faktori su zabranjeni, a slova 1 i 2 se pojavljuju u parovima. Taj niz lema završava teoremom u kojemu dokazujemo da je takav niz slika s obzirom na supstituciju $a \mapsto 11$, $b \mapsto 22$ nekog niza koji zadovoljava Markovljevo svojstvo. Sličan rezultat ustanovit ćemo i za obostrano beskonačne nizove.

4.6.1 Od $L(s) < 3$ do Markovljevog svojstva

Prisjetimo se da je Lagrangeov broj beskonačne riječi $s = a_0 a_1 a_2 \ldots$ na alfabetu $\mathbb{Z}$, gdje je $a_i \in \mathbb{N}$ za $i > 0$, definiran sa $\limsup_{i\to\infty} \lambda_i(s)$ pri čemu je za $i \geq 2$

$$\lambda_i(s) = [a_{i-1}, a_{i-2}, \ldots, a_1]^{-1} + [a_i, a_{i+1}, a_{i+2}, \ldots] = [a_i, a_{i-1}, \ldots, a_1] + [a_{i+1}, a_{i+2}, \ldots]^{-1},$$

a za $i = 1$ je $\lambda_1(s) = [a_1, a_2, \ldots]$. Kad želimo naznačiti indeks i, podcrtat ćemo slovo na tom položaju, tj. $\underline{a_i}$. Podsjetimo se da je konačna riječ w faktor od s ako je $s = uwt$ za neku konačnu riječ u i neku beskonačnu riječ t.

U čitavom ovom odjeljku promatramo beskonačnu riječ $s = a_0 a_1 a_2 \ldots$ koja ima Lagrangeov broj strogo manji od 3. To znači da postoji prirodan broj I i realan broj $\zeta < 3$ tako da za svaki $i \geq I$ vrijedi $\lambda_i(s) \leq \zeta$.

Lema 4.45. *Neka je u konačna ili beskonačna riječ na alfabetu $\mathbb{N}$ i neka je $n \geq 2$ prirodan broj. Tada je $[0nu] + [01(n-1)u] = 1$.*

Dokaz. Ako je u prazna riječ, tvrdnju je lako provjeriti. Za nepraznu u, neka je $x = [u]$. Tada lema slijedi iz jednakosti

$$\frac{1}{n+\frac{1}{x}} + \frac{1}{1+\frac{1}{n-1+\frac{1}{x}}} = \frac{1}{n+\frac{1}{x}} + \frac{1}{1+\frac{x}{(n-1)x+1}} = \frac{x}{nx+1} + \frac{(n-1)x+1}{(n-1)x+1+x} = 1. \quad \square$$

Lema 4.46. *(i) Ako je $i \geq I$, onda je $a_i \in \{1, 2\}$.*

(ii) Ako je $s = a_0 \ldots 1\underline{2}1 \ldots$, gdje je i položaj podcrtanog slova, onda je $i < I$.

(iii) Ako je $s = a_0 \dots 21\underline{2} \dots$, gdje je i položaj podcrtanog slova, onda je $i < I$.

Dokaz. (i) Budući je $\lambda_i(s) = [a_i, a_{i-1}, \dots, a_1] + [a_{i+1}, a_{i+2}, \dots]^{-1} \geq a_i$ i $\lambda_i(s) < 3$ za $i \geq I$, vidimo da je $a_i \in \{1, 2\}$ za $i \geq I$.

(ii) Imamo $s = a_0\widetilde{u}1\underline{2}1v$ te je $\lambda_i(s) = 2 + [1u]^{-1} + [1v]^{-1}$. Očito je $[1u] < 2$ i $[1v] < 2$, pa su im inverzi veći od $\frac{1}{2}$ i stoga $\lambda_i(s) > 3$. Zato je nužno $i < I$.

(iii) Imamo $s = a_0\widetilde{u}21\underline{2}v$. Pretpostavimo suprotno tvrdnji leme, tj. da je $i \geq I$. Tada je $\lambda_i(s) < 3$. Prema upravo dokazanom dijelu (ii), iza ovog pojavka faktora 212 ne slijedi 1, pa je $s = a_0\widetilde{u}21\underline{2}2v'$ i $\lambda_i(s) = 2 + [012u] + [02v']$. Bilo iz Teorema 1.8, bilo iz Teorema 4.34 i njegova korolara, dobivamo $[012u] \geq [012]$ i $[02v'] > [021]$. Stoga je $\lambda_i(s) \geq 2 + [012] + [021] = 3$ što je kontradikcija s našom pretpostavkom. Dakle, mora biti $i < I$. □

Lema 4.47. *Neka s ima faktorizaciju $s = a_0\widetilde{u}11\underline{2}2v$ pri čemu je u neprazna i $[u] \leq [v]$ (odnosno $s = a_0\widetilde{u}2\underline{2}11v$ pri čemu je u neprazna i $[v] \leq [u]$) te je i pozicija podcrtanog slova. Tada je $i < I$.*

Dokaz. Promatrajmo slučaj $s = a_0\widetilde{u}11\underline{2}2v$ za nepraznu u i $[u] \leq [v]$. Tada je $\lambda_i(s) = 2 + [011u] + [02v]$. Budući je u konačna riječ, a v beskonačna, prema Korolaru 4.35 je $u \leq_{alt} v$ i stoga $02u \leq_{alt} 02v$. Iz istog korolara je $[02u] \leq [02v]$ te zato $\lambda_i(s) \geq 2 + [011u] + [02u] = 3$ po Lemi 4.45. Dakle, $i < I$.

Slučaj $s = a_0\widetilde{u}2\underline{2}11v$ pokazuje se vrlo slično koristeći $[02v] \leq [02u]$. □

Iz ove leme izravno dobivamo sljedeći korisni rezultat.

Korolar 4.48. *Ako s ima faktor $111\underline{2}22$ ili $22\underline{2}111$, gdje je i položaj podcrtanog slova, onda je $i < I$.*

Blok unutar s je svaki konačan faktor w takav da je w potencija istog slova, označimo ga sa x, te je $s = uwv$, pri čemu u ne završava sa x i v ne počinje sa x.

Lema 4.49. *Postoji faktorizacija $s = a_0vt$ takva da je $|v| \in \{I-2, I-1\}$ i riječ t nema blok neparne duljine.*

Dokaz. Najprije dokazujemo svojstvo $(*)$: s nema faktorizaciju $s = a_0\widetilde{u}21^m2v$, gdje je m neparan i $|u| \geq I-3$. Pretpostavimo, naprotiv, da s ima takvu faktorizaciju. Neka je m minimalan. Ako je $m = 1$, imamo $s = a_0\widetilde{u}21\underline{2}v$ što je u kontradikciji s Lemom 4.46.iii jer je indeks podcrtanog slova barem I. Stoga je $m \geq 3$.

Ne može biti $s = a_0\widetilde{u}21^m\underline{2}1\dots$ po Lemi 4.46.ii jer bi postojao faktor $1\underline{2}1$ s indeksom podcrtanog slova barem I. Ne može biti ni $s = a_0\widetilde{u}21^m\underline{2}22\dots$ jer bi to bilo u kontradikciji s Korolarom 4.48 budući da je naznačeni indeks veći ili jednak I. Stoga je $s = a_0\widetilde{u}21^m221^p\dots$, gdje je $p \geq 1$ uzet maksimalan, pa može biti beskonačan, tj. ili je $s = a_0\widetilde{u}21^m221^p2\dots$ ili je $s = a_0\widetilde{u}21^m221^\infty$. Imamo $s = a_0\widetilde{u}21^{m-2}11\underline{2}21^p\dots$. Kako je podcrtano slovo indeksa barem I, po Lemi 4.47 je $[1^p\dots] < [1^{m-2}2u]$. Budući da je $m-2$ neparan, mora biti $p \leq m-2$ jer bi u protivnom bilo $1^p\dots = 1^{m-2}1\dots >_{alt} 1^{m-2}2u$ i zato $[1^p\dots] > [1^{m-2}2u]$ po Korolaru 4.35. To znači da je p konačan i $s = a_0\widetilde{u}21^m221^p2\dots$. Ako je p paran, onda je $p < m-2$ i $1^p\dots = 1^p2\dots >_{alt} 1^p11^{m-3-p}2u = 1^{m-2}2u$ što nam opet po istom korolaru daje kontradikciju s dobivenim rezultatom. Zato je p neparan, a to je u suprotnosti s minimalnošću od m. Time je $(*)$ dokazano.

Sada dokazujemo svojstvo $(**)$: s nema faktorizaciju $s = a_0\widetilde{u}12^m1\ldots$, gdje je m neparan i $|u| \geq I-2$. Dokaz je sličan prethodnom, pa ćemo ga zapisati nešto sažetije. Pretpostavimo da s ima navedenu faktorizaciju. Neka je m minimalan. Po Lemi 4.46.ii ne može biti $m = 1$ jer je inače $s = a_0\widetilde{u}1\underline{2}1\ldots$. Zato je $m \geq 3$. Po Lemi 4.46.iii ne može biti $s = a_0\widetilde{u}12^m1\underline{2}\ldots$, a po Korolaru 4.48 ne može biti $s = a_0\widetilde{u}12^{m-1}\underline{2}111\ldots$. Stoga je $s = a_0\widetilde{u}12^m112^p\ldots$, gdje je $p \geq 1$ maksimalan i možda beskonačan. Tada je $s = a_0\widetilde{u}12^{m-2}2\underline{2}112^p\ldots$, pa po Lemi 4.47 mora biti $[2^{m-2}1u] < [2^p\ldots]$ jer je podcrtani položaj indeksa većeg ili jednakog I. Ako je $p > m-2$, dobivamo kontradikciju s Korolarom 4.35 koristeći definiciju $<_{alt}$ i činjenicu da je $m-2$ neparan. Zato je $p \leq m-2$. Ako je p paran, onda je $p < m-2$ što opet vodi na kontradikciju jer unutar s iza 2^p slijedi 1 budući da je p maksimalan. Slijedi da je p neparan što je u suprotnosti s minimalnošću od m. Tako je i $(**)$ dokazano.

Iz $(*)$ i $(**)$ zaključujemo da uz faktorizaciju $s = a_0vt$, $|v| = I-2$, beskonačna riječ t nema blokova neparne duljine, osim možebitno prvog bloka u t. U tom slučaju v produljimo za jedno slovo čime je lema dokazana. $\square$

Lema 4.50. *Postoje cijeli brojevi $L \geq I$ i M takvi da za svaku faktorizaciju $s = a_0\widetilde{u}1122v$ (odnosno $s = a_0\widetilde{u}2211v$) takvu da je $|u| \geq L$, imamo $u = w2u'$, $v = w1v'$ (odnosno $u = w1u'$, $v = w2v'$) za neku riječ w duljine najviše M. Brojeve L i M možemo omeđiti funkcijama koje ovise samo o ζ i I.*

Dokaz. Tvrdimo da ako je $s = a_0\widetilde{(wu')}11\underline{2}2wv'$ ili $s = a_0\widetilde{(wu')}2\underline{2}11wv'$, uz $|w| = n+1$, onda vrijedi $\lambda_i(s) > 3 - 2^{-n+1}$. Zaista, neka je u prvom slučaju $a = [wv']$, $b = [wu']$ te u drugom slučaju $b = [wv']$, $a = [wu']$. Tada je u oba slučaja $\lambda_i(s) = 2 + [02a] + [011b]$, pa imamo

$$3-\lambda_i(s) = 1 - \frac{1}{2+\frac{1}{a}} - \frac{1}{1+\frac{1}{1+\frac{1}{b}}} = 1 - \frac{a}{2a+1} - \frac{b+1}{2b+1} = \frac{b-a}{(2a+1)(2b+1)} < b-a < 2^{-n+1},$$

pri čemu smo u zadnjem koraku koristili Teorem 1.8.b.

Znamo da je $\lambda_i(s) \leq \zeta < 3$ za $i \geq I$, pa prethodna tvrdnja povlači da za svaku faktorizaciju $s = a_0\widetilde{(wu')}11\underline{2}2wv'$, gdje je $i \geq I$ indeks podcrtane 2 te $|w| = n+1$, mora vrijediti $3 - 2^{-n+1} < \lambda_i(s) \leq \zeta$. Zato je $2^{n-1} < \frac{1}{3-\zeta}$, pa je $n+1$ omeđen nekom konstantom M koja ovisi samo o ζ.

Prema Lemi 4.49, možemo pisati $s = ms_1$, gdje je $|m| \in \{I, I-1\}$ i s_1 ima blokove samo parne duljine.

Neka je sada $s = a_0\widetilde{u}1122v$ i $|u| \geq I + M$. Iz upravo pokazanog svojstva je $u = wyu'$, $v = wxv'$, gdje je w duljine najviše M i $\{x, y\} = \{1, 2\}$. Nadalje, po Lemi 4.47 je $[u] > [v]$. Želimo pokazati da je $x = 1$, $y = 2$. Pretpostavimo suprotno, tj. da je $u = w1u'$, $v = w2v'$. Kako je $[w1u'] > [w2v']$, w mora biti neparne duljine. Imamo $s = a_0\widetilde{u}'1\widetilde{w}1122w2v'$. Primijetimo da je $|a_0\widetilde{u}'| \geq I \geq |m|$ jer iz $|u| \geq I+M$, $u = wyu'$ i $|wy| \leq M+1$ slijedi $|u'| \geq I-1$. Stoga je $1\widetilde{w}1122w2$ faktor od s_1, pa budući da w kao riječ neparne duljine ima barem jedan blok neparne duljine, lako vidimo da i s_1 mora imati blok neparne duljine što nije moguće. Ova kontradikcija pokazuje da je $x = 1$, $y = 2$, pa možemo uzeti $L = I + M$.

Slučaj $s = a_0\widetilde{u}2211v$ se razmatra analogno. $\square$

Prisjetimo se da je χ supstitucija koja a zamjenjuje sa 11 te b zamjenjuje sa 22.

Teorem 4.51. *Neka je dana beskonačna riječ $s = a_0a_1a_2\ldots$ na $\mathbb{Z}$, $a_i > 0$ za $i > 0$, kojoj je Lagrangeov broj strogo manji od 3. Tada postoji faktorizacija $s = a_0ws'$ i beskonačna riječ t na $\{a, b\}$ koja zadovoljava Markovljevo svojstvo uz parametre K i N, tako da je $s' = \chi(t)$. Ako pretpostavimo da je $\lambda_i(s) \le \zeta < 3$ za svaki $i \ge I$, onda K, N, $|w|$ možemo omeđiti funkcijama koje ovise samo o ζ i I.*

Dokaz. Ako beskonačna riječ na $\{1, 2\}$ ima blokove samo parne duljine, onda je ta riječ u slici od χ. Neka je L kao u Lemi 4.50, posebno, $L \ge I$. Zapišimo $s = a_0\widetilde{w}s'$, gdje je odabrano $|a_0\widetilde{w}| \in \{L, L+1\}$ tako da, prema Lemi 4.46.i i Lemi 4.49, s' ima samo slova 1 i 2 te blokove samo parne duljine. Tada je $s' = \chi(t)$ za neku beskonačnu riječ t na $\{a, b\}$. Primijenimo Lemu 4.50 i neka je M kao u toj lemi.

Pokazat ćemo da t zadovoljava Markovljevo svojstvo. Neka je $t = \widetilde{u}xyv$ sa $\{x, y\} = \{a, b\}$ i $|u| \ge K = \lceil (M + 1)/2\rceil$. Budući je $K \ge (M + 1)/2$, vrijedi $2|u| \ge M + 1$. Promotrimo slučaj $x = a$, $y = b$ jer je drugi slučaj sličan. Tada je $s = a_0\widetilde{w}s' = a_0\widetilde{w}\chi(t) = a_0\widetilde{w}\chi(\widetilde{u})1122\chi(v)$. Kako je $\chi(\widetilde{u}) = \widetilde{\chi(u)}$, imamo $s = a_0(\widetilde{\chi(u)w})1122\chi(v)$. Zbog $|w| \ge L-1$ i $|u| \ge 1$ je duljina od $\chi(u)w$ barem L, pa Lema 4.50 povlači $\chi(u)w = p2u'$ i $\chi(v) = p1v'$ za neku konačnu riječ p duljine najviše M. Duljina $2|u|$ od $\chi(u)$ je veća od M, pa p mora biti parne duljine jer bi inače i $p2$ i $p1$ bili u slici od χ što bi značilo da p ima zadnje slovo i 2 i 1, a to nije moguće. Zato je p u slici u χ, tj. $p = \chi(m)$ za neku riječ m duljine najviše $M/2$. Budući da je $|\chi(u)| \ge M + 1 \ge |p2|$ te $\chi(u)w = p2u'$, dobivamo $\chi(u) = p2u'_1 = \chi(m)22\chi(u'')$. Također, $\chi(v) = \chi(m)11\chi(v'')$. Ovo povlači da je $u = mbu''$, $v = mav''$, gdje je m duljine naviše $M/2$. Uzmemo li $N = \lceil M/2\rceil$, završili smo dokaz. □

4.6.2 Obostrano beskonačni nizovi

Teorem 4.52. *Ako je Markovljev supremum obostrano beskonačnog niza $s = (a_n)_{n\in\mathbb{Z}}$ strogo manji od 3, onda je s do na pomak jednak $\chi(t)$ za neki obostrano beskonačni niz $t \in \{a, b\}^{\mathbb{Z}}$ koji zadovoljava Markovljevo svojstvo.*

Dokaz. Imamo $M(s) = \sup_{i\in\mathbb{Z}} \lambda_i(s) < 3$. Uzmimo neki realni broj ζ takav da je $M(s) < \zeta < 3$ i neki prirodan broj I takav da je $M(s) + 2^{-I+2} \le \zeta$.

Fiksirajmo neki cijeli broj k i promatrajmo beskonačnu riječ $t_k = (b_n)_{n\in\mathbb{N}_0}$ definiranu sa $b_n = a_{k+n}$. Pokažimo da za $i \ge I$ imamo $\lambda_i(t_k) \le \zeta$. Zaista,

$$\lambda_i(t_k) = [b_i, \ldots, b_1] + [b_{i+1}, b_{i+2}, \ldots]^{-1} = [a_{k+i}, \ldots, a_{k+1}] + [a_{k+i+1}, a_{k+i+2}, \ldots]^{-1}$$
$$\lambda_{k+i}(s) = [a_{k+i}, \ldots, a_{k+1}, a_k, a_{k-1}, \ldots] + [a_{k+i+1}, a_{k+i+2}, \ldots]^{-1}.$$

Iz definicije je $\lambda_{k+i}(s) \le M(s)$, a razvoji u verižne razlomke realnih brojeva $[a_{k+i}, \ldots, a_{k+1}]$ i $[a_{k+i}, \ldots, a_{k+1}, a_k, a_{k-1}, \ldots]$ se podudaraju u prvih $i \ge I$ parcijalnih kvocijenata, pa je po Teoremu 1.8.b razlika ta dva broja manja od 2^{-I+2}. Stoga je $\lambda_i(t_k) \le M(s) + 2^{-I+2} \le \zeta$.

Iz Teorema 4.51 slijedi da je $t_k = w_k\chi(u_k)$ za neku riječ w_k i neku beskonačnu riječ u_k koja zadovoljava Markovljevo svojstvo uz parametre K i N. Štoviše, duljine od w_k te brojevi K i N omeđeni su konstantama neovisnim o k. Budući da to vrijedi za svaki k, sada nije teško pokazati da je $s = \chi(t)$ za neki obostrano beskonačni niz t koji zadovoljava Markovljevo svojstvo uz parametar N. □

Ovaj teorem ima i obrat koji nećemo dokazivati.

Teorem 4.53. *Ako je obostrano beskonačni niz s jednak $\chi(t)$ za neki obostrano beskonačni niz $t \in \{a, b\}^{\mathbb{Z}}$ koji zadovoljava Markovljevo svojstvo, onda je $M(s) < 3$.*

4.7 Markovljev teorem

U ovom potpoglavlju najprije dokazujemo Markovljev teorem za aproksimacije: ako je x iracionalan realan broj kojemu je Lagrangeov broj $L(x) < 3$, onda je verižni razlomak od x (od nekog mjesta) periodski i periodski uzorak mu je Christoffelova riječ na alfabetu $\{11, 22\}$. Zatim pokazujemo da se supremum iz Definicije 1.1 Lagrangeova broja postiže za takav x. Koristeći prije izloženu teoriju indefininitnih binarnih kvadratnih formi, dokazujemo i Markovljev teorem za kvadratne forme.

4.7.1 Markovljev teorem za aproksimacije

Lema 4.54. *Neka je dana beskonačna riječ $s \in \mathbb{Z}^{\mathbb{N}_0}$ uz $a_i > 0$ za $i > 0$ za koju postoje $I \in \mathbb{N}$ i $\zeta < 3$ tako da vrijedi $\lambda_i(s) < \zeta$ za svaki $i \geq I$. Tada je $s = u\chi(w)^\infty$ za neku donju Christoffelovu riječ w i neku riječ u čija duljina je omeđena funkcijom koja ovisi samo o I i ζ.*

Dokaz. Po Teoremu 4.51 imamo $s = a_0 m s'$, $s' = \chi(t)$, gdje t zadovoljava Markovljevo svojstvo uz parametre K i N, a duljina od m i brojevi K i N su omeđeni nekim funkcijama koje ovise samo o I i ζ. Prema Teoremu 4.31 je $t = vw^\infty$, gdje je w donja Christoffelova riječ, a duljina od v je omeđena nekom funkcijom koja ovisi samo o K i N. Stoga je $s = a_0 m\chi(v)\chi(w)^\infty$ čime je lema dokazana. $\square$

Teorem 4.55 (Markovljev teorem za aproksimacije)**.** *Neka je x iracionalan realan broj. Tada je Lagrangeov broj $L(x)$ strogo manji od 3 ako i samo ako je x ekvivalentan broju $x_w = [\chi(\widetilde{w})^\infty]$ (ili, što je isto, broju $[\chi(w)^\infty]$) za neku donju Christoffelovu riječ w. U slučaju da to vrijedi, neka je m Markovljev broj $\mu(w)_{12} = \frac{1}{3}\operatorname{Tr}(\mu(w))$ pridružen w. Tada je*

$$L(x) = L(x_w) = \sqrt{9 - \frac{4}{m^2}}, \quad x_w = \frac{p-s}{2m} + \frac{1}{2}\sqrt{9 - \frac{4}{m^2}},$$

gdje je $\mu(w) = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right]$ i zato $m = q = \frac{1}{3}(p+s)$.

Dokaz. Budući da su w i $\widetilde{w}$ konjugati prema Teoremu 4.40, i njihove slike po χ su konjugati, pa su $[\chi(\widetilde{w})^\infty]$ i $[\chi(w)^\infty]$ ekvivalentni realni brojeva. Iz Propozicije 1.13 slijedi da imaju isti Lagrangeov broj.

Ako je x ekvivalentan x_w, iz istog rezultata slijedi da je $L(x) = L(x_w)$, a po Teoremu 4.44 je $L(x_w) = \sqrt{9 - \frac{4}{m^2}} < 3$, gdje je m Markovljev broj kao u izreci teorema. Imamo $\mu(w)_{12} = \frac{1}{3}\operatorname{Tr}(\mu(w))$ po Lemi 4.17. Nadalje, po Teoremu 4.37 (gdje y odgovara x_w) je $x_w = \frac{p-s}{2m} + \frac{1}{2}\sqrt{9 - \frac{4}{m^2}}$.

Obratno, neka je x takav da je $L(x) < 3$. Neka je $s = a_0 a_1 a_2 \ldots$ beskonačna riječ takva da x ima razvoj u verižni razlomak $[a_0, a_1, a_2, \ldots]$. Tada je $L(s) = L(x) < 3$, pa postoje I i ζ kao u Lemi 4.54. Zaključujemo da je $s = u\chi(w)^\infty$ za neku donju Christoffelovu riječ w, a jer su w i $\widetilde{w}$ konjugati, također imamo $s = v\chi(\widetilde{w})^\infty$ te je zato x ekvivalentan x_w. $\square$

Markovljev teorem se često iskazuje kao niz sve boljih aproksimacija sa iznimkama pri čemu je prva iznimka skup brojeva ekvivalentnih zlatnom rezu. Takav oblik teorema dan je u sljedećem rezultatu.

Korolar 4.56. *Neka je M konačan skup Markovljevih brojeva takav da za svaka dva Markovljeva broja n i p, iz $n < p$ i $p \in M$, slijedi $n \in M$. Neka je W konačan skup svih donjih Christoffelovih riječi kojima su pridruženi Markovljevi brojevi iz M, tj. W je skup svih donjih Christoffelovih riječi w takvih da je $\mu(w)_{12} \in M$. Neka je m najmanji Markovljev broj koji nije u M. Tada za svaki iracionalan broj x koji nije ekvivalentan nijednom x_w, $w \in W$, postoji beskonačno mnogo racionalnih brojeva $\frac{p}{q}$ takvih da je*

$$\left|x - \frac{p}{q}\right| < \frac{1}{\sqrt{9 - \frac{4}{m^2}}\, q^2}.$$

Prije dokaza korolara, dat ćemo nekoliko primjera. Za $M = \emptyset$ imamo $W = \emptyset$ i $m = 1$, pa je $\sqrt{9 - \frac{4}{m^2}} = \sqrt{5}$. Dobivamo da svaki iracionalan realan broj x ima beskonačno mnogo racionalnih aproksimacija za koje vrijedi $\left|x - \frac{p}{q}\right| < \frac{1}{\sqrt{5}q^2}$. To je upravo Hurwitzov rezultat koji smo dokazali u Teoremu 1.9.f.

Za $M = \{1\}$ je $W = \{a\}$ i $m = 2$, pa je $\sqrt{9 - \frac{4}{m^2}} = \sqrt{8}$. Nadalje, $x_a = \frac{1+\sqrt{5}}{2}$, tj. zlatni rez. Dobivamo da svaki iracionalan broj x koji nije ekvivalentan x_a, tj. nisu mu od nekog mjesta nadalje samo 1 parcijalni kvocijenti u verižnom razlomku, ima beskonačno mnogo racionalnih aproksimacija za koje vrijedi $\left|x - \frac{p}{q}\right| < \frac{1}{\sqrt{8}q^2}$.

Konačno, za $M = \{1, 2\}$ je $W = \{a, b\}$, $m = 5$, $\sqrt{9 - \frac{4}{m^2}} = \frac{\sqrt{221}}{5}$ te $x_b = 1 + \sqrt{2}$. Dobivamo da svaki iracionalan broj x koji nije ekvivalentan ni x_a ni x_b, tj. nema u razvoju u verižni razlomak od nekog mjesta nadalje samo 1 ili samo 2, nužno ima beskonačno mnogo racionalnih aproksimacija za koje vrijedi $\left|x - \frac{p}{q}\right| < \frac{5}{\sqrt{221}q^2}$.

Za dokaz korolara trebat ćemo nekoliko pomoćnih tvrdnji.

Lema 4.57. *Neka su $\frac{p_n}{q_n}$ konvergente realnog broja x. Ako je $\lambda_{n+1}(x) > \gamma$, onda je $\left|x - \frac{p_n}{q_n}\right| < \frac{1}{\gamma q_n^2}$. Ako je $\lambda_{n+1}(x) < \gamma$, onda je $\left|x - \frac{p_n}{q_n}\right| > \frac{1}{\gamma q_n^2}$.*

Dokaz. Rezultat slijedi direktno iz dokaza Propozicije 1.12, točnije, jednakosti (1.16). ☐

Korolar 4.58. *Ako je $\gamma > 0$ i $L(x) > \gamma$, onda x ima beskonačno mnogo racionalnih aproksimacija $\frac{p}{q}$ takvih da je $\left|x - \frac{p}{q}\right| < \frac{1}{\gamma q^2}$.*

Dokaz. Budući da je $L(x) = \limsup_n \lambda_n(x)$, postoji beskonačno mnogo $n \in \mathbb{N}$ takvih da je $\lambda_{n+1}(x) > \gamma$ te primijenimo prethodnu lemu. ☐

Lema 4.59. *Neka je x iracionalan broj takav da je $L(x) < 3$ (tako da je x ekvivalentan x_v za neku donju Christoffelovu riječ v te je $L(x) = \sqrt{9 - \frac{4}{m^2}}$ za Markovljev broj m pridružen riječi v). Tada postoji beskonačno mnogo racionalnih aproksimacija $\frac{p}{q}$ od x takvih da je $\left|x - \frac{p}{q}\right| < \frac{1}{L(x)q^2}$.*

Dokaz. Ovaj će rezultat slijediti direktno iz Teorema 4.62 koji ćemo uskoro dokazati. ☐

Dokaz Korolara 4.56. Ako je riječ w duljine n, onda je svaki element u $\mu(w)$ veći ili jednak 2^{n-1}. Zaista, svi elementi matrica $\mu(a)$ i $\mu(b)$ su prirodni brojevi, a budući je $\mu(w)$ umnožak n takvih matrica i $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]^n = \left[\begin{smallmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{smallmatrix}\right]$, vrijedi navedena tvrdnja. Dakle, ako je $\mu(w)_{12} \in M$ onda je duljina od w omeđena. Stoga je W konačan skup.

Neka je x iracionalan broj koji nije ekvivalentan nijednom x_w, $w \in W$. Stavimo $\gamma = \sqrt{9 - \frac{4}{m^2}}$, pa je $\gamma < 3$. Ako je $L(x) \geq 3$, onda je $L(x) > \gamma$ te po Korolaru 4.58 znamo da postoji beskonačno mnogo racionalnih aproksimacija $\frac{p}{q}$ od x za koje je $\left|x - \frac{p}{q}\right| < \frac{1}{\gamma q^2}$. Ako je $L(x) < 3$, onda je po Teoremu 4.55 x ekvivalentan x_v za neku donju Christoffelovu riječ v kojoj je pridružen Markovljev broj n i vrijedi $L(x) = \sqrt{9 - \frac{4}{n^2}}$. Po pretpostavci $v \notin W$, pa $n \notin M$ i zato je $n \geq m$. Ako je $n > m$, onda je $L(x) > \gamma$ i po Korolaru 4.58 postoji beskonačno mnogo racionalnih aproksimacija kao gore. Ako je $n = m$, koristimo Lemu 4.59. □

Pokažimo primjerom da zaključak Leme 4.59 općenito ne mora vrijediti ako ne zahtijevamo $L(x) < 3$.

Primjer. Slijedeći Perrona, uzmimo prirodne brojeve $b > c$ po volji i definirajmo $x = [\overline{b, c}]$ i $y = [\overline{c, b}]$. Očito su x i y ekvivalentni brojevi te je $L(x) = L(y) = b + 2[0, \overline{c, b}]$. Znamo da su konvergente s parnim (odnosno neparnim) indeksom manje (odnosno veće) od vrijednosti verižnog razlomka kojoj konvergiraju. Za $n \in \mathbb{N}$ paran je

$$\begin{aligned} \lambda_n(x) &= b + [0, c, b, c, b, \dots, c] + [0, \overline{c, b}] > b + [0, \overline{c, b}] + [0, \overline{c, b}] = L(x), \\ \lambda_n(y) &= c + [0, b, c, b, c, \dots, b] + [0, \overline{b, c}] < L(y), \end{aligned}$$

dok za n neparan vrijedi

$$\begin{aligned} \lambda_n(x) &= c + [0, b, c, b, c, \dots, b, c] + [0, \overline{b, c}] < L(x), \\ \lambda_n(y) &= b + [0, c, b, c, b, \dots, c, b] + [0, \overline{c, b}] < b + [0, \overline{c, b}] + [0, \overline{c, b}] = L(y). \end{aligned}$$

Stoga iz Leme 4.57 zaključujemo da x ima beskonačno mnogo racionalnih aproksimacija za koje je $\left|x - \frac{p}{q}\right| < \frac{1}{L(x)q^2}$, a y ima samo konačno mnogo takvih aproksimacija.

4.7.2 Dobre i loše aproksimacije

U ovom odjeljku promatramo proizvoljan realan broj x za koji je $L(x) < 3$, pa je prema Teoremu 4.55 x ekvivalentan nekom broju x_w za donju Christoffelovu riječ w i vrijedi $L(x) = L(x_w) = \sqrt{9 - \frac{4}{m^2}}$, gdje je m Markovljev broj pridružen w. Podsjetimo se da je razvoj u verižni razlomak $x_w = [\chi(\widetilde{w})^\infty]$. Nadalje, x ima razvoj oblika $[a_0 s]$, gdje je $s = (a_1, a_2, \dots)$ periodski niz s periodskim uzorkom $p = \chi(w)$ i također $\widetilde{p} = \chi(\widetilde{w})$.

Prema Teoremu 4.40 su donja Christoffelova riječ w i njoj obratna riječ $\widetilde{w}$ koja je gornja Christoffelova riječ istog nagiba, uvijek konjugati, tj. $w = hk$, $\widetilde{w} = kh$ za neke riječi h, k. Stoga su $\chi(w) = \chi(h)\chi(k)$ i $\chi(\widetilde{w}) = \chi(k)\chi(h)$ također konjugati i konjugirajuće riječi $\chi(h)$, $\chi(k)$ su parne duljine jer χ preslikava slova a, b na riječi 11, 22 koje su parne duljine.

Lema 4.60. *Neka je s čisto periodski niz koji ima periodski uzorak $p = \chi(w)$ za neku donju Christoffelovu riječ w. Tada je ili $s = u\widetilde{p}^\infty$, gdje je u sufiks neparne duljine od $\widetilde{p}$, ili je $s = up^\infty$, gdje je u sufiks parne duljine od p.*

Dokaz. Budući da su p i $\widetilde{p}$ konjugati, s također ima periodski uzorak $\widetilde{p}$. Stoga možemo pisati $s = u_1 p^\infty = u_2 \widetilde{p}^\infty$, gdje je u_1 sufiks od p i u_2 sufiks od $\widetilde{p}$. Po primjedbama prije ove leme, zaključujemo da su u_1 i u_2 iste parnosti te dobivamo traženu tvrdnju. □

Lema 4.61. *Neka je s periodski, ali ne čisto periodski niz koji ima periodski uzorak $p = \chi(w)$ za neku donju Christoffelovu riječ w. Tada mora vrijediti jedan od sljedećih slučajeva, gdje su $u, u', v \in \mathbb{N}^*$ i $a, b \in \mathbb{N}$.*

(i) $s = ubv\widetilde{p}^\infty$, $\widetilde{p} = u'av$, *gdje je ili* v *parne duljine i* $b < a$ *ili je* v *neparne duljine i* $b > a$*;*

(ii) $s = ubvp^\infty$, $p = u'av$, *gdje je ili* v *parne duljine i* $b > a$ *ili je* v *neparne duljine i* $b < a$*.*

Dokaz. Neka je niz t najveći periodski sufiks od s. Tada je t pravi sufiks od s i t ima periodske uzorke p i $\widetilde{p}$. Neka je a slovo takvo da je at čisto periodski i b slovo koje prethodi t u s. Nužno je $b \neq a$. Nadalje, $at = av_1p^\infty = av_2\widetilde{p}^\infty$, gdje je av_1 sufiks od p i av_2 je sufiks od $\widetilde{p}$. Prema primjedbi prije prethodne leme, duljine od v_1 i v_2 su iste parnosti. Imamo $p = u_1'av_1$, $\widetilde{p} = u_2'av_2$ i $s = ubt = ubv_1p^\infty = ubv_2\widetilde{p}^\infty$.

Pretpostavimo da je $b < a$. Ako je $|v_i|$ paran, onda vrijedi slučaj (i), prvi dio. Inače vrijedi slučaj (ii), drugi dio.

Pretpostavimo sada da je $a < b$. Ako je $|v_i|$ paran, onda vrijedi slučaj (ii), prvi dio. Inače vrijedi slučaj (i), drugi dio. □

U skladu s Lemama 4.60 i 4.61, možemo klasificirati brojeve $x = [a_0s]$ u dvije nedisjunktne klase. Kažemo da je x broj *tipa I* ako vrijedi jedan od sljedećih uvjeta:

1. $s = up^\infty$, u je sufiks parne duljine od p;
2. $s = ubvp^\infty$, $p = u'av$, riječ v je parne duljine, a, b su slova takva da je $b > a$;
3. $s = ubvp^\infty$, $p = u'av$, riječ v je neparne duljine, a, b su slova takva da je $b < a$.

Za x tipa I kažemo da je n *dobar* ako je $n + 1$ indeks posljednjeg slova nekog p u nizu $a_0s = \dots p^\infty$ (indeks od a_0 je 0). Kažemo da je n *loš* ako je $n + 1$ indeks bilo kojeg drugog slova u nekom p^∞. Dakle, ako s nije čisto periodski, zanemarujemo konačno mnogo indeksa slova na početku niza.

Kažemo da je x broj *tipa II* ako vrijedi jedan od sljedećih uvjeta

1. $s = u\widetilde{p}^\infty$, u je sufiks neparne duljine od $\widetilde{p}$;
2. $s = ubv\widetilde{p}^\infty$, $\widetilde{p} = u'av$, riječ v je parne duljine, a, b su slova takva da je $b < a$;
3. $s = ubv\widetilde{p}^\infty$, $\widetilde{p} = u'av$, riječ v je neparne duljine, a, b su slova takva da je $b > a$;

Za x tipa II kažemo da je n *dobar* ako je $n + 1$ indeks prvog slova nekog $\widetilde{p}$ u nizu $a_0s = \dots \widetilde{p}^\infty$. Kažemo da je n *loš* ako je $n+1$ indeks bilo kojeg drugog slova u nekom $\widetilde{p}^\infty$.

Teorem 4.62 (Bombieri)**.** *Neka je x iracionalan broj za koji je $L(x) < 3$ i neka je $[a_0s]$ razvoj od x u verižni razlomak kojemu su $\frac{p_n}{q_n}$ konvergente. Osim za konačno $n \in \mathbb{N}$ vrijedi: ako je n dobar, onda je $\left|x - \frac{p_n}{q_n}\right| < \frac{1}{L(x)q_n^2}$; ako je n loš, onda je $\left|x - \frac{p_n}{q_n}\right| > \frac{1}{L(x)q_n^2}$.*

Lema 4.59 slijedi iz ovog teorema jer je zbog periodičnosti broj dobrih n beskonačan.

Primijetimo da je svaki x kao u teoremu nužno tipa I ili tipa II zbog prethodnih dviju lema, ali može biti i jednog i drugog tipa, kao primjerice zlatni rez.

U sljedećoj lemi uspoređujemo vrijednosti brojeva $\lambda_n(s)$ i $\lambda_n(t)$, gdje je t obostrano beskonačni niz, a s je (jednostrano beskonačni) niz koji se podudara sa t u pozitivnim indeksima, osim eventualno na konačno mnogo mjesta.

Lema 4.63. *Neka je $t = (b_n)_{n\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$, $n_0 \geq 0$ te $s = (a_n)_{n\in\mathbb{N}_0} \in \mathbb{Z}^{\mathbb{N}_0}$ uz $a_n > 0$ za $n > 0$. Pretpostavimo da za sve $n > n_0$ vrijedi $a_n = b_n$. Neka je $n > n_0$.*

Ako je $n_0 = 0$ i n neparan; ili je $n_0 \geq 1$, $n - n_0$ neparan i $a_{n_0} > b_{n_0}$; ili je $n_0 \geq 1$, $n - n_0$ paran i $a_{n_0} < b_{n_0}$; onda je $\lambda_n(s) < \lambda_n(t)$.

Ako je $n_0 = 0$ i n paran; ili je $n_0 \geq 1$, $n - n_0$ paran i $a_{n_0} > b_{n_0}$; ili je $n_0 \geq 1$, $n - n_0$ neparan i $a_{n_0} < b_{n_0}$; onda je $\lambda_n(s) > \lambda_n(t)$.

Dokaz. Pretpostavimo da je $n_0 = 0$ i n je neparan. Ako je $n = 1$, onda je

$$\lambda_1(s) = [a_1, a_2, \ldots] = [b_1, b_2, \ldots] < [b_1, b_2, \ldots] + [b_0, b_{-1}, \ldots]^{-1} = \lambda_1(t).$$

Ako je $n \geq 3$, onda je $[b_{n-1}, b_{n-2}, \ldots, b_1] > [b_{n-1}, b_{n-2}, \ldots, b_1, b_0, b_{-1}, \ldots]$ jer je riječ $b_{n-1}b_{n-2}\ldots b_1$ prefiks parne duljine beskonačne riječi $b_{n-1}b_{n-2}\ldots b_1 b_0 b_{-1}\ldots$ (usp. Korolar 4.35). Stoga je

$$\begin{aligned}\lambda_n(s) &= [a_{n-1}, a_{n-2}, \ldots, a_1]^{-1} + [a_n, a_{n+1}, \ldots] = [b_{n-1}, b_{n-2}, \ldots, b_1]^{-1} + [b_n, b_{n+1}, \ldots] \\ &< [b_{n-1}, b_{n-2}, \ldots, b_1, b_0, b_{-1}, \ldots]^{-1} + [b_n, b_{n+1}, \ldots] = \lambda_n(t).\end{aligned}$$

Pretpostavimo sada da je $n_0 \geq 1$, $n - n_0$ neparan i $a_{n_0} > b_{n_0}$. Tada je

$$\begin{aligned}[a_{n-1}, a_{n-2}, \ldots, a_1] &= [a_{n-1}, \ldots, a_{n_0+1}, a_{n_0}, a_{n_0-1}, \ldots, a_1] \\ &= [b_{n-1}, \ldots, b_{n_0+1}, a_{n_0}, a_{n_0-1}, \ldots, a_1] \\ &> [b_{n-1}, \ldots, b_{n_0+1}, b_{n_0}, b_{n_0-1}, \ldots, b_1, b_0, b_{-1}, \ldots]\end{aligned}$$

jer je $a_{n_0} > b_{n_0}$ i riječ $b_{n-1}\ldots b_{n_0+1}$ je parne duljine. Zato je

$$\begin{aligned}\lambda_n(s) &= [a_{n-1}, a_{n-2}, \ldots, a_1]^{-1} + [a_n, a_{n+1}, \ldots] = [a_{n-1}, a_{n-2}, \ldots, a_1]^{-1} + [b_n, b_{n+1}, \ldots] \\ &< [b_{n-1}, b_{n-2}, \ldots]^{-1} + [b_n, b_{n+1}, \ldots] = \lambda_n(t).\end{aligned}$$

Svi drugi slučajevi su slični, pa ih preskačemo. □

Dokaz Teorema 4.62. Koristimo oznake koje smo prije uveli u ovom odjeljku. Znamo da je $a_0 s = (a_n)_{n\in\mathbb{N}_0}$ periodski niz s periodskim uzorkom p i periodskim uzorkom $\widetilde{p}$. Neka je $t = (b_n)_{n\in\mathbb{Z}}$ (čisto) periodski obostrano beskonačni niz koji se podudara sa $a_0 s$ za dovoljno velike indekse n. Tada t ima iste periodske uzorke, odnosno jednak je ${}^\infty p^\infty$ i ${}^\infty \widetilde{p}^\infty$ do na pomake. Ako je i indeks prvog slova u nekom $\widetilde{p}$ unutar t ili indeks zadnjeg slova u nekom p unutar t, onda smo u dokazu Teorema 4.39 pokazali da je $\lambda_i(t) = L(t)$. Također znamo da je $L(t) = L(x) = \sqrt{9 - \frac{4}{m^2}}$.

Promatrat ćemo samo slučaj kada je x broj tipa II jer je drugi slučaj potpuno analogan.

Uzmimo najprije da je $s = u\widetilde{p}^\infty$, gdje je u sufiks neparne duljine od $\widetilde{p}$. Primijetimo da je $a_0 s = a_0 u \widetilde{p}$, pri čemu je duljina od $a_0 u$ paran broj.

Pretpostavimo da je n dobar, tj. $i = n + 1$ je indeks prvog slova nekog $\widetilde{p}$ u nizu $a_0 s$. Taj je indeks paran jer je u neparne duljine i $\widetilde{p}$ je parne duljine. Iz Leme 4.63, uzevši $n_0 = 0$, $a_0 s$ umjesto s i $n + 1$ umjesto n, slijedi $\lambda_{n+1}(a_0 s) > \lambda_{n+1}(t) = L(x)$ te prema Lemi 4.57 imamo $\left|x - \frac{p_n}{q_n}\right| < \frac{1}{L(x) q_n^2}$.

Pretpostavimo sada da je n loš. Postoje dva slučaja, ovisno o parnosti od n. Pretpostavimo da je $n + 1$ paran i veći od duljine riječi p. Tada $a_{n+1}a_{n+2}\ldots$ ima prefiks oblika $\chi(v)$ za neki konjugat $v \neq \widetilde{w}$ od w i $a_n a_{n-1} \ldots a_1$ ima prefiks $\chi(\widetilde{v})$. Po Teoremu 4.40 je $w <_{lex} \widetilde{v}$, $v <_{lex} \widetilde{w}$, pa budući je χ rastuća za leksikografski uređaj, vrijedi $p = \chi(w) <_{lex} \chi(\widetilde{v})$ i $\chi(v) <_{lex} \chi(\widetilde{w}) = \widetilde{p}$. Stoga je prema Lemi 4.36, $p <_{alt} \chi(\widetilde{v})$ i $\chi(v) <_{alt} \widetilde{p}$. Budući da ove riječi imaju istu duljinu, koja je paran broj, i zbog prethodno navedenog svojstva prefiksa, vidimo da je $p^\infty <_{alt} a_n a_{n-1} \dots a_1$ i $a_{n+1} a_{n+2} \dots <_{alt} \widetilde{p}^\infty$. Slijedi $\lambda_{n+1}(a_0 s) = [a_n, a_{n-1}, \dots, a_1]^{-1} + [a_{n+1}, a_{n+2}, \dots] < [p^\infty]^{-1} + [\widetilde{p}^\infty] = L(x)$. Zato po Lemi 4.57 dobivamo $\left|x - \frac{p_n}{q_n}\right| > \frac{1}{L(x) q_n^2}$. Ako je, s druge strane, $n+1$ neparan, onda iz Leme 4.63, uzevši $n_0 = 0$, $a_0 s$ umjesto s i $n+1$ umjesto n, dobivamo direktno $\lambda_{n+1}(a_0 s) < \lambda_{n+1}(t) \le L(x)$ i slično zaključimo.

Uzmimo sada da je $s = ubv\widetilde{p}^\infty$, gdje je $\widetilde{p} = u'av$, v je parne duljine, a, b su slova takva da je $b < a$.

Pretpostavimo najprije da je n dobar, tj. $i = n+1$ je indeks prvog slova nekog $\widetilde{p}$ u nizu $a_0 s$. Označimo li sa n_0 indeks od b unutar $a_0 s = a_0 ubv\widetilde{p}^\infty$, vidimo da je broj $n + 1 - n_0 = |v| + 1 + l|\widetilde{p}|$ neparan. Nadalje, $a_{n_0} < b_{n_0}$ zbog $b < a$. Stoga po Lemi 4.63 vrijedi $\lambda_{n+1}(a_0 s) > \lambda_{n+1}(t) = L(x)$ i završavamo primjenom Leme 4.57.

Pretpostavimo sada da je n loš. Ako je $n+1$ iste parnosti kao indeks prvog slova u nekog $\widetilde{p}$, onda argument nastavljamo kao prije uzevši n dovoljno velik. Ako je suprotne parnosti, onda iz leme dobivamo $\lambda_{n+1}(a_0 s) < \lambda_{n+1}(t) \le L(x)$ jer je $n + 1 - n_0$ paran i $a_{n_0} < b_{n_0}$.

Svi ostali slučajevi dokazuju se slično, pa ih preskačemo. □

Kao primjer promotrimo zlatni rez $x = \frac{1+\sqrt{5}}{2}$ kojemu je razvoj u verižni razlomak $[a_0 s] = [1^\infty]$, pa je $a_0 = 1$ i $s = p^\infty = 1(\widetilde{p})^\infty$ uz $p = \widetilde{p} = 11 = \chi(a)$. Zato je x istovremeno tipa I i tipa II (slučajevi 1. u obje definicije, uz u prazan i u duljine 1, respektivno). Dobri n su neparni brojevi i isti su u oba slučaja. Budući da konvergente od x zadovoljavaju $\frac{p_n}{q_n} = \frac{F_{n+2}}{F_{n+1}}$, uz pomak indeksa dobivamo da za paran n vrijedi $\left|x - \frac{F_{n+1}}{F_n}\right| < \frac{1}{\sqrt{5} F_n^2}$ dok za neparan n imamo $\left|x - \frac{F_{n+1}}{F_n}\right| > \frac{1}{\sqrt{5} F_n^2}$. Primijetimo da je u prvom slučaju $x < \frac{F_{n+1}}{F_n}$, a u drugom $x > \frac{F_{n+1}}{F_n}$.

4.7.3 Markovljev teorem za kvadratne forme

Lema 4.64. *Neka je $s = (\ell_n)_{n\in\mathbb{Z}}$ obostrano beskonačan niz prirodnih brojeva kojemu je Markovljev supremum $M(s)$ strogo manji od 3. Tada je za neku donju Christoffelovu riječ w, niz s do na pomak jednak ${}^\infty\chi(w)^\infty$.*

Dokaz. Tvrdnja slijedi iz Teorema 4.52 i 4.33. □

Neka je w donja Christoffelova riječ i $\mu(w) = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right]$, gdje je $m = q = \frac{1}{3}(p+s)$ Markovljev broj pridružen w. Definirajmo sada pridružene *Markovljeve kvadratne forme* $f_w^\pm(x, y) = mx^2 \pm (p - s)xy - ry^2$.

Kao i prije, forme koje promatramo u nastavku imaju realne koeficijente i kažemo da je jedna forma višekratnik druge ako im je kvocijent realan broj različit od nule.

Teorem 4.65 (Markovljev teorem za kvadratne forme)**.** *Neka je $f(x, y)$ indefinitna binarna kvadratna forma. Pretpostavimo da infimum $m(f)$ od $|f|$ na $\mathbb{Z}^2 \setminus \{(0,0)\}$ i diskriminanta $d(f)$ zadovoljavaju nejednakost $\sqrt{d(f)} < 3m(f)$. Tada je f ekvivalentna višekratniku neke Markovljeve forme f_w^- ili f_w^+. Neka je m Markovljev broj pridružen donjoj Christoffelovoj riječi w. Imamo $d(f_w^\pm) = 9 - 4m^2$, $m(f_w^\pm) = f_w^\pm(1, 0) = m$, pa se navedeni infimum od $f_w^\pm$ i f postiže i*

$$\frac{\sqrt{d(f)}}{m(f)} = \frac{\sqrt{d(f_w^\pm)}}{m(f_w^\pm)} = \sqrt{9 - \frac{4}{m^2}}.$$

Dokaz. Iz pretpostavke je $m(f) \neq 0$, pa je posebno $f(x, y) \neq 0$ za cijele brojeve x, y koji nisu oba jednaki nula. Prema Teoremu 2.7 znamo da je f ekvivalentna nekoj reduciranoj formi koja je po Propoziciji 2.8 element u lancu $(f_n)_{n\in\mathbb{Z}}$ reduciranih formi. Možemo stoga uzeti da je $f = f_0$. Koristeći oznake iz tog potpoglavlja 2.3, definiramo $s = (\ell_n)_{n\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$, gdje je $f_0(1, 0) > 0$, a $\vartheta_0 = [0, \ell_0, \ell_1, \ell_2, \dots] \in (0, 1)$ i $\varphi_0 = -[\ell_{-1}, \ell_{-2}, \ell_{-3}, \dots] \in (-\infty, -1)$ su korijeni od $f_0(x, 1)$. Za svaki cijeli broj k, definiramo pomaknuti niz $s_k = (\ell_{n+k})_{n\in\mathbb{Z}}$. Iz Teorema 2.12 i pretpostavke teorema koji dokazujemo je $M(s) < 3$. Prema Lemi 4.64 slijedi da je $s_k = {}^{\infty}\chi(w)^{\infty}$ za neki $k \in \mathbb{Z}$, gdje je w donja Christoffelova riječ. Znamo da je f ekvivalentna f_k, a iz (2.14) su korijeni od $f_k(x, 1)$ jednaki $\vartheta_k = (-1)^k[\chi(w)^{\infty}]^{-1}$ i $\varphi_k = (-1)^{k+1}[\chi(\widetilde{w})^{\infty}]$. Po Teoremu 4.37 je $[\chi(w)^{\infty}]$ korijen od $rx^2 + (s - p)x - q$, dok je $[\chi(\widetilde{w})^{\infty}]$ korijen od $qx^2 + (s - p)x - r$. Stoga lako dobivamo da su ϑ_k i φ_k korijeni od $qx^2 + (-1)^k(p - s)x - r$ iz čega slijedi da je f_k proporcionalna formi $qx^2 - (p - s)xy - ry^2 = f_w^-(x, y)$ ili formi $qx^2 + (p - s)xy - ry^2 = f_w^+(x, y)$.

Imamo $d(f_w^\pm) = (p - s)^2 + 4qr = (p + s)^2 - 4ps + 4qr = 9m^2 - 4$ jer je $p + s = 3m$ i matrica $\mu(w)$ ima determinantu 1. Iz Teorema 4.39 je $M(s) = M(s_k) = \sqrt{9 - \frac{4}{m^2}}$ te Teorem 2.12 daje $\frac{\sqrt{d(f_w^\pm)}}{m(f_w^\pm)} = \sqrt{9 - \frac{4}{m^2}}$, pa zaključujemo

$$m(f_w^\pm) = \frac{\sqrt{9m^2 - 4}}{\sqrt{9 - \frac{4}{m^2}}} = m = f_w^\pm(1, 0),$$

čime je dokaz gotov. □

Napomena. Tvrdnju teorema možemo još precizirati. Lako se vidi da je uvijek $p > s$, pa je f ekvivalentna pozitivnom višekratniku forme $(-1)^k qx^2 + (p - s)xy + (-1)^{k+1} ry^2$ jer je samo takav višekratnik reduciran (usporedite (2.10)). Dakle, forme f_w^+ i $-f_w^-$ su reducirane te je f ekvivalentna pozitivnom višekratniku od f_w^+ ili negativnom višekratniku od f_w^-. Pokažimo da reducirane forme f_w^+ i $-f_w^-$ nisu ekvivalentne osim za neprave donje Christoffelove riječi $w \in \{a, b\}$. Naime, ako su ϑ i φ korijeni od $f_w^+(x, 1)$ kao u (2.6), onda su korijeni od $-f_w^-(x, 1)$ jednaki $-\vartheta$ i $-\varphi$. Te dvije forme su po Teoremu 2.9 ekvivalentne ako i samo ako su u istom lancu, a iz oblika od ϑ u dokazu Teorema 4.65 vidimo kako to znači da je $\chi(w)$ riječ konjugirana sama sebi s konjugirajućim riječima neparne duljine, tj. $\chi(w) = uv = vu$, gdje je $|u|$, pa i $|v|$ neparno. Budući da je $\chi(w)$ riječ na $\{11, 22\}$, slijedi da su zadnji blok u riječi u i prvi blok u riječi v neparnih duljina. Lako se vidi da je to moguće samo ako $\chi(w)$ ima sva slova ista, a to znači da je w nužno a ili b. Također se može pokazati da je $f_w^- \sim -f_w^-$ i $-f_w^+ \sim f_w^+$. No, da smo koristili širu definiciju ekvivalencije kvadratnih formi, dosta bi bilo promatrati višekratnike formi oblika f_w^+ jer je $f_w^-(x, y) = f_w^+(x, -y)$, a navedena supstitucija očito ima matricu s determinantom -1.

Kao u verziji s aproksimacijama, i Markovljev teorem za kvadratne forme često se iskazuje kao niz sve boljih nejednakosti koje povezuju $d(f)$ i $m(f)$ uz iznimke kada te nejednakosti ne vrijede. To je sljedeći rezultat.

Korolar 4.66. *Neka je M konačan skup Markovljevih brojeva takav da za svaka dva Markovljeva broja n i p, iz $n < p$ i $p \in M$, slijedi $n \in M$. Neka je W konačan skup svih donjih Christoffelovih riječi kojima su pridruženi Markovljevi brojevi iz M, tj. W je skup svih donjih Christoffelovih riječi w takvih da je $\mu(w)_{12} \in M$. Neka je m najmanji Markovljev broj koji nije u M. Tada za svaku indefinitnu binarnu kvadratnu formu f koja nije ekvivalentna višekratniku nijedne forme $f_w^\pm$, $w \in W$, vrijedi $\sqrt{d(f)} \geq \sqrt{9 - \frac{4}{m^2}}\, m(f)$.*

Dokaz. Ako je $\sqrt{d(f)} \geq 3m(f)$, onda zaključak vrijedi jer je $\sqrt{9 - \frac{4}{m^2}} < 3$. Stoga pretpostavimo da je $\sqrt{d(f)} < 3m(f)$. Po Teoremu 4.65 je f ekvivalentna nekom višekratniku f_v^- ili f_v^+ za donju Christoffelovu riječ v s pridruženim Markovljevim brojem n te vrijedi $\frac{\sqrt{d(f)}}{m(f)} = \sqrt{9 - \frac{4}{n^2}}$. Po uvjetu korolara mora biti $n \notin M$, pa je $n \geq m$ i zato $\sqrt{9 - \frac{4}{n^2}} \geq \sqrt{9 - \frac{4}{m^2}}$ iz čega slijedi tvrdnja. □

Sada dajemo nekoliko primjera. Za $M = \emptyset$ imamo $W = \emptyset$ i $m = 1$, pa je $\sqrt{9 - \frac{4}{m^2}} = \sqrt{5}$ te dobivamo rezultat Korkina i Zolotareva koji smo već izveli dobivši nejednakost (2.7). Dakle, za svaku indefinitnu binarnu kvadratnu formu je $\sqrt{d(f)} \geq \sqrt{5}m(f)$.

Isti autori dokazali su i idući rezultat koji slijedi iz prethodnog korolara kad je $M = \{1\}$, pa zato $W = \{a\}$, $m = 2$ i $\sqrt{9 - \frac{4}{m^2}} = \sqrt{8}$. Sjetimo se da je $\mu(a) = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right]$, pa je $f_a^{\pm} = x^2 \pm xy - y^2$. Ako forma f nije ekvivalentna višekratniku formi $f_a^{\pm} = x^2 \pm xy - y^2$, onda vrijedi $\sqrt{d(f)} \geq \sqrt{8}m(f)$. Ove primjere Markov navodi kao motivaciju svoga istraživanja čiji su rezultati objavljeni u radovima iz 1879. i 1880. godine.

Za $M = \{1, 2\}$ je $W = \{a, b\}$, $m = 5$, $\sqrt{9 - \frac{4}{m^2}} = \frac{\sqrt{221}}{5}$. Zbog $\mu(b) = \left[\begin{smallmatrix} 5 & 2 \\ 2 & 1 \end{smallmatrix}\right]$ je $f_b^{\pm} = 2x^2 \pm 4xy - 2y^2$. Stoga za formu f koja nije ekvivalentna višekratniku nijedne od formi $x^2 \pm xy - y^2$, $2x^2 \pm 4xy - 2y^2$, vrijedi $\sqrt{d(f)} \geq \frac{\sqrt{221}}{5}m(f)$.

Bilo iz Teorema 4.55 i 4.65 bilo, uzevši u obzir karakterizaciju spektara u Propoziciji 3.1 i Lemu 3.2, direktno iz Leme 4.64 i Teorema 4.39, slijedi karakterizacija spektara na njihovom početnom dijelu.

Teorem 4.67. *Za Lagrangeov i Markovljev spektar vrijedi*

$$\mathcal{L} \cap (0, 3) = \mathcal{M} \cap (0, 3) = \left\{ \sqrt{9 - \frac{4}{m^2}} \ : m \text{ Markovljev broj} \right\}.$$

4.7.4 Primjeri

Markovljevi brojevi manji od 500 su

$$1,\ 2,\ 5,\ 13,\ 29,\ 34,\ 89,\ 169,\ 194,\ 233,\ 433$$

kao što se vidi iz početnog dijela stabla Markovljevih trojki na str. 60. Pripadne Markovljeve konstante $\sqrt{9 - \frac{4}{m^2}} = \frac{\sqrt{9m^2 - 4}}{m}$ su redom

$$\sqrt{5},\ \sqrt{8},\ \frac{\sqrt{221}}{5},\ \frac{\sqrt{1517}}{13},\ \frac{\sqrt{7565}}{29},\ \frac{\sqrt{10400}}{34} = \frac{\sqrt{2600}}{17},\ \frac{\sqrt{71285}}{89},$$
$$\frac{\sqrt{257045}}{169},\ \frac{\sqrt{338720}}{194} = \frac{\sqrt{84680}}{97},\ \frac{\sqrt{488597}}{233},\ \frac{\sqrt{1687397}}{433}.$$

Jedanaest donjih Christoffelovih riječi koje odgovaraju navedenim Markovljevim brojevima dobivamo iz usporedbe sa stablom na str. 53 i to su

$$a,\ b,\ ab,\ a^2b,\ ab^2,\ a^3b,\ a^4b,\ ab^3,\ a^2bab,\ a^5b,\ abab^2.$$

Slike ovih donjih Christoffelovih riječi po homomorfizmu μ su matrice

$$\mu(a)=\begin{bmatrix}2&1\\1&1\end{bmatrix},\ \mu(b)=\begin{bmatrix}5&2\\2&1\end{bmatrix},\ \mu(ab)=\begin{bmatrix}12&5\\7&3\end{bmatrix},\ \mu(a^2b)=\begin{bmatrix}31&13\\19&8\end{bmatrix},$$
$$\mu(ab^2)=\begin{bmatrix}70&29\\41&17\end{bmatrix},\ \mu(a^3b)=\begin{bmatrix}81&34\\50&21\end{bmatrix},\ \mu(a^4b)=\begin{bmatrix}212&89\\131&55\end{bmatrix},\ \mu(ab^3)=\begin{bmatrix}408&169\\239&99\end{bmatrix},$$
$$\mu(a^2bab)=\begin{bmatrix}463&194\\284&119\end{bmatrix},\ \mu(a^5b)=\begin{bmatrix}555&233\\343&144\end{bmatrix},\ \mu(abab^2)=\begin{bmatrix}1045&433\\613&254\end{bmatrix}.$$

Odgovarajuće kvadratne iracionalnosti $x_w=[\chi(\widetilde{w})^\infty]=\frac{p-s}{2m}+\frac{1}{2}\sqrt{9-\frac{4}{m^2}}=\frac{p-s+\sqrt{9m^2-4}}{2m}$ su redom

$$x_a=\frac{1+\sqrt{5}}{2},\ x_b=1+\sqrt{2},\ x_{ab}=\frac{9+\sqrt{221}}{10},\ x_{a^2b}=\frac{23+\sqrt{1517}}{26},$$
$$x_{ab^2}=\frac{53+\sqrt{7565}}{58},\ x_{a^3b}=\frac{15+\sqrt{650}}{17},\ x_{a^4b}=\frac{157+\sqrt{71285}}{178},\ x_{ab^3}=\frac{309+\sqrt{257045}}{338},$$
$$x_{a^2bab}=\frac{172+\sqrt{84680}}{97},\ x_{a^5b}=\frac{411+\sqrt{488597}}{466},\ x_{abab^2}=\frac{791+\sqrt{1687397}}{866}.$$

Konačno, dajemo Markovljeve kvadratne forme $f_w^+(x,y)=mx^2+(p-s)xy-ry^2$, a forme f_w^- se dobiju promjenom predznaka koeficijenta uz xy. Radi preglednosti, ispuštamo $+$ u oznakama formi

$$f_a=x^2+xy-y^2,\ f_b=2x^2+4xy-2y^2,\ f_{ab}=5x^2+9xy-7y^2,\ f_{a^2b}=13x^2+23xy-19y^2,$$
$$f_{ab^2}=29x^2+53xy-41y^2,\ f_{a^3b}=34x^2+60xy-50y^2,\ f_{a^4b}=89x^2+157xy-131y^2,$$
$$f_{ab^3}=169x^2+309xy-239y^2,\ f_{a^2bab}=194x^2+344xy-284y^2,$$
$$f_{a^5b}=233x^2+411xy-343y^2,\ f_{abab^2}=433x^2+791xy-613y^2.$$

4.7.5 Dodatni rezultati

Pokazat ćemo još neke rezultate vezane uz Markovljeve kvadratne forme.

Propozicija 4.68. *Neka je w donja Christoffelova riječ i m njoj pridruženi Markovljev broj. Tada su za neparan m koeficijenti od f_w, a za paran m koeficijenti od $\frac{1}{2}f_w$ relativno prosti cijeli brojevi.*

Dokaz. Imamo $s=3m-p$, pa je $p-s=2p-3m$. Također je $1=\det\big(\mu(w)\big)=p(3m-p)-mr$, pa zato $p^2+1=m(3p-r)$ iz čega slijedi da su p i m relativno prosti. Ako neki prosti broj l dijeli $p-s$ i m, onda l dijeli $2p$. Budući da l ne može dijeliti p jer su p i m relativno prosti, mora biti $l=2$. To se može dogoditi samo ako je m paran. Posebno, $p-s$ i m su relativno prosti ako je m neparan.

Obratno, za paran $m=2n$, imamo da je $p-s$ paran te je p neparan jer su p i m relativno prosti. Modulo 4 je $2n(3p-r)=p^2+1\equiv 2$, tako da su n i $3p-r$ neparni, dok je r paran. Nadalje, najveći zajednički djelitelj od m i $p-s$ je 2 jer je n neparan. Time smo dokazali propoziciju. $\square$

Za donju Christoffelovu riječ w, matrica $\mu(w)$ ima poseban oblik koji je opisan u sljedećem teoremu.

Teorem 4.69. *Neka je w prava donja Christoffelova riječ sa standardnom faktorizacijom $w = w_1w_2$. Neka su m, m_1, m_2 redom Markovljevi brojevi pridruženi riječima w, w_1, w_2. Tada postoje jedinstveni prirodni brojevi u i v takvi da je $m_1u \equiv m_2 \pmod m$, $0 < u < \frac{m}{2}$ i $u^2 + 1 = mv$. Vrijedi*

$$\mu(w) = \begin{bmatrix} 2m + u & m \\ 2m - u - v & m - u \end{bmatrix}.$$

Korolar 4.70. *Imamo $f_w^{\pm}(x, y) = mx^2 \pm (m + 2u)xy - (2m - u - v)y^2$.*

Znamo da prema Teoremu 4.16 i Lemi 4.17, možemo pisati

$$\mu(w) = \begin{bmatrix} a_w & m_w \\ c_w & 3m_w - a_w \end{bmatrix},$$

gdje je m_w Markovljev broj pridružen donjoj Christoffelovoj riječi w.

Lema 4.71. *Ako je $w = w_1w_2$ standardna faktorizacija donje Christoffelove riječi w, onda je $m_{w_1} = -a_wm_{w_2} + m_wa_{w_2}$ i $m_{w_2} = -a_{w_1}m_w + m_{w_1}a_w$.*

Dokaz. Kako je $\mu(w) = \mu(w_1)\mu(w_2)$ i te matrice imaju determinantu 1, vrijedi

$$\mu(w_1) = \mu(w)\mu(w_2)^{-1} = \begin{bmatrix} a_w & m_w \\ c_w & 3m_w - a_w \end{bmatrix} \begin{bmatrix} 3m_{w_2} - a_{w_2} & -m_{w_2} \\ -c_{w_2} & a_{w_2} \end{bmatrix} \quad \text{i}$$
$$\mu(w_2) = \mu(w_1)^{-1}\mu(w) = \begin{bmatrix} 3m_{w_1} - a_{w_1} & -m_{w_1} \\ -c_{w_1} & a_{w_1} \end{bmatrix} \begin{bmatrix} a_w & m_w \\ c_w & 3m_w - a_w \end{bmatrix}.$$

Slijedi da je $m_{w_1} = -a_wm_{w_2} + m_wa_{w_2}$ i $m_{w_2} = (3m_{w_1} - a_{w_1})m_w - m_{w_1}(3m_w - a_w) = -a_{w_1}m_w + m_{w_1}a_w$. $\square$

Definiramo *indeks* donje Christoffelove riječi w kao broj $I_w = \frac{a_w}{m_w}$.

Lema 4.72. *Ako je $w = w_1w_2$ prava donja Christoffelova riječ sa svojom standardnom faktorizacijom, onda je $2 = I_a \le I_{w_1} < I_w < I_{w_2} \le I_b = \frac{5}{2}$.*

Dokaz. Direktno vidimo da je $I_a = 2$ i $I_b = \frac{5}{2}$. Neka je $w = w_1w_2$ standardna faktorizacija od w. Tada je po prethodnoj lemi $m_{w_1} = -a_wm_{w_2} + m_wa_{w_2}$ i $m_{w_2} = -a_{w_1}m_w + m_{w_1}a_w$. Podijelimo li prvu jednakost sa $m_wm_{w_2}$ te drugu jednakost sa $m_wm_{w_1}$, dobivamo

$$0 < \frac{a_{w_2}}{m_{w_2}} - \frac{a_w}{m_w}, \qquad 0 < \frac{a_w}{m_w} - \frac{a_{w_1}}{m_{w_1}}.$$

Stoga je $I_{w_1} < I_w < I_{w_2}$, pa možemo zaključiti da tvrdnja leme vrijedi indukcijom po duljini riječi iz $I_a \le I_{w_1}$ i $I_{w_2} \le I_b$. $\square$

Dokaz Teorema 4.69. Imamo $m = m_w$, $m_1 = m_{w_1}$, $m_2 = m_{w_2}$. Neka je $a = a_w$, $c = c_w$. Zbog Leme 4.71 je $m_2 = -a_{w_1}m + m_1a$, pa stoga $m_1a \equiv m_2 \pmod m$. Prema Lemi 4.72 imamo $2 < \frac{a}{m} < \frac{5}{2} = 2 + \frac{1}{2}$, pa je $2m < a < 2m + \frac{1}{2}m$. Zato je $a = 2m + u$ za neki cijeli broj u takav da je $0 < u < \frac{1}{2}m$. Kako je $a \equiv u \pmod m$, dobivamo $m_1u \equiv m_2 \pmod m$, pa kvadriranjem slijedi $m_1^2u^2 \equiv m_2^2 \pmod m$. Prema Teoremu 4.16 je $\{m, m_1, m_2\}$ Markovljeva trojka te iz Markovljeve jednadžbe $m^2 + m_1^2 + m_2^2 = 3mm_1m_2$ vidimo da je $m_1^2 \equiv -m_2^2 \pmod m$, pa zato $m_1^2u^2 \equiv -m_1^2 \pmod m$. Po Korolaru 4.25 su članovi iste Markovljeve trojke u parovima relativno prosti, pa je m_1 invertibilan modulo m i dobivamo $u^2 \equiv -1 \pmod{m}$. Stoga je $u^2+1 = vm$ za neki prirodan broj v. Konačno, $3m - a = m - u$ te imamo

$$\begin{aligned}
1 = \det \begin{bmatrix} 2m+u & m \\ c & m-u \end{bmatrix} &= (2m+u)(m-u) - mc = 2m^2 - um - u^2 - mc, \\
mc = 2m^2 - um - u^2 - 1 &= 2m^2 - um - vm, \\
c = 2m - u - v&.
\end{aligned}$$

Time smo pokazali da matrica $\mu(w)$ ima traženi oblik. $\square$

Napomenimo da primjeri kvadratnih formi koje je dao Markov nisu isti kao forme koje smo promatrali, ali su, dakako, njima ekvivalentne forme. Originalna Markovljeva forma, uz oznake iz Teorema 4.69 je $mx^2 + (3m-2u)xy + (v-3u)y^2$. Lako se vidi da je ta forma reducirana, a zaista je ekvivalentna f_w^- jer je

$$\begin{aligned}
f_w^-(x+2y, y) &= m(x+2y)^2 - (m+2u)(x+2y)y - (2m-u-v)y^2 \\
&= mx^2 + (4m - m - 2u)xy + (4m - 2m - 4u - 2m + u + v)y^2 \\
&= mx^2 + (3m-2u)xy + (v-3u)y^2.
\end{aligned}$$

Eulerove *kontinuante* možemo promatrati kao polinome koji su definirani za svaki $k \in \mathbb{N}_0$ i proizvoljne $n_1, \ldots, n_k$ u nekom komutativnom prstenu na sljedeći način: $K_{-1} = 0$, $K_0 = 1$, a za $k \geq 1$ je

$$K_k(n_1, \ldots, n_k) = K_{k-1}(n_1, \ldots, n_{k-1})n_k + K_{k-2}(n_1, \ldots, n_{k-2}).$$

Uobičajeno je ispustiti indeks k i pisati $K(n_1, \ldots, n_k)$ umjesto $K_k(n_1, \ldots, n_k)$. Usporedbom s (1.2), vidimo da je $K(n_1, \ldots, n_k)$ naprosto brojnik verižnog razlomka $[n_1, \ldots, n_k]$ formalno sređenog "odozdo prema gore".

Jednakost (1.3) možemo zapisati pomoću ovih polinoma kao

$$\begin{bmatrix} K(n_1, \ldots, n_k) & K(n_1, \ldots, n_{k-1}) \\ K(n_2, \ldots, n_k) & K(n_2, \ldots, n_{k-1}) \end{bmatrix} = \begin{bmatrix} n_1 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} n_k & 1 \\ 1 & 0 \end{bmatrix}. \tag{4.6}$$

Ovu je formulu lako dokazati indukcijom po k, množeći matrice slijeva nadesno. Množeći iste matrice zdesna nalijevo, dobivamo dualnu rekurzivnu formulu

$$K(n_1, \ldots, n_k) = n_1 K(n_2, \ldots, n_k) + K(n_3, \ldots, n_k).$$

Transponirajući pak isti produkt i koristeći simetričnost matrica koje su faktori, dobivamo $K(n_1, \ldots, n_k) = K(n_k, \ldots, n_1)$. Primijetimo također da je $K(1, n-1) = K(1)(n-1) + K_0 = n = K(n)$ te za $k \in \mathbb{N}$ iz dualne rekurzivne formule dobivamo

$$\begin{aligned}
K(1, n-1, n_1, \ldots, n_k) &= K(n-1, n_1, \ldots, n_k) + K(n_1, \ldots, n_k) \\
&= (n-1)K(n_1, \ldots, n_k) + K(n_2, \ldots, n_k) + K(n_1, \ldots, n_k) \\
&= nK(n_1, \ldots, n_k) + K(n_2, \ldots, n_k) = K(n, n_1, \ldots, n_k).
\end{aligned}$$

U idućem rezultatu ćemo za proizvoljnu riječ $h = n_1 \ldots n_k$ na $\mathbb{N}$ pisati $K(h)$ za broj $K(n_1, \ldots, n_k)$.

Teorem 4.73 (Frobenius). *Neka je $w = ahb$ prava donja Christoffelova riječ, neka je $g = \chi(h)$ i neka su m, u, v brojevi definirani u Teoremu 4.69. Tada je $m = K(2g2)$, $u = K(2g) = K(g2)$, $v = K(g)$.*

Uočimo da je po Teoremu 4.6 riječ h, pa zato i g palindrom.

Za primjer uzmimo donju Christoffelovu riječ nagiba $\frac{2}{3}$, tj. $w = aabab$ za koju je $h = aba$, $g = 112211$, $m = K(21122112) = 194$, $u = K(2112211) = 75$, $v = K(112211) = 29$ i $75^2 + 1 = 194 \cdot 29$.

Dokaz. Po Teoremu 4.16 i Lemi 4.17 znamo da je $m = \mu(w)_{12}$. Kako je $\mu(a) = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]^2$ i $\mu(b) = \left[\begin{smallmatrix} 2 & 1 \\ 1 & 0 \end{smallmatrix}\right]^2$, imamo iz jednakosti (4.6) i Teorema 4.69 da je

$$\begin{bmatrix} 2m + u & m \\ 2m - u - v & m - u \end{bmatrix} = \begin{bmatrix} K(11g22) & K(11g2) \\ K(1g22) & K(1g2) \end{bmatrix}.$$

Sada je $m = K(11g2) = K(2g2)$ prema zadnjoj primjedbi prije ovog teorema.

Nadalje, $m - u = K(1g2)$, odnosno $K(11g2) = K(1g2) + u$. Iz dualne rekurzivne formule je $K(11g2) = K(1g2) + K(g2)$, pa zaključujemo da je $u = K(g2) = K(2g)$ jer je g palindrom.

Konačno, vrijedi $2m - u - v = K(1g22)$ tako da je

$$\begin{aligned} v &= 2m - u - K(1g22) = 2K(11g2) - K(g2) - K(1g22) \\ &= 2K(1g2) + 2K(g2) - K(g2) - 2K(1g2) - K(1g) \\ &= K(g2) - K(1g) = K(g11) - K(g1) = K(g), \end{aligned}$$

pri čemu smo koristili da je $K(g2) = K(g11)$ što se lako pokaže. $\square$

Poglavlje 5

Slutnja o jedinstvenosti

Već smo spomenuli da je u Frobeniusovom članku iz 1913. po prvi puta izložena slutnja koju ćemo najprije iznijeti u nekoliko ekvivalentnih formulacija.

I. Preslikavanje $w \mapsto m_w$ koje donjoj Christoffelovoj riječi w pridružuje Markovljev broj $m_w = \frac{1}{3}\operatorname{Tr}(\mu(w)) = \mu(w)_{12}$ je injekcija (vidi kod Teorema 4.16 za oznake).

II. Svaki Markovljev broj pojavljuje se točno jednom kao maksimum u Markovljevoj trojci. Ovu tvrdnju možemo iskazati i drugačije: centralni elementi u čvorovima stabla Markovljevih trojki su svi različiti (vidi str. 60).

III. Iracionalni brojevi α i β takvi da je $L(\alpha) = L(\beta) < 3$ nužno su ekvivalentni.

Ekvivalentnost verzija I i II slutnje slijedi direktno iz Teorema 4.16 i Korolara 4.23, a ekvivalentnost verzija I i III iz Teorema 4.55. Pokazat ćemo sada neke parcijalne rezultate vezane uz slutnju o jedinstvenosti.

5.1 Jedinstvenost Markovljevih brojeva s određenom faktorizacijom

Označimo sa $\mathbb{M}$ skup Markovljevih brojeva.

Propozicija 5.1. *Neka je $m \in \mathbb{M}$. Svaki neparan prosti djelitelj p od m je oblika $p \equiv 1 \pmod 4$. Ako je m neparan, onda je $m \equiv 1 \pmod 4$. Ako je m paran, onda je $m \equiv 2 \pmod{32}$.*

Dokaz. Neka je m sadržan u Markovljevoj trojci $\{m, m_2, m_3\}$. Tada iz Markovljeve jednadžbe slijedi $m_2^2 \equiv -m_3^2 \pmod m$. Za svaki neparni prosti djelitelj p od m je $m_2^2 \equiv -m_3^2 \pmod p$, a kako su m_2 i m_3 relativno prosti s m, zaključujemo da je -1 kvadratni ostatak modulo p, pa je $p \equiv 1 \pmod 4$. Time je dokazan slučaj kada je m neparan.

Neka je $m = 2^k t$ paran broj, gdje je t neparan. Tada je po upravo dokazanom $t = 4s+1$. Kako je m paran, to su m_2 i m_3 neparni, pa je $m_2 = 4a+1$ i $m_3 = 4b+1$. Ako je $k \geq 2$, onda je $m \equiv 0 \pmod 4$, $m_2^2 \equiv m_3^2 \equiv 1 \pmod 4$ i zato

$$2 \equiv m^2 + m_2^2 + m_3^2 = 3mm_2m_3 \equiv 0 \pmod 4,$$

što nije moguće. Stoga je $k = 1$, $m = 8s + 2$ te možemo pretpostaviti da je $s \geq 1$. Markovljeva jednadžba sada glasi

$$(8s+2)^2 + (4a+1)^2 + (4b+1)^2 = 3(8s+2)(4a+1)(4b+1),$$

što daje

$$8s(8s + 4 - 3m_2m_3) = 16(6ab + a - a^2 + b - b^2).$$

Lijeva je strana različita od nule jer je m_2m_3, pa stoga i izraz u zagradi, neparan. Zato je

$$0 \neq 6ab + a - a^2 + b - b^2 \equiv 0 \pmod 2,$$

te zaključujemo da 32 dijeli $8s$. Stoga je $s = 4s'$ i zato $m = 32s' + 2$. □

Neka je $\{m, m_2, m_3\}$ prava Markovljeva trojka s maksimumom $m \geq 5$. Kao što smo već više puta naveli, brojevi m, m_2, m_3 su po Korolaru 4.25 u parovima relativno prosti i vrijedi $m^2 + m_2^2 + m_3^2 = 3mm_2m_3$, pa zato $m_2^2 \equiv -m_3^2 \pmod m$. Modulo m imamo ekvivalentnost kongruencija

$$m_3u \equiv m_2 \quad \Leftrightarrow \quad m_3^2u \equiv m_2m_3 \quad \Leftrightarrow \quad -m_2^2u \equiv m_2m_3 \quad \Leftrightarrow \quad m_2u \equiv -m_3.$$

Ako je u_0 rješenje jedne od jednadžbi $m_2u \equiv m_3 \pmod m$ i $m_2u \equiv -m_3 \pmod m$, onda je $m - u_0$ rješenje druge. Stoga točno jedna od jednadžbi $m_2u \equiv m_3 \pmod m$ i $m_3u \equiv m_2 \pmod m$ ima rješenje u intervalu $0 < u < \frac{m}{2}$. Takav jedinstveni cijeli broj u koji se pojavljuje u Teoremu 4.69 i zadovoljava

$$m_2u \equiv \pm m_3 \pmod m, \qquad 0 < u < \frac{m}{2}, \qquad u^2 \equiv -1 \pmod m \tag{5.1}$$

zove se *karakteristični broj* u pridružen Markovljevoj trojci $\{m, m_2, m_3\}$ s maksimumom $m \geq 5$.

Tako smo dobili nužan uvjet da neki broj bude Markovljev.

Lema 5.2. *Neka je $m \in \mathbb{M}$, $m \geq 5$. Tada kongruencija $x^2 \equiv -1 \pmod m$ ima rješenje u koje zadovoljava $0 < u < \frac{m}{2}$.*

Pokažimo da m i u zajedno određuju čitavu Markovljevu trojku $\{m, m_2, m_3\}$.

Propozicija 5.3. *Neka je $m \in \mathbb{M}$, $m \geq 5$ i pretpostavimo da su $\{m, m_2, m_3\}$ i $\{m, n_2, n_3\}$ Markovljeve trojke s maksimumom m i pripadnim karakterističnim brojevima u_1 i u_2. Ako je $u_1 = u_2$, onda je $\{m_2, m_3\} = \{n_2, n_3\}$, tj. trojke su jednake.*

Dokaz. Neka je $u = u_1 = u_2$. Tada je

$$m_2u \equiv \pm m_3 \pmod m, \qquad n_2u \equiv \pm n_3 \pmod m,$$

pa zato $u \equiv \pm m_2^{-1}m_3 \equiv \pm n_2^{-1}n_3 \pmod m$. Slijedi $m_2n_3 \equiv \pm m_3n_2 \pmod m$, tj. $m \mid m_2n_3 - m_3n_2$ ili $m \mid m_2n_3 + m_3n_2$. Pretpostavimo da je $m \mid m_2n_3 - m_3n_2$ jer se drugi slučaj pokazuje analogno.

Koristit ćemo više puta činjenicu dokazanu u Korolaru 4.25 da su brojevi u Markovljevoj trojci u parovima relativno prosti. Neka je p neparan prosti djelitelj od m. Tada $p \mid m_2n_3 - m_3n_2$, ali $p \nmid m_2n_3 + m_3n_2$ jer bi inače $p \mid 2m_2n_3$ i zato $p \mid m_2$ ili $p \mid n_3$ što nije moguće.

Iz

$$\frac{m^2 + m_2^2 + m_3^2}{m_2m_3} = 3m = \frac{m^2 + n_2^2 + n_3^2}{n_2n_3}$$

slijedi

$$m^2(m_2m_3 - n_2n_3) = (m_2n_2 - m_3n_3)(m_2n_3 - m_3n_2). \tag{5.2}$$

Ako je $m_2 m_3 = n_2 n_3$, onda imamo $m_2 n_2 = m_3 n_3$ ili $m_2 n_3 = m_3 n_2$. Uzmimo najprije da vrijedi prvi slučaj. Budući su m_2 i m_3 relativno prosti, vrijedi $m_2 \mid n_3$, a jer su n_2 i n_3 prosti, dobivamo i $n_3 \mid m_2$, pa je zato $m_2 = n_3$ i $m_3 = n_2$. Isto zaključivanje vodi nas u drugom slučaju do $m_2 = n_2$ i $m_3 = n_3$, pa je u oba slučaja $\{m_2, m_3\} = \{n_2, n_3\}$. Zato možemo pretpostaviti da su svi faktori u 5.2 različiti od nule.

Broj p dijeli $m_2 n_3 - m_3 n_2$, ali ne i $m_2 n_2 - m_3 n_3$. U protivnom bi množeći kongruencije $m_2 n_3 \equiv m_3 n_2 \pmod{p}$ i $m_2 n_2 \equiv m_3 n_3 \pmod{p}$ dobili $m_2^2 n_2 n_3 \equiv m_3^2 n_2 n_3 \pmod{p}$ te stoga $m_2^2 \equiv m_3^2 \pmod{p}$. No također vrijedi $p \mid m_2^2 + m_3^2$, odnosno $m_2^2 \equiv -m_3^2 \pmod{p}$ što daje $p \mid 2m_2^2$ i zato $p \mid m_2$, a to je u kontradikciji s $p \mid m$.

Dakle, zbog (5.2) vidimo da za sve neparne proste brojeve p, iz $p^k \mid m$ slijedi $p^{2k} \mid m_2 n_3 - m_3 n_2$. Ako je m neparan, to znači da $m^2 \mid m_2 n_3 - m_3 n_2$ i stoga $m_2 n_3 = m_3 n_2$ jer je $m > m_2, m_3, n_2, n_3$. Kao i prije, ovo povlači $\{m_2, m_3\} = \{n_2, n_3\}$. Ako je m paran, onda je $m_2, m_3, n_2, n_3 \equiv 1 \pmod{4}$ po Propoziciji 5.1, pa je $m_2 n_3 - m_3 n_2 \equiv 0 \pmod{4}$. Istim zaključivanjem kao prije, iz ovoga slijedi $m^2 \mid m_2 n_3 - m_3 n_2$ što vodi do $\{m_2, m_3\} = \{n_2, n_3\}$. □

Radi sažetosti ćemo Markovljev broj m zvati *jedincatim* ako je m maksimum u samo jednoj Markovljevoj trojci. Općenitije ćemo reći da za podskup S skupa prirodnih brojeva vrijedi *svojstvo jedincatosti* ako je svaki Markovljev broj u S jedincat. Slutnja o jedinstvenosti može se sada iskazati kao tvrdnja da čitav $\mathbb{N}$ ima svojstvo jedincatosti.

Iz Propozicije 5.3 neposredno dobivamo sljedeći rezultat o jedincatosti.

Korolar 5.4. *Neka je $m \in \mathbb{M}$, $m \geq 5$. Ako kongruencija $x^2 \equiv -1 \pmod{m}$ ima jedinstveno rješenje u intervalu $0 < x < \frac{m}{2}$, onda je m jedincati Markovljev broj.*

S obzirom na prethodni korolar, vidimo da će nam biti od interesa broj rješenja kongruencije $x^2 \equiv -1 \pmod{m}$.

Lema 5.5. *Neka je p neparan prost broj, k prirodan broj te c cijeli broj koji nije djeljiv sa p. Tada kongruencija $x^2 \equiv c \pmod{p^k}$ ima najviše dva rješenja.*

Dokaz. Za dva rješenja x_1, x_2 ove kongruencije, imamo $(x_1 - x_2)(x_1 + x_2) \equiv x_1^2 - x_2^2 \equiv 0 \pmod{p^k}$. Ne može biti $x_1 - x_2 \equiv x_1 + x_2 \equiv 0 \pmod{p}$ jer bi inače $p \mid 2x_1$, pa i $p \mid c$. Stoga je $x_1 \equiv x_2 \pmod{p^k}$ ili $x_1 \equiv -x_2 \pmod{p^k}$ iz čega slijedi tvrdnja leme. □

Teorem 5.6. *Neka je $m = p_1^{k_1} \cdots p_t^{k_t} \in \mathbb{M}$ ili $m = 2p_1^{k_1} \cdots p_t^{k_t} \in \mathbb{M}$, gdje su $p_1, \dots, p_t$ neparni prosti brojevi. Tada je m maksimum u najviše 2^{t-1} Markovljevih trojki.*

Dokaz. Možemo pretpostaviti da je $m \geq 5$ i da su $p_1, \dots, p_t$ različiti prosti brojevi. Po lemi kongruencija $x^2 \equiv -1 \pmod{p_i^{k_i}}$ ima najviše dva rješenja, a $x^2 \equiv -1 \pmod{2}$ očito ima točno jedno rješenje. Iz Kineskog teorema o ostatcima slijedi da i u jednom i u drugom slučaju $x^2 \equiv -1 \pmod{m}$ ima najviše 2^t rješenja. Rješenja $0 < u < m$ dolaze u parovima $u, m - u$, pa polovica njih zadovoljava $0 < u < \frac{m}{2}$. Tvrdnja teorema slijedi sada iz Propozicije 5.3. □

Korolar 5.7. *Svaki Markovljev broj m oblika $m = p^k$ ili $m = 2p^k$, gdje je p neparan prost broj, je jedincat.*

Izvest ćemo pomoću Leme 5.5 još jedan rezultat o jedincatosti Markovljevih brojeva. Promotrimo pravu Markovljevu trojku $\{m, m_2, m_3\}$ različitu od $\{5, 2, 1\}$, pri čemu je $m > m_2 > m_3$. Pokažimo da vrijedi

$$m > 2m_2, \qquad m_2 > 2m_3. \tag{5.3}$$

Ove nejednakost očito vrijede za trojke $\{13,5,1\}$ i $\{29,5,2\}$ s najmanjim maksimumom. Zbog rekurzivnog pravila za izgradnju stabla Markovljevih trojki na str. 60, vidimo da je dovoljno provjeriti kako iz činjenice da trojka $\{m,m_2,m_3\}$ zadovoljava (5.3), slijedi da analogne nejednakosti zadovoljavaju i trojke $\{3mm_2-m_3,m,m_2\}$ i $\{3mm_3-m_2,m,m_3\}$. To je očito jer iz (5.3) slijedi

$$3mm_2-m_3>3m-m_3>2m>4m_2 \qquad \text{i} \qquad 3mm_3-m_2\geq 3m-m_2>2m>4m_3.$$

Propozicija 5.8. *Svaki Markovljev broj oblika* $m=\dfrac{2^\ell p^k\pm 2}{3}$, *gdje je* p *neparan prost broj i* $\ell\in\{0,1,2,3\}$, *je jedincat.*

Dokaz. Možemo pretpostaviti da je $m\geq 13$. Promotrimo trojku $\{m,m_2,m_3\}$ sa $m>m_2>m_3$. Markovljeva jednadžba povlači da je

$$(m_2-m_3)^2+m^2=m_2m_3(3m-2), \qquad (m_2+m_3)^2+m^2=m_2m_3(3m+2). \tag{5.4}$$

Slučaj 1. Imamo da je m neparan, pa zato $m=\frac{p^k\pm 2}{3}$.

Najprije uzimamo da je $m=\frac{p^k+2}{3}$. Tada iz (5.4) slijedi

$$(m_2-m_3)^2+m^2=m_2m_3p^k\equiv 0 \pmod{p^k}.$$

Nadalje, m je relativno prost sa p^k jer je p neparan. Zbog (5.3) je

$$0<m_2-m_3<\frac{m}{2}-1<\frac{p^k}{2}.$$

Lema 5.5 povlači da je m_2-m_3 jedinstveno rješenje kongruencije $x^2\equiv -m^2 \pmod{p^k}$ u intervalu $0<x<\frac{p^k}{2}$. Neka je $\{m,n_2,n_3\}$ druga Markovljeva trojka. Zbog jedinstvenosti rješenja je $m_2-m_3=n_2-n_3$, a iz (5.4) je $m_2m_3=n_2n_3$. Stoga su $\{m_2,-m_3\}$, $\{n_2,-n_3\}$ korijeni iste kvadratne jednadžbe $x^2+(m_3-m_2)x-m_2m_3=0$, pa zaključujemo $\{m_2,m_3\}=\{n_2,n_3\}$.

Neka je sad $m=\frac{p^k-2}{3}$. Tada je

$$(m_2+m_3)^2+m^2=m_2m_3p^k\equiv 0 \pmod{p^k}$$

te po (5.3)

$$0<m_2+m_3<\frac{3m+2}{2}=\frac{p^k}{2},$$

pa jedincatost od m dobivamo kao i prije.

Slučaj 2. Broj m je paran.

Prema Propoziciji 5.1 je $m=2+32t$ i $m_2,m_3\equiv 1 \pmod 4$. Stoga je

$$3m-2=4+96t=4(1+24t), \quad \text{pa u tom slučaju imamo} \quad m=\frac{4p^k+2}{3},$$
$$3m+2=8+96t=8(1+12t), \quad \text{pa u tom slučaju imamo} \quad m=\frac{8p^k-2}{3}.$$

U prvom slučaju po (5.4) imamo

$$\left(\frac{m_2-m_3}{2}\right)^2+\left(\frac{m}{2}\right)^2=m_2m_3p^k\equiv 0 \pmod{p^k}, \qquad 0<\frac{m_2-m_3}{2}<\frac{m}{4}<\frac{3m-2}{8}=\frac{p^k}{2}.$$

Stoga je $\frac{m_2-m_3}{2}$ jedinstveno rješenje kongruencije $x^2 \equiv -(\frac{m}{2})^2 \pmod{p^k}$ u intervalu $0 < x < \frac{p^k}{2}$ te kao i prije zaključujemo da je m jedincat.

U završnom slučaju je $3m+2 = 8p^k$ i vrijedi

$$\left(\frac{m_2+m_3}{2}\right)^2 + \left(\frac{m}{2}\right)^2 = m_2m_3(2p^k) \equiv 0 \pmod{2p^k}, \qquad 0 < \frac{m_2+m_3}{2} < \frac{3m+2}{8} = p^k.$$

Budući da Lema 5.5 također vrijedi za modul $2p^k$, opet dobivamo jedincatost m. □

Primjer. Za Markovljev broj $F_{15} = 610 = 2 \cdot 5 \cdot 61$ jedincatost ne možemo dobiti pomoću Korolara 5.7, ali zbog $610 = \frac{8 \cdot 229 - 2}{3}$ iz Propozicije 5.8 slijedi da je 610 jedincat.

Prirodno se pojavljuju dva moguća smjera napada na slutnju o jedinstvenosti:

1. Povećati skup S koji ima svojstvo jedincatosti.

2. Suziti interval $0 < x < \frac{m}{2}$ koji treba provjeriti, a u kojemu mora postojati rješenje kongruencije $x^2 \equiv -1 \pmod{m}$ ako je m Markovljev broj.

Vezano uz prvi smjer, najbolji trenutni rezultati za Markovljeve brojeve kojima faktorizacija u proste brojeve ima određeni oblik su Buttonovi:

(a) Neka je $m = p^k q^\ell \in \mathbb{M}$, gdje su p i q neparni prosti brojevi. Ako je $q^\ell > \frac{2}{3}p^{3k}$, onda je m jedincat.

(b) Neka je m neparan Markovljev broj oblika $m = Np^k$, gdje je p prost i $N \leq 10^{35}$. Tada je m jedincat.

Nažalost, čak ni opći slučaj $m = pq$ za različite proste brojeve p i q još nije dostižan. Više rezultata ovog tipa može se naći u Aignerovoj knjizi koju slijedimo u ovom potpoglavlju.

5.2 Rast Markovljevih brojeva

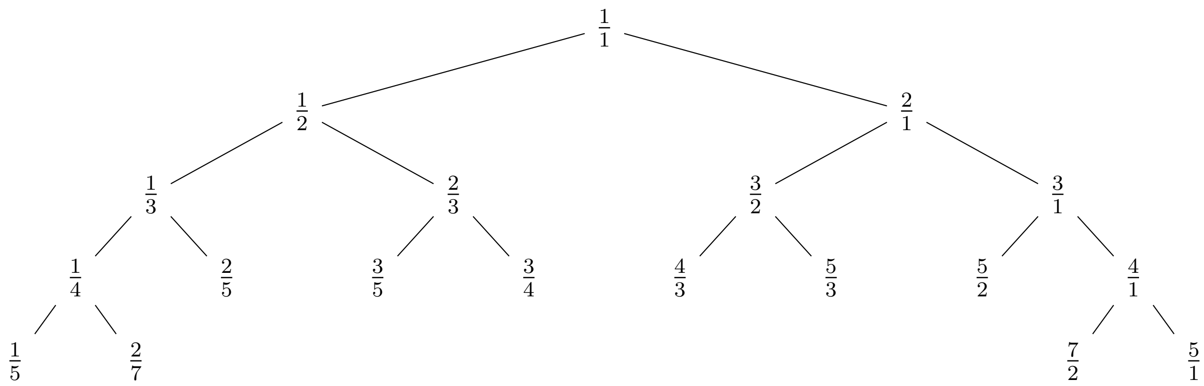
U nastavku ćemo Markovljeve brojeve pridružene donjim Christoffelovim riječima numerirati pomoću nagiba tih riječi. Svaka donja Christoffelova riječ w jednoznačno je određena svojim nagibom koji ćemo označiti $\nu(w) = \frac{|w|_b}{|w|_a}$. Zamijenimo li u stablu Christoffelovih parova na str. 53 svaki par (u, v) sa $\nu(uv)$, dobivamo takozvano *Stern–Brocotovo stablo* u kojemu se svaki pozitivan racionalan broj pojavljuje točno jedanput i to u svom skraćenom obliku.

Iz konstrukcije stabla Christoffelovih parova znamo da čvor (w_1, w_2) ima lijevo dijete (w_1, w_1w_2) i desno dijete (w_1w_2, w_2). Lako je provjeriti da za $\nu(w_1) < \nu(w_2)$ imamo

$$\nu(w_1) < \nu(w_1w_1w_2) < \nu(w_1w_2) < \nu(w_1w_2w_2) < \nu(w_2).$$

Budući da za početni čvor (a, b) vrijedi $\nu(a) = 1 < \nu(b) = \infty$, rekurzivno zaključujemo da za svaki čvor vrijedi $\nu(w_1) < \nu(w_2)$. Stoga lijevo dijete čvora (w_1, w_2) i svi njegovi potomci imaju nagib pridružene donje Christoffelove riječi manji od $\nu(w_1w_2)$, a desno dijete i svi njegovi potomci veći nagib. Drugim riječima, u Stern–Brocotovom stablu lijevo dijete i svi njegovi potomci su manji od danog čvora, a desno dijete i svi njegovi potomci su veći.

Za $t \in \mathbb{Q}_{>0} \cup \{\frac{0}{1}, \frac{1}{0}\}$ sa w_t označavamo donju Christoffelovu riječ nagiba t, a sa $m_t = \mu(w_t)_{12}$ pridruženi Markovljev broj. Jasno, slutnja o jedinstvenosti kaže da je $t \mapsto m_t$ ne samo surjekcija, nego i injekcija.



Početni dio Stern–Brocotova stabla

Pogledajmo dvije izvanjske grane u stablu Markovljevih trojki. Vidimo da lijeva grana ima čvorove redom $(1, m_{\frac{1}{n}}, m_{\frac{1}{n-1}})$ za $n \geq 1$. Iz pravila konstrukcije tog stabla je

$$m_{\frac{1}{n}} = 3 \cdot 1 \cdot m_{\frac{1}{n-1}} - m_{\frac{1}{n-2}}, \qquad \text{uz} \quad m_{\frac{1}{0}} = 2,\ m_{\frac{1}{1}} = 5.$$

Sada je lako indukcijom pokazati da je $m_{\frac{1}{n}} = F_{2n+3}$ za $n \geq 0$, gdje je kao i prije $F_0 = 0$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$ Fibonaccijev niz. Stoga ovu lijevu granu zovemo *Fibonaccijeva grana*.

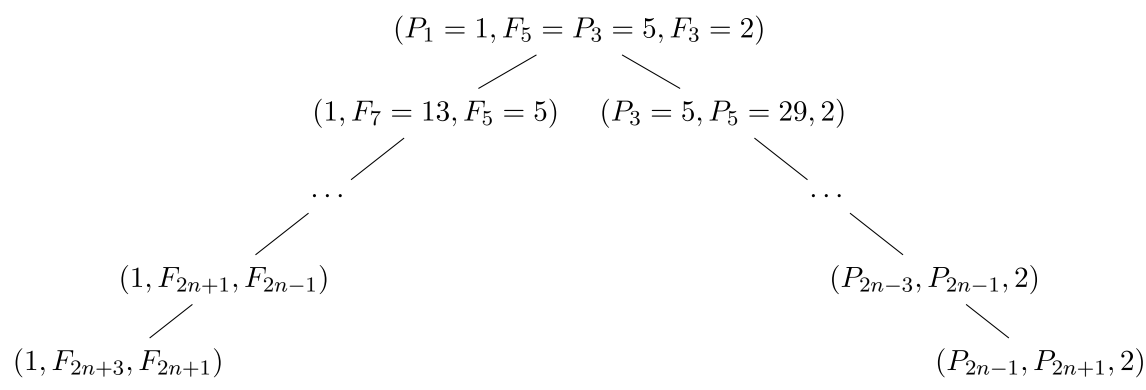
Slično dobivamo da desna grana ima redom čvorove $(m_{\frac{n-1}{1}}, m_{\frac{n}{1}}, 2)$ za $n \geq 1$. Vrijedi

$$m_{\frac{n}{1}} = 3 \cdot 2 \cdot m_{\frac{n-1}{1}} - m_{\frac{n-2}{1}}, \qquad \text{uz} \quad m_{\frac{0}{1}} = 1,\ m_{\frac{1}{1}} = 5.$$

Indukcijom se pokaže da je $m_{\frac{n}{1}} = P_{2n+1}$ za $n \geq 0$, gdje je $P_0 = 0$, $P_1 = 1$, $P_n = 2P_{n-1} + P_{n-2}$ niz Pellovih brojeva. Zato desnu granu stabla Markovljevih trojki zovemo *Pellova grana*.

Posebno smo dobili da su Fibonaccijevi i Pellovi brojevi s neparnim indeksom nužno Markovljevi brojevi. Primijetimo da je

$$\frac{F_{n+1}}{F_n} = [\underbrace{1, 1, \ldots, 1}_{n}] \to \frac{1+\sqrt{5}}{2}, \quad \frac{P_{n+1}}{P_n} = [\underbrace{2, 2, \ldots, 2}_{n}] \to 1+\sqrt{2} \quad \text{kad } n \to +\infty.$$



Fibonaccijeva i Pellova grana (s istim početnim čvorom) u stablu Markovljevih trojki

Prisjetimo se da smo u Teoremu 4.69 pokazali da je za svaku pravu donju Christoffelovu riječ w matrica $\mu(w) = \left[\begin{smallmatrix} 2m+u & m \\ 2m-u-v & m-u \end{smallmatrix}\right]$, a u Lemi 4.72 smo vidjeli da uz standardnu faktorizaciju $w = w'w''$ imamo $2 \le I_{w'} < I_w < I_{w''} \le \frac{5}{2}$, gdje je $I_w = \frac{\mu(w)_{11}}{\mu(w)_{12}}$ indeks od w. Pišemo I_t za I_{w_t} ako je t pozitivan racionalan broj ili $\frac{0}{1}$ ili $\frac{1}{0}$. Usporedbom navedenog rezultata i pravila konstrukcije stabla Christoffelovih parova s maloprije pokazanim rezultatom o uređaju na čvorovima Stern–Brocotova stabla, lako se pokazuje da za $t_1, t_2 \in \mathbb{Q}$ iz $t_1 < t_2$ slijedi $I_{t_1} < I_{t_2}$. Dovoljno je uspeti se od t_1 i t_2 do prvog zajedničkog pretka t u Stern–Brocotovom stablu i onda provjeriti da je $I_{t_1} \le I_t \le I_{t_2}$, pri čemu $I_{t_1} = I_t$ vrijedi samo ako je $t_1 = t$ i slično za t i t_2.

Teorem 5.9. *Indeks je strogo rastuća funkcija na $\mathbb{Q}_{>0} \cup \{\frac{0}{1}, \frac{1}{0}\}$.*

Korolar 5.10. *Matrice $\mu(w_t)$ za $t \in \mathbb{Q}_{>0} \cup \{\frac{0}{1}, \frac{1}{0}\}$ su sve različite.*

Slutnja o jedinstvenosti kaže da vrijedi i više, tj. ne samo da su matrice $\mu(w_t)$ različite, nego su im i tragovi različiti.

Imamo za $t \in \mathbb{Q}_{>0}$

$$\mu(w_t) = \begin{bmatrix} 2m_t + u_t & m_t \\ 2m_t - u_t - v_t & m_t - u_t \end{bmatrix}, \qquad 2 < I_t = 2 + \frac{u_t}{m_t} < 2 + \frac{1}{2}. \tag{5.5}$$

Korolar 5.4 kaže da ako $x^2 \equiv -1 \pmod{m_t}$ ima jedinstveno rješenje u $0 < x < \frac{m_t}{2}$, onda je m_t jedincat. Tu tvrdnju vidimo sada direktno iz oblika matrice u (5.5). Naime, ako m_t jednoznačno određuje rješenje u_t, onda je zbog $u_t^2 + 1 = m_t v_t$ matrica $\mu(w_t)$ potpuno određena, a budući da su ove matrice različite, m_t je jedincat.

Sljedeću lemu za Fibonaccijevu i Pellovu granu lako je dokazati indukcijom koristeći jednakosti $\mu(w_{\frac{1}{n}}) = \mu(a)\mu(w_{\frac{1}{n-1}})$ i $\mu(w_{\frac{n}{1}}) = \mu(w_{\frac{n-1}{1}})\mu(b)$.

Lema 5.11. *Za $n \in \mathbb{N}_0$ vrijedi*

$$\mu(w_{\frac{1}{n}}) = \mu(a^n b) = \begin{bmatrix} 3F_{2n+3} - F_{2n+2} & F_{2n+3} \\ 3F_{2n+2} - F_{2n+1} & F_{2n+2} \end{bmatrix}, \qquad I_{\frac{1}{n}} = 3 - \frac{F_{2n+2}}{F_{2n+3}},$$
$$\mu(w_{\frac{n}{1}}) = \mu(ab^n) = \begin{bmatrix} P_{2n+2} & P_{2n+1} \\ P_{2n+1} + P_{2n} & P_{2n+1} - P_{2n} \end{bmatrix}, \qquad I_{\frac{n}{1}} = \frac{P_{2n+2}}{P_{2n+1}}.$$

Iz činjenica o parnim i neparnim konvergentama ili iz Teorema 5.9, imamo da je niz $(I_{\frac{1}{n}})_{n\in\mathbb{N}_0}$ monotono padajući, a niz $(I_{\frac{n}{1}})_{n\in\mathbb{N}_0}$ monotono rastući i vrijedi

$$I_{\frac{1}{n}} = 3 - \frac{F_{2n+2}}{F_{2n+3}} \to 3 - \frac{2}{1+\sqrt{5}} = \frac{7-\sqrt{5}}{2} = 2.381\ldots, \qquad I_{\frac{n}{1}} = \frac{P_{2n+2}}{P_{2n+1}} \to 1 + \sqrt{2} = 2.414\ldots.$$

Teorem 5.9 povlači da je za svaki $t \in \mathbb{Q}_{>0}$ indeks $I_t = 2 + \frac{u_t}{m_t}$ u puno manjem intervalu $\left(\frac{7-\sqrt{5}}{2}, 1+\sqrt{2}\right)$ od intervala $(2, \frac{5}{2})$ što smo dosad znali. Time je sužen i interval u kojemu kongruencija $x^2 \equiv -1 \pmod{m}$ mora imati rješenje ako je m Markovljev broj.

Korolar 5.12. *Za pozitivan racionalan broj t je*

$$0.381\, m_t < \frac{3-\sqrt{5}}{2} m_t < u_t < (\sqrt{2}-1) m_t < 0.415\, m_t. \tag{5.6}$$

Primjer. Pogledajmo Markovljev broj $m_{\frac{2}{5}} = 1325$. Kongruencija $x^2 \equiv -1 \pmod{1325}$ ima dva rješenja $x_1 = 182$ i $x_2 = 507$ u intervalu $0 < x < \frac{1325}{2}$, pa Korolar 5.4 nije dovoljno dobar da bi jamčio jedincatost. No, budući da samo $x_2 = 507$ leži unutar ograda iz (5.6), zaključujemo da je 1325 jedincat.

Korolar 5.13. *Neka je (m_r, m_t, m_s) Markovljeva trojka takva da je $w_t = w_r w_s$ za $r, t, s \in \mathbb{Q}_{>0}$. Tada je*

$$\frac{11 - \sqrt{5} - \sqrt{8}}{2} m_r m_s < m_t < 3 m_r m_s, \tag{5.7}$$

gdje je $\frac{11-\sqrt{5}-\sqrt{8}}{2} = 2.967\ldots$. Za $r = \frac{0}{1}$ imamo trojke $(1, F_{2n+3}, F_{2n+1})$ kod kojih je $F_{2n+3} \geq \frac{5}{2} F_{2n+1}$, a za $s = \frac{1}{0}$ trojke $(P_{2n-1}, P_{2n+1}, 2)$ kod kojih je $P_{2n+1} \geq 5 P_{2n-1}$.

Dokaz. Zbog pravila konstrukcije Markovljevih trojki (usporedi str. 60) znamo da je čvor (m_r, m_t, m_s) dijete čvora kojemu pripadna trojka u neuređenom obliku izgleda $\{3m_r m_s - m_t, m_r, m_s\}$. Stoga je $3m_r m_s - m_t > 0$, odnosno vrijedi gornja ograda u (5.7).

Iz $\mu(w_t) = \mu(w_r)\mu(w_s)$ imamo prema (5.5)

$$\begin{aligned} m_t &= (2m_r + u_r)m_s + m_r(m_s - u_s) \\ &> 3m_r m_s + \frac{3-\sqrt{5}}{2} m_r m_s - (\sqrt{2} - 1) m_r m_s = \frac{11 - \sqrt{5} - \sqrt{8}}{2} m_r m_s, \end{aligned}$$

pri čemu smo koristili ograde $u_r > \frac{3-\sqrt{5}}{2} m_r$ i $u_s < (\sqrt{2} - 1) m_s$ iz (5.6).

Zadnja tvrdnja slijedi iz

$$\frac{F_{2n+3}}{F_{2n+1}} = 1 + \frac{F_{2n+2}}{F_{2n+1}} \geq 1 + \frac{F_4}{F_3} = \frac{5}{2} \quad \text{i} \quad \frac{P_{2n+1}}{P_{2n-1}} = 1 + \frac{2P_{2n}}{P_{2n-1}} \geq 1 + \frac{2P_2}{P_1} = 5. \qquad \square$$

Sada ćemo se pozabaviti rastom niza Markovljevih brojeva. Prvo pitanje na koje ćemo odgovoriti je koliko se duboko moramo spustiti u stablu Markovljevih trojki da bismo bili sigurni kako smo pokrili sve Markovljeve brojeve manje od nekog zadanog broja.

Propozicija 5.14. *Među svim Markovljevim brojevima generiranim u n-tom retku stabla Markovljevih trojki, najmanji je $m_{\frac{1}{n}} = F_{2n+3}$.*

Dokaz. U drugom retku stabla pojavljuju se novi brojevi $m_{\frac{1}{2}} = F_7 = 13$ i $m_{\frac{2}{1}} = P_5 = 29$, a u trećem retku brojevi

$$m_{\frac{1}{3}} = F_9 = 34 < m_{\frac{3}{1}} = P_7 = 169 < m_{\frac{2}{3}} = 194 < m_{\frac{3}{2}} = 433.$$

Tvrdnja, dakle, vrijedi za $n = 2, 3$, pa pretpostavimo da vrijedi za neki n. Prijeđemo li sa n na $n+1$, vidimo da je $F_{2n+5} = 3F_{2n+3} - F_{2n+1}$. Neka je m' neki Markovljev broj koji se prvi put pojavljuje u $n+1$-om retku i generiran je Markovljevom trojkom (a, m, b) ili (b, m, a), tako da je $m' = 3am - b$, ali m' nije u Fibonaccijevoj grani, tj. $a \geq 2$. Tada je po pretpostavci indukcije

$$3am - 3F_{2n+3} \geq 3am - 3m = 3(a-1)m \geq 3m > m > b - F_{2n+1}.$$

Stoga je $m' = 3am - b > 3F_{2n+3} - F_{2n+1} = F_{2n+5}$. $\square$

Može se pokazati da je od Markovljevih brojeva generiranih u n-tom retku stabla idući po veličini upravo P_{2n+1}.

Korolar 5.15. *Svi Markovljevi brojevi $m \leq 10^N$ pojavljuju se u prvih n redaka Markovljevog stabla, gdje je*

$$n = \left\lceil \frac{N - \log_{10} \frac{11+5\sqrt{5}}{2\sqrt{5}}}{\log_{10} \frac{3+\sqrt{5}}{2}} \right\rceil \approx 2.392\, N. \tag{5.8}$$

Dokaz. Prema prethodnoj propoziciji, najmanji novi broj u $n+1$-om retku je F_{2n+5}, pa je n najmanji broj takav da je $F_{2n+5} > 10^N$. Pomoću Binetove formule

$$F_k = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k\right), \tag{5.9}$$

lako je pokazati da vrijedi

$$\frac{11+5\sqrt{5}}{2\sqrt{5}}\left(\frac{3+\sqrt{5}}{2}\right)^n < F_{2n+5} < \frac{11+5\sqrt{5}}{2\sqrt{5}}\left(\frac{3+\sqrt{5}}{2}\right)^n + 0.02. \tag{5.10}$$

Usporedbom sa zahtijevanom nejednakošću, dobivamo (5.8). □

Želimo sada ocijeniti koliko brzo raste niz Markovljevih brojeva, tj. zanima nas asimptotsko ponašanje funkcije

$$g(x) = |\{m \in \mathbb{M} \,:\, m \leq x\}|.$$

koja daje kardinalitet odgovarajućeg skupa. Naći ćemo ocjenu za funkciju

$$\bar{g}(x) = |\{t \in \mathbb{Q} \,:\, m_t \leq x\}|,$$

gdje smo uključili i $m_{\frac{0}{1}} = 1$ te $m_{\frac{1}{0}} = 2$.

Jasno da vrijedi $g(x) \leq \bar{g}(x)$, a jednakost $g(x) = \bar{g}(x)$ za sve x je samo još jedna varijanta iskaza slutnje o jedinstvenosti.

Lema 5.16. *Za svaki $t \in \mathbb{Q}_{>0} \cup \{\frac{0}{1}, \frac{1}{0}\}$ je*

$$\frac{\ln(3m_t)}{1.8} < |w_t| < \frac{\ln(\rho\, m_t)}{0.9}, \tag{5.11}$$

gdje je $\rho = \frac{11-\sqrt{5}-\sqrt{8}}{2}$.

Dokaz. Nejednakosti je lako provjeriti za početnu Markovljevu trojku $(1, 5, 2)$. Za induktivni korak koristimo nejednakost (5.7) iz Korolara 5.13. Neka je (m_r, m_t, m_s) Markovljeva trojka za koju je $r \neq \frac{0}{1}$ i $s \neq \frac{1}{0}$. Prema (5.7) je $3m_t < (3m_r)(3m_s)$ i $\rho m_t > (\rho m_r)(\rho m_s)$. To nam induktivno daje

$$\frac{\ln(3m_t)}{1.8} < \frac{\ln(3m_r) + \ln(3m_s)}{1.8} < |w_r| + |w_s| = |w_t|,$$
$$\frac{\ln(\rho m_t)}{0.9} > \frac{\ln(\rho m_r) + \ln(\rho m_s)}{0.9} > |w_r| + |w_s| = |w_t|.$$

Treba još pokazati da (5.11) vrijedi za Fibonaccijeve brojeve $m_{\frac{1}{n}} = F_{2n+3}$ i Pellove brojeve $m_{\frac{n}{1}} = P_{2n+1}$ ako je $n \geq 2$. No, direktno iz rekurzije za te brojeve je

$$\left| F_{2n+3} - \frac{2+\sqrt{5}}{\sqrt{5}}\left(\frac{3+\sqrt{5}}{2}\right)^n \right| < 0.5 \qquad \text{i} \qquad \left| P_{2n+1} - \frac{1+\sqrt{2}}{\sqrt{8}}(3+\sqrt{8})^n \right| < 0.5,$$

pa treba samo provjeriti

$$\frac{\ln\left(\frac{3(1+\sqrt{2})}{\sqrt{8}}\right) + n\ln(3+\sqrt{8})}{1.8} < n+1 < \frac{\ln\left(\frac{\rho(2+\sqrt{5})}{\sqrt{5}}\right) + n\ln\left(\frac{3+\sqrt{5}}{2}\right)}{0.9},$$

što nije teško vidjeti. □

Definiramo $d(x) = |\{t : |w_t| \leq x\}|$ za $x > 0$.

Propozicija 5.17. *Za $x \geq 1$ je*

$$d\Big(\frac{\ln(3x)}{1.8}\Big) \leq \bar{g}(x) \leq d\Big(\frac{\ln(\rho x)}{0.9}\Big).$$

Dokaz. Ako je $t \in \mathbb{Q}_{>0} \cup \{\frac{0}{1}, \frac{1}{0}\}$ i vrijedi $m_t \leq x$, onda je po prethodnoj lemi

$$|w_t| < \frac{\ln(\rho m_t)}{0.9} \leq \frac{\ln(\rho x)}{0.9},$$

pa za odgovarajuće kardinalitete vrijedi $\bar{g}(x) \leq d\big(\frac{\ln(\rho x)}{0.9}\big)$, a analogno se pokazuje donja ograda. $\square$

Donjih Christoffelovih riječi w duljine $|w| = n \geq 2$ ima koliko i uređenih parova $(|w|_a, |w|_b)$ relativno prostih prirodnih brojeva za koje je $|w|_a + |w|_b = n$, tj. upravo $\varphi(n)$, gdje je φ Eulerova funkcija. Postoje dvije riječi duljine 1, pa je $d(x) = 1 + \sum_{k \leq x} \varphi(k)$ za $x \geq 1$. Poznata je asimptotska ocjena ove sume (vidi primjerice Dujellinu skriptu iz teorije brojeva). Imamo

$$d(x) = \frac{3}{\pi^2} x^2 + \mathcal{O}(x \ln x).$$

Sada iz Propozicije 5.17 dobivamo sljedeći rezultat.

Teorem 5.18. *Vrijedi*

$$C_1 \leq \liminf_{n\to\infty} \frac{\bar{g}(n)}{(\ln n)^2} \leq \limsup_{n\to\infty} \frac{\bar{g}(n)}{(\ln n)^2} \leq C_2, \tag{5.12}$$

gdje možemo uzeti $C_1 = \frac{3}{\pi^2 (1.8)^2} = 0.093\ldots$ i $C_2 = \frac{3}{\pi^2(0.9)^2} = 0.375\ldots$.

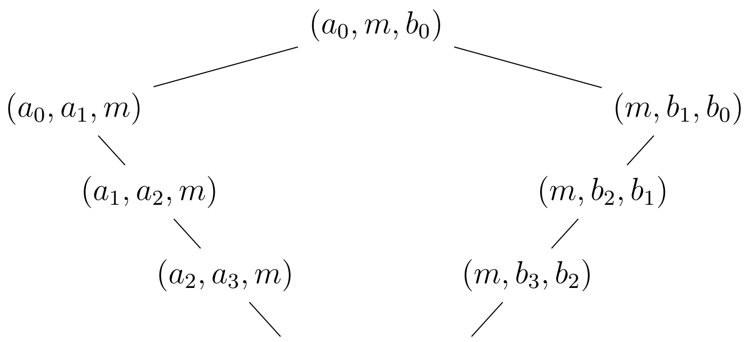
U prethodnom teoremu dobivamo znatno bolje konstante C_1 i C_2 ako u (5.11) umjesto $|w_t| = |w_t|_a + |w_t|_b$ promatramo $|w_t|_a + 2|w_t|_b$.

Zagier je dokazao da limes u (5.12) postoji i jednak je $\lim_{n\to\infty} \frac{\bar{g}(n)}{(\ln n)^2} = 0.180\ldots$. On je umjesto funkcija $\ln(3x)$, odnosno $\ln(\rho x)$ iz (5.11), koristio funkciju $\ln \frac{3x+\sqrt{9x^2-4}}{2} = \operatorname{Arch} \frac{3x}{2}$ i onda gledao podstabla stabla Markovljevih trojki koja počinju dovoljno daleko od korijena. Naime, naša indukcija za ograde u Lemi 5.16 zasniva se na tome da jednadžbu $x^2 + y^2 + z^2 = 3xyz$ (za $x \geq y, z$) zamijenimo prilično lošom aproksimacijom $x^2 = 3xyz$, odnosno $3x = (3y)(3z)$ ili $\ln(3x) = \ln(3y) + \ln(3z)$. No za funkciju $f(x) = \ln \frac{3x+\sqrt{9x^2-4}}{2}$ je jednakost $f(x) = f(y) + f(z)$ ekvivalentna $x^2 + y^2 + z^2 = 3xyz + \frac{4}{9}$ što znači da uz pomoć funkcije f možemo dobiti puno jače nejednakosti nego prije.

5.3 Jedinstvenost u dijelovima stabla

Slutnja o jedinstvenosti kaže da su svi centralni brojevi u stablu Markovljevih trojki različiti. Prirodan način da se približimo dokazu slutnje je promatrati samo dio toga stabla te pokazati da u tom podgrafu tvrdnja vrijedi. Promatrat ćemo dva tipa podgrafa.

Prvi se za proizvoljni fiksirani $t \in \mathbb{Q}_{>0}$ sastoji od svih potomaka trojke (a_0, m, b_0) koji sadrže $m = m_t$ kao jednu od koordinata. Evo kako taj podgraf izgleda.



Tako dobivamo obostrano beskonačan niz $\mathcal{S}_t$ centralnih elemenata u Markovljevim trojkama koji raste u oba smjera,

$$\mathcal{S}_t : \quad \ldots > a_3 > a_2 > a_1 > m_t < b_1 < b_2 < b_3 < \ldots .$$

Primjerice, niz $\mathcal{S}_{\frac{1}{1}}$ ima članove

$$\ldots > 2897 > 194 > 13 > 5 < 29 < 433 < 6466 < \ldots .$$

Sljedeći rezultat pokazuje da svaki niz $\mathcal{S}_t$ ima svojstvo jedincatosti.

Propozicija 5.19. *Markovljevi brojevi u nizu $\mathcal{S}_t : \ldots, a_2, a_1, m, b_1, b_2, \ldots$ su svi različiti. Preciznije, ako je $a_0 < b_0$ u trojci (a_0, m, b_0), onda grane alterniraju*

$$m < a_1 < b_1 < a_2 < b_2 < a_3 < b_3 < \ldots .$$

Dokaz. Iz rekurzije kojom se dobiva stablo Markovljevih trojki imamo

$$\begin{aligned} a_1 &= (3m)a_0 - b_0, \qquad & a_n &= (3m)a_{n-1} - a_{n-2} \quad \text{za} \quad n \geq 2, \\ b_1 &= (3m)b_0 - a_0, \qquad & b_n &= (3m)b_{n-1} - b_{n-2} \quad \text{za} \quad n \geq 2. \end{aligned} \tag{5.13}$$

Standardne metode rješavanja homogenih linearnih rekurzija s konstantnim koeficijentima, npr. koristeći funkcije izvodnice, daju

$$a_n = \delta_a \left(\frac{3m + \sqrt{9m^2 - 4}}{2} \right)^n + \bar{\delta}_a \left(\frac{3m - \sqrt{9m^2 - 4}}{2} \right)^n, \tag{5.14}$$

gdje je

$$\delta_a = \frac{a_0(3m + \sqrt{9m^2 - 4}) - 2b_0}{2\sqrt{9m^2 - 4}}, \qquad \bar{\delta}_a = \frac{a_0(-3m + \sqrt{9m^2 - 4}) + 2b_0}{2\sqrt{9m^2 - 4}},$$

te

$$b_n = \delta_b \left(\frac{3m + \sqrt{9m^2 - 4}}{2} \right)^n + \bar{\delta}_b \left(\frac{3m - \sqrt{9m^2 - 4}}{2} \right)^n, \tag{5.15}$$

gdje je

$$\delta_b = \frac{b_0(3m + \sqrt{9m^2 - 4}) - 2a_0}{2\sqrt{9m^2 - 4}}, \qquad \bar{\delta}_b = \frac{b_0(-3m + \sqrt{9m^2 - 4}) + 2a_0}{2\sqrt{9m^2 - 4}}.$$

Napomenimo da konjugate kvadratnih iracionalnosti u ovom potpoglavlju označavamo sa $\bar{\cdot}$, a ne $\cdot'$ iz tipografskih razloga.

Budući da je $a_0, b_0 < \frac{m}{2}$ po (5.3) i $m < \frac{1}{2}\sqrt{9m^2-4}$, za drugi član u (5.14) dobivamo

$$\frac{3m-\sqrt{9m^2-4}}{2} \leq \frac{15-\sqrt{221}}{2} < 0.067, \quad |\bar{\delta}_a| < \frac{0.067\, a_0}{\sqrt{9m^2-4}} + \frac{b_0}{\sqrt{9m^2-4}} < 0.27,$$

pa zaključujemo da je a_n cijeli broj najbliži $\delta_a\left(\frac{3m+\sqrt{9m^2-4}}{2}\right)^n$. Slično je b_n cijeli broj najbliži $\delta_b\left(\frac{3m+\sqrt{9m^2-4}}{2}\right)^n$.

Bez smanjenja općenitosti možemo pretpostaviti da je $a_0 < b_0$. Pokažimo da je tada $a_n < b_n < a_{n+1}$ iz čega odmah slijedi jedincatost za $\mathcal{S}_t$. Po (5.14) i (5.15) je

$$\frac{a_{n+1}}{a_n} = \frac{3m+\sqrt{9m^2-4}}{2}\big(1+\varepsilon_1(n)\big), \tag{5.16}$$

$$\frac{b_n}{a_n} = \frac{b_0(3m+\sqrt{9m^2-4})-2a_0}{a_0(3m+\sqrt{9m^2-4})-2b_0}\big(1+\varepsilon_2(n)\big), \tag{5.17}$$

gdje su za $n \geq 1$ greške $\varepsilon_1(n)$ i $\varepsilon_2(n)$ po apsolutnoj vrijednosti manje od 0.03. Kako je $a_0 < b_0$, imamo $\frac{b_n}{a_n} > 1$ i direktan račun pokazuje da je desna strana u (5.16) više nego dvostruko veća od desne strane u (5.17). Tako je dokazano $\frac{a_{n+1}}{a_n} > \frac{b_n}{a_n} > 1$, pa time i propozicija. $\square$

Na kraju promatramo podgraf koji se sastoji od dviju vanjskih grana koje izlaze iz čvora (a, m_t, b) kao na slici.

$$\begin{array}{ccccccccc}
 & & & & (a,m_t,b) & & & & \\
 & & & \swarrow & & \searrow & & & \\
 & & & (a,u_1,m_t) & & (m_t,v_1,b) & & & \\
 & & \swarrow & & & & \searrow & & \\
 & & (a,u_2,u_1) & & & & (v_1,v_2,b) & & \\
 & \swarrow & & & & & & \searrow & \\
 & (a,u_3,u_2) & & & & & & (v_2,v_3,b) & \\
 \swarrow & & & & & & & & \searrow
\end{array}$$

Niz centralnih Markovljevih brojeva u lijevoj grani je $\{m_t = u_0 < u_1 < u_2 < \dots\}$, a u desnoj grani je $\{m_t = v_0 < v_1 < v_2 < \dots\}$. Tako dobivamo još jedan obostrano beskonačan niz

$$\mathcal{R}_t : \quad \dots > u_3 > u_2 > u_1 > m_t < v_1 < v_2 < v_3 < \dots,$$

pa se kao u Propoziciji 5.19 postavlja pitanje ima li $\mathcal{R}_t$ svojstvo jedincatosti te kakav je poredak elemenata u $\mathcal{R}_t$.

Slijedeći rad Bugeauda, Reutenauera i Sikseka, odgovorit ćemo na ovo pitanje za $t = \frac{1}{1}$, tj. u slučaju kada gledamo Fibonaccijevu i Pellovu granu u stablu Markovljevih trojki.

Koristit ćemo Bakerovu teoriju linearnih formi u logaritmima, ali prije toga nam trebaju još neke leme iz diofantskih aproksimacija. Prvu lemu dajemo u verziji kako su je formulirali Dujella i Pethő.

Lema 5.20 (Baker–Davenport)**.** *Neka je M prirodan broj te κ, μ, A, B realni brojevi takvi da je $\kappa > 0$, $A > 0$, $B > 1$. Neka su p, q prirodni brojevi takvi da vrijedi*

$$|q\kappa - p| \leq \alpha, \tag{5.18}$$
$$\|\mu q\| - M\alpha = \varepsilon > 0, \tag{5.19}$$

za neki realan broj $\alpha > 0$, pri čemu je $\|x\|$ udaljenost od x do njemu najbližeg cijelog broja. Tada nejednadžba

$$0 < m\kappa - n + \mu < AB^{-m} \tag{5.20}$$

nema rješenja u cijelim brojevima m i n takvima da je

$$\frac{\ln\left(\frac{Aq}{\varepsilon}\right)}{\ln B} \leq m \leq M.$$

Dokaz. Pretpostavimo da vrijedi (5.20) za $0 \leq m \leq M$. Množenjem sa q i pregrupiranjem članova, dobivamo

$$0 < \big((mp - nq) + \mu q\big) + m(q\kappa - p) < qAB^{-m}.$$

Stoga je

$$\begin{aligned} qAB^{-m} &> \left|\big((mp - nq) + \mu q\big) + m(q\kappa - p)\right| \\ &\geq |(mp - nq) + \mu q| - m|q\kappa - p| \\ &\geq \|\mu q\| - m\alpha \geq \|\mu q\| - M\alpha = \varepsilon > 0, \end{aligned}$$

pa je $\ln(qA) - m \ln B > \ln \varepsilon$ što povlači $m < \frac{\ln(qA/\varepsilon)}{\ln B}$. $\square$

Neka su sada K_1, K_2 realna kvadratna polja i φ_1, φ_2 redom njihove fundamentalne jedinice izabrane tako da su veće od 1. Dakle, ako je $K_i = \mathbb{Q}(\sqrt{d_i})$, onda $2\varphi_i$ odgovara minimalnom rješenju Pellove jednadžbe $x^2 - d_i y^2 = \pm 4$ u prirodnim brojevima.

Neka su $\delta_i \in K_i$, $\delta_i > 0$ te

$$u_m = \delta_1 \varphi_1^m + \bar{\delta}_1 \overline{\varphi}_1^m, \qquad v_n = \delta_2 \varphi_2^n + \bar{\delta}_2 \overline{\varphi}_2^n,$$

gdje je $\bar{\cdot}$ konjugiranje u K_i. Koristit ćemo oznake

$$\delta = \min\{\delta_1, \delta_2\}, \quad \delta' = |\bar{\delta}_1| + |\bar{\delta}_2|, \quad \varphi = \min\{\varphi_1, \varphi_2\}.$$

Želimo riješiti jednadžbu $u_m = v_n$ za $m, n \in \mathbb{N}_0$ te stoga pretpostavljamo da ta jednakost u nastavku vrijedi. Zato imamo

$$|\delta_1 \varphi_1^m - \delta_2 \varphi_2^n| = |\bar{\delta}_1 \overline{\varphi}_1^m - \bar{\delta}_2 \overline{\varphi}_2^n| \leq |\bar{\delta}_1| + |\bar{\delta}_2| = \delta' \tag{5.21}$$

jer je $|\overline{\varphi}_1|, |\overline{\varphi}_2| < 1$ budući je norma od φ_i jednaka $\varphi_i \overline{\varphi}_i = \pm 1$ i po pretpostavci je $\varphi_i > 1$.

Lema 5.21. *Ako je $\delta_1 \varphi_1^m \geq \frac{3}{2} \delta_2 \varphi_2^n$, onda je $m \leq \frac{\ln(3\delta'/\delta)}{\ln \varphi}$, a ako je $\delta_2 \varphi_2^n \geq \frac{3}{2} \delta_1 \varphi_1^m$, onda je $n \leq \frac{\ln(3\delta'/\delta)}{\ln \varphi}$.*

Dokaz. Pretpostavimo da je $\delta_1 \varphi_1^m \geq \frac{3}{2} \delta_2 \varphi_2^n$. Tada (5.21) daje $\frac{3}{2} \delta_2 \varphi_2^n \leq \delta_1 \varphi_1^m \leq \delta_2 \varphi_2^n + \delta'$, odnosno $\frac{1}{2} \delta_2 \varphi_2^n \leq \delta'$. Nadalje, $\delta_1 \varphi_1^m = \delta_2 \varphi_2^n + (\delta_1 \varphi_1^m - \delta_2 \varphi_2^n) \leq 3\delta'$. Zaključujemo da je $m \ln \varphi_1 + \ln \delta_1 \leq \ln(3\delta')$ te slijedi ograda na m jer je $\varphi_1 \geq \varphi > 1$ i $\delta_1 \geq \delta$. Drugi slučaj pokazuje se analogno. $\square$

Lema 5.22. *Ako je*

$$\delta_2\varphi_2^n < \delta_1\varphi_1^m < \frac{3}{2}\delta_2\varphi_2^n, \tag{5.22}$$

onda m i n zadovoljavaju nejednakost (5.20) *uz*

$$\kappa = \frac{\ln\varphi_1}{\ln\varphi_2}, \qquad \mu = \frac{\ln(\delta_1/\delta_2)}{\ln\varphi_2}, \qquad A = \frac{3\delta'}{2\delta_1\ln\varphi_2}, \qquad B = \varphi_1.$$

Dokaz. Iz (5.22) i (5.21) dobivamo

$$0 < \frac{\delta_1\varphi_1^m}{\delta_2\varphi_2^n} - 1 \leq \frac{\delta'}{\delta_2\varphi_2^n}.$$

Za $x > 0$ je $\ln(1+x) < x$, pa vrijedi

$$0 < \ln\left(\frac{\delta_1\varphi_1^m}{\delta_2\varphi_2^n}\right) = \ln\left(1 + \frac{\delta_1\varphi_1^m}{\delta_2\varphi_2^n} - 1\right) < \frac{\delta_1\varphi_1^m}{\delta_2\varphi_2^n} - 1 \leq \frac{\delta'}{\delta_2\varphi_2^n} < \frac{3\delta'}{2\delta_1\varphi_1^m},$$

gdje zadnja nejednakost slijedi iz uvjeta leme. Sada dobivamo

$$0 < m\ln\varphi_1 - n\ln\varphi_2 + \ln\left(\frac{\delta_1}{\delta_2}\right) < \frac{3\delta'}{2\delta_1\varphi_1^m},$$

odnosno nakon dijeljenja sa $\ln\varphi_2$, traženu nejednakost

$$0 < m\kappa - n + \mu < AB^{-m}. \qquad \square$$

Da bismo iskazali rezultat o linearnim formama u logaritmima koji nam treba, moramo najprije uvesti sljedeće oznake. Neka je $\mathbb{L}$ polje algebarskih brojeva stupnja D, neka su $\alpha_1, \ldots, \alpha_k$ nenul elementi od $\mathbb{L}$ i $b_1, \ldots, b_k$ racionalni cijeli brojevi. Stavimo $B' = \max\{|b_1|, \ldots, |b_k|\}$ i

$$\Lambda = \alpha_1^{b_1} \cdots \alpha_k^{b_k} - 1.$$

Za algebarski broj α čiji je minimalni polinom nad $\mathbb{Z}$ oblika $P(X) = a\prod_{i=1}^d (X - \alpha^{(i)})$, pišemo $h(\alpha)$ za njegovu *apsolutnu logaritamsku visinu*

$$h(\alpha) = \frac{1}{d}\Big(\ln|a| + \sum_{i=1}^{d} \max\{\ln|\alpha^{(i)}|, 0\}\Big).$$

Neka su $A_1, \ldots, A_k$ realni brojevi takvi da je $A_j \geq \max\{Dh(\alpha_j), |\ln\alpha_j|, 0.16\}$ za $j \in \{1, \ldots, k\}$.

Teorem 5.23 (Matveev)**.** *Ako je $\Lambda \neq 0$ i $\mathbb{L}$ realno polje, onda vrijedi*

$$\ln|\Lambda| > -1.4 \cdot 30^{k+3} k^{4.5} D^2 A_1 \cdots A_k (1 + \ln D)(1 + \ln B').$$

U prethodnom teoremu ključan je faktor $1 + \ln B'$.

Pokažimo sada da Fibonaccijeva i Pellova grana nemaju drugih zajedničkih elemenata osim 1, 2 i 5.

Teorem 5.24. *Za $k, \ell \geq 2$, Markovljevi brojevi $m_{\frac{1}{k}}$ i $m_{\frac{\ell}{1}}$ su različiti.*

Dokaz. Koristeći rekurzije za Fibonaccijeve i Pellove brojeve, dobiva se

$$u_m = F_{2m+3} = \delta_1\varphi_1^m + \bar{\delta}_1\overline{\varphi}_1^m, \qquad v_n = P_{2n+1} = \delta_2\varphi_2^n + \bar{\delta}_2\overline{\varphi}_2^n, \tag{5.23}$$

gdje je

$$\begin{aligned} \varphi_1 &= \frac{3+\sqrt{5}}{2} = \Big(\frac{1+\sqrt{5}}{2}\Big)^2, \qquad & \varphi_2 &= 3+2\sqrt{2} = (1+\sqrt{2})^2, \\ \delta_1 &= \frac{25+11\sqrt{5}}{10}, & \delta_2 &= \frac{10+7\sqrt{2}}{4}. \end{aligned} \tag{5.24}$$

Približne vrijednosti ovih brojeva su

$$\begin{aligned} \varphi_1 &= 2.6180, \quad & \overline{\varphi}_1 &= 0.3819, \quad & \varphi_2 &= 5.8284, \quad & \overline{\varphi}_2 &= 0.1715, \\ \delta_1 &= 4.9596, & \bar{\delta}_1 &= 0.0403, & \delta_2 &= 4.9748, & \bar{\delta}_2 &= 0.0251. \end{aligned}$$

Neka je $\Lambda = \frac{\delta_2\varphi_2^n}{\delta_1\varphi_1^m} - 1$. Pretpostavimo da je $u_m = v_n$ za $(m,n) \neq (0,0)$. Iz te jednakosti slijedi

$$\frac{\delta_2\varphi_2^n}{\delta_1\varphi_1^m} - 1 = \frac{\bar{\delta}_1\overline{\varphi}_1^m - \bar{\delta}_2\overline{\varphi}_2^n}{\delta_1\varphi_1^m}.$$

Kako su $\bar{\delta}_1, \overline{\varphi}_1, \bar{\delta}_2, \overline{\varphi}_2$ iz intervala $(0,1)$, a $\delta_1 > 2$, dobivamo $|\Lambda| < \varphi_1^{-m}$, odnosno

$$\ln|\Lambda| \leq -m\ln\varphi_1.$$

Sada primijenimo Teorem 5.23 kako bismo omeđili $|\Lambda|$ odozdo. Imamo

$$k = 3, \quad \mathbb{L} = \mathbb{Q}[\sqrt{2},\sqrt{5}], \quad D = 4,$$

$$\alpha_1 = \frac{\delta_2}{\delta_1}, \quad \alpha_2 = \varphi_2, \quad \alpha_3 = \varphi_1, \quad b_1 = 1, \quad b_2 = n, \quad b_3 = -m.$$

Prema maloprije pokazanom je $\frac{\delta_2\varphi_2^n}{\delta_1\varphi_1^m} < 2$, pa numerički račun uz upotrebu $\varphi_2 > \varphi_1$ povlači $n \leq m$. Stoga za naš izbor imamo $B' = \max\{|b_1|,|b_2|,|b_3|\} = m$. Iz Teorema 5.23 slijedi

$$\ln|\Lambda| > -1.4\cdot 30^6\cdot 3^{4.5}\cdot 16\cdot A_1A_2A_3(1+\ln 4)(1+\ln m)$$

uz $A_1 = \ln(320\delta_2^2) \leq 9$, $A_2 = 2\ln\varphi_2 \leq 3.6$, $A_3 = 2\ln\varphi_1 \leq 2$.

Kombinirajući s gornjom ogradom za $\ln|\Lambda|$ koju smo dobili, proizlazi

$$-0.96m > -3.55\cdot 10^{14}(1+\ln m),$$

a ta nejednakost ne vrijedi za $m > 10^{20}$. Tako smo koristeći Bakerovu teoriju pokazali da je $m \leq 10^{20}$, a sada ćemo pomoću Baker–Davenportove redukcije znatno smanjiti ovu gornju ogradu.

Neka je $K_1 = \mathbb{Q}[\sqrt{5}]$, $K_2 = \mathbb{Q}[\sqrt{2}]$, pa su φ_1, φ_2 redom jedinice u K_1, K_2. Ako vrijedi jedan od dva disjunktna uvjeta u Lemi 5.21, onda mora biti m ili $n \leq \frac{\ln(3\delta'/\delta)}{\ln\varphi}$, ali lako se izračuna da je ovaj broj ≤ -3 što nije moguće. Zato je $\delta_1\varphi_1^m < \frac{3}{2}\delta_2\varphi_2^n$ i $\delta_2\varphi_2^n < \frac{3}{2}\delta_1\varphi_1^m$.

Primijetite da ne možemo imati $\Lambda = 0$, tj. $\delta_1\varphi_1^m = \delta_2\varphi_2^n$ jer bi ta jednakost povlačila jednakost oblika $a + b\sqrt{5} = c + d\sqrt{2}$ za neke pozitivne racionalne brojeve a, b, c, d, što nije moguće. Stoga imamo $\delta_2\varphi_2^n < \delta_1\varphi_1^m$ ili $\delta_1\varphi_1^m < \delta_2\varphi_2^n$.

Pretpostavimo da je $\delta_2\varphi_2^n < \delta_1\varphi_1^m$. Tada po Lemi 5.22 znamo da vrijedi nejednakost 5.20 iz Leme 5.20. Možemo uzeti $M = 10^{20}$ i želimo odabrati vrijednosti za p, q i α

tako da su uvjeti Leme 5.20 ispunjeni. Za to je potreban numerički račun koji se može obaviti primjerice u programu Pari. Ovdje je $\kappa = \frac{\ln\varphi_1}{\ln\varphi_2}$. Radeći sa preciznošću od 1000 decimalnih mjesta, neka je κ_0 broj s pomičnom točkom koji aproksimira κ tako da je sigurno $|\kappa-\kappa_0| \le 10^{-900}$. Neka je $\frac{p}{q}$ konvergenta razvoja u verižni razlomak od κ_0 koju ćemo kasnije odabrati. Uzimamo $\alpha = \frac{1}{q} + \frac{q}{10^{900}}$, pa vrijedi (5.18) jer je

$$|q\kappa - p| \le |q\kappa - q\kappa_0| + |q\kappa_0 - p| \overset{(1.14)}{\le} \frac{q}{10^{900}} + \frac{1}{q} = \alpha.$$

Konačno još trebamo odabrati $\frac{p}{q}$ tako da je $\varepsilon = \|\mu q\| - M\alpha$ pozitivno. Pokazuje se da to vrijedi ako za $\frac{p}{q}$ uzmemo 43. konvergentu verižnog razlomka od κ_0, tj.

$$p = 387952129646429739199, \qquad q = 710561840528321688446 < 7.2\cdot 10^{20}.$$

Prema Lemi 5.20 mora biti

$$m < \frac{\ln\left(\frac{Aq}{\varepsilon}\right)}{\ln B} < 48.$$

Stoga je $m \le 47$ te zbog (5.22) dobivamo da je $n \le 25$.

Pretpostavimo li da vrijedi obratna nejednakost, tj. $\delta_1\varphi_1^m < \delta_2\varphi_2^n$, sličnim računom dobijemo $n \le 26$, $m \le 48$.

Dakle, reducirali smo problem na rješavanje jednadžbe $u_m = v_n$ za $0 \le m, n \le 50$ što se lako napravi na računalu. Jedina mogućnost koja se dobiva je $m = n = 0$, a ta je isključena. □

Spojimo li nizove $(u_m)_{m\in\mathbb{N}}$ i $(v_n)_{n\in\mathbb{N}}$ u jedan niz, dobivamo

$$13, \mathbf{29}, 34, 89, \mathbf{169}, 233, 610, \mathbf{985}, 1597, 4181, \mathbf{5741}, 10946, 28657, \mathbf{33461}, 75025, \mathbf{195025}, \ldots$$

pri čemu smo Pellove brojeve, tj. elemente niza $(v_n)_n$ podebljali. Pišući redom za svako pojavljivanje $u_m = F_{2m+5}$ slovo a i za svako pojavljivanje $v_n = P_{2n+3}$ slovo b dobivamo beskonačnu riječ w koja počinje

$$w = abaabaabaabaabab\ldots.$$

Pokazat ćemo da je w *sturmovska* riječ (beskonačni analogon donje Christoffelove riječi), tj. riječ koja odozdo diskretizira par $\big((0,-1),\ell\big)$ u vertikalnom položaju, gdje je ℓ polupravac koji leži na pravcu $y = Cx + D$ za

$$C = \frac{\ln\varphi_1}{\ln\varphi_2} = 0.5459\ldots \quad \text{i} \quad D = \frac{\ln\left(\frac{\delta_1\varphi_1}{\delta_2\varphi_2}\right)}{\ln\varphi_2} = -0.4557\ldots,$$

pri čemu su vrijednosti $\varphi_1, \varphi_2, \delta_1, \delta_2$ dane u (5.24). Definicije i ilustracije ovih pojmova pogledajte kod Teorema 4.40 i na Slici 2.a.

Lema 5.25. *Neka su $U = a\mathbb{N}_0 + b$ i $V = c\mathbb{N}_0 + d$ dva disjunktna skupa pri čemu je*

$$a > 0, \quad c > 0, \quad d - c < b < d. \tag{5.25}$$

Uredimo u rastućem poretku elemente od $U \cup V = \{e_1 < e_2 < \ldots\}$ te zapišemo slovo a kad je $e_i \in U$ i slovo b kad je $e_i \in V$. Dobivena riječ $h \in \{a,b\}^{\mathbb{N}}$ diskretizira odozdo par $\big((0,-1),\ell\big)$ u vertikalnom položaju, gdje je ℓ polupravac $ax + b = cy + d$, $x \ge 0$.

Dokaz. Možemo zamijeniti U i V sa $A\mathbb{N}_0 + B$ i $\mathbb{N}_0$, gdje je $A = \frac{a}{c}$, $B = \frac{b-d}{c}$. Relativne pozicije elemenata očito ostaju iste. Iz (5.25) je $-1 < B < 0$. Promatrajmo polupravac $y = Ax + B$, $x \geq 0$ i njegova sjecišta s pravcima cjelobrojne rešetke, tj. pravcima $x = k$, $y = l$, $k, l \in \mathbb{Z}$. Slično kao kod Christoffelovih riječi, definiramo reznu riječ označavajući redom navedena sjecišta sa a ili b, ovisno o tome je li pravac rešetke oblika $x = k$ ili $y = l$, vidi Sliku 1.d. Zaista je rezna riječ koju smo dobili upravo h jer ortogonalno projicirajući spomenuta sjecišta na y os, dobivamo realne brojeve $Ak+B$ ili l u ovisnosti o kojem se od dva slučaja radi. Na kraju se još treba prisjetiti bijekcije između tipa koraka diskretizacije odozdo i tipa sjecišta za reznu riječ, vidi Sliku 1.e. Tako smo dobili da je h diskretizacija odozdo od $\big((-1,0);\, y = Ax + B, x \geq 0\big)$, odnosno $\big((-1,0);\, ax + b = cy + d, x \geq 0\big)$. $\square$

Uzmimo sada $U = \{u_m : m \geq 1\}$ i $V = \{v_n : n \geq 1\}$, gdje su kao i u (5.23) $u_m = \delta_1\varphi_1^m + \bar{\delta}_1\overline{\varphi}_1^m$, $v_n = \delta_2\varphi_2^n + \bar{\delta}_2\overline{\varphi}_2^n$ uz sve vrijednosti dane u (5.24). Stavimo li $u'_m = \delta_1\varphi_1^m$ i $v'_n = \delta_2\varphi_2^n$, vidimo da će vrijediti $u_m - 1 < u'_m < u_m$ i $v_n - 1 < v'_n < v_n$, pa skupovi $U' = \{u'_m : m \geq 1\}$ i $V' = \{v'_n : n \geq 1\}$ generiraju stapanjem istu riječ na $\{a,b\}^{\mathbb{N}}$ kao i riječ w koju generiraju U i V. No, i skupovi $U'' = \{\ln u'_m : m \geq 1\} = (\ln\varphi_1)\mathbb{N}_0 + \ln(\delta_1\varphi_1)$ i $V'' = \{\ln v'_n : n \geq 1\} = (\ln\varphi_2)\mathbb{N}_0 + \ln(\delta_2\varphi_2)$ očito stapanjem generiraju istu riječ w. Skupovi U i V su disjunktni zbog Teorema 5.24, pa su disjunktni U' i V', a onda i U'' i V''. Provjerimo da su za U'' i V'' ispunjeni uvjeti (5.25), pa Lema 5.25 konačno povlači da w diskretizira odozdo par $\big((0,-1), \ell\big)$ u vertikalnom položaju, gdje je ℓ polupravac na pravcu $(\ln\varphi_1)x + \ln(\delta_1\varphi_1) = (\ln\varphi_2)y + \ln(\delta_2\varphi_2)$, odnosno $y = Cx + D$ uz prije definirane C i D.

Poglavlje 6

Struktura spektara iznad 3

6.1 Hallova zraka

Pokazat ćemo da Lagrangeov i Markovljev spektar sadrže sve brojeve veće od neke konstante. Prvi rezultat tog tipa dao je 1947. Hall po kojemu se beskonačni interval sadržan u $\mathcal{L}$ ili $\mathcal{M}$ naziva *Hallova zraka*. On je dokazao sljedeći teorem.

Teorem 6.1. *Svaki realan broj može se napisati u obliku* $a + [0, b_1, b_2, \ldots] + [0, c_1, c_2, \ldots]$*, gdje je* a *cijeli broj te su parcijalni kvocijenti* b_i *i* c_i $(i \in \mathbb{N})$ *svi manji ili jednaki* 4.

Prethodni je rezultat neposredna posljedica idućeg budući da je segment iz Teorema 6.2 dulji od 1, pa sadrži svaki realni broj modulo 1.

Teorem 6.2. *Svaki realan broj u segmentu* $[\sqrt{2}-1, 4\sqrt{2}-4] = [0.414\ldots, 1.656\ldots]$ *može se napisati u obliku* $[0, b_1, b_2, \ldots] + [0, c_1, c_2, \ldots]$*, gdje su za sve* $i \in \mathbb{N}$ *parcijalni kvocijenti* b_i *i* c_i *manji ili jednaki* 4.

Pokazat ćemo da postojanje Hallove zrake jednostavno slijedi iz Teorema 6.1. Mnogo je teže pitanje koji je najmanji broj sadržan u Hallovoj zraci. Dat ćemo nekoliko poboljšanja početnog rezultata, a konačni je odgovor dao Freiman 1975., no njegov dokaz sadrži previše računa da bismo ga ovdje iznosili.

Označimo za svaki prirodan broj $k \geq 2$ sa $C(k)$ skup *iracionalnih* realnih brojeva $\alpha \in (0,1)$ takvih da u razvoju u verižni razlomak od α nijedan parcijalni kvocijent nije strogo veći od k. Kao što je uobičajeno, sumu $A + B$ dvaju podskupova A i B od $\mathbb{R}$ definiramo kao $\{a + b : a \in A, b \in B\}$.

Budući da je u $C(4)$ najmanji broj jednak $[0, \overline{4, 1}] = \frac{1}{2}(\sqrt{2}-1)$, a najveći broj je $[0, \overline{1, 4}] = 2\sqrt{2} - 2$, vidimo da je Teorem 6.2 ekvivalentan

$$C(4) + C(4) = [\sqrt{2}-1,\ 4\sqrt{2}-4]. \tag{6.0}$$

U dokazu (6.0) koristit ćemo da se skup $C(4)$ može dobiti iz $[\frac{1}{2}(\sqrt{2}-1), 2\sqrt{2}-2]$ izbacivanjem odgovarajućeg beskonačnog skupa disjunktnih otvorenih intervala. Slijedeći Halla, pokazat ćemo da $C(4)$ pripada klasi *Cantorovih skupova* koji se dobivaju sljedećim postupkom. Uzmemo zatvoreni interval $A = [x, x+a]$ u skupu realnih brojeva i izbacimo srednji otvoreni interval $A_{12} = (x + a_1, x + a_1 + a_{12})$, pa preostaju dva zatvorena intervala $A_1 = [x, x + a_1]$ i $A_2 = [x + a_1 + a_{12}, x + a]$. U drugom koraku ovog postupka izbacimo po jedan otvoreni interval iz sredine A_1 i A_2. Taj se proces nastavlja tako da u n-tom koraku izbacimo 2^{n-1} srednjih otvorenih intervala. Skup koji dobivamo u limesu naziva se Cantorov skup.

Kako bismo dobili $C(4)$ kao Cantorov skup najprije uzmemo $A = [\frac{1}{2}(\sqrt{2}-1), 2\sqrt{2}-2]$. U definiranju subdivizija od A pomoću kojih se dobiva $C(4)$, promatrat ćemo tri tipa intervala.

Intervali prvog tipa su segmenti s krajevima oblika

$$[0, a_1, \dots, a_r, 1, \overline{4, 1}] \quad \text{i} \quad [0, a_1, \dots, a_r, 4, \overline{1, 4}] \tag{6.1}$$

Dakle, interval prvog tipa ima za paran r lijevi kraj $[0, a_1, \dots, a_r, 4, \overline{1, 4}]$, a za neparan r mu je lijevi kraj $[0, a_1, \dots, a_r, 1, \overline{4, 1}]$. Stoga je i sam A interval prvog tipa uz $r = 0$.

Intervali drugog tipa imaju krajeve

$$[0, a_1, \dots, a_r, 2, \overline{4, 1}] \quad \text{i} \quad [0, a_1, \dots, a_r, 4, \overline{1, 4}]. \tag{6.2}$$

Intervali trećeg tipa imaju krajeve

$$[0, a_1, \dots, a_r, 3, \overline{4, 1}] \quad \text{i} \quad [0, a_1, \dots, a_r, 4, \overline{1, 4}]. \tag{6.3}$$

U procesu subdivizije iz intervala (6.1) prvog tipa uklonimo otvoreni interval s krajevima

$$[0, a_1, \dots, a_r, 1, \overline{1, 4}] \quad \text{i} \quad [0, a_1, \dots, a_r, 2, \overline{4, 1}], \tag{6.4}$$

iz intervala (6.2) drugog tipa uklonimo otvoreni interval s krajevima

$$[0, a_1, \dots, a_r, 2, \overline{1, 4}] \quad \text{i} \quad [0, a_1, \dots, a_r, 3, \overline{4, 1}], \tag{6.5}$$

a iz intervala (6.3) trećeg tipa uklonimo otvoreni interval s krajevima

$$[0, a_1, \dots, a_r, 3, \overline{1, 4}] \quad \text{i} \quad [0, a_1, \dots, a_r, 4, \overline{4, 1}]. \tag{6.6}$$

Tako od svakog intervala prvog tipa dobijemo interval prvog tipa (uz $a_{r+1} = 1$) s jedne strane i interval drugog tipa s druge, od svakog intervala drugog tipa dobijemo interval prvog tipa (uz $a_{r+1} = 2$) s jedne strane i interval trećeg tipa s druge strane, a od svakog intervala trećeg tipa dobijemo interval prvog tipa (uz $a_{r+1} = 3$) s jedne strane i interval prvog tipa (uz $a_{r+1} = 4$) s druge strane.

Ovaj proces subdivizije započinjemo sa segmentom A i nastavljamo prema gornjim pravilima. Dobiveni Cantorov skup je očito $C(4)$. Kako bismo dokazali (6.0), moramo proučiti sume dvaju Cantorovih skupova i u tu svrhu trebamo dokazati četiri leme. Ako je I interval u skupu realnih brojeva, sa $|I|$ označavamo duljinu tog intervala.

Lema 6.3. *Neka je I bilo koji od otvorenih intervala uklonjenih tijekom Cantorova postupka kojim dobivamo $C(4)$. Neka je M zatvoreni interval iz kojeg je I uklonjen tijekom tog postupka te neka su M_1 i M_2 redom lijevi i desni intervali koji ostaju nakon uklanjanja I iz M. Tada je*

$$|M_i| \geq |I| \qquad \textit{za } i \in \{1, 2\}. \tag{6.7}$$

Dokaz. Neka su $\frac{p_j}{q_j} = [0, a_1, \dots, a_j]$ odgovarajuće konvergente nekog verižnog razlomka. Ako su μ i ν pozitivni realni brojevi, onda smo u Teoremu 1.8.b vidjeli da je

$$\begin{aligned}\big|[0, a_1, \dots, a_r, \mu] - [0, a_1, \dots, a_r, \nu]\big| &= \left| \frac{\mu p_r + p_{r-1}}{\mu q_r + q_{r-1}} - \frac{\nu p_r + p_{r-1}}{\nu q_r + q_{r-1}} \right| \\ &= \frac{|\mu - \nu|}{|(\mu q_r + q_{r-1})(\nu q_r + q_{r-1})|} = \frac{|\mu - \nu|}{|q_{r-1}^2 (1 + \frac{\mu q_r}{q_{r-1}})(1 + \frac{\nu q_r}{q_{r-1}})|}\end{aligned}$$

Koristeći ovu formulu uz odgovarajući izbor μ i ν, možemo izračunati duljine intervala (6.4), (6.5), (6.6) koji su uklonjeni iz (6.1), (6.2), (6.3), tim redom, te duljine intervala koji preostaju nakon uklanjanja. Račun nam onda pokazuje da je omjer duljine uklonjenog intervala i duljine preostalog intervala, bilo lijevog bilo desnog, uvijek manji od 1. Ilustrirajmo to samo u jednom slučaju, gledajući navedeni omjer za interval prvog tipa (6.1) i jedan od intervala koji iz njega dobijemo kad uklonimo odgovarajući interval (6.4). Treba, dakle, pokazati da je sljedeći kvocijent manji od 1

$$\left|\frac{[0,a_1,\ldots,a_r,2,\overline{4,1}]-[0,a_1,\ldots,a_r,1,\overline{1,4}]}{[0,a_1,\ldots,a_r,1,\overline{4,1}]-[0,a_1,\ldots,a_r,1,\overline{1,4}]}\right| = \frac{\frac{|[2,\overline{4,1}]-[1,\overline{1,4}]|}{|q_{r-1}^2(1+[2,\overline{4,1}]Q_r)(1+[1,\overline{1,4}]Q_r)|}}{\frac{|[1,\overline{4,1}]-[1,\overline{1,4}]|}{|q_{r-1}^2(1+[1,\overline{4,1}]Q_r)(1+[1,\overline{1,4}]Q_r)|}}$$
$$= \frac{|[2,\overline{4,1}]-[1,\overline{1,4}]|}{|[1,\overline{4,1}]-[1,\overline{1,4}]|}\cdot\frac{1+[1,\overline{4,1}]Q_r}{1+[2,\overline{4,1}]Q_r} = \frac{0.378+\varepsilon_1}{0.621+\varepsilon_2}\cdot\frac{1+(1.207+\varepsilon_3)Q_r}{1+(2.207+\varepsilon_4)Q_r},$$

gdje smo označili $Q_r = \frac{q_r}{q_{r-1}} > 1$, a ε_i su pozitivni brojevi manji od 10^{-3}. Sada je očito da tražena tvrdnja vrijedi, a jednako se pokazuje i za preostalih pet kvocijenata. Time je dokazano (6.7). $\square$

Lema 6.4. *Neka su A_1 i A_2 proizvoljni omeđeni zatvoreni intervali realnih brojeva. Otvoreni interval I uklonjen je iz sredine A_1 te ostaju zatvoreni intervali A_{11} lijevo i A_{12} desno. Ako je $|A_2| \geq |I|$, onda je*

$$A_1 + A_2 = (A_{11} \cup A_{12}) + A_2.$$

Dokaz. Neka suma $A_1 + A_2$ ima lijevi kraj L. Tada je $A_{11} + A_2$ zatvoreni interval kojemu je lijevi kraj također L, a desni kraj mu je $L + |A_{11}| + |A_2|$. Suma $A_{12} + A_2$ je zatvoreni interval s istim desnim krajem kao i $A_1 + A_2$ i s lijevim krajem $L + |A_{11}| + |I|$. Dakle, ta dva intervala pokrivaju čitav $A_1 + A_2$ ako i samo ako se preklapaju, tj. ako i samo ako je

$$L + |A_{11}| + |I| \leq L + |A_{11}| + |A_2|.$$

Time je lema dokazana. $\square$

Lema 6.5. *Neka je B unija u parovima disjunktnih omeđenih zatvorenih intervala A_1, A_2, …, A_r kojih je konačno mnogo. Otvoreni interval I uklonjen je iz sredine A_1 te ostaju zatvoreni intervali A_{r+1} lijevo i A_{r+2} desno. Neka je $B^* = \bigcup_{i=2}^{r+2} A_i$. Ako je*

$$|A_i| \geq |I| \qquad \text{za } 2 \leq i \leq r+2, \tag{6.8}$$

onda je $B^* + B^* = B + B$.

Dokaz. Stavimo $A_1^* = A_{r+1} \cup A_{r+2}$ te $A_i^* = A_i$ za $2 \leq i \leq r$. Imamo $B^* = \bigcup_{i=1}^{r} A_i^*$. Iz (6.8) i Leme 6.4 slijedi

$$A_1 + A_1 = A_1^* + A_1^* \qquad \text{te} \qquad A_1 + A_i = A_1^* + A_i^* \quad \text{za } 2 \leq i \leq r. \tag{6.9}$$

Zato je

$$B + B = \bigcup_{j=1}^{r}(A_1 + A_j) \cup \bigcup_{2\leq i\leq j\leq r}(A_i^* + A_j^*) = \bigcup_{1\leq i\leq j\leq r}(A_i^* + A_j^*) = B^* + B^*,$$

gdje druga jednakost slijedi iz (6.9). $\square$

Lema 6.6. *Ako je $K_1 \supseteq K_2 \supseteq K_3 \supseteq \dots$ niz zatvorenih intervala od kojih svaki sadrži svoga sljedbenika, onda je*

$$\bigcap_{i=1}^{\infty} K_i + \bigcap_{i=1}^{\infty} K_i = \bigcap_{i=1}^{\infty} (K_i + K_i). \tag{6.10}$$

Dokaz. Za svaki $i \in \mathbb{N}$ skup $K_i + K_i$ sadrži lijevu stranu od (6.10), pa desna strana u (6.10) očito sadrži lijevu. Kako bismo dokazali obratnu inkluziju, uzmimo proizvoljan broj t iz $\bigcap_{i=1}^{\infty}(K_i + K_i)$.

Neka je $V_i = \{(x_1, x_2) : x_1, x_2 \in K_i\}$ za $i \in \mathbb{N}$ te neka je $T = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 = t\}$. Budući da je V_i zatvoren i omeđen skup, a T je zatvoren, to je skup $T \cap V_i$ zatvoren i omeđen, a ujedno i neprazan jer je $t \in K_i + K_i$. Također $T \cap V_i$ sadrži $T \cap V_{i+1}$ za $i \in \mathbb{N}$ iz čega po Cantorovom teoremu o presjeku (silazni niz nepraznih omeđenih zatvorenih podskupova od $\mathbb{R}^n$ ima neprazan presjek) slijedi da je $\bigcap_{i=1}^{\infty}(T \cap V_i) = T \cap \bigcap_{i=1}^{\infty} V_i$ neprazan skup. Ako je (t_1, t_2) proizvoljni element tog skupa, onda su t_1 i t_2 u $\bigcap_{i=1}^{\infty} K_i$ i vrijedi $t_1 + t_2 = t$. Stoga je t element od $\bigcap_{i=1}^{\infty} K_i + \bigcap_{i=1}^{\infty} K_i$ čime je lema dokazana. $\square$

Dokaz (6.0). Stavimo $S_0 = [\frac{1}{2}(\sqrt{2}-1), 2(\sqrt{2}-1)]$. Izgradit ćemo niz omeđenih zatvorenih skupova $S_0, S_1, S_2, \dots$ sa sljedećim svojstvima

$$S_i \supseteq S_{i+1} \quad \text{za } i \in \mathbb{N}_0, \tag{6.11}$$

$$\bigcap_{i=0}^{\infty} S_i = C(4), \tag{6.12}$$

$$S_i + S_i = S_{i+1} + S_{i+1} \quad \text{za } i \in \mathbb{N}_0. \tag{6.13}$$

Vidjeli smo prije da se $C(4)$ može dobiti iz S_0 uklanjanjem beskonačnog skupa u parovima disjunktnih otvorenih intervala, točnije, skupa svih otvorenih intervala oblika (6.4), (6.5) ili (6.6), pri čemu je svaki $a_i \leq 4$. Možemo ovaj skup otvorenih intervala urediti po duljini intervala u padajućem poretku. Ako nekoliko intervala ima istu duljinu, možemo ih poredati po volji, primjerice slijeva nadesno u S_0. Neka je dobiveni niz intervala $D_0, D_1, D_2, \dots$ i definirajmo S_i kao S_{i-1} iz kojeg je izbačen D_{i-1}, tj. $S_i = S_{i-1} \setminus D_{i-1}$ za $i \in \mathbb{N}$. Očito niz $S_0, S_1, S_2, \dots$ ima svojstva (6.11) i (6.12).

Kako bismo pokazali da vrijedi i (6.13), primjenjujemo Lemu 6.5 uzastopce sa $B = S_i$, $B^* = S_{i+1}$, $I = D_i$ za $i \in \mathbb{N}_0$. Iduća analiza pokazuje da je uvjet (6.8) zadovoljen u svakoj primjenu te leme.

Za svaki $i \in \mathbb{N}_0$ je S_{i+1} dobiven od S_i uklanjanjem D_i iz nekog zatvorenog intervala, nazovimo ga H_i, koji je sadržan u S_i. Uklanjanjem D_i, od H_i ostaju dva zatvorena intervala, H_{i1} lijevo i H_{i2} desno od D_i. Ako (6.8) vrijedi za $B = S_{i-1}$, $B^* = S_i$, $I = D_{i-1}$, onda je kako bismo dokazali (6.8) za $B = S_i$, $B^* = S_{i+1}$, $I = D_i$, zbog $|D_{i-1}| \geq |D_i|$ dovoljno provjeriti da vrijedi

$$|H_{ij}| \geq |D_i| \qquad \text{za } j \in \{1, 2\}. \tag{6.14}$$

U Lemi 6.3 uzimamo $I = D_i$ tako da su M_1 i M_2 zatvoreni intervali u Cantorovom postupku koji su susjedni D_i s lijeve i s desne strane, redom. Ako H_{ij} sadrži ili je jednak M_j, onda (6.14) slijedi iz (6.7).

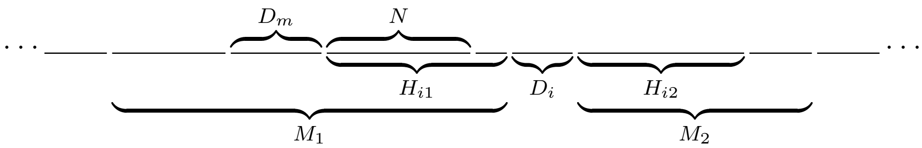
Ako je H_{ij} strogo sadržan u M_j, onda postoji neki interval D_m, sa $m < i$, koji je susjedan H_{ij} na strani suprotnoj od D_i i taj D_m ima neprazan presjek sa M_j. Stoga je u Cantorovom procesu kojim dobivamo $C(4)$ interval D_m uklonjen tek nakon D_i. Uklanjanje

D_m u Cantorovom procesu ostavlja dva zatvorena intervala, svaki od njih susjedan jednom kraju od D_m. Neka je N jedan od ta dva intervala koji se nalazi između D_m i D_i. Budući da H_{ij} mora sadržavati N, imamo

$$|H_{ij}| \geq |N| \geq |D_m| \geq |D_i|,$$

gdje druga nejednakost slijedi iz (6.7) uz $I = D_m$, a treća nejednakost vrijedi zbog $m < i$. Time je završen dokaz (6.14), pa time i (6.13).

Prethodni je slučaj ilustriran na sljedećoj skici za $j = 1$.



Budući da je $S_0 + S_0 = [\sqrt{2} - 1, 4\sqrt{2} - 4]$, svojstvo (6.13) povlači $[\sqrt{2} - 1, 4\sqrt{2} - 4] = \bigcap_{i=0}^{\infty}(S_i + S_i)$. Sada (6.11), (6.12) i Lema 6.6 dokazuju (6.0). $\square$

Lako je iskoristiti Teorem 6.1 za pokazati da Lagrangeov, pa onda i Markovljev spektar sadrži sve brojeve veće od nekog konkretnog broja. Najjednostavniji rezultat tog tipa je sljedeći.

Teorem 6.7. *Lagrangeov spektar, pa zato i Markovljev spektar sadrži sve brojeve veće od* 6*.*

Dokaz. Neka je $\lambda > 6$. Tada prema Teoremu 6.1 možemo pisati

$$\lambda = a + [0, b_1, b_2, \ldots] + [0, c_1, c_2, \ldots], \tag{6.15}$$

pri čemu nijedan od parcijalnih kvocijenata b_i i c_i nije veći od 4. Budući da je $\lambda > 6$, mora biti $a \geq 5$. Definirajmo sada

$$\alpha = [a_0, a_1, a_2, \ldots] = [b_{k_1}, \ldots, b_1, a, c_1, \ldots, c_{k_1}, b_{k_2}, \ldots, b_1, a, c_1, \ldots, c_{k_2}, b_{k_3}, \ldots], \tag{6.16}$$

gdje je $(k_i)_{i \in \mathbb{N}}$ strogo rastući niz prirodnih brojeva. Po definiciji je $L(\alpha) = \limsup_n \lambda_n(\alpha)$, a da bismo odredili taj $\limsup$ dovoljno je promatrati one n za koje je $a_n = a \geq 5$. Za takve n je

$$\lambda_n = a + [0, b_1, b_2, \ldots] + [0, c_1, c_2, \ldots] \to \lambda \quad \text{kad } n \to \infty,$$

pa je $L(\alpha) = \lambda$. $\square$

Pažljivija upotreba Teorema 6.1 daje bolji rezultat.

Teorem 6.8. *Lagrangeov spektar, pa zato i Markovljev spektar sadrži sve brojeve veće od* $4 + [0, \overline{1, 4}] + [0, 1, 5, \overline{1, 4}] = 5.6819\ldots$*.*

Dokaz. Neka je $\mu = 4 + [0, \overline{1, 4}] + [0, 1, 5, \overline{1, 4}]$ i uzmimo $\lambda > \mu$. Po Teoremu 6.1, može se λ zapisati u obliku (6.15) pri čemu nijedan od b_i i c_i nije veći od 4. Budući da je $\lambda > \mu > 4 + [0, \overline{1, 4}] + [0, \overline{1, 4}] = 5.656\ldots$, mora biti $a \geq 5$. Definiramo α sa (6.16) tako da je

$$\limsup_{a_n = a} \lambda_n(\alpha) = \lambda \qquad \text{i} \qquad \limsup_{a_n \neq a} \lambda_n \leq 4 + [0, 1, a, \overline{1, 4}] + [0, \overline{1, 4}].$$

Stoga ako je $a \geq 6$, sigurno vrijedi $\limsup_n \lambda_n(\alpha) = \lambda$, a ako je $a = 5$, onda imamo $\limsup_n \lambda_n(\alpha) = \lambda$ jer je $\lambda > \mu$. $\square$

Čini se da je ograda od oko 5.682 iz prethodnog teorema najbolje što se može dobiti koristeći ovu Hallovu metodu. Novu ideju dali su Freiman i Judin 1966. godine.

Teorem 6.9. *Markovljev spektar sadrži sve brojeve veće od*

$$4+[0,\overline{1,3}]+[0,3,4,\overline{1,3}]=4+\frac{1}{2}(\sqrt{21}-3)+\frac{1}{50}(\sqrt{21}+11)=5.1029\dots.$$

Dokaz Teorema 6.9 zasniva se na sljedećem rezultatu koji je sličan Teoremu 6.2, ali koristi manji skup verižnih razlomaka.

Teorem 6.10. *Svaki realan broj u segmentu* $[5-\sqrt{21},\sqrt{21}-3]=[0.417\dots,1.582\dots]$ *može se napisati u obliku* $[0,b_1,b_2,\dots]+[0,c_1,c_2,\dots]$*, gdje su za sve* $i\in\mathbb{N}$ *parcijalni kvocijenti* b_i *i* c_i *manji ili jednaki* 4 *te nijedan par* b_i,b_{i+1} *ili* c_i,c_{i+1} *nije jednak* $1,4$ *ili* $2,4$.

Dokaz. Neka D označava skup svih $\alpha=[0,a_1,a_2,\dots]$ takvih da nijedan parcijalni kvocijent a_i nije veći od 4 i nijedan par a_i,a_{i+1} nije jednak $1,4$ ili $2,4$. Prvo ćemo pokazati da je D Cantorov skup dobiven iz intervala $A=\big[[0,4,\overline{1,3}],[0,\overline{1,3}]\big]$.

U subdivizijama kojima se iz A dobiva D treba promatrati dva tipa intervala. Za paran r definiramo ih sa

$$\begin{array}{lll}\text{(prvi tip)} & \big[[0,a_1,\dots,a_r,4,\overline{1,3}],[0,a_1,\dots,a_r,\overline{1,3}]\big] & \text{za } r=0 \text{ ili } a_r\in\{3,4\},\\ \text{(drugi tip)} & \big[[0,a_1,\dots,a_r,\overline{3,1}],[0,a_1,\dots,a_r,\overline{1,3}]\big] & \text{za } a_r\in\{1,2\}.\end{array}$$

Za paran r iz intervala koji je prvog tipa redom uklanjamo intervale

$$\begin{gathered}([0,a_1,\dots,a_r,2,\overline{3,1}],[0,a_1,\dots,a_r,1,\overline{1,3}]),\\ ([0,a_1,\dots,a_r,3,4,\overline{1,3}],[0,a_1,\dots,a_r,2,\overline{1,3}]),\\ ([0,a_1,\dots,a_r,4,4,\overline{1,3}],[0,a_1,\dots,a_r,3,\overline{1,3}]).\end{gathered}$$

Za paran r iz intervala koji je drugog tipa redom uklanjamo intervale

$$\begin{gathered}([0,a_1,\dots,a_r,2,\overline{3,1}],[0,a_1,\dots,a_r,1,\overline{1,3}]),\\ ([0,a_1,\dots,a_r,3,4,\overline{1,3}],[0,a_1,\dots,a_r,2,\overline{1,3}]).\end{gathered}$$

Na očit način se analogno definiraju intervali prvog i drugog tipa za neparan r. Tako se svaki interval prvog tipa razdijeli na dva intervala prvog tipa i dva intervala drugog tipa, a svaki interval drugog tipa se razdijeli na jedan interval prvog tipa i dva intervala drugog tipa. Subdivizija se nastavlja po navedenim pravilima i lako je vidjeti da Cantorov skup koji dobijemo kao rezultat mora biti upravo D.

Računski se provjeri da intervali u postupku subdivizije zadovoljavaju nejednakost (6.7) iz Leme 6.3, tj. nijedan uklonjeni interval u pojedinom koraku subdivizije nije duljinom veći od bilo lijevog bilo desnog susjednog zatvorenog intervala koje dobijemo. Stoga možemo koristiti Leme 6.5 i 6.6 kao u dokazu (6.0) pa slijedi $D+D=[5-\sqrt{21},\sqrt{21}-3]$ čime je teorem dokazan. □

Dokaz Teorema 6.9. Neka je $\eta=4+[0,\overline{1,3}]+[0,3,4,\overline{1,3}]$. Za svaki $\lambda\in[9-\sqrt{21},1+\sqrt{21}]$ možemo pisati

$$\lambda=4+[0,a_1,a_2,\dots]+[0,a_{-1},a_{-2},\dots],$$

pri čemu brojevi $\alpha_d=[0,a_1,a_2,\dots]$ i $\alpha_\ell=[0,a_{-1},a_{-2},\dots]$ leže u skupu D iz Teorema 6.10. Definiramo obostrano beskonačni niz prirodnih brojeva $A=\dots,a_{-2},a_{-1},a_0=4,a_1,a_2,\dots$, pa vrijedi $\lambda_0(A)=\lambda$.

Budući da α_d i α_ℓ leže u D, ako je k paran pozitivan broj mora biti

$$\lambda_k(A) \leq 4 + [0, 3, 4, \overline{1, 3}] + [0, \overline{1, 3}] = \eta$$

i jednakost se može postići samo za $k = 2$, a ako je k neparan pozitivan broj imamo

$$\lambda_k(A) \leq 4 + [0, \overline{3, 4}] + [0, \overline{1, 3}] < \eta.$$

Simetrično se pokazuje da iste nejednakosti vrijede za negativne k, pa je $M(A) = \lambda$ uz uvjet da je $\lambda \geq \eta$. Time smo pokazali da svaki broj iz segmenta $[\eta, 1 + \sqrt{21}] = [5.103\ldots, 5.582\ldots]$ pripada Markovljevom spektru.

Da popunimo rupu između ovog rezultata i onoga iz Teorema 6.8, koristimo Teorem 6.10 kako bismo proizvoljni $\lambda \in [10-\sqrt{21}, 2+\sqrt{21}]$ prikazali u obliku $\lambda = 5+\alpha_\ell+\alpha_d$, gdje su $\alpha_d = [0, a_1, a_2, \ldots]$ i $\alpha_\ell = [0, a_{-1}, a_{-2}, \ldots]$ iz D. Neka je $A = \ldots, a_{-2}, a_{-1}, a_0 = 5, a_1, a_2, \ldots$, pa vrijedi $\lambda_0(A) = \lambda$. Nije teško vidjeti da je ovdje uvijek $M(A) = \lambda_0(A) = \lambda$, pa zaključujemo da Markovljev spektar sadrži segment $[10-\sqrt{21}, 2+\sqrt{21}] = [5.418\ldots, 6.582\ldots]$ te je dokaz teorema završen. □

Preskačemo neka neznatna poboljšanja. Daljnji napredak je ovisio o tome da se pokaže kako skup $C(3) + C(3)$ sadrži dugački interval. Budući je u $C(3)$ najmanji broj $[0, \overline{3, 1}] = \frac{-3+\sqrt{21}}{6}$, a najveći broj $[0, \overline{1, 3}] = \frac{-3+\sqrt{21}}{2}$, imamo

$$C(3) + C(3) \subset [0.52752\ldots, 1.58258\ldots].$$

No, lako se vidi da $C(3) + C(3)$ ne sadrži interval duljine 1 jer je najveća suma koju možemo dobiti ako oba pribrojnika počinju sa $[0, 3, \ldots]$ jednaka $2[0, 3, \overline{3, 1}] < 0.61279$, a najmanja suma koju dobivamo ako barem jedan pribrojnik počinje sa $[0, 1, \ldots]$ ili $[0, 2, \ldots]$ je $[0, \overline{3, 1}] + [0, 2, \overline{1, 3}] > 0.62201$. Zato $C(3) + C(3)$ ima rupu čija je duljina veća od 0.009. Drugim riječima, u Cantorovu procesu pojavljuje se otvoreni interval kojega treba ukloniti, ali je njegova duljina veća od duljine jednog od dva zatvorena intervala koji ostaju što znači da postupak opisan u dokazu (6.0) nije moguće provesti.

Freiman je dao precizan opis $C(3) + C(3)$ kao beskonačne unije zatvorenih intervala i našao je najdulji interval

$$\big[[0, \overline{3, 1}] + [0, 2, \overline{1, 3}], [0, \overline{1, 3}] + [0, 1, 2, \overline{1, 3}]\big] \approx [0.62202, 1.52753]$$

koji je sadržan u $C(3) + C(3)$. Freiman i, neovisno od njega, Schecker, koristeći ovakav rezultat, pokazali su da Lagrangeov spektar sadrži sve brojeve veće od $\sqrt{21} = 4.5825\ldots$.

Konačno je 1975. Freiman dao najbolji mogući rezultat koji daje točan početak Hallove zrake, a to je zahtijevalo "herojske izračune" (citat prema knjizi autora Cusick i Flahive koju smo pratili u ovom potpoglavlju).

Teorem 6.11. *Lagrangeov spektar, pa zato i Markovljev spektar sadrži sve brojeve veće ili jednake*

$$\begin{aligned}\mu &= 4 + [0\,3\,2\,1\,1\,\overline{3\,1\,3\,1\,2\,1}] + [0\,4\,3\,2\,2\,\overline{3\,1\,3\,1\,2\,1}] \\ &= \frac{2221564096 + 283748\sqrt{462}}{491993569} = 4.5278295661\ldots.\end{aligned}$$

Interval (ν, μ) *ne sadrži nijedan element iz Markovljevog spektra, gdje je*

$$\nu = 4 + [0\,3\,1\,3\,1\,2\,1\,1\,3\,3\,\overline{3\,1\,3\,1\,2\,1}] + [0\,3\,1\,3\,1\,3\,\overline{4\,4\,4\,3\,2\,3}] = 4.5278295384\ldots.$$

6.2 Rupe u spektrima

Po Teoremu 3.5 Markovljev spektar $\mathcal{M}$ je zatvoren skup, a iz Teorema 3.7 slijedi da je i Lagrangeov spektar $\mathcal{L}$ zatvoren skup. Stoga je komplement pojedinog spektra u $\mathbb{R}$ prebrojiva unija u parovima disjunktnih otvorenih intervala koje nazivamo maksimalnim rupama u spektru. Dakle, *maksimalna rupa* je otvoreni interval realnih brojeva koji ne sadrži nijedan element spektra, a oba kraja su mu elementi spektra. Zbog $\mathcal{L} \subset \mathcal{M}$ je svaka maksimalna rupa u $\mathcal{M}$ ujedno i rupa u $\mathcal{L}$, ali vidjet ćemo da postoje maksimalne rupe u $\mathcal{L}$ koje sadrže elemente iz $\mathcal{M}$. U nastavku ćemo pokazati kako naći neke rupe koje su istovremeno maksimalne u oba spektra. Primjer takve rupe je $(\sqrt{5}, \sqrt{8})$, no jer smo potpuno okarakterizirali elemente spektara manje od 3 i zbog postojanja Hallove zrake, zanimat će nas samo rupe koje se nalaze unutar intervala $(3, 4.528)$. Prve rezultate u tom smjeru dao je Perron 1921.

Propozicija 6.12. *Interval* $(\sqrt{12}, \sqrt{13}) = (3.46410\ldots, 3.60555\ldots)$ *je maksimalna rupa u Lagrangeovom i Markovljevom spektru. Imamo* $M(A) = \sqrt{13}$ *samo za* $A = \overline{3} = {}_\infty(3)_\infty$ *i* $M(A) = \sqrt{12}$ *samo ako* A *sadrži jedino* 1 *i* 2 *te za svaki* $n \in \mathbb{N}$ *sadrži uzorak* $(21)_n$.

Dokaz. Kako bismo pokazali da ne postoji obostrano beskonačan niz $A = (a_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$ takav da je $\sqrt{12} < M(A) < \sqrt{13}$, možemo pretpostaviti da je $a_i \leq 3$ za svaki i jer bi očito vrijedilo $M(A) > 4 > \sqrt{13}$ ako bi neki a_i bio barem 4. Ako se 3 ne pojavljuje u A, onda iz Teorema 1.8.a slijedi

$$M(A) \leq [\overline{2,1}] + [0, \overline{1,2}] = \sqrt{12},$$

čime je dokazana i tvrdnja propozicije o tome za koje A je $M(A) = \sqrt{12}$.

Ako za neki n vrijedi $a_n = 3$ te $a_{n-1} = 1$ ili $a_{n+1} = 1$, onda imamo

$$M(A) \geq \lambda_n(A) \geq [3, 1, \overline{1,3}] + [0, \overline{3,1}] = 3.82202\ldots > \sqrt{13}.$$

Slično, ako za neki n vrijedi $a_n = 3$ te $a_{n-1} = 2$ ili $a_{n+1} = 2$, onda je

$$M(A) \geq \lambda_n(A) \geq [3, 2, \overline{1,3}] + [0, \overline{3,1}] = 3.62202\ldots > \sqrt{13}.$$

Konačno, za $A = \overline{3}$ je $M(A) = [\overline{3}] + [0, \overline{3}] = \sqrt{13}$. $\square$

Lema 6.13. *Ako obostrano beskonačan niz* $A = (a_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$ *sadrži uzorak* $b_1, b_2, \ldots, b_k$, *gdje je* k *neparan prirodan broj, onda je*

$$M(A) \geq [\overline{b_1, b_2, \ldots, b_k}] + \frac{1}{[\overline{b_1, b_2, \ldots, b_k}]} \quad i \quad M(A) \geq [\overline{b_k, b_{k-1}, \ldots, b_1}] + \frac{1}{[\overline{b_k, b_{k-1}, \ldots, b_1}]}.$$

Dokaz. Za $m \in \mathbb{Z}$ stavimo $x_m = [a_m, a_{m+1}, \ldots]$ i $y_m = [a_{m-1}, a_{m-2}, \ldots]$. Ako je $x_m \geq y_m$, vrijedi

$$\lambda_m(A) = x_m + y_m^{-1} \geq x_m + x_m^{-1},$$

a ako je $y_m \geq x_m$, vrijedi

$$\lambda_{m-1}(A) = x_m^{-1} + y_m \geq x_m + x_m^{-1}.$$

Stoga za svaki $m \in \mathbb{Z}$ imamo $M(A) \geq x_m + x_m^{-1}$. Izaberimo n takav da je $x_{n-k} = [b_1, b_2, \ldots, b_k, x_n]$. Budući da je k neparan, mora biti $\max\{x_{n-k}, x_n\} \geq [\overline{b_1, b_2, \ldots, b_k}]$. Funkcija $x \mapsto x + x^{-1}$ je rastuća na $(1, +\infty)$, pa smo dobili prvu donju ogradu na $M(A)$ iz leme. Druga donja ograda na $M(A)$ dobiva se analogno korištenjem činjenice da je $M(A) \geq y_m + y_m^{-1}$ za svaki m. $\square$

Perron je iskoristio ovu lemu za dobivanje sljedećeg rezultata.

Propozicija 6.14. *Interval* $(\sqrt{13}, \frac{1}{22}(65+9\sqrt{3})) = (3.60555\ldots, 3.66311\ldots)$ *je maksimalna rupa u Lagrangeovom i Markovljevom spektru. Imamo* $M(A) = \frac{1}{22}(65+9\sqrt{3})$ *samo za* $A = \overline{1\,2}\,3\,3\,\overline{2\,1}$ *i ovaj je broj gomilište skupa* $\mathcal{M}$.

Dokaz. Već je iz dokaza Propozicije 6.12 jasno da je $\sqrt{13}$ lijevi kraj neke rupe. Zato trebamo samo provjeriti da je i desni kraj dan u ovoj propoziciji ispravan.

Neka je $\alpha_0 = \frac{1}{22}(65+9\sqrt{3})$ i neka je $f(x) = x + x^{-1}$ što je, kako smo rekli, rastuća funkcija na $(1, +\infty)$. Pretpostavimo da je $A = (a_i)_{i\in\mathbb{Z}} \in \mathbb{N}^{\mathbb{Z}}$ takav da je $\sqrt{13} < M(A) \leq \alpha_0$. Očito je $a_i \leq 3$ za sve $i \in \mathbb{Z}$ i $A \neq \overline{3}$. U A se ne mogu pojaviti uzorci $3, 1$ i $1, 3$ jer bi u protivnom prema Lemi 6.13 imali za neki a_i da je

$$M(A) \geq f([\overline{3, 1, a_i}]) > f([3, \overline{2, 1}]) = \alpha_0.$$

Sličnim računom dobivamo da se u A ne mogu pojaviti ni uzorci $3, 2, 3$ i $3, 2, 2$.

Zato se mora pojaviti uzorak $3, 2, 1$. Ako se za neki $n \geq 1$ pojavljuje u A uzorak $3, (2, 1)_n, a_i, a_{i+1}$, gdje je par a_i, a_{i+1} različit od $2, 1$, onda taj par mora biti ili $1, 1$ ili $1, 2$ ili $2, 2$ ili $2, 3$. No, iz Leme 6.13 bi u sva četiri ova slučaja slijedilo

$$M(A) \geq f([\overline{3, (2, 1)_n, a_i, a_{i+1}}]) > f([3, \overline{2, 1}]) = \alpha_0.$$

Dakle, mora biti $A = \overline{1\,2}\,3_k\,\overline{2\,1}$ za neki $k \in \mathbb{N}$. Samo je $k = 2$ moguće jer imamo

$$\begin{aligned} &M(A) \geq [3, \overline{2, 1}] + [0, \overline{2, 1}] > f([3, \overline{2, 1}]) = \alpha_0 \quad \text{za } k = 1 \quad \text{i} \\ &M(A) \geq [3, \overline{2, 1}] + [0, 3, 3, \ldots] > [3, \overline{2, 1}] + [0, 3, \overline{2, 1}] = f([3, \overline{2, 1}]) = \alpha_0 \quad \text{za } k \geq 3. \end{aligned}$$

Time smo pokazali da je $A = \overline{1\,2}\,3\,3\,\overline{2\,1}$ jedini niz za koji je $\sqrt{13} < M(A) \leq \alpha_0$.

Preostalo je još samo pokazati da α_0 nije izolirana točka u $\mathcal{M}$. Definiramo nizove $A_t = \overline{(1\,2)_t\,3\,3\,2\,(1\,2)_t}$ za $t \in \mathbb{N}$, pa očito $M(A_t) \to \alpha_0$ kad $t \to \infty$. $\square$

Prethodne propozicije možemo iskazati i na sljedeći način.

Za sve obostrano beskonačne nizove A iz $\{1, 2, 3\}^{\mathbb{Z}}$ u kojima nije dopušten uzorak 3 je $M(A) \leq \sqrt{12}$, a jednakost vrijedi samo ako A sadrži uzorak $(21)_n$ za svaki $n \in \mathbb{N}$. Od svih $A \in \mathbb{N}^{\mathbb{Z}}$ u kojima se pojavljuje uzorak 3, minimalni $M(A)$ postiže se samo za $A = \overline{3}$ i tada je $M(A) = \sqrt{13}$.

Za sve $A \in \{1, 2, 3\}^{\mathbb{Z}}$ u kojima nisu dopušteni uzorci 32 i 31 (te 23 i 13, ali u nastavku nećemo posebno spominjati obrtanje uzoraka), imamo $M(A) \leq \sqrt{13}$, a jednakost vrijedi samo za $A = \overline{3}$. Od svih $A \in \mathbb{N}^{\mathbb{Z}}$ u kojima se pojavljuje neki od uzoraka 32 i 31, minimalni $M(A)$ postiže se samo za $A = \overline{1\,2}\,3\,3\,\overline{2\,1}$.

Generalizacija ovog Perronovog postupka postala je standardna metoda naslućivanja i dokazivanja maksimalnih rupa u spektrima. Pokažimo da će ova metoda uvijek dati krajeve intervala određene obostrano beskonačnim nizovima koji su periodski ulijevo i udesno.

Teorem 6.15 (Cusick)**.** *Za* $t \in \mathbb{N}$ *uzmemo po volji* $T \subset \{1, 2, 3, 4\}^t$ *i za* $a_1, \ldots, a_n \in \mathbb{N}$ *definiramo*

$$C(T; a_1, \ldots, a_n) = \{[0, a_1, \ldots, a_n, c_1, c_2, \ldots] : (c_{i+1}, c_{i+2}, \ldots, c_{i+t}) \in T \text{ za svaki } i \geq 0\}.$$

Ako je $\alpha = \max C(T; a_1, \ldots, a_n)$ *ili* $\alpha = \min C(T; a_1, \ldots, a_n)$*, onda* α *ima periodski verižni razlomak (od nekog mjesta nadalje).*

Dokaz. Provest ćemo dokaz za $\alpha = \max C(T; a_1, \dots, a_n) = [0, a_1, a_2, \dots]$, a analogno se radi u slučaju kada je α minimum tog skupa.

Budući da je za $i > n$ svaki $a_i \leq 4$, po Dirichletovom principu su barem dva od $4^t + 1$ nizova

$$a_{n+2kt+1}, \dots, a_{n+2kt+t} \qquad \text{za} \quad k \in \{0, 1, \dots, 4^t\}$$

jednaki, tj. postoje $i \geq n$ i $d \geq t$ takvi da je

$$(a_{i+1}, \dots, a_{i+t}) = (a_{i+2d+1}, \dots, a_{i+2d+t}). \tag{6.17}$$

Stavimo

$$\begin{aligned} \alpha' &= [0, a'_1, a'_2, \dots] = [0, a_1, \dots, a_{i+2d}, a_{i+1}, a_{i+2}, \dots], \\ \alpha'' &= [0, a''_1, a''_2, \dots] = [0, a_1, \dots, a_{i+t}, a_{i+2d+t+1}, a_{i+2d+t+2}, \dots], \end{aligned}$$

tj. α' i α'' smo dobili iz α pomakom odgovarajućih podnizova udesno, odnosno ulijevo. Zbog $i \geq n$ vidimo da je $a''_j = a'_j = a_j$ za $1 \leq j \leq n$. Također, (6.17) povlači da se svaki niz duljine t parcijalnih kvocijenta u α', odnosno u α'' pojavljuje u razvoju od α. Stoga su $\alpha', \alpha'' \in C(T; a_1, \dots, a_n)$. Iz $\alpha = \max C(T; a_1, \dots, a_n)$ slijedi $\alpha' \leq \alpha$ i $\alpha'' \leq \alpha$. U kombinaciji sa (6.17) i usporedbom verižnih razlomaka od α i α', odnosno od α i α'' dobivamo za $i + t$ neparan

$$\begin{aligned} \alpha' \leq \alpha \quad &\Rightarrow \quad [a_{i+t+1}, a_{i+t+2}, \dots] \leq [a_{i+2d+t+1}, a_{i+2d+t+2}, \dots], \\ \alpha'' \leq \alpha \quad &\Rightarrow \quad [a_{i+2d+t+1}, a_{i+2d+t+2}, \dots] \leq [a_{i+t+1}, a_{i+t+2}, \dots], \end{aligned} \tag{6.18}$$

dok za paran $i + t$ obje nejednakosti na desnoj strani (6.18) mijenjaju smjer. No, u oba slučaja dobivamo $[a_{i+t+1}, a_{i+t+2}, \dots] = [a_{i+2d+t+1}, a_{i+2d+t+2}, \dots]$, pa α ima periodski razvoj u verižni razlomak. $\square$

Za korištenje Perronove metode određivanja maksimalnih rupa potrebni su ekstremalni elementi spektara uz uvjet isključenosti ili uključenosti određenih konačnih nizova. Pri tome nam može biti od koristi sljedeći Bumbyjev rezultat koji poopćuje Lemu 6.13.

Lema 6.16. *Za $A = (a_i)_{i \in Z} \in \mathbb{N}^{\mathbb{Z}}$ i fiksirani $n \in \mathbb{N}$, neka je $x = [a_0, a_{-1}, a_{-2}, \dots]$, $y = [a_n, a_{n+1}, a_{n+2}, \dots]$ te $f(z) = [0, a_1, \dots, a_{n-1}, z]$. Ako je $(a_1, \dots, a_{n-1})$ simetričan, onda je $x \geq y$ ako i samo ako je $\lambda_0(A) \geq \lambda_n(A)$. Nadalje, ako je $(a_1, \dots, a_{n-1})$ simetričan i $x \geq y$, onda je*

$$\begin{aligned} \lambda_n(A) \leq y + f(y) \leq x + f(x) \leq \lambda_0(A) \quad &\text{za neparan } n, \\ y + f(y) \leq \lambda_n(A) \leq \lambda_0(A) \leq x + f(x) \quad &\text{za paran } n. \end{aligned}$$

Dokaz. Najprije primijetimo da je $\lambda_0(A) = x + f(y)$ te, zbog simetrije danog bloka, $\lambda_n(A) = y + f(x)$. Nejednakost $x \geq y$ povlači za neparan n da je $f(x) \leq f(y)$, pa je $\lambda_n(A) \leq y + f(y)$ i $x + f(x) \leq \lambda_0(A)$ za neparan n. Za paran n vrijede obrnute nejednakosti.

Iz definicije od $f(z)$ imamo $f(z) = \frac{az+b}{cz+d}$, gdje su a, b, c, d neki nenegativni cijeli brojevi koji zadovoljavaju $ad - bc = \pm 1$. Stoga za derivaciju te funkcije vrijedi $|f'(z)| = |cz + d|^{-2} \leq 1$ za svaki $z \geq 1$ te dobivamo da su $z + f(z)$ i $z - f(z)$ rastuće funkcije u z na $[1, +\infty)$. Dakle, $x \geq y$ vrijedi ako i samo ako je $x - f(x) \geq y - f(y)$, odnosno $\lambda_0(A) = x + f(y) \geq y + f(x) = \lambda_n(A)$.

Za $x \geq y$ je $y + f(y) \leq x + f(x)$, pa za neparan n slijedi

$$\lambda_n(A) \leq y + f(y) \leq x + f(x) \leq \lambda_0(A),$$

što smo i trebali dokazati. Tvrdnja za paran n dobiva se analogno. $\square$

Propozicija 6.17. *Interval* $(L(\overline{2\,1\,2\,2}), L(\overline{1\,1\,2})) = (\frac{\sqrt{480}}{7}, \sqrt{10}) = (3.12984\ldots, 3.16227\ldots)$ *je maksimalna rupa u oba spektra.*

Dokaz. Najprije vidimo da je $L(\overline{2\,1\,2\,2}) = \frac{\sqrt{480}}{7} < \sqrt{10} = L(\overline{1\,1\,2})$.

Želimo pokazati da je $\min\{M(A) : 121$ se pojavljuje u $A\} = L(\overline{1\,1\,2})$. Pretpostavimo da se 121 pojavljuje u A, ali se 1212 ne pojavljuje u A i $\widetilde{A}$. Zbog Leme o kompaktnosti nije nikakvo smanjenje općenitosti pretpostaviti da je $M(A) = \lambda_0(A)$, a očito je dovoljno gledati $A \in \{1,2\}^{\mathbb{Z}}$ jer inače imamo $M(A) \geq \sqrt{13}$. Stoga je $(a_{-2}, a_{-1}, a_0, a_1, a_2) = (1,1,2,1,1)$, pa imamo $M(A) = \lambda_0(A) = [2,1,1,u] + [0,1,1,v]$, gdje je $0 < u \leq v < 1$. Primijenimo Lemu 6.16 uz $n = 3$, pa iz $\lambda_0(A) \geq \lambda_3(A)$ zbog simetrije od $(1,1)$ slijedi

$$[2,1,1,u] \geq v \qquad \text{i} \qquad M(A) = \lambda_0(A) \geq [2,1,1,u] + [0,1,1,2,1,1,u].$$

Vrijedi $[2,1,1,u] \geq v \geq u$, pa je $u \leq [\overline{2\,1\,1}]$ jer bi za $u > [\overline{2\,1\,1}]$ imali

$$u \leq [2,1,1,u] < [2,1,1,\overline{2,1,1}] = [\overline{2,1,1}] < u$$

što ne može biti. Zato je

$$M(A) \geq [2,1,1,u] + [0,1,1,2,1,1,u] \geq [\overline{2,1,1}] + [0,1,1,\overline{2,1,1}] = L(\overline{1\,1\,2}).$$

Ovdje smo iskoristili činjenicu da je $f(x) = x + [0,1,1,x]$ rastuća funkcija i da je $[2,1,1,u] \geq [\overline{2,1,1}]$.

Ako se 1212 pojavljuje u A ili $\widetilde{A}$, lako se vidi da je

$$M(A) \geq [2,1,2,\overline{2,1}] + [0,1,\overline{1,2}] = 3.28\ldots > L(\overline{1\,1\,2}).$$

Preostaje još dokazati da je $\max\{M(A) : A \in \{1,2\}^{\mathbb{Z}}$ i 121 se ne pojavljuje u $A\} = L(\overline{2\,1\,2\,2})$. Za takav niz A imamo $M(A) = \lambda_0(A) = [2,u] + [0,2,v]$ za neke $u, v \in (0,1)$. Budući da je 121 isključen, vrijedi $[2,u] \leq [\overline{2,1,2,2}]$. Koristeći Lemu 6.16 uz $n = 2$, iz $M(A) = \lambda_0(A) \geq \lambda_2(A)$ dobivamo $[2,u] \geq v$. Stoga je

$$M(A) \leq [2,u] + [0,2,2,u] \leq [\overline{2,1,2,2}] + [0,\overline{2,2,1,2}] = L(\overline{2\,1\,2\,2}).$$

Zaključujemo da je dani interval maksimalna rupa u oba spektra. $\square$

Hightower i Gbur su koristeći Perronovu metodu pokazali da je interval iz prethodne propozicije prvi element u beskonačnom nizu maksimalnih rupa

$$\big(M(A_{2n-1}), M(B_{2n})\big) \quad \text{i} \quad \big(M(B_{2n+1}), M(A_{2n})\big),$$

gdje je $B_m = \overline{1_m\,2}$ za $m \geq 2$ i

$$A_1 = \overline{2\,1\,2\,2}, \quad A_2 = \overline{1\,2\,1\,2\,2\,1\,2\,1}, \quad A_m = \overline{2\,1\,2\,2}\,1_{m-1}\,2\,1_m\,2\,1_{m-1}\,\overline{2\,2\,1\,2} \text{ za } m \geq 3$$

te da je svaki kraj gomilište oba spektra.

U idućoj tablici navodimo četiri maksimalne rupe poredane po veličini. Ostale rupe su prema Cusickovim rezultatima duljine manje od 0.025. Niz $A_0 \in \mathbb{N}^{\mathbb{Z}}$ naveden u drugom stupcu ima svojstvo da je $M(A_0)$ maksimalan po svim obostrano beskonačnim nizovima iz $\{1,2,3\}^{\mathbb{Z}}$ koji ne sadrže uzorke navedene u prvom stupcu (i njihove obrate). Niz $A_1 \in \mathbb{N}^{\mathbb{Z}}$ je takav da je $M(A_1)$ minimalan po svim nizovima iz $\mathbb{N}^{\mathbb{Z}}$ koji sadrže barem jedan od uzoraka iz prvog stupca. Duljina rupe je $M(A_1) - M(A_0)$.

uzorci	A_0	$M(A_0)$	A_1	$M(A_1)$	duljina
3	$\overline{1\,2}$	$\sqrt{12} = 3.464\ldots$	$\overline{3}$	$\sqrt{13} = 3.605\ldots$	$0.141\ldots$
$32, 31$	$\overline{3}$	$\sqrt{13} = 3.605\ldots$	$\overline{1\,2\,3\,3\,2\,1}$	$\frac{65+9\sqrt{3}}{22} = 3.663\ldots$	$0.057\ldots$
$3, 21212$	$\overline{1\,2\,1\,1\,1\,2}$	$\frac{\sqrt{1365}}{11} = 3.358\ldots$	$\overline{2\,1\,2\,1\,2}$	$\frac{\sqrt{290}}{5} = 3.405\ldots$	$0.047\ldots$
$3, 121$	$\overline{2\,1\,2\,2}$	$\frac{\sqrt{480}}{7} = 3.129\ldots$	$\overline{1\,2\,1}$	$\sqrt{10} = 3.162\ldots$	$0.032\ldots$

Cusick je pokazao da za svaku maksimalnu rupu (a, b) u $\mathcal{M}$ i niz $A \in \mathbb{N}^{\mathbb{Z}}$ takav da je $M(A) = \lambda_0(A) = a$ vrijedi da je A periodski od nekog mjesta ulijevo i od nekog mjesta udesno. Gayfulin je dokazao analogni rezultat kod $\mathcal{L}$. Nisu poznati takvi općeniti rezultati za desne krajeve maksimalnih rupa.

Kažemo da je iracionalan broj α *dostižan* ako nejednadžba $\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{L(\alpha)q^2}$ ima beskonačno mnogo rješenja. Lema 4.59 kaže da je svaki α za koji vrijedi $L(\alpha) < 3$ dostižan, a primjer nakon te leme pokazuje da postoje i dostižni α_1 i nedostižni α_2 takvi da je $L(\alpha_1) = L(\alpha_2) > 3$. Budući da je po Propoziciji 1.12 $L(\alpha) = \limsup_{i\to\infty} \lambda_i(\alpha)$, očito će α biti dostižan ako i samo ako vrijedi $\lambda_i(\alpha) \geq L(\alpha)$ za beskonačno mnogo prirodnih brojeva i. Pokažimo sada rezultat koji je dokazao Gayfulin 2017.

Teorem 6.18. *Ako $\lambda \in \mathcal{L}$ nije lijevi kraj nijedne maksimalne rupe u Lagrangeovom spektru, onda postoji dostižan realan broj α takav da je $L(\alpha) = \lambda$.*

Dokaz. Iz Teorema 3.7 slijedi da je $\mathcal{L}$ topološki zatvarač skupa

$$\{L(\alpha) \,:\, \alpha \in \mathbb{R} \text{ kvadratna iracionalnost}\}.$$

Stoga za $\lambda \in \mathcal{L}$ koji nije lijevi kraj nijedne maksimalne rupe u Lagrangeovom spektru mora vrijediti jedan od sljedeća dva slučaja:

(1) Postoji kvadratna iracionalnost γ takva da je $L(\gamma) = \lambda$.

(2) Postoji niz kvadratnih iracionalnosti $(\gamma_n)_{n\in\mathbb{N}}$ takav da je $\lim_{n\to\infty} L(\gamma_n) = \lambda$ i niz $(L(\gamma_n))_{n\in\mathbb{N}}$ je padajući.

Slučaj 1. Bez smanjenja općenitosti možemo možemo pretpostaviti da γ ima razvoj u verižni razlomak $\gamma = [0, \overline{P}]$, gdje je $P = (c_1, \ldots, c_n)$. Postoji $1 \leq j \leq n$ tako da je

$$\lim_{m\to\infty} \lambda_{j+mn}(\gamma) = L(\gamma) = [c_j, c_{j+1}, \ldots, c_n, \overline{P}] + [0, c_{j-1}, \ldots, c_1, \overline{\widetilde{P}}].$$

Ekvivalentni brojevi imaju isti Lagrangeov broj, pa za svaki konačan niz prirodnih brojeva R vrijedi $L(\gamma) = L([0, R, \overline{P}])$. Stavimo $\gamma_R = [0, R, \overline{P}]$ i neka je duljina $|R| = l$. Tada je

$$\lambda_{j+mn+l}(\gamma_R) = [c_j, c_{j+1}, \ldots, c_n, \overline{P}] + [0, c_{j-1}, \ldots, c_1, \underbrace{\widetilde{P}, \ldots, \widetilde{P}}_{m \text{ puta}}, \widetilde{R}].$$

Neka je R po volji tako da je $[0, c_{j-1}, \ldots, c_1, \widetilde{R}] > [0, c_{j-1}, \ldots, c_1, \overline{\widetilde{P}}]$. Tada je

$$[0, c_{j-1}, \ldots, c_1, \underbrace{\widetilde{P}, \ldots, \widetilde{P}}_{2m \text{ puta}}, \widetilde{R}] > [0, c_{j-1}, \ldots, c_1, \overline{\widetilde{P}}]$$

za svaki $m \in \mathbb{N}$. Stoga je

$$\lambda_{j+2mn+l}(\gamma_R) > L(\gamma) = L(\gamma_R)$$

za svaki $m \in \mathbb{N}$. Budući da je

$$\lim_{m\to\infty} \lambda_{j+mn+l}(\gamma_R) = L(\gamma) = L(\gamma_R) = \lambda,$$

teorem je dokazan u prvom slučaju.

Slučaj 2. Označimo periodski uzorak u verižnom razlomku od γ_n sa P_n. Možemo pretpostaviti bez smanjenja općenitosti da nijedan element u nijednom uzorku P_n nije veći od $L(\gamma) + 1$. Također, kao u prethodnom slučaju, možemo uzeti da period u γ_n počinje odmah nakon parcijalnog kvocijenta 0. Trebat će nam dvije pomoćne tvrdnje.

Tvrdnja 1. Neka je γ_n kvadratna iracionalnost $[0, \overline{P_n}]$ s periodskim uzorkom $P_n = (c_{n,1}, c_{n,2}, \ldots, c_{n,\ell(n)})$ duljine $\ell(n)$. Promotrimo j takav da je $1 \le j \le \ell(n)$ i

$$\begin{aligned}\lim_{m\to\infty} \lambda_{j+m\ell(n)}(\gamma_n) &= \lim_{m\to\infty} \Big([c_{n,j}, c_{n,j+1}, \ldots, c_{n,\ell(n)}, \overline{P_n}] + [0, c_{n,j-1}, \ldots, c_{n,1}, \underbrace{\widetilde{P_n}, \ldots, \widetilde{P_n}}_{m \text{ puta}}]\Big)\\ &= L(\gamma_n).\end{aligned}$$

Za proizvoljni $\varepsilon > 0$ postoji $N(\varepsilon) \in \mathbb{N}$ takav da za sve konačne ili beskonačne nizove R i S prirodnih brojeva, vrijedi

$$[c_{n,j}, c_{n,j+1}, \ldots, c_{n,\ell(n)}, \underbrace{P_n, \ldots, P_n}_{N(\varepsilon) \text{ puta}}, R] + [0, c_{n,j-1}, \ldots, c_{n,1}, \underbrace{\widetilde{P_n}, \ldots, \widetilde{P_n}}_{N(\varepsilon) \text{ puta}}, S] > L(\gamma_n) - \varepsilon.$$

Tvrdnja 1. slijedi izravno iz svojstava verižnih razlomaka, točnije Teorema 1.8.b.

Iduća tvrdnja pokazuje da duljina od P_n teži u beskonačnost sa n.

Tvrdnja 2. Neka je $(\gamma_n)_{n\in\mathbb{N}}$ niz kvadratnih iracionalnosti $\gamma_n = [0, \overline{P_n}]$ s periodskim uzorcima $P_n = (c_{n,1}, c_{n,2}, \ldots, c_{n,\ell(n)})$ duljine $\ell(n)$. Neka je λ iracionalan broj koji nije kvadratna iracionalnost takav da je $\lim_{n\to\infty} L(\gamma_n) = \lambda$ i niz $(L(\gamma_n))_n$ je padajući. Tada $\ell(n) \to \infty$ kad $n \to \infty$. Štoviše, postoji beskonačan podniz $(P'_n)_n$ od $(P_n)_n$ takav da se za svaki $n \in \mathbb{N}$ prvih n elemenata u uzorcima $P'_n, P'_{n+1}, \ldots$ podudaraju.

Za dokaz tvrdnje pretpostavimo najprije da $\ell(n)$ ne teži u beskonačnost. Tada postoji prirodan broj M takav da za beskonačno $n \in \mathbb{N}$ imamo $\ell(n) = M$. Kako ima samo konačno nizova duljine M s elementima iz $\{1, 2, \ldots, \lfloor L(\gamma)\rfloor + 1\}$, to postoji uzorak P' duljine M koji se u $(P_n)_{n\in\mathbb{N}}$ pojavljuje beskonačno mnogo puta. Neka su $(i_n)_{n\in\mathbb{N}}$ indeksi takvi da je $P_{i_n} = P'$. Tada je

$$\lambda = \lim_{n\to\infty} L(\gamma_n) = \lim_{n\to\infty} L([0, \overline{P_n}]) = \lim_{n\to\infty} L([0, \overline{P_{i_n}}]) = L([0, \overline{P'}]) > \lambda,$$

što daje kontradikciju.

Preostaje još pokazati drugi dio u tvrdnji. Budući da su svi elementi u P_n-ovima ograničeni, to za svaki $m \in \mathbb{N}$ postoji niz $c'_1, \ldots, c'_m$ i beskonačan niz indeksa $(i_n)_{n\in\mathbb{N}}$ takav da je $(c_{i_n,1}, c_{i_n,2}, \ldots, c_{i_n,m}) = (c'_1, c'_2, \ldots, c'_m)$ za svaki $n \in \mathbb{N}$. Konstrukcija podniza $(P'_n)_n$ je sada jasna, pa ispuštamo detalje.

Završimo sada dokaz Teorema 6.18. Promotrimo niz $(P'_n)_{n\in\mathbb{N}}$ iz Tvrdnje 2 i odgovarajući niz kvadratnih iracionalnosti $\widehat{\gamma}_n = [0, \overline{P'_n}]$. Stavimo $\varepsilon_n = \frac{1}{3}(L(\widehat{\gamma}_n) - \lambda)$ i $N_n = N(\varepsilon_n)$. Definirajmo iracionalan broj

$$\widehat{\gamma} = [0, \underbrace{P'_1, \ldots, P'_1}_{2N_1+1 \text{ puta}}, \underbrace{P'_2, \ldots, P'_2}_{2N_2+1 \text{ puta}}, \ldots, \underbrace{P'_n, \ldots, P'_n}_{2N_n+1 \text{ puta}}, \ldots].$$

Tvrdnja 1 povlači da postoji beskonačno mnogo $j \in \mathbb{N}$ takvih da je $\lambda_j(\widehat{\gamma}) > \lambda$. Tvrdnja 2 povlači da je $L(\widehat{\gamma}) = \lambda$, pa smo konstruirali traženi broj. □

Gayfulin je pokazao da broj $\lambda_0 = [3\,3\,3\,2\,1\,\overline{1\,2}] + [0\,2\,1\,\overline{1\,2}]$ pripada $\mathcal{L}$, ali ne postoji nijedan dostižan α takav da je $L(\alpha) = \lambda_0$. Naravno, prema prethodnom teoremu, λ_0 je lijevi kraj jedne maksimalne rupe u $\mathcal{L}$ što je već otprije bilo poznato. Isti autor dao je karakterizaciju lijevih krajeva rupa u $\mathcal{L}$ koji imaju svojstva kao λ_0.

6.3 Razlika spektara

Budući da se Lagrangeov i Markovljev spektar podudaraju na $(0, 3]$ i na Hallovoj zraci, dugo nakon Markovljevih radova mislilo se da je to isti skup. No Freiman je 1968. dao primjer realnog broja koji je u $\mathcal{M}$, ali nije u $\mathcal{L}$

$$\sigma = M(S) = \lambda_0(S) = 3.1181201781599\ldots, \qquad S = \overline{2_2\,1\,2_2\,1_2\,2_2}\,1_2\,\overline{2_2\,1_2\,2_2\,1\,2_2}, \tag{6.19}$$

pri čemu dva elementa 1_2 koji nisu u periodima imaju indekse 17 i 18 u $S \in \mathbb{N}^{\mathbb{Z}}$. Mi ćemo pokazati kasniji Freimanov rezultat za što definiramo sljedeće brojeve

$$\begin{aligned} \alpha_\infty &= \lambda_0(A_\infty) = [2\,\overline{1_2\,2_3\,1\,2}] + [0\,1\,2_3\,1_2\,2\,1\,\overline{2}] = 3.2930442654\ldots, \\ \alpha_n &= \lambda_0(A_n) = [2\,\overline{1_2\,2_3\,1\,2}] + [0\,1\,2_3\,1_2\,2\,1\,2_n\,\overline{1\,2\,1_2\,2_3}] \quad \text{za } n \geq 4. \end{aligned}$$

Trebat će nam sljedeće leme.

Lema 6.19. *Ako obostrano beskonačni niz $B \in \{1, 2\}^{\mathbb{Z}}$ sadrži bilo koji od sljedećih uzoraka*

$$\begin{array}{lrcl}
1) & 2\ 1 & 2^* & 1\ 2 \\
2) & 2\ 1 & 2^* & 1\ 1\ 1 \\
3) & 1\ 2\ 1 & 2^* & 1\ 1 \\
4) & 2\ 2\ 2\ 1 & 2^* & 1\ 1\ 2\ 2\ 1 \\
5) & 2\ 1\ 2\ 2\ 2\ 1 & 2^* & 1\ 1\ 2\ 2\ 2 \\
6) & 1\ 1\ 2\ 2\ 2\ 1 & 2^* & 1\ 1\ 2\ 2\ 2\ 2 \\
7) & 1\ 1\ 2\ 2\ 2\ 1 & 2^* & 1\ 1\ 2\ 2\ 2\ 1\ 1 \\
8) & 1\ 1\ 1\ 2\ 2\ 2\ 1 & 2^* & 1\ 1\ 2\ 2\ 2\ 1\ 2 \\
9) & 2\ 1\ 1\ 2\ 2\ 2\ 1 & 2^* & 1\ 1\ 2\ 2\ 2\ 1\ 2\ 2 \\
10) & 2\ 2\ 1\ 1\ 2\ 2\ 2\ 1 & 2^* & 1\ 1\ 2\ 2\ 2\ 1\ 2\ 1 \\
11) & 1\ 1\ 2\ 1\ 1\ 2\ 2\ 2\ 1 & 2^* & 1\ 1\ 2\ 2\ 2\ 1\ 2\ 1\ 1\ 2 \\
12) & 2\ 2\ 1\ 2\ 1\ 1\ 2\ 2\ 2\ 1 & 2^* & 1\ 1\ 2\ 2\ 2\ 1\ 2\ 1\ 1\ 2\ 1,
\end{array}$$

onda je $\lambda_j(B) > \alpha_\infty + 10^{-8}$, pri čemu je b_j element označen zvjezdicom.

Dokaz. Ako se pojavljuje blok 1), onda je

$$\lambda_j(B) \geq [2\,1\,2\,\overline{2\,1}] + [0\,1\,2\,\overline{2\,1}] > [2\,1\,2] + [0\,1\,2] = \frac{10}{3} > \alpha_\infty + 10^{-8}.$$

Ako se pojavljuje blok 2), onda je

$$\lambda_j(B) \geq [2\,1\,1\,1\,\overline{1\,2}] + [0\,1\,2\,\overline{2\,1}] = 3.31491832\ldots > \alpha_\infty + 10^{-8}.$$

Sličnim računom provjeravamo i sve ostale slučajeve. $\square$

Lema 6.20. *Ako obostrano beskonačni niz $B \in \{1,2\}^{\mathbb{Z}}$ sadrži bilo koji od sljedećih uzoraka*

$$\begin{array}{ll}
13) & 1^* \\
14) & 2\ 2^* \\
15) & 1\ 1\ 2^*\ 1\ 1 \\
16) & 2\ 2\ 1\ 2^*\ 1\ 1\ 2\ 1 \\
17) & 1\ 2\ 2\ 1\ 2^*\ 1\ 1\ 2 \\
18) & 2\ 2\ 2\ 2\ 1\ 2^*\ 1\ 1\ 2\ 2\ 2,
\end{array}$$

onda je $\lambda_j(B) < \alpha_\infty - 10^{-5}$, pri čemu je b_j element označen zvjezdicom.

Dokaz. Primjerice, ako se pojavljuje blok 15), imamo

$$\lambda_j(B) \le 2 + 2[0\,1\,1\,\overline{1,2}] = 3.267949\ldots < \alpha_\infty - 10^{-5}.$$

Sličnim računom provjeravamo i sve ostale slučajeve. □

Teorem 6.21. *Za svaki $n \ge 4$ je $\alpha_n \in \mathcal{M} \setminus \mathcal{L}$ te $\alpha_\infty \in \mathcal{M} \setminus \mathcal{L}$. Vrijedi $\lim_{n\to\infty} M(A_n) = \alpha_\infty$. Imamo $M(A_\infty) = \lambda_i(A_\infty)$ samo za $i = 0$, dok za $n \ge 4$ jednakost $M(A_n) = \lambda_i(A_n)$ vrijedi samo za $i \in \{0, -17-n\}$.*

Dokaz. Neka je $B = (b_i)_{i\in\mathbb{Z}}$ niz u $\{1,2\}^{\mathbb{Z}}$ takav da je za neki m

$$\lambda_m(B) \ge \alpha_\infty - 10^{-8} \quad \text{te} \tag{6.20}$$

$$\lambda_i(B) \le \alpha_\infty + 10^{-8} \quad \text{za sve } i \text{ takve da je } |i| \ge m. \tag{6.21}$$

Po Lemi 6.19 i (6.21) nijedan od nizova 1) do 12) nije dozvoljen u B za $|j| \ge m$. Isto tako, Lema 6.20 i (6.20) povlače da nizovi 13) do 18) nisu dozvoljeni za $j = m$.

Zabrana nizova 13) i 14) daje $b_m = 2$ i $b_{m-1} = b_{m+1} = 1$. Nadalje, iz zabrane nizova 1) i 15) te obrtanjem B ako je potrebno, dobivamo da je $b_{m-2} = 2$ i $b_{m+2} = 1$. Cijeli postupak određivanja vrijednosti b_{m+i} uzastopnim korištenjem Lema 6.19 i 6.20 dan je u idućoj tablici. Drugi redak u tablici navodi broj odgovarajućeg zabranjenog podniza iz tih lema pomoću kojeg određujemo pojedini b_{m+i}. Tako dobivamo b_{m+i} za $-10 \le i \le 11$.

i	0	1	-1	2	-2	3	-3	4	-4
br.	13	14	14	1,15	1,15	2	3	16	17
b_{m+i}	2	1	1	1	2	2	2	2	2

i	5	-5	-6	6	7	-7	8	-8	9	10	-9	-10	11
br.	4	18	5	6	7	8	9	10	1	2	11	3	12
b_{m+i}	2	1	1	1	2	2	1	1	1	2	2	2	2

Stoga je

$$B = \ldots 2\,2\,1\,2\,1\,1\,2\,2\,2\,1\,2^*\,1\,1\,2\,2\,2\,1\,2\,1\,1\,2\,2\ldots = \ldots 2\,2\,1\,2\,1_2\,2_3\,1\,2^*\,1_2\,2_3\,1\,2\,1_2\,2\,2\ldots,$$

gdje je b_m naznačen zvjezdicom. Iz prethodne tablice vidimo da su samo podnizovi 1) do 12) korišteni u određivanju b_{m+i} za $5 \le i \le 11$. Zato (6.21) i Lema 6.19 osiguravaju da su, počnemo li s b_{m+7} umjesto b_m, određeni b_i za $12 \leq i \leq 18$ te nastavivši tako dalje dobivamo da je B oblika

$$B = \ldots 2\,2\,1^*\,\overline{2\,1_2\,2_3\,1}, \tag{6.22}$$

gdje zvjezdica označava b_{m-8}.

Pretpostavimo sada da je za neki $n \geq 4$ ili $n = \infty$ broj α_n element od $\mathcal{L}$ i da je B niz za koji je $\alpha_n = L(B) = \lim_{i\to\infty} \lambda_{k(i)}(B)$. Uzimajući ako je potrebno podniz od $(k(i))_{i\in\mathbb{N}}$, možemo pretpostaviti da je $(k(i))_i$ monoton niz. Budući da je $|\alpha_n - \alpha_\infty| < 10^{-8}$, postoji $I \in \mathbb{N}$ takav da za sve $i \geq I$ nejednakosti (6.20) i (6.21) vrijede za svaki $m = k(i)$. Stoga B ima oblik iz (6.22), pri čemu zvjezdica označava proizvoljni $b_{k(i)-8}$ za $i \geq I$. Kako je $(k(i))_{i\in\mathbb{N}}$ monoton, time smo dobili da je $\alpha_n = L(B) = L(\overline{2\,1_2\,2_3\,1}) < \alpha_\infty - 10^{-8}$ što nije istina. Zaključujemo da nijedan od brojeva α_n ne leži u $\mathcal{L}$.

Treba još dokazati drugu tvrdnju ovog teorema. Pokažimo da je $M(A_\infty) = \lambda_i(A_\infty)$ samo za $i = 0$. Znamo da je

$$A_\infty = \overline{2}\,1\,2\,1_2\,2_3\,1\,2^*\,\overline{1_2\,2_3\,1\,2} \qquad (* \text{ označava poziciju s indeksom } 0).$$

Budući da je $\alpha_\infty = \lambda_0(A_\infty) = [2, 1\ldots] + [0, 1, 2\ldots]$ što daje najveću vrijednost uz verižne razlomke kojima odabiremo naznačeni broj parcijalnih kvocijenata na početku, dovoljno je gledati gdje se u A_∞ pojavljuje $2\,1\,2^*\,1$ ili $1\,2^*\,1\,2$. Vidimo da je to za zvjezdicu na mjestima s indeksom $7k$, $k \in \{-1, 0\} \cup \mathbb{N}$. Svi $\lambda_i(A_\infty)$ imaju isti pribrojnik $[\overline{2\,1_2\,2_3\,1}]$ koji odgovara desnom dijelu niza, a drugi pribrojnik koji odgovara lijevom dijelu ima početne parcijalne kvocijente

za $i = -7$	$0\,1\,2\,2\,2\,2\ldots$		
za $i = 0$	$0\,1\,2\,2\,2\,1\,1\,2\,1\,2\,2\,2\,2\ldots$		
za $i = 7k,\ k \geq 1$	$0\,1\,2\,2\,2\,1\,1\,2\,1\,2\,2\,2\,1\ldots$		
indeks gdje se razlikuju		5	12

pa je očito $M(A_\infty) = \lambda_i(A_\infty)$ samo za $i = 0$. Slično za α_n treba gledati $\lambda_i(A_n)$ samo za $i \in \{0, 7k, -n-10-7k : k \in \mathbb{N}\}$, pa dobivamo odgovarajući zaključak teorema pri čemu detalje ispuštamo. □

Sljedeći teorem daje neke jednostavne dovoljne uvjete da broj α bude u $\mathcal{M}$, ali ne u $\mathcal{L}$.

Teorem 6.22. *Ako je $\alpha \in \mathcal{M}$ i vrijede oba uvjeta*

(1) α je izolirana točka u $\mathcal{M}$,

(2) ako je $M(A) = \alpha$ za neki $A \in \mathbb{N}^{\mathbb{Z}}$, onda je $L(A) \neq M(A)$,

onda $\alpha \notin \mathcal{L}$.

Dokaz. Pretpostavimo da je $\alpha \in \mathcal{L}$. Iz uvjeta (1) i Teorema 3.7 slijedi da postoji barem jedan čisto periodski obostrano beskonačan niz A takav da je $M(A) = \alpha$. No, budući je A čisto periodski, imamo $L(A) = M(A)$, što je u kontradikciji s uvjetom (2). □

Računom sličnim onom koji je dan u dokazu Teorema 6.21 dobiva se da je σ definiran u (6.19) izolirana točka u $\mathcal{M}$ te da je S jedini obostrano beskonačni niz s Markovljevim supremumom jednakim σ. Tako Teorem 6.22 pokriva originalni Freimanov primjer broja u $\mathcal{M} \setminus \mathcal{L}$. S druge strane, α_∞ iz Teorema 6.21 je limes niza $(\alpha_n)_n$ elemenata iz $\mathcal{M}$, pa uvjet (1) iz Teorema 6.22 očito nije nužan.

Iz metričkih rezultata Matheusa i Moreire objavljenih 2018/19. slijedi da skup $\mathcal{M} \setminus \mathcal{L}$ ima neprebrojivo mnogo elemenata. Ovi su autori također našli brojeve iz $\mathcal{M} \setminus \mathcal{L}$ veće od $\sqrt{12} = 3.464\ldots$ (štoviše, neprebrojivo elemenata u intervalu $(3.7, 3.71)$) i time opovrgnuli slutnju Cusicka koji je mislio da se spektri na $[\sqrt{12}, +\infty)$ podudaraju (vidi Propoziciju 6.12 za podsjećanje gdje se pojavljuje $\sqrt{12}$). Zajedno s još dvojicom koautora isti su autori nedavno konstruirali padajući niz brojeva iz $\mathcal{M}$ koji konvergira prema 3 te pokazali da prva četiri člana toga niza leže u $\mathcal{M} \setminus \mathcal{L}$. Kada bi uspjeli dokazati da ni ostali članovi niza nisu iz $\mathcal{L}$, slijedilo bi da $\mathcal{M} \setminus \mathcal{L}$ nije zatvoren skup jer je, kao što smo vidjeli u Teoremu 1.16, broj 3 u Lagrangeovom spektru.

Poglavlje 7

Metrički rezultati

Sada ćemo proučiti Lebesgueovu mjeru i Hausdorffovu dimenziju određenih dijelova spektara.

7.1 Dio spektra s Lebesgueovom mjerom nula

Dokazat ćemo sljedeći Hallov rezultat.

Teorem 7.1. *Dio Markovljevog spektra ispod* $\sqrt{10} = L(\overline{1\,1\,2}) = 3.16227\ldots$, *odnosno* $\mathcal{M} \cap (0, \sqrt{10})$ *ima Lebesgueovu mjeru nula.*

Dokaz. Prema Propoziciji 6.17, za $M(A) < \sqrt{10}$ trebamo promatrati samo nizove $A \in \mathbb{N}^{\mathbb{Z}}$ u kojima je svaki element iz $\{1, 2\}$ i ne pojavljuje se blok 121. Neka je U skup svih iracionalnih brojeva u intervalu $[0, 1]$ kojima je svaki parcijalni kvocijent (osim nultog) jednak 1 ili 2 i uzorak 121 se ne pojavljuje u nizu parcijalnih kvocijenata. Dokazat ćemo da je suma $U + U$ mjere nula.

Podsjetimo se rezultata iz Teorema 1.8.b koji će nam trebati za određivanje duljina intervala u nastavku. Neka su dani $\alpha = [b_0, b_1, \ldots, b_n, \alpha_{n+1}]$ i $\beta = [b_0, b_1, \ldots, b_n, \beta_{n+1}]$ te neka su $\frac{p_{n-1}}{q_{n-1}}$ i $\frac{p_n}{q_n}$ zadnje dvije konvergente od $[b_0, b_1, \ldots, b_n]$. Tada je

$$|\alpha - \beta| = \frac{|\alpha_{n+1} - \beta_{n+1}|}{q_n^2(\alpha_{n+1} + r)(\beta_{n+1} + r)}, \tag{7.1}$$

gdje je $r = [0, b_n, b_{n-1}, \ldots, b_1]$.

Budući da U leži u intervalu od $[0\,\overline{2\,1\,2\,2}]$ do $[0\,\overline{1\,2\,2\,2}]$, provodimo postupak uzastopne subdivizije za dobivanje Cantorovog skupa. Slično smo radili kod određivanja Hallove zrake. Prva subdivizija daje nam dva intervala

$$\begin{aligned} \big[[0\,2\,\overline{1\,2\,2\,2}], [0\,2\,\overline{2\,1\,2\,2}]\big] &= [0.3693\ldots, 0.4220\ldots] \qquad \text{i} \\ \big[[0\,1\,\overline{1\,2\,2\,2}], [0\,1\,\overline{2\,2\,2\,1}]\big] &= [0.5855\ldots, 0.7077\ldots]. \end{aligned}$$

Općenito, u ovom postupku imamo dva tipa intervala. Intervali prvog tipa su segmenti s krajevima oblika

$$[0, b_1, \ldots, b_{n-1}, 1, \overline{1, 2, 2, 2}] \quad \text{i} \quad [0, b_1, \ldots, b_{n-1}, 1, \overline{2, 2, 2, 1}]$$

i to u ovom poretku ako je je n neparan, a u obratnom ako je n paran.

Primijetimo da bi $b_n = 2$, $b_{n-1} = 1$ zbog zabrane uzorka 121 nužno povlačilo da je $b_{n+1} = 2$, pa ćemo za $b_n = 2$ promatrati samo slučaj $b_{n-1} = 2$. Intervali drugog tipa su segmenti s krajevima oblika

$$[0, b_1, \dots, b_{n-2}, 2, 2, \overline{1, 2, 2, 2}] \quad \text{i} \quad [0, b_1, \dots, b_{n-2}, 2, 2, \overline{2, 1, 2, 2}].$$

Ako je $I_1 \cup I_2$ subdivizija intervala
I prvog tipa koji ima krajeve $[0, b_1, \dots, b_{n-1}, 1, \overline{1, 2, 2, 2}]$ i $[0, b_1, \dots, b_{n-1}, 1, \overline{2, 2, 2, 1}]$, onda
I_1 ima krajeve $[0, b_1, \dots, b_{n-1}, 1, 1, \overline{2, 2, 2, 1}]$ i $[0, b_1, \dots, b_{n-1}, 1, 1, \overline{1, 2, 2, 2}]$, a
I_2 ima krajeve $[0, b_1, \dots, b_{n-1}, 1, 2, \overline{2, 1, 2, 2}]$ i $[0, b_1, \dots, b_{n-1}, 1, 2, \overline{2, 2, 1, 2}]$.
Možemo izračunati omjere duljina ovih intervala

$$\frac{|I_1|}{|I|} = \frac{(13 - \sqrt{120})(\sqrt{120} + 6 + 7r)}{7(\sqrt{120} + 1 + 7r)}, \qquad \frac{|I_2|}{|I|} = \frac{(12 - \sqrt{120})(\sqrt{120} + 6 + 12r)}{30(\sqrt{120} + 8 + 8r)},$$

gdje je $r = [0, 1, b_{n-1}, \dots, b_1]$. Budući da su $\frac{|I_1|}{|I|}$ i $\frac{|I_2|}{|I|}$ redom padajuća i rastuća funkcija u r, iz $[0, 1, 1] = \frac{1}{2} \leq r \leq \frac{5}{7} = [0, 1, 2, 2]$ dobivamo

$$\frac{|I_1|}{|I|} \leq \frac{33 - \sqrt{120}}{57} < 0.387, \qquad \frac{|I_2|}{|I|} \leq \frac{1083 - 92\sqrt{120}}{2085} < 0.037.$$

Potpuno analogno, ako je $I_1 \cup I_2$ subdivizija intervala
I drugog tipa s krajevima $[0, b_1, \dots, b_{n-2}, 2, 2, \overline{1, 2, 2, 2}]$ i $[0, b_1, \dots, b_{n-2}, 2, 2, \overline{2, 1, 2, 2}]$, onda
I_1 ima krajeve $[0, b_1, \dots, b_{n-2}, 2, 2, 1, \overline{2, 2, 2, 1}]$ i $[0, b_1, \dots, b_{n-2}, 2, 2, 1, \overline{1, 2, 2, 2}]$, a
I_2 ima krajeve $[0, b_1, \dots, b_{n-2}, 2, 2, 2, \overline{2, 1, 2, 2}]$ i $[0, b_1, \dots, b_{n-2}, 2, 2, 2, \overline{1, 2, 2, 2}]$.
Možemo izračunati omjere duljina ovih intervala

$$\frac{|I_1|}{|I|} = \frac{(55 - 5\sqrt{120})(\sqrt{120} + 8 + 7r)}{\sqrt{120} + 1 + 7r}, \qquad \frac{|I_2|}{|I|} = \frac{(12 - \sqrt{120})(\sqrt{120} + 6 + 12r)}{6(\sqrt{120} + 8 + 8r)},$$

gdje je $r = [0, 2, 2, b_{n-2}, \dots, b_1]$. Budući da su $\frac{|I_1|}{|I|}$ i $\frac{|I_2|}{|I|}$ redom padajuća i rastuća funkcija u r, iz $[0, 2, 2] = \frac{2}{5} \leq r \leq \frac{3}{7} = [0, 2, 2, 1]$ dobivamo

$$\frac{|I_1|}{|I|} \leq \frac{510 - 35\sqrt{120}}{377} < 0.336, \qquad \frac{|I_2|}{|I|} \leq \frac{55 - 4\sqrt{120}}{65} < 0.173.$$

Nastavljajući ovaj postupak, za subdiviziju $I_1 \cup I_2$ intervala I prvog tipa dobivamo daljnje subdivizije $I_{11} \cup I_{12}$ od I_1 i $I_{21} \cup I_{22}$ od I_2 te omjere

$$\frac{|I_{11}|}{|I|} < 0.387 \cdot 0.387 < 0.150, \qquad \frac{|I_{12}|}{|I|} < 0.387 \cdot 0.037 < 0.015,$$
$$\frac{|I_{21}|}{|I|} < 0.037 \cdot 0.336 < 0.013, \qquad \frac{|I_{22}|}{|I|} < 0.037 \cdot 0.173 < 0.007.$$

Daljnje subdivizije intervala drugog tipa daju omjere

$$\frac{|I_{11}|}{|I|} < 0.336 \cdot 0.387 < 0.131, \qquad \frac{|I_{12}|}{|I|} < 0.336 \cdot 0.037 < 0.013,$$
$$\frac{|I_{21}|}{|I|} < 0.173 \cdot 0.336 < 0.059, \qquad \frac{|I_{22}|}{|I|} < 0.173 \cdot 0.173 < 0.030.$$

Stoga iz svakog segmenta I nakon dva koraka ovog procesa dobivamo četiri podsegmenta čija suma duljina je manja od $0.233|I|$.

Pogledajmo kako ovaj proces dobivanja Cantorovog skupa utječe na skup $U+U$. U svakom koraku imamo uniju skupova od kojih je svaki suma dva intervala iz postupka, primjerice $I+I'$. Podjelom svakog od ovih intervala na četiri intervala, npr. I na I_i, a I' na I'_i za $i \in \{1,\ldots,4\}$, dobivamo novu sumu

$$\bigcup_{i=1}^{4} I_i + \bigcup_{j=1}^{4} I'_j = \bigcup_{(i,j)\in\{1,\ldots,4\}^2} (I_i + I'_j)$$

za koju vrijedi (Lebesgueovu mjeru također označavamo $|\cdot|$)

$$\begin{aligned}\frac{\left|\bigcup_{(i,j)}(I_i+I'_j)\right|}{|I+I'|} &\leq \frac{\sum_{(i,j)}|I_i+I'_j|}{|I+I'|} = \frac{\sum_{(i,j)}|I_i|+|I'_j|}{|I|+|I'|} = \frac{4\big(\sum_i |I_i| + \sum_j |I'_j|\big)}{|I|+|I'|} \\ &< 4\,\frac{0.233|I|+0.233|I'|}{|I|+|I'|} = 0.932.\end{aligned}$$

Kada n puta ponovimo ovu subdiviziju, ukupna duljina intervalâ smanjila se barem $(0.932)^n$ puta, što znači da sumu duljina intervala koji pokrivaju skup $U+U$ možemo učiniti po volji malom, a to dokazuje da dani skup ima mjeru nula. $\square$

Napomenimo da smo u prethodnom dokazu činjenicu da je $M(A)$ element Markovljevog spektra koristili samo da ograničimo nizove A na one koji sadrže jedino 1 i 2 i ne sadrže 121. Puno svojstvo ovog supremuma, tj. da je $M(A) = \sup_{i\in\mathbb{Z}} \lambda_i(A)$ nismo koristili.

Granica intervala iz Teorema 7.1 je više puta podizana, a zadnju poznatu verziju dao je Bumby 1982. koji je pokazao da je $\mathcal{M}\cap(0, 3.33437\ldots)$ Lebesgueove mjere nula. Pritom je koristio dosta izračuna na računalu.

7.2 Hausdorffova dimenzija

Za razlikovanje između skupova Lebesgueove mjere nula, kakvi su mnogi skupovi koji se pojavljuju u teoriji brojeva, treba nam finija mjera. Hausdorffova ideja sastojala se u tome da mjeri veličinu skupa pokrivajući ga beskonačnom prebrojivom familijom skupova ograničenog dijametra i zatim gleda što se događa kada maksimalni dijametar skupova iz tog pokrivača teži u 0. Iznesimo najprije definiciju i osnovna svojstva Hausdorffove mjere i dimenzije uglavnom izostavljajući dokaze.

Dijametar nepraznog skupa U u $\mathbb{R}^n$ je supremum udaljenosti točaka iz U te ga označavamo $\operatorname{diam} U$. Familiju skupova $\{U_i : i \in \mathbb{N}\}$ zovemo δ-pokrivač skupa $E \subset \mathbb{R}^n$ ako je $E \subset \bigcup_{i\in\mathbb{N}} U_i$ i $0 < \operatorname{diam}(U_i) < \delta$ za svaki i. Za realne brojeve $s \geq 0$ i $\delta > 0$ definiramo

$$\mathcal{H}^s_\delta(E) = \inf\Big\{ \sum_{i=1}^{\infty} (\operatorname{diam} U_i)^s \;:\; \{U_i : i\in\mathbb{N}\} \text{ je } \delta\text{-pokrivač od } E \Big\}. \tag{7.2}$$

Kada se δ smanjuje, sve je manje dopuštenih pokrivača od E u (7.2), pa se infimum $\mathcal{H}^s_\delta$ povećava i zato ima limes kad δ teži u 0. Stavimo

$$\mathcal{H}^s(E) = \lim_{\delta\to 0} \mathcal{H}^s_\delta(E) = \sup_{\delta>0} \mathcal{H}^s_\delta(E)$$

i taj broj koji je iz $[0, +\infty]$ zovemo *s-dimenzionalna Hausdorffova mjera* skupa E. Pokazuje se da je $\mathcal{H}^s$ zaista (vanjska) mjera. Nije teško vidjeti da je $\mathcal{H}^s(\emptyset) = 0$, da za $E \subset F$ imamo $\mathcal{H}^s(E) \leq \mathcal{H}^s(F)$ i da za prebrojivu familiju skupova $\{E_i : i \in \mathbb{N}\}$ vrijedi

$$\mathcal{H}^s\Big(\bigcup_{i=1}^{\infty} E_i\Big) \leq \sum_{i=1}^{\infty} \mathcal{H}^s(E_i),$$

a nešto je teže vidjeti da za familiju u parovima disjunktnih Borelovih skupova $\{E_i\}$ u zadnjoj nejednakosti vrijedi jednakost.

Može se pokazati da je za podskupove od $\mathbb{R}^n$ n-dimenzionalna Hausdorffova mjera samo konstantni višekratnik n-dimenzionalne Lebesgueove mjere, tj. uobičajenog n-dimenzionalnog volumena, a navedena konstanta je recipročna vrijednost volumena n-dimenzionalne kugle dijametra 1. Budući da će nas zanimati samo skupovi realnih brojeva, dovoljno je spomenuti da se $\mathcal{H}^1$ i Lebesgueova mjera na $\mathbb{R}$ podudaraju što je očito i iz definicija tih dviju mjera. Također je jasno da je $\mathcal{H}^0(E)$ broj elemenata skupa E.

Lako je vidjeti da za $E \subset \mathbb{R}^n$ i preslikavanje $f : E \to \mathbb{R}^m$ koje za neke $c > 0$ i $\alpha > 0$ zadovoljava Hölderov uvjet s eksponentom α

$$|f(x) - f(y)| \leq c|x - y|^{\alpha} \qquad \text{za sve } x, y \in E \tag{7.3}$$

vrijedi za svaki $s > 0$

$$\mathcal{H}^{s/\alpha}(f(E)) \leq c^{s/\alpha}\mathcal{H}^s(E).$$

Funkcije koja zadovoljavaju uvjet (7.3) su očito neprekidne i zovemo ih α-Hölder neprekidnim. Za $\alpha = 1$, kada je f Lipschitzovo preslikavanje, tj. $|f(x) - f(y)| \leq c|x - y|$ za sve $x, y \in E$, imamo

$$\mathcal{H}^s(f(E)) \leq c^s\mathcal{H}^s(E). \tag{7.4}$$

Zbog teorema srednje vrijednosti su diferencijabilne funkcije s omeđenom derivacijom nužno Lipschitzove, pa za takve funkcije vrijedi (7.4). Ako je f izometrija, tj. $|f(x) - f(y)| = |x - y|$, onda je $\mathcal{H}^s(f(E)) = \mathcal{H}^s(E)$, pa su Hausdorffove mjere primjerice translacijski invarijantne, tj. $\mathcal{H}^s(E + z) = \mathcal{H}^s(E)$ za svaki $z \in \mathbb{R}^n$.

Za $t > s \geq 0$, $\delta < 1$ i δ-pokrivač $\{U_i\}$ od E je

$$\sum_{i=1}^{\infty}(\operatorname{diam} U_i)^t \leq \delta^{t-s}\sum_{i=1}^{\infty}(\operatorname{diam} U_i)^s,$$

pa iz (7.2) slijedi $\mathcal{H}^t_\delta(E) \leq \delta^{t-s}\mathcal{H}^s_\delta(E)$. Pustimo li $\delta \to 0$, dobivamo sljedeću tvrdnju. Ako za neki $s \geq 0$ vrijedi $\mathcal{H}^s(E) < +\infty$, onda je $\mathcal{H}^{s+\varepsilon}(E) = 0$ za svaki $\varepsilon > 0$, a ako za neki $s > 0$ vrijedi $\mathcal{H}^s(E) > 0$, onda je $\mathcal{H}^{s-\varepsilon}(E) = +\infty$ za svaki $\varepsilon \in (0, s]$. Ova kritična vrijednost u kojoj $\mathcal{H}^s(E)$ skoči sa $+\infty$ na 0 zove se *Hausdorffova dimenzija* skupa E, označavamo je $\operatorname{HD}(E)$ i definirana je za svaki skup $E \subset \mathbb{R}^n$. Dakle,

$$\operatorname{HD}(E) = \inf\{s \geq 0 : \mathcal{H}^s(E) = 0\} = \sup\{s : \mathcal{H}^s = +\infty\},$$

odnosno

$$\mathcal{H}^s(E) = \begin{cases} +\infty & \text{ako je } 0 \leq s < \operatorname{HD}(E), \\ 0 & \text{ako je } s > \operatorname{HD}(E), \end{cases}$$

dok za $s = \operatorname{HD}(E)$ poprima $\mathcal{H}^s(E)$ neku vrijednost iz $[0, +\infty]$. Osnovna svojstva Hausdorffove dimenzije većinom nije teško dobiti iz definicije, a sabrana su u sljedećoj propoziciji.

Propozicija 7.2. *Neka su $E, F, E_1, E_2, \ldots$ podskupovi od $\mathbb{R}^n$. Vrijedi*

(1) Ako je $E \subset F$, onda je $\mathrm{HD}(E) \leq \mathrm{HD}(F)$;

(2) $\mathrm{HD}(E) \leq n$;

(3) Ako je (n-dimenzionalna) Lebesgueova mjera od E pozitivna, onda je $\mathrm{HD}(E) = n$;

(4) Ako je E konačan ili prebrojiv skup, onda je $\mathrm{HD}(E) = 0$;

(5) Ako se E i F razlikuju u prebrojivom skupu, onda je $\mathrm{HD}(E) = \mathrm{HD}(F)$;

(6) $\mathrm{HD}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sup\{\mathrm{HD}(E_i) : i \in \mathbb{N}\}$;

(7) $\mathrm{HD}(E \times F) \geq \mathrm{HD}(E) + \mathrm{HD}(F)$ (moguća je i stroga nejednakost);

(8) Ako je $f : E \to F$ α-Hölder neprekidna, odnosno vrijedi (7.3), onda je $\alpha \cdot \mathrm{HD}(f(E)) \leq \mathrm{HD}(E)$;

(9) Ako je $f : F \to \mathbb{R}$ funkcija klase C^1 na otvorenom skupu F, onda za $E \subset F$ vrijedi $\mathrm{HD}(f(E)) \leq \mathrm{HD}(E)$.

Postoje skupovi koji su neprebrojivi i Hausdorffove dimenzije nula, npr. skup Liouvilleovih brojeva, tj. svih $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ za koje $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^w}$ ima rješenje u racionalnim brojevima $\frac{p}{q}$ za svaki $w \in \mathbb{R}$. Postoje i neprebrojivi skupovi realnih brojeva Lebesgueove mjere nula čija je dimenzija maksimalna, tj. jedan, primjerice skup loše aproksimabilnih brojeva. Mi ćemo izračunati Hausdorffovu dimenziju u jednom jednostavnom primjeru.

Primjer. Pokažimo da za Cantorov trijadski skup

$$C = \left\{ \sum_{i=1}^{\infty} \frac{a_i}{3^i} \ : \ a_i \in \{0, 2\} \text{ za svaki } i \in \mathbb{N} \right\}$$

i $s = \frac{\ln 2}{\ln 3} = 0.6309\ldots$ imamo $\mathrm{HD}(C) = s$ i $\frac{1}{2} \leq \mathcal{H}^s(C) \leq 1$.

Znamo da je $C = \bigcap_{k=0}^{\infty} E_k$, gdje je E_k unija 2^k zatvorenih intervala, svaki duljine 3^{-k}. Primjerice, $E_1 = \{[0, \frac{1}{3}], [\frac{2}{3}, 1]\}$, $E_2 = \{[0, \frac{1}{9}], [\frac{2}{9}, \frac{1}{3}], [\frac{2}{3}, \frac{7}{9}], [\frac{8}{9}, 1]\}$.

Uzmemo li intervale koji čine E_k kao 3^{-k}-pokrivač od C, dobivamo ogradu $\mathcal{H}^s_{3^{-k}}(C) \leq 2^k(3^{-k})^s = 1$ za $s = \frac{\ln 2}{\ln 3}$. Pustimo $k \to \infty$, pa slijedi $\mathcal{H}^s(C) \leq 1$, odakle je $\mathrm{HD}(C) \leq s$.

Kako bismo dokazali da je $\mathcal{H}^s(C) \geq \frac{1}{2}$, pokazat ćemo da je $\sum(\operatorname{diam} U_i)^s \geq \frac{1}{2} = 3^{-s}$ za svaki pokrivač $\{U_i\}$ od C. Očito bez smanjenja općenitosti možemo pretpostaviti da su $\{U_i\}$ sve intervali i, povećavajući ih neznatno ako je potrebno, otvoreni. Koristeći kompaktnost od C, vidimo da prethodnu nejednakost treba provjeriti samo kad je $\{U_i\}$ konačna familija intervala u $[0, 1]$. Neka je za proizvoljni U_i cijeli broj k takav da vrijedi $3^{-(k+1)} \leq \operatorname{diam} U_i < 3^{-k}$. Tada U_i može sjeći najviše jedan od intervala koji čine E_k jer su rupe među tim intervalima duljine barem 3^{-k}. Ako je $j \geq k$, onda po konstrukciji U_i siječe najviše $2^{j-k} = 2^j 3^{-sk} \leq 2^j 3^s (\operatorname{diam} U_i)^s$ intervala koji čine E_j. Izaberemo j dovoljno velik tako da je $3^{-(j+1)} \leq \operatorname{diam} U_i$ za sve i. Budući da $\{U_i\}$ siječe svih 2^j osnovnih intervala duljine 3^{-j} koji čine E_j, brojeći intervale dobivamo $2^j \leq \sum_i 2^j 3^s (\operatorname{diam} U_i)^s$ što povlači $\sum_i (\operatorname{diam} U_i)^s \geq \frac{1}{2}$. Pokazali smo da je $\mathcal{H}^s(C) \geq \frac{1}{2}$, pa je $\mathrm{HD}(C) \geq s$.

Već iz prethodnog primjera vidimo da je lakše dobiti gornje, nego donje ograde za Hausdorffovu dimenziju. Za dobivanje gornje ograde dovoljno je naći prikladan δ-pokrivač za svaki $\delta > 0$ što je često direktno dostupno iz prirode skupa kojeg promatramo. Za donju ogradu je potrebno ocijeniti sve moguće pokrivače i to obično zahtijeva znatno složenije metode.

Od fraktalnih dimenzija trebat će nam još *(gornja) box dimenzija* koja se za $E \subset \mathbb{R}^n$ definira kao

$$\overline{\dim}(E) = \lim_{\delta \to 0} \frac{\ln N_E(\delta)}{\ln(1/\delta)},$$

gdje je $N_E(\delta)$ najmanji broj kutija brida duljine δ (tj. sukladnih $[0,\delta]^n$) potrebnih da se prekrije E. Odmah je jasno da možemo uključiti i kutije s kraćim bridom od δ, ali to neće promijeniti $N_E(\delta)$. Budući da smo u definiciji Hausdorffove dimenzije dopustili puno širu klasu pokrivača, nije teško pokazati da ako je box dimenzija skupa E definirana, onda mora vrijediti $\overline{\dim}(E) \geq \mathrm{HD}(E)$.

7.3 Moreirin teorem

Izložit ćemo u nastavku osnovni rezultat i skicu dokaza sadržane u članku koji je Moreira 2018. objavio u Annals of Mathematics.

Teorem 7.3. *Za $t \in \mathbb{R}$ skupovi $\mathcal{L} \cap (-\infty, t)$ i $\mathcal{M} \cap (-\infty, t)$ imaju istu dimenziju, tj.*

$$d(t) = \mathrm{HD}\left(\mathcal{L} \cap (-\infty, t)\right) = \mathrm{HD}\left(\mathcal{M} \cap (-\infty, t)\right) = \overline{\dim}\big(\mathcal{L} \cap (-\infty, t)\big) = \overline{\dim}\big(\mathcal{M} \cap (-\infty, t)\big).$$

Funkcija $d(t)$ je neprekidna (ne strogo) rastuća surjektivna funkcija sa $\mathbb{R}$ na $[0,1]$ i vrijedi

(i) $d(t) = \min\{1, 2D(t)\}$, *gdje smo stavili* $D(t) = \mathrm{HD}\left(L^{-1}(-\infty, t)\right) = \mathrm{HD}\left(L^{-1}(-\infty, t]\right)$ *i $D(t)$ je neprekidna funkcija sa $\mathbb{R}$ u $[0,1]$;*

(ii) $\max\{t \in \mathbb{R} : d(t) = 0\} = 3$, *tj.* $d(3+\varepsilon) > 0$ *za svaki* $\varepsilon > 0$*;*

(iii) postoji $\delta > 0$ tako da je $d(\sqrt{12} - \delta) = 1$.

Primijetimo da je $d(\sqrt{12}) = 1$, pa je $\mathcal{L} \cap (-\infty, \sqrt{12})$ pune Hausdorffove mjere iako je $\sqrt{12} = 3.4641\ldots$ puno prije početka Hallove zrake $4.5278\ldots$.

Pokažimo najprije jednu direktnu posljedicu ovog teorema. Iako je $d(t)$ neprekidna, ona nije Hölder neprekidna.

Korolar 7.4. *Funkcija $d(t)$ nije α-Hölder neprekidna ni za jedan $\alpha > 0$.*

Dokaz. Prema Teoremu 7.3, za svaki $\varepsilon > 0$ funkcija d preslikava $\mathcal{L} \cap [3, 3+\varepsilon]$ u netrivijalni interval $[0, d(3+\varepsilon)]$.

Pretpostavimo da je za neki $\alpha > 0$ funkcija d α-Hölder neprekidna. Iz Propozicije 7.2.8 slijedi da je

$$0 < \alpha = \alpha \cdot \mathrm{HD}\left([0, d(3+\varepsilon)]\right) \leq \mathrm{HD}(\mathcal{L} \cap [3, 3+\delta]) = d(3+\varepsilon)$$

za svaki $\varepsilon > 0$.

Budući da je je $\mathcal{L} \cap (-\infty, 3)$ prebrojiv skup, imamo po Propoziciji 7.2.4 da je $d(3) = \mathrm{HD}\left(\mathcal{L} \cap (-\infty, 3)\right) = 0$. No, Teorem 7.3 kaže da je d neprekidna, pa je $\lim_{\varepsilon \to 0} d(3+\varepsilon) = d(3) = 0$. Zajedno s prethodnom nejednakošću dobili smo $0 < \alpha \leq \lim_{\varepsilon \to 0} d(3+\varepsilon) = 0$, što je kontradikcija. $\square$

Dokaz Teorema 7.3 zasniva se na aproksimaciji dijelova spektara iznuta i izvana sumama regularnih Cantorovih skupova. Pri tome se koristi duboki rezultat Moreire i Yoccoza koji povlači da suma dvaju neesencijalnoafinih regularnih Cantorovih skupova ima Hausdorffovu dimenziju jednaku minimumu između 1 i sume njihovih Hausdorffovih dimenzija. Definirajmo najprije spomenute pojmove.

Kažemo da je $K \subset \mathbb{R}$ *regularni Cantorov skup klase* C^k, $k \geq 1$, ako

(i) postoje u parovima disjunktni kompaktni intervali I_1, $I_2, \ldots$, I_r takvi da je $K \subset I_1 \cup \ldots \cup I_r$ i rub svakog od I_j je sadržan u K;

(ii) postoji realna funkcija ψ klase C^k definirana na okolini od $I_1 \cup I_2 \cup \ldots \cup I_r$ koja je ekspandirajuća, tj. $|\psi'(x)| > 1$ za sve x iz domene, i svaki $\psi(I_j)$ je konveksna ljuska unije nekih od intervala I_s te vrijedi

   (ii.1) za svaki $j \in \{1, \ldots, r\}$ i dovoljno veliki n je $\psi^n(K \cap I_j) = K$,

   (ii.2) $K = \bigcap_{n \in \mathbb{N}} \psi^{-n}(I_1 \cup I_2 \cup \ldots \cup I_r)$.

Primjerice, Cantorov trijadski skup C je regularan jer za $\psi : [0, \frac{1}{3}] \cup [\frac{2}{3}, 1] \to [0, 1]$ definiranu sa

$$\psi(x) = \begin{cases} 3x, & \text{ako je } 0 \leq x \leq \frac{1}{3} \\ 3x - 2, & \text{ako je } \frac{2}{3} \leq x \leq 1 \end{cases}$$

očito imamo $|\psi'(x)| = 3 > 1$, $\psi([0, \frac{1}{3}]) = \psi([\frac{2}{3}, 1]) = [0, 1]$ i $C = \bigcap_{n \in \mathbb{N}} \psi^{-n}([0, \frac{1}{3}] \cup [\frac{2}{3}, 1])$.

Također, za $A \geq 2$ su $C(A) = \{[0, a_1, a_2, \ldots] : 1 \leq a_i \leq A \text{ za svaki } i \in \mathbb{N}\}$ regularni Cantorovi skupovi pridruženi Gaussovom preslikavanju $g : (0, 1] \to [0, 1]$, $g(x) = \left\{\frac{1}{x}\right\} = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$ koje djeluje kao pomak na nizu parcijalnih kvocijenata verižnog razlomka, tj.

$$g([0, a_1, a_2, a_3, \ldots]) = [0, a_2, a_3, a_4, \ldots].$$

Tako za $A = 2$ imamo $I_1 = \big[[0, 2, \overline{1, 2}], [0, 2, \overline{2, 1}]\big]$ i $I_2 = \big[[0, 1, \overline{1, 2}], [0, 1, \overline{2, 1}]\big]$ te je $C(2) = \bigcap_{n \in \mathbb{N}} g^{-n}(I_1 \cup I_2)$. Hensley je pokazao da je

$$\mathrm{HD}(C(2)) = 0.531\ldots, \qquad \mathrm{HD}(C(3)) = 0.705\ldots, \qquad \mathrm{HD}(C(4)) = 0.788\ldots.$$

Za regularan Cantorov skup kažemo da je *afin* ako je $\psi|_{I_j}$ linearno preslikavanje za svaki j. Cantorov trijadski skup očito je afin dok skupovi $C(A)$ nisu.

Skupovi $C(A)$ su poseban slučaj tzv. Gauss-Cantorovih skupova. Neka je dan konačan skup $B = \{\beta_1, \ldots, \beta_m\}$, $m \geq 2$, gdje su riječi $\beta_j \in \mathbb{N}^{r_j}$, $r_j \in \mathbb{N}$, $1 \leq j \leq m$ i β_i nije prefiks β_j za $i \neq j$. *Gauss-Cantorov skup* $K(B) \subset [0, 1]$ pridružen B je

$$K(B) = \{[0, \gamma_1, \gamma_2, \ldots] : \gamma_i \in B \text{ za svaki } i\}.$$

To je regularni Cantorov skup. Naime, za svaku riječ $\beta_j \in \mathbb{N}^{r_j}$ iz B stavimo

$$I_j = I(\beta_j) = \{[0, \beta_j, a_1, a_2, \ldots] : a_i \in \mathbb{N} \text{ za svaki } i\}$$

i $\psi|_{I_j} = g^{r_j}$, gdje je g Gaussovo preslikavanje. Sada se lako provjere svi potrebni uvjeti.

Skupu B kao gore pridružujemo skup obostrano beskonačnih nizova prirodnih brojeva

$$\Sigma(B) = \{(\gamma_i)_{i \in \mathbb{Z}} : \gamma_i \in B\} \subset \Sigma = \mathbb{N}^{\mathbb{Z}} = \Sigma^- \times \Sigma^+ = \mathbb{N}^{\mathbb{Z}_{<0}} \times \mathbb{N}^{\mathbb{N}_0}.$$

Označimo li sa $\pi^+ : \Sigma \to \Sigma^+$ prirodnu projekciju kojom od obostrano beskonačnog niza dobivamo jednostrano beskonačni niz koji je njegov desni dio, imamo $K(B) = \{[0,\gamma] : \gamma \in \pi^+(\Sigma(B))\}$.

Ako je $\beta = (b_1, b_2, \ldots, b_{n-1}, b_n)$, stavljamo kao prije $\widetilde{\beta} = (b_n, b_{n-1}, \ldots, b_2, b_1)$ za obratnu riječ te za skup konačnih nizova B definiramo $\widetilde{B} = \{\widetilde{\beta} : \beta \in B\}$.

Propozicija 7.5. *Za svaki konačni skup B konačnih nizova je* $\mathrm{HD}(K(B)) = \mathrm{HD}(K(\widetilde{B}))$.

Skica dokaza. Za $\beta = (b_1, b_2, \ldots, b_n)$ označimo sa $\frac{p_n(\beta)}{q_n(\beta)} = [0, b_1, \ldots, b_n]$ n-tu konvergentu od $[0,\beta]$. Iz (1.4) je $\frac{q_{n-1}(\beta)}{q_n(\beta)} = [0, b_n, \ldots, b_1]$, pa je $q_n(\beta) = q_n(\widetilde{\beta})$. Neka je sada $\beta \in B^k = \{\beta_1, \ldots, \beta_m\}^k$ za neki $k \in \mathbb{N}$ i $\beta \in \mathbb{N}^n$. Duljina intervala $I(\beta) = \{[0, \beta, a_1, a_2 \ldots] : a_i \in \mathbb{N}\}$ je

$$s(\beta) = |I(\beta)| = \left| \frac{p_n(\beta)}{q_n(\beta)} - \frac{p_n(\beta) + p_{n-1}(\beta)}{q_n(\beta) + q_{n-1}(\beta)} \right| = \frac{1}{q_n(\beta)(q_n(\beta) + q_{n-1}(\beta))},$$

pa je

$$\frac{1}{2(q_n(\beta))^2} < s(\beta) < \frac{1}{(q_n(\beta))^2}.$$

Zato je $\frac{1}{2} < \frac{s(\widetilde{\beta})}{s(\beta)} < 2$.

To znači da su $K(B)$ i $K(\widetilde{B})$ Cantorovi skupovi konstruirani tako da im odgovarajući intervali u svakom koraku konstrukcije imaju usporedive duljine, pa slijedi da su i Hausdorffove dimenzije od $K(B)$ i $K(\widetilde{B})$ jednake. $\square$

Kažemo da je regularni Cantorov skup klase C^2

$$K = \bigcap_{n \in \mathbb{N}} \psi^{-n}(I_1 \cup \ldots \cup I_r)$$

neesencijalnoafin ako ne postoji globalni konjugat $h \circ \psi \circ h^{-1}$ takav da su sve grane

$$(h \circ \psi \circ h^{-1})|_{h(I_j)}, \quad j \in \{1, \ldots, r\}$$

linearna (tj. afina) preslikavanja intervala realnih brojeva. Ovo svojstvo možemo ekvivalentno iskazati na sljedeći način. Neka je $p \in K$ periodska točka od ψ perioda k, tj. $\psi^k(p) = p$, i neka je $h : I \to I$ difeomorfizam konveksne ljuske I od $I_1 \cup \ldots \cup I_r$ takav da je $h \circ \psi^k \circ h^{-1}$ linearno na $h(J)$, gdje je J komponenta povezanosti domene od ψ^k koja sadrži p. Poincaré je pokazao da takav difeomorfizam koji linearizira jednu granu od ψ uvijek postoji. Tada je K neesencijalnoafin ako i samo ako je $(h \circ \psi \circ h^{-1})''(x) \neq 0$ za neki $x \in h(K)$.

Propozicija 7.6. *Gauss-Cantorovi skupovi su neesencijalnoafini.*

Dokaz. Osnovna ideja dokaza je koristiti činjenicu da za razlomljenu linearnu funkciju (Möbiusovu funkciju) koja nije linearna druga derivacija nigdje ne iščezava.

Neka je $B = \{\beta_1, \beta_2, \ldots, \beta_m\}$, $\beta_j = (b_1^{(j)}, b_2^{(j)}, \ldots, b_{r_j}^{(j)}) \in \mathbb{N}^{r_j}$, $1 \leq j \leq m$. Za svaki $j \leq m$, neka je $x_j = [0, \beta_j, \beta_j, \beta_j, \ldots] \in I_j = I(\beta_j) = \{[0, \beta_j, \alpha] : \alpha \geq 1\}$ fiksna točka grane $\psi|_{I_j} = g^{r_j}$. Primijetimo da, jer β_i nije prefiks β_j za $i \neq j$, svi x_j, $1 \leq j \leq m$, moraju biti različiti.

Očito je $\frac{1}{g^r(x)}$ $r+1$-vi potpuni kvocijent od $x = [0, a_1, a_2, \ldots]$. Označimo li $\frac{p_k^{(j)}}{q_k^{(j)}} = [0, b_1^{(j)}, b_2^{(j)}, \ldots, b_k^{(j)}]$ za $1 \le j \le m$, $1 \le k \le r_j$, imamo

$$\frac{1}{\psi|_{I_j}(x)} = \frac{1}{g^{r_j}(x)} = \frac{1}{\begin{bmatrix} p_{r_j}^{(j)} & p_{r_j-1}^{(j)} \\ q_{r_j}^{(j)} & q_{r_j-1}^{(j)} \end{bmatrix}^{-1} x},$$

odnosno

$$\psi|_{I_j}(x) = \frac{-q_{r_j}^{(j)} x + p_{r_j}^{(j)}}{q_{r_j-1}^{(j)} x - p_{r_j-1}^{(j)}},$$

pa je x_j kao fiksna točka ovog preslikavanja pozitivni korijen od $q_{r_j-1}^{(j)} x^2 + (q_{r_j}^{(j)} - p_{r_j-1}^{(j)})x - p_{r_j}^{(j)}$. Posebno, x_j je kvadratna iracionalnost, pa zato hiperbolička fiksna točka razlomljene linearne transformacije $\psi|_{I_j}$. Nije teško pokazati (primjerice koristeći matrični zapis ovih transformacija) da za svaki $j \le m$ postoji razlomljena linearna transformacija $\alpha_j(x) = \frac{a_j x + b_j}{c_j x + d_j}$ takva da je $\alpha_j(x_j) = x_j$, $\alpha_j'(x_j) = 1$ i $\alpha_j \circ (\psi|_{I_j}) \circ \alpha_j^{-1}$ je linearno preslikavanje.

Pokažemo li sada da $\alpha_1 \circ (\psi|_{I_2}) \circ \alpha_1^{-1}$ nije linearna, gotovi smo jer druga derivacija razlomljene linearne transformacije koja nije linearna nema nultočaka. Pretpostavimo suprotno, tj. da $\alpha_1 \circ (\psi|_{I_2}) \circ \alpha_1^{-1}$ jest linearna funkcija. Budući da je i $\alpha_1 \circ (\psi|_{I_1}) \circ \alpha_1^{-1}$ također linearna, ove dvije funkcije imaju zajedničku fiksnu točku u ∞, pa je $\alpha_1^{-1}(\infty) = -\frac{d_1}{c_1}$ fiksna točka od $\psi|_{I_2}$ i $\psi|_{I_1}$ što znači da je $\alpha_1^{-1}(\infty)$ zajednički korijen od

$$q_{r_1-1}^{(1)} x^2 + (q_{r_1}^{(1)} - p_{r_1-1}^{(1)})x - p_{r_1}^{(1)} \qquad \text{i} \qquad q_{r_2-1}^{(2)} x^2 + (q_{r_2}^{(2)} - p_{r_2-1}^{(2)})x - p_{r_2}^{(2)}.$$

Budući da ovi polinomi iz $\mathbb{Z}[x]$ imaju iracionalne korijene x_1, odnosno x_2, oni su ireducibilni u $\mathbb{Q}[x]$. Kako im se jedan korijen podudara, to se i preostali korijeni (koji su pozitivni), tj. x_1 i x_2 podudaraju, što nije moguće. □

Sljedeći teorem je poseban slučaj Moreirine dimenzijske formule. Dokaz ne navodimo jer izlazi daleko izvan okvira ovih predavanja.

Teorem 7.7. *Neka su K i K' regularni Cantorovi skupovi klase C^2. Ako je K neesencijalnoafin, onda je* $\mathrm{HD}(K + K') = \min\{1, \mathrm{HD}(K) + \mathrm{HD}(K')\}$.

Korolar 7.8. *Za svaki par konačnih skupova B, B' konačnih nizova prirodnih brojeva je*

$$\mathrm{HD}(K(B) + K(B')) = \min\{1, \mathrm{HD}(K(B)) + \mathrm{HD}(K(B'))\}.$$

Korolar 7.9. *Za svaki konačni skup B konačnih nizova prirodnih brojeva je*

$$\mathrm{HD}(K(B) + K(\widetilde{B})) = \min\{1, 2\,\mathrm{HD}(K(B))\}.$$

Označimo kao prije za $\alpha = (a_1, a_2, \ldots, a_n) \in \mathbb{N}^n$ veličinu intervala $I(\alpha) = \{x \in [0,1] : x = [0, a_1, a_2, \ldots, a_n, \alpha_{n+1}], \alpha_{n+1} \ge 1\}$ sa $s(\alpha) = |I(\alpha)|$ te uvedimo $r(\alpha) = \lfloor -\ln s(\alpha) \rfloor$ tako da manji intervali daju veću vrijednost od r. Za $r \in \mathbb{N}$ definiramo

$$P_r = \{\alpha = (a_1, a_2, \ldots, a_n) \,:\, r(\alpha) \ge r,\, r(a_1, a_2, \ldots, a_{n-1}) < r\}.$$

Imamo $\Sigma = \mathbb{N}^{\mathbb{Z}} = \Sigma^- \times \Sigma^+$, gdje je $\Sigma^- = \mathbb{N}^{\mathbb{Z}_{<0}}$, $\Sigma^+ = \mathbb{N}^{\mathbb{N}_0}$. Neka je $\pi^+ : \Sigma \to \Sigma^+$ projekcija, a $\sigma : \Sigma \to \Sigma$ pomak $\sigma((a_n)_{n\in\mathbb{Z}}) = (a_{n+1})_{n\in\mathbb{Z}}$. Kao i prije, za $\underline{\vartheta} = (a_n)_{n\in\mathbb{Z}} \in \Sigma$

definiramo $\alpha_n = [a_n, a_{n+1}, a_{n+2}, \dots]$ i $\beta_n = [0, a_{n-1}, a_{n-2}, \dots]$ te je $M(\underline{\vartheta}) = \sup\{\alpha_n + \beta_n : n \in \mathbb{Z}\}$ i $\mathcal{M} = \{M(\underline{\vartheta}) : \underline{\vartheta} \in \Sigma\}$. U nastavku ćemo raditi sa $\Sigma_t = \{\underline{\vartheta} \in \Sigma : M(\underline{\vartheta}) \leq t\}$ za $t \in [3, +\infty)$, a taj skup je očito invarijantan na obrtanje (zrcaljenje nizova oko nulte pozicije) i σ.

Budući za $\underline{\vartheta} = (a_n)_{n\in\mathbb{Z}} \in \Sigma$ vrijedi $\alpha_n + \beta_n > \alpha_n \geq a_n$ za svaki n, to vrijedi $M(\underline{\vartheta}) > \sup\{a_n : n \in \mathbb{Z}\}$, pa iz $M(\underline{\vartheta}) \leq t$ slijedi $a_n \leq \lfloor t \rfloor$ za svaki n. Za $t \in [3, +\infty)$ i $r \in \mathbb{N}$ stavimo $T = \lfloor t \rfloor$ i

$$C(t, r) = \{\alpha = (a_1, \dots, a_n) \in P_r : K_t \cap I(\alpha) \neq \emptyset\},$$

gdje je $K_t = \{[0, \gamma] : \gamma \in \pi^+(\Sigma_t)\}$. Kako je Σ_t invarijantan na obrtanje i σ, to je K_t invarijantan na Gaussovo preslikavanje g i vrijedi $\mathcal{M} \cap (-\infty, t) \subset (\mathbb{N} \cap [1, T]) + K_t + K_t$.

Definiramo kardinalitet $N(t, r) = |C(t, r)|$ te primjećujemo da za $r \leq s$ vrijedi $N(t, r) \leq N(t, s)$ i za $t \leq t'$ vrijedi $N(t, r) \leq N(t', r)$.

Budući da znamo eksplicitno zapisati $s(\alpha)$ preko odgovarajućih nazivnika konvergenti u razvoju od α (vidi Teorem 1.8.a), pomoću svojstava kontinuanti lako se dobiva da za proizvoljne konačne nizove α, β prirodnih brojeva i prirodne brojeve $k_1, k_2 \leq T$ imamo $r(\alpha\beta k_1 k_2) \geq r(\alpha) + r(\beta)$. Dakle, ako su $C(t, r) = \{\alpha_1, \alpha_2, \dots, \alpha_u\}$ i $C(t, s) = \{\beta_1, \beta_2, \dots, \beta_v\}$, onda K_t možemo prekriti sa $T^2 uv = T^2 N(t, r) N(t, s)$ intervala

$$I(\alpha_i \beta_j k_1 k_2), \qquad 1 \leq i \leq u, \quad 1 \leq j \leq v, \quad 1 \leq k_1, k_2 \leq T,$$

koji zadovoljavaju $r(\alpha_i \beta_j k_1 k_2) \geq r + s$ za sve i, j, k. Zamjenjujući, ako je potrebno, neke od ovih intervala većim intervalima $I(\gamma)$ u P_{r+s}, zaključujemo da je $N(t, r + s) \leq T^2 N(t, r) N(t, s)$ i stoga za sve r i s imamo

$$\ln(T^2 N(t, r + s)) \leq \ln(T^2 N(t, r)) + \ln(T^2 N(t, s)).$$

Iz subaditivnosti sada slijedi postojanje limesa

$$\lim_{m\to\infty} \frac{1}{m} \ln(T^2 N(t, m)) = \inf_{m\in\mathbb{N}} \frac{1}{m} \ln(T^2 N(t, m)) = \lim_{m\to\infty} \frac{1}{m} \ln(N(t, m)).$$

Ovaj ćemo limes označiti sa $D(t)$ i direktno iz definicije vidi se da je to upravo box dimenzija od K_t. Primijetimo da je $D(t)$ rastuća funkcija, a u dokazu Teorema 7.3 vidjet ćemo da je neprekidna i da vrijedi $\mathrm{HD}(L^{-1}(-\infty, t)) = D(t)$.

Lema 7.10. *Funkcija $D(t)$ je neprekidna zdesna, tj. za svaki $t_0 \in [3, +\infty)$ i $\eta > 0$ postoji $\delta > 0$ tako da za $t_0 \leq t \leq t_0 + \delta$ vrijedi $D(t_0) \leq D(t) \leq D(t_0 + \delta) < D(t_0) + \eta$.*

Dokaz. Pretpostavimo suprotno, tj. da postoje t_0 i $\eta > 0$ takvi da je $D(t) > D(t_0) + \eta$ za svaki $t > t_0$. Iz definicije od $D(t)$ (vidjeli smo da je limes koji se ondje pojavljuje baš infimum) to znači da postoji $r_0 \in \mathbb{N}$ takav da je $\frac{1}{r} \ln N(t, r) > D(t_0) + \eta$ za sve $r \geq r_0$ i $t > t_0$. S druge strane imamo $C(t, r) \subset C(s, r)$ za $t \leq s$ i (argumentirajući slično kao kod leme o kompaktnosti) $C(t_0, r) = \bigcap_{t>t_0} C(t, r)$, pa bi kad $r \to \infty$ i $t \to t_0$ prethodna nejednakost povlačila $D(t_0) \geq D(t_0) + \eta$ što nije moguće. $\square$

Lema 7.11. *Neka su dani $t \in (3, +\infty)$ i $\eta \in (0, 1)$. Tada postoje $\delta > 0$ i Gauss-Cantorov skup $K(B)$ takvi da je $\Sigma(B) \subset \Sigma_{t-\delta}$ i $\mathrm{HD}(K(B)) > (1 - \eta)D(t)$.*

Skica dokaza. U najkraćim crtama ćemo prepričati dosta tehnički dokaz ove leme u kojemu se konstruiraju Gauss-Cantorovi skupovi koji se približavaju Σ_t iznutra.

Fiksiramo $r_0 \in \mathbb{N}$ dovoljno velik tako da je

$$\left|\frac{\ln N(t,r)}{r} - D(t)\right| < \frac{\eta}{80}D(t) \qquad \text{za sve } r \geq r_0.$$

Stavimo $B_0 = C(t, r_0)$, $N_0 = N(t, r_0) = |B_0|$, $k = 8N_0^2\lceil 80/\eta \rceil$ i

$$\widehat{B} = \{\beta = (\beta_1, \ldots, \beta_k) \, : \, \beta_j \in B_0, \, 1 \leq j \leq k \text{ i } K_t \cap I(\beta) \neq \emptyset\} \subset B_0^k.$$

Pokazuje se da $\widehat{B}$ ima značajan kardinalitet u smislu

$$|\widehat{B}| > 2N_0^{(1-\frac{\eta}{40})k}.$$

Posebno, pomoću ove informacije se može pokazati da $\mathrm{HD}(K(\widehat{B}))$ nije daleko od $D(t)$, tj.

$$\mathrm{HD}(K(\widehat{B})) > \Big(1 - \frac{\eta}{20}\Big)D(t).$$

Nažalost, budući da nemamo kontrolu nad vrijednostima Markovljevog supremuma M na elementima od $\Sigma(\widehat{B})$, nemamo garanciju da je $\Sigma(\widehat{B}) \subset \Sigma_{t-\delta}$ za neki $\delta > 0$.

Ovaj problem rješavamo uz pomoć *lijevo-dobrih* i *desno-dobrih* pozicija, odnosno indeksa. Kažemo da je $j \in \{1, \ldots, k\}$ desno-dobra pozicija od $\beta = (\beta_1, \ldots, \beta_k) \in \widehat{B}$ ako postoje dva elementa $\beta^{(s)} = \beta_1\beta_2 \ldots \beta_{j-1}\beta_j^{(s)}\beta_{j+1}^{(s)} \ldots \beta_k^{(s)} \in \widehat{B}$, $s \in \{1, 2\}$ takva da je

$$[0, \beta_j^{(1)}] < [0, \beta_j] < [0, \beta_j^{(2)}].$$

Slično, $j \in \{1, \ldots, k\}$ je lijevo-dobra pozicija od $\beta = (\beta_1, \ldots, \beta_k) \in \widehat{B}$ ako postoje dva elementa $\beta^{(s)} = \beta_1^{(s)}\beta_2^{(s)} \ldots \beta_{j-1}^{(s)}\beta_j^{(s)}\beta_{j+1}\beta_{j+2} \ldots \beta_k \in \widehat{B}$, $s \in \{3, 4\}$ takva da je

$$[0, \widetilde{\beta_j^{(3)}}] < [0, \widetilde{\beta_j}] < [0, \widetilde{\beta_j^{(4)}}].$$

Kažemo da je j *dobra* pozicija ako je i desno-dobra i lijevo-dobra.

Budući da postoje najviše dva izbora za $\beta_j \in B_0$ kada su $\beta_1, \ldots, \beta_{j-1}$ fiksirani i j nije desno dobra pozicija, uz malo računa dobivamo da podskup

$$\mathcal{E} = \{\beta \in \widehat{B} \, : \, \beta \text{ ima barem } 9k/10 \text{ dobrih pozicija}\}$$

izvrsnih riječi u $\widehat{B}$ ima kardinalitet $|\mathcal{E}| > \frac{1}{2}|\widehat{B}| > N_0^{(1-\frac{\eta}{40})k}$.

Očekujemo da se vrijednosti od M na $\Sigma(\mathcal{E})$ smanje jer izvrsne riječi imaju puno dobrih pozicija. Također, Hausdorffova dimenzija od $K(\mathcal{E})$ nije daleko od $D(t)$ zbog gornje ocjene od $|\mathcal{E}|$. No, nema razloga da bude $\Sigma(\mathcal{E}) \subset \Sigma_{t-\delta}$ za neki $\delta > 0$ jer proizvoljnim spajanjem riječi iz $\mathcal{E}$ ne moramo dobiti niz u Σ_t.

Zato je iduća ideja izgraditi $\Sigma(B) \subset \Sigma_{t-\delta}$ koristeći $\mathcal{E}$ sljedećim kombinatornim argumentom. Budući da $\beta = (\beta_1, \ldots, \beta_k) \in \mathcal{E}$ ima $\frac{9k}{10}$ dobrih pozicija, možemo naći dobre pozicije $1 \leq i_1 \leq i_2 \leq \ldots \leq i_{\lceil 2k/5 \rceil} \leq k-1$ takve da je $i_s + 2 \leq i_{s+1}$ za $s \in \{1, \ldots, \lceil 2k/5 \rceil - 1\}$ i da su za sve s iz tog skupa $i_s + 1$ također dobre pozicije. Kako je $k = 8N_0^2\lceil 80/\eta \rceil$, iz Dirichletovog principa dobivamo da možemo izabrati pozicije $j_1 \leq \ldots \leq j_{3N_0^2}$ i riječi $\gamma_{j_1}, \gamma_{j_1+1}, \ldots, \gamma_{j_{3N_0^2}}, \gamma_{j_{3N_0^2}+1} \in B_0$ takve da je $j_s + 2\lceil 80/\eta \rceil \leq j_{s+1}$ za sve $s < 3N_0^2$ i podskup

$$X = \{(\beta_1, \ldots, \beta_k) \in \mathcal{E} \, : \, j_s, j_s{+}1 \text{ dobre pozicije}, (\beta_{j_s}, \beta_{j_s+1}) = (\gamma_{j_s}, \gamma_{j_s+1}) \text{ za sve } s \leq 3N_0^2\}$$

skupa izvrsnih riječi s propisanim podriječima $\gamma_{j_s}, \gamma_{j_s+1}$ na dobrim pozicijama $j_s, j_s + 1$ ima kardinalitet

$$|X| > N_0^{(1-\frac{\eta}{20})k}.$$

Sada pretvorimo X u alfabet B za traženi $\Sigma(B)$ pomoću projekcija $\pi_{a,b} : X \to B_0^{j_b - j_a}$, $\pi_{a,b}(\beta_1\beta_2 \dots \beta_k) = (\beta_{j_a+1}, \beta_{j_a+2}, \dots, \beta_{j_b})$. Naime, argument s prebrajanjem pokazuje da možemo uzeti neke $1 \le a < b \le 3N_0^2$ takve da je $(\gamma_{j_a}, \gamma_{j_a+1}) = (\gamma_{j_b}, \gamma_{j_b+1})$ i slika $\pi_{a,b}(X)$ projekcije $\pi_{a,b}$ je dovoljno velika, odnosno

$$|\pi_{a,b}(X)| > N_0^{(1-\frac{\eta}{4})(j_b - j_a)}.$$

Tako smo dobili alfabet $B = \pi_{a,b}(X)$ čiji se elementi ispravno spajaju (jer je $\gamma_{j_a} = \gamma_{j_b}$ i $\gamma_{j_a+1} = \gamma_{j_b+1}$), Hausdorffova dimenzija od $K(B)$ je $\mathrm{HD}(K(B)) > (1-\eta)D(t)$ (jer je $|B| > N_0^{(1-\frac{\eta}{4})(j_b-j_a)}$ i $j_b - j_a > 2\lceil 80/\eta \rceil$) te imamo $\Sigma(B) \subset \Sigma_{t-\delta}$ za neki $\delta > 0$ (jer svojstva dobrih pozicija sile vrijednosti od M na $\Sigma(B)$ da se smanje). Time smo završili skicu dokaza. □

Lema 7.12. *Neka je X konačan skup konačnih nizova prirodnih brojeva. Vrijedi*

$$\mathrm{HD}\left(L(\Sigma(X))\right) = \mathrm{HD}\left(M(\Sigma(X))\right) = \min\{1, 2\,\mathrm{HD}(K(X))\}.$$

Dokaz. Neka je T najveći element u nizovima u X, a R duljina najdulje riječi u X. Odmah vidimo da je

$$L(\Sigma(X)) \subset M(\Sigma(X)) \subset \bigcup_{\substack{1 \le a \le T \\ 1 \le i,j \le R}} \left(a + g^i(K(X)) + g^j(K(\widetilde{X}))\right),$$

gdje smo inkluzije dobili analogno zaključivanju kod Leme o kompaktnosti. Stoga po svojstvima Hausdorffove dimenzije iz Propozicije 7.2 i po Korolaru 7.9 zaključujemo da je

$$\mathrm{HD}\left(L(\Sigma(X))\right) \le \mathrm{HD}\left(M(\Sigma(X))\right) \le \min\{1, 2\,\mathrm{HD}(K(X))\}.$$

Neka je $\varepsilon > 0$ po volji. Pokazat ćemo da postoje regularni Cantorovi skupovi K, K' definirani iteracijama Gaussovog preslikavanja g takvi da je

$$\min\{\mathrm{HD}(K), \mathrm{HD}(K')\} > \mathrm{HD}(K(X)) - \varepsilon \qquad \text{i} \qquad K + K' \subset L(\Sigma(X)) \subset M(\Sigma(X)).$$

Budući da po Teoremu 7.7 znamo

$$\mathrm{HD}(K + K') = \min\{1, \mathrm{HD}(K) + \mathrm{HD}(K')\} > \min\{1, 2\,\mathrm{HD}(K(X))\} - 2\varepsilon$$

i $\varepsilon > 0$ je bio proizvoljan, tvrdnja leme odmah slijedi.

Za prirodan broj n je $X^n = \{(\gamma_1, \gamma_2, \dots, \gamma_n) : \gamma_j \in X, \ j \in \{1, \dots, n\}\}$ standardni Kartezijev produkt te vrijedi $\Sigma(X^n) = \Sigma(X)$ i $K(X^n) = K(X)$. Zamijenimo li X sa X^n za dovoljno veliki n, možemo bez smanjenja općenitosti pretpostaviti da za svaki $A \subset X$ takav da je $|A| \le 2$ vrijedi $\mathrm{HD}(K(X \setminus A)) > \mathrm{HD}(K(X)) - \varepsilon$. Isto tako, pretpostavljamo da za svaki $\widetilde{A} \subset \widetilde{X}$ sa $|\widetilde{A}| \le 2$ vrijedi $\mathrm{HD}(K(\widetilde{X} \setminus \widetilde{A})) > \mathrm{HD}(K(\widetilde{X})) - \varepsilon = \mathrm{HD}(K(X)) - \varepsilon$.

Uvedimo uređaj na X i $\widetilde{X}$ koristeći verižne razlomke. Za $\gamma, \widehat{\gamma} \in X$ (odnosno $\gamma, \widehat{\gamma} \in \widetilde{X}$) definiramo $\gamma < \widehat{\gamma}$ ako i samo ako je $[0, \gamma] < [0, \widehat{\gamma}]$.

Pretpostavimo (koristeći lemu o kompaktnosti) da se maksimum od $M(\Sigma(X))$ postiže u

$$\underline{\widehat{\vartheta}} = (\dots, \widehat{\gamma}_{-1}, \widehat{\gamma}_0, \widehat{\gamma}_1, \dots), \qquad \widehat{\gamma}_i \in X, \ i \in \mathbb{Z}$$

i to na poziciji koja pripada riječi $\widehat{\gamma}_0$. Neka je $X^* = X \setminus \{\min X, \max X\}$ i $\widetilde{X}^* = \widetilde{X} \setminus \{\min \widetilde{X}, \max \widetilde{X}\}$. Očekujemo da se vrijednosti od L i M na $(\widetilde{\widetilde{X}^*})^{\mathbb{Z}_{<0}} \times (X^*)^{\mathbb{N}_0}$ smanje jer smo uklonili minimalne i maksimalne elemente od X i $\widetilde{X}$ (usporedi Teorem 1.8.b). Tako će $K(X^*)$ i $K(\widetilde{X}^*)$ biti skoro pa traženi Cantorovi skupovi, ali budući da ne mora vrijediti $K(X^*)+K(\widetilde{X}^*) \subset L(\Sigma(X))$, moramo još kontrolirati pozicije na kojima se lim sup postiže te uzeti odgovarajuće "kopije" od $K(X^*)$ i $K(\widetilde{X}^*)$.

Za svaki prirodan broj m, označimo sa G^m skup nizova

$$(\ldots, \gamma_{-m-2}, \gamma_{-m-1}, \widehat{\gamma}_{-m}, \widehat{\gamma}_{-m+1}, \ldots, \widehat{\gamma}_{-1}, \widehat{\gamma}_0, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_{m-1}, \widehat{\gamma}_m, \gamma_{m+1}, \gamma_{m+2}, \ldots),$$

gdje su $\gamma_k \in X^*$ za $k \geq m+1$ i $\widetilde{\gamma}_k \in \widetilde{X}^*$ za $k \leq -m-1$. Tada nije teško pokazati da za dovoljno veliki m postoji $\eta > 0$ takav da se za svaki $\underline{\vartheta} \in G^m$, $\sup(\alpha_n + \beta_n) = M(\underline{\vartheta})$ postiže samo za neke vrijednosti indeksa n koje su pozicije elemenata unutar bloka

$$\tau = (\widehat{\gamma}_{-m}, \widehat{\gamma}_{-m+1}, \ldots, \widehat{\gamma}_{-1}, \widehat{\gamma}_0, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_{m-1}, \widehat{\gamma}_m)$$

od $\underline{\vartheta}$ te za n izvan toga bloka vrijedi $\alpha_n + \beta_n < M(\underline{\vartheta}) - \eta$. Također, za proizvoljni $\underline{\vartheta} \in G^m$ i $\underline{\vartheta}^* \in \Sigma(X^*)$ imamo $M(\underline{\vartheta}^*) < M(\underline{\vartheta}) - \eta$ ako je m dovoljno velik.

Fiksirajmo sada takvu dovoljno veliku vrijednost $m \in \mathbb{N}$ te uzmimo $\gamma^{(0)} \in X$ tako da je $\widetilde{\gamma^{(0)}} \in \widetilde{X}^*$. Svakom $x = [0, \gamma_1(x), \gamma_2(x), \gamma_3(x), \ldots] \in K(X^*)$, $\gamma_i \in X^*$, pridružujemo element $\underline{\Theta}(x) \in G^m$ dan sa

$$\underline{\Theta}(x) = (\ldots, \gamma^{(0)}, \gamma^{(0)}, \gamma^{(0)}, \widehat{\gamma}_{-m}, \widehat{\gamma}_{-m+1}, \ldots, \widehat{\gamma}_{-1}, \widehat{\gamma}_0, \widehat{\gamma}_1, \ldots, \widehat{\gamma}_{m-1}, \widehat{\gamma}_m, \gamma_1(x), \gamma_2(x), \gamma_3(x), \ldots).$$

Za svaku poziciju n koja odgovara bloku τ od $\underline{\Theta}(x)$ pišemo $\lambda_n(x) = \alpha_n(\underline{\Theta}(x)) + \beta_n(\underline{\Theta}(x))$, a budući da $\beta_n(\underline{\Theta}(x))$ ne ovisi o x, vidimo da su za različite vrijednosti od n funkcije λ_n različite racionalne funkcije u x. Ovo povlači da su, osim za konačno mnogo vrijednosti od x, vrijednosti $\lambda_n(x)$ za navedene n-ove sve različite. Neka je $x^\# = [0, \gamma_1^\#, \gamma_2^\#, \gamma_3^\#, \ldots]$ jedna od tih vrijednosti tako da su $\lambda_n(x^\#)$ za n-ove koji odgovaraju pozicijama unutar bloka τ sve različiti brojevi. Budući da se $\sup(\alpha_n + \beta_n) = M(\underline{\Theta}(x^\#))$ postiže za vrijednosti n koje odgovaraju pozicijama unutar τ od $\underline{\Theta}(x^\#)$, neka je n_0 pozicija u τ za koju je $M(\underline{\Theta}(x^\#)) = \alpha_{n_0}(\underline{\Theta}(x^\#)) + \beta_{n_0}(\underline{\Theta}(x^\#))$. Za N dovoljno velik, uzevši

$$\begin{aligned}\tau^\# &= ((\gamma^{(0)})^N, \tau, \gamma_1^\#, \gamma_2^\#, \ldots, \gamma_n^\#) \\ &= (a_{-N_1}, \ldots, a_{-1}, a_0, a_1, \ldots, a_{N_2}), \ a_i \in \mathbb{N} \quad (a_0 \text{ na poziciji } n_0 \text{ u } \tau),\end{aligned}$$

vrijedi: ako je $\underline{\vartheta} = (\ldots, \gamma_{-2}, \gamma_{-1}, \tau^\#, \gamma_1, \gamma_2, \ldots)$, $\gamma_k \in X^*$, $\widetilde{\gamma_{-k}} \in \widetilde{X}^*$ za sve $k \in \mathbb{N}$, onda imamo

$$M(\underline{\vartheta}) = [a_0, a_1, \ldots, a_{N_2}, \gamma_1, \gamma_2, \gamma_3, \ldots] + [0, a_{-1}, a_{-2}, \ldots, a_{-N_1}, \widetilde{\gamma_{-1}}, \widetilde{\gamma_{-2}}, \widetilde{\gamma_{-3}}, \ldots].$$

Sada definiramo

$$\begin{aligned}K &= \{[a_0, a_1, a_2, \ldots, a_{N_2}, \gamma_1, \gamma_2, \gamma_3, \ldots] : \gamma_j \in X^*, \ j \in \mathbb{N}\}, \\ K' &= \{[0, a_{-1}, a_{-2}, \ldots, a_{-N_1}, \widetilde{\gamma}_1', \widetilde{\gamma}_2', \widetilde{\gamma}_3', \ldots] : \widetilde{\gamma}_j' \in \widetilde{X}^*, \ j \in \mathbb{N}\}\end{aligned}$$

te imamo $K + K' \subset L(\Sigma(X))$. Da bi se to pokazalo, za dane

$$\begin{aligned}x &= [a_0, a_1, a_2, \ldots, a_{N_2}, \gamma_1, \gamma_2, \gamma_3, \ldots] \in K \quad \text{i} \\ y &= [0, a_{-1}, a_{-2}, \ldots, a_{-N_1}, \widetilde{\gamma}_1', \widetilde{\gamma}_2', \widetilde{\gamma}_3', \ldots] \in K',\end{aligned}$$

definirajući za svaki $m \in \mathbb{N}$

$$\tau^{(m)} = (\gamma'_m, \gamma'_{m-1}, \dots, \gamma'_1, \tau^{\#}, \gamma_1, \gamma_2, \dots, \gamma_m),$$

stavimo

$$\underline{\Theta}_L(x,y) = (\dots, \gamma^{(0)}, \gamma^{(0)}, \gamma^{(0)}, \tau^{(1)}, \tau^{(2)}, \tau^{(3)}, \dots),$$
$$\underline{\Theta}_M(x,y) = (\dots, \gamma'_3, \gamma'_2, \gamma'_1, \tau^{\#}, \gamma_1, \gamma_2, \gamma_3, \dots).$$

Sada se iz svojstava koja smo osigurali ovom konstrukcijom pokaže da je

$$L(\underline{\Theta}_L(x,y)) = M(\underline{\Theta}_M(x,y)) = x+y.$$

Primijetimo još da su K i K' difeomorfni redom $K(X^*)$ i $K(\widetilde{X}^*)$, pa je

$$\mathrm{HD}(K) = \mathrm{HD}(K(X^*)) > \mathrm{HD}(K(X)) - \varepsilon \qquad \text{i}$$
$$\mathrm{HD}(K') = \mathrm{HD}(K(\widetilde{X}^*)) > \mathrm{HD}(K(\widetilde{X})) - \varepsilon = \mathrm{HD}(K(X)) - \varepsilon,$$

što smo i htjeli. □

Dokaz Teorema 7.3. Neka su $t \in (3,+\infty)$ i $\eta \in (0,1)$ po volji. Iz Leme 7.11 dobivamo $\delta > 0$ i $K(B)$ takve da je $\Sigma(B) \subset \Sigma_{t-\delta}$ i $\mathrm{HD}(K(B)) > (1-\eta)D(t)$. Primijenimo Lemu 7.12 za $X = B$. Dobivamo

$$\begin{aligned}
\min\{1, 2(1-\eta)D(t)\} &\le \min\{1, 2\,\mathrm{HD}(K(B))\} && \text{(Lema 7.11)}\\
&\le \mathrm{HD}\big(L(\Sigma(B))\big) && \text{(Lema 7.12)}\\
&\le \mathrm{HD}(\mathcal{L} \cap (-\infty, t-\delta]) && (\Sigma(B) \subset \Sigma_{t-\delta})\\
&\le \mathrm{HD}(\mathcal{L} \cap (-\infty, t)) && \text{(Propozicija 7.2.1)}\\
&\le \mathrm{HD}(\mathcal{M} \cap (-\infty, t)) && \text{(Teorem 3.6)}\\
&\le \overline{\dim}(\mathcal{M} \cap (-\infty, t)) && (\mathrm{HD}(E) \le \overline{\dim}(E))\\
&\le \min\{1, \overline{\dim}(K_t + K_t)\} && (\mathcal{M} \cap (-\infty, t) \subset \{1,2,\dots,T\} + K_t + K_t)\\
&\le \min\{1, 2\overline{\dim}(K_t)\} && (\overline{\dim}(E+E) \le 2\overline{\dim}(E))\\
&\le \min\{1, 2D(t)\}. && \text{(iz definicije od } D(t))
\end{aligned}$$

Dakle, označimo li $d(t) = \mathrm{HD}(\mathcal{L} \cap (-\infty, t))$, vrijedi

$$d(t) = \mathrm{HD}(\mathcal{M} \cap (-\infty, t)) = \overline{\dim}(\mathcal{L} \cap (-\infty, t)) = \overline{\dim}(\mathcal{M} \cap (-\infty, t)) = \min\{1, 2D(t)\}.$$

Pokažimo sada da je $\mathrm{HD}(L^{-1}(-\infty, t)) = D(t)$. Primijetimo da ovdje gledamo prasliku za Lagrangeov broj koji je funkcija definirana na $\mathbb{R} \setminus \mathbb{Q}$, a ne na $\mathbb{N}^{\mathbb{Z}}$, no to je jasno i iz konteksta (određivanje Hausdorffove dimenzije skupa). Koristeći notaciju iz dokaza Leme 7.12, neka su $x \in K$, $y \in K'$ te za svaki $z = [0, \alpha_1, \alpha_2, \dots] \in K(B^*)$, $\alpha_i \in B^*$, definiramo

$$\begin{aligned}
h(z) = h_{x,y}(z) = [0, \alpha_{1!}, \tau^{(1)}, \alpha_{2!}, \tau^{(2)}, \alpha_3, \alpha_4, \alpha_5, \alpha_{3!}, \tau^{(3)}, \alpha_7, \dots, \alpha_{4!}, \tau^{(4)},&\\
\alpha_{25}, \alpha_{26}, \dots, \alpha_{5!}, \tau^{(5)}, \dots, \alpha_{r!}, \tau^{(r)}, \alpha_{r!+1}, \dots].&
\end{aligned}$$

Kao i u dokazu navedene leme, dobiva se $L(h(z)) = x + y$. S druge strane, za proizvoljni $\rho > 0$ imamo $|z - z'| = \mathcal{O}(|h(z) - h(z')|^{1-\rho})$ za $|z - z'|$ mali, pa je $\mathrm{HD}(L^{-1}(x+y)) \ge \mathrm{HD}(K(B^*)) > \mathrm{HD}(K(B)) - \varepsilon$. Kao prije, imamo

$$\begin{aligned}
\mathrm{HD}(L^{-1}(-\infty, t)) &\ge \mathrm{HD}(L^{-1}(-\infty, t-\delta]) \ge \mathrm{HD}\big(L^{-1}(L(\Sigma(B)))\big)\\
&\ge \mathrm{HD}(L^{-1}(K + K')) \ge \mathrm{HD}(L^{-1}(x+y))\\
&> \mathrm{HD}(K(B)) - \varepsilon > (1-\eta)D(t) - \varepsilon.
\end{aligned}$$

Kako su η i ε proizvoljni, $\mathrm{HD}(L^{-1}(-\infty,t)) \geq D(t)$. Da dobijemo obratnu nejednakost, uzmimo $w \in L^{-1}(-\infty,t)$ po volji. Imamo $\limsup_{n\to\infty}(\alpha_n(w)+\beta_n(w)) = L(w) < t$, pa postoji $n_0 \in \mathbb{N}$ takav da za $n \geq n_0$ vrijedi $\alpha_n(w)+\beta_n(w) < t$. To povlači da je

$$L^{-1}(-\infty,t) \subset \bigcup_{n\in\mathbb{N}}(g^{-n}(K_t)),$$

gdje je g Gaussovo preslikavanje, iz čega je

$$\mathrm{HD}(L^{-1}(-\infty,t)) \leq \mathrm{HD}(K_t) \leq \overline{\dim}(K_t) = D(t).$$

Dakle, imamo $\mathrm{HD}(L^{-1}(-\infty,t)) = D(t)$.

U Lemi 7.10 smo dobili da je $D(t)$ neprekidna zdesna, pa vrijedi

$$D(t) = \mathrm{HD}(L^{-1}(-\infty,t)) \leq \mathrm{HD}(L^{-1}(-\infty,t]) \leq \lim_{s\to t+}\mathrm{HD}(L^{-1}(-\infty,s)) = \lim_{s\to t+} D(s) = D(t).$$

Zato je $\mathrm{HD}(L^{-1}(-\infty,t]) = \mathrm{HD}(L^{-1}(-\infty,t)) = D(t)$ te zaključujemo da je

$$d(t) = \min\{1, 2\,\mathrm{HD}(L^{-1}(-\infty,t))\} = \min\{1, 2\,\mathrm{HD}(L^{-1}(-\infty,t])\}.$$

Funkcija $D(t)$ je neprekidna i slijeva (pa je neprekidna) jer po Lemi 7.11 za dane $t \in [3,+\infty)$ i $\eta \in (0,1)$ postoji $\delta > 0$ tako da je $D(t-\delta) \geq \mathrm{HD}(K(B)) > (1-\eta)D(t)$ ($\Sigma(B) \subset \Sigma_{t-\delta}$ povlači $K(B) \subset K_{t-\delta}$), pa je $\lim_{s\to t-} D(s) = D(t)$.

Još treba pokazati tvrdnje (ii) i (iii).

Zbog $3 = [2,\overline{1}] + [0,2,\overline{1}]$, koristeći Teorem 1.8.b, lako se vidi da za svaki $m \in \mathbb{N}$ i $B_m = \{2\,1_{2m}\,2,\ 2\,1_{2m+2}\,2\}$ imamo $\Sigma(B_m) \subset \Sigma_{3+2^{-m}}$, pa je

$$d(3+2^{-m}) = \mathrm{HD}(\mathcal{M} \cap (-\infty, 3+2^{-m})) \geq \mathrm{HD}(\Sigma(B_m)) > \mathrm{HD}(K(B_m)) > 0.$$

Zato je $d(3+\varepsilon) > 0$ za svaki $\varepsilon > 0$.

U Propoziciji 6.12 smo vidjeli da je $\Sigma_{\sqrt{12}} = \{1,2\}^{\mathbb{Z}}$, pa je

$$D(\sqrt{12}) = \mathrm{HD}(K_{\sqrt{12}}) = \mathrm{HD}(K(\{1,2\})) = \mathrm{HD}(C(2)) = 0.531\ldots > \frac{1}{2}$$

te zbog neprekidnosti od $D(t)$ slijedi da postoji $\delta > 0$ takav da je $D(\sqrt{12}-\delta) > \frac{1}{2}$ i imamo $d(\sqrt{12}-\delta) = \min\{1, 2D(\sqrt{12}-\delta)\} = 1$. $\square$

Koristeći tehniku dokaza iz Leme 7.12, Moreira je u istom radu dokazao da je $\mathcal{L}'$ (topološki) savršen skup, tj. $\mathcal{L}'' = \mathcal{L}'$, gdje smo sa S' označili skup svih gomilišta skupa $S \subset \mathbb{R}$. Dokaz koristi činjenicu da je element Lagrangeovog spektra pridružen beskonačnom nizu $\underline{\vartheta}$ gomilište beskonačno mnogo $\lambda_n(\underline{\vartheta})$ što ne mora vrijediti za elemente Markovljevog spektra. Zato je pitanje je li $\mathcal{M}'' = \mathcal{M}'$ još uvijek otvoreno.

Spomenimo na kraju da su Matheus i Moreira u dva članka pokazali

$$0.353 < \mathrm{HD}(\mathcal{M} \setminus \mathcal{L}) < 0.987$$

uz još bolje heurističke ograde zasnovane na nekim numeričkim metodama aproksimacije Hausdorffove dimenzije Gauss-Cantorovih skupova. Iz $\mathrm{HD}(\mathcal{M}\setminus\mathcal{L}) < 1$ slijedi da $\mathcal{M}\setminus\mathcal{L}$ ima praznu nutrinu $\mathrm{int}(\mathcal{M}\setminus\mathcal{L})$, pa zbog zatvorenosti od $\mathcal{L}$ slijedi da je $\mathrm{int}(\mathcal{L}) = \mathrm{int}(\mathcal{M})$.

Poglavlje 8

Neki analogoni spektara

Klasični spektri prirodno se mogu interpretirati u terminima simboličke dinamike, pa se uvodi sljedeća definicija.

Ako je dano preslikavanje $\psi : X \to X$ i funkcija $f : X \to \mathbb{R}$, pridruženi *dinamički Markovljev i Lagrangeov spektar* definiraju se kao

$$\mathcal{M}(f,\psi) = \{\sup_{n\in\mathbb{N}} f(\psi^n(x)) \, : \, x \in X\} \qquad \text{i} \qquad \mathcal{L}(f,\psi) = \{\limsup_{n\to\infty} f(\psi^n(x)) \, : \, x \in X\}.$$

Za klasične spektre je $X = \mathbb{N}^{\mathbb{Z}}$, ψ je pomak, a f je funkcija λ_0. Sva pitanja koja smo promatrali u klasičnom slučaju, proučavana su i u slučaju dinamičkih spektara (primjerice, minimalni element u spektru, postojanje Hallove zrake, zatvorenost spektra ili određivanje Hausdorffove dimenzije dijela spektra).

Istaknimo još ukratko samo dva smjera u kojima su gledani analogoni spektara. Napomenimo da kod njih ostavljamo 'recipročnu' definiciju spektara jer je takva uobičajena u literaturi.

Imaginarna kvadratna polja

Neka je $d \in \mathbb{N}$ kvadratno slobodan i $\mathcal{O}_d$ prsten algebarskih cijelih brojeva u polju $\mathbb{Q}(\sqrt{-d})$. Označimo za $p, q \in \mathcal{O}_d$ sa $n(p,q)$ normu ideala generiranog sa p, q. Neka je za $\vartheta \in \mathbb{C}$

$$\nu_d(\vartheta) = \liminf_{(p,q)\in\mathcal{O}_d^2\setminus\{(0,0)\}} \frac{|q(q\vartheta - p)|}{n(p,q)}.$$

Tada nejednadžba

$$\left|\vartheta - \frac{p}{q}\right| < \nu_d(\vartheta)\frac{n(p,q)}{|q|^2}$$

ima beskonačno mnogo rješenja u $p, q \in \mathcal{O}_d$ takvih da je $n(p,q) < 2\sqrt{d}$. Skup brojeva $\mathcal{L}_d = \{\nu_d(\vartheta) \, : \, \vartheta \in \mathbb{C} \setminus \mathbb{Q}(\sqrt{-d})\}$ je *Lagrangeov spektar imaginarnog kvadratnog polja* $\mathbb{Q}(\sqrt{-d})$.

Hurwitzovu konstantu $\sup \mathcal{L}_d$ dobio je za $d = 1$ (Gaussovo polje) Ford 1925. pokazavši da je jednaka $\frac{1}{\sqrt{3}}$. Za još neke d odredili su ih Ford, Perron, Hofreiter, Poitou, Vulakh, a poznati su i drugi i viši minimumi za neke od vrijednosti d. Iz Vulakhovih rezultata slijedi i da postoji Hallova zraka (ovdje je to konačni interval s lijevim krajem u 0). Za $d \in \{1, 3, 5, 6, 11\}$ je određen čitav diskretni dio Lagrangeovog spektra i pokazano je da se podudara sa Markovljevim spektrom koji se ovdje definira

$$\mathcal{M}_d = \{\nu_d(f) \,:\, f(x,y) = ax^2 + bxy + cy^2,\ a,b,c \in \mathbb{C},\ \delta(f) = b^2 - 4ac \neq 0\},$$
$$\nu_d(f) = \inf_{(x,y)\in\mathcal{O}_d^2\setminus\{(0,0)\}} \left|\frac{f(x,y)}{n(x,y)\sqrt{\delta(f)}}\right|.$$

Kvadratni Lagrangeov spektar

Neka je $\alpha \in \mathbb{R}$ fiksirana kvadratna iracionalnost i $\mathcal{E}_\alpha$ skup kvadratnih iracionalnosti čiji razvoj u verižni razlomak se od nekog mjesta nadalje podudara sa α ili α' (konjugat u odgovarajućem realnom kvadratnom polju). Parkkonen i Paulin definirali su

$$c_\alpha(\xi) = \liminf_{\beta\in\mathcal{E}_\alpha \,:\, |\beta-\beta'|\to 0} 2\frac{|\xi-\beta|}{|\beta-\beta'|}$$

za realne brojeve ξ koji nisu u $\mathbb{Q}\cup\mathcal{E}_\alpha$. Sa

$$Sp_\alpha = \{c_\alpha(\xi) \,:\, \xi \in \mathbb{R}\setminus(\mathbb{Q}\cup\mathcal{E}_\alpha)\}$$

definiran je *kvadratni Lagrangeov* spektar za α.

Parkkonen i Paulin dokazali su da je Sp_α zatvoren podskup od $[0,(1+\sqrt{2})\sqrt{3}]$, a Bugeaud je pokazao da je Sp_α sadržan u $[0,\frac{1}{2}]$. Pejković je dokazao da je $Sp_\alpha \subset [0,\frac{3}{\sqrt{5}}-1] = [0, 0.341\ldots]$ što je ujedno i najbolje moguće jer za $\alpha = \frac{1+\sqrt{5}}{2}$ gornja ograda leži u spektru. Lin je dokazao postojanje Hallove zrake, tj. da za svaki α postoji neki (mali) $\psi(\alpha) > 0$ takav da je $[0,\psi(\alpha)] \subset Sp_\alpha$.

Parkkonen i Paulin te Bugeaud dobili su i rezultate vezane uz nearhimedski kvadratni Lagrangeov spektar, tj. onaj koji se pojavljuje kod polja formalnih redova potencija s koeficijentima iz konačnog polja.

Najvažnija literatura

Poglavlje 1.

- M. Aigner, Markov's theorem and 100 years of the uniqueness conjecture, Springer, Cham, 2013. (§1)
- Y. Bugeaud, Approximation by algebraic numbers, Cambridge University Press, Cambridge, 2004. (§1)
- A.M. Rockett, P. Szüsz, Continued fractions, World Scientific Publishing Co., River Edge, 1992. (§4)

Poglavlje 2.

- J. Buchmann, U. Vollmer, Binary quadratic forms. An algorithmic approach, Springer, Berlin, 2007. (§5,6)
- D.A. Buell, Binary quadratic forms. Classical theory and modern computations, Springer-Verlag, New York, 1989. (§2,3)
- L.E. Dickson, Introduction to the Theory of Numbers, The University of Chicago Press, 1929. (§7,11)
- A. Dujella, Teorija brojeva, Zagreb, 2019. (§5)
- A.M. Rockett, P. Szüsz, isto. (§4)

Poglavlje 3.

- T.W. Cusick, M.E. Flahive, The Markoff and Lagrange spectra, American Mathematical Society, Providence, 1989. (§1,3)

Poglavlje 4.

- C. Reutenauer, From Christoffel words to Markoff numbers, Oxford University Press, Oxford, 2019. (§1-11)

Poglavlje 5.

- M. Aigner, isto. (§3.3,4.3,10.1,10.2)
- Y. Bugeaud, C. Reutenauer, S. Siksek, A Sturmian sequence related to the uniqueness conjecture for Markoff numbers, Theoret. Comput. Sci. 410 (2009), 2864–2869.

Poglavlje 6.

- T.W. Cusick, M.E. Flahive, isto. (§1,3,4,5)
- G.A. Frejman, Diofantovy približenija i geometrija čisel (zadača Markova), Kalininskij gosudarstvennyj universitet, Kalinin, 1975.
- D. Gayfulin, Attainable numbers and the Lagrange spectrum, Acta Arith. 179 (2017), no. 2, 185–199.

Poglavlje 7.

- T.W. Cusick, M.E. Flahive, isto. (§6)
- K. Falconer, Fractal geometry. Mathematical foundations and applications. Third edition. John Wiley & Sons, Chichester, 2014. (§3)
- C. Matheus, The Lagrange and Markov spectra from the dynamical point of view. Ergodic theory and dynamical systems in their interactions with arithmetics and combinatorics, 259–291, Lecture Notes in Math., 2213, Springer, Cham, 2018.
- C.G. Moreira, Geometric properties of the Markov and Lagrange spectra, Ann. of Math. 188 (2018), no. 1, 145–170.

Poglavlje 8.

- A.V. Malyšev, Markov and Lagrange spectra (a survey of the literature). (Russian) Studies in number theory (LOMI), 4. Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 67 (1977), 5–38, 225. (translation in J. Soviet Math. (1981) 16, 767–788)
- L.Ya. Vulakh, Diophantine approximation in imaginary quadratic fields. Int. J. Number Theory 6 (2010), no. 4, 731–766.
- Y. Bugeaud, Nonarchimedean quadratic Lagrange spectra and continued fractions in power series fields, preprint (i ondje spomenuti radovi)