

Möbiusove transformacije in periodični verižni ulomki

Seminar

Nejc Zajc

Fakulteta za matematiko in fiziko

Oddelek za matematiko

17. marec 2020

1 Uvod

Z verižnimi ulomki oblike

$$b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \ddots}},$$

se večinoma prvič srečamo pri teoriji števil, kjer so njihovi členi naravna števila. A ko to ne velja več in so členi poljubna kompleksna števila, hitro opazimo potrebo po novih pristopih. V članku si bomo ogledali pristop z Möbiusovimi transformacijami. To so funkcije oblike

$$g(z) = \frac{az + b}{cz + d}, \quad (1)$$

kjer so a, b, c in d kompleksna števila, za katera velja $ad - bc \neq 0$.

Njihov pomen pri obravnavi verižnih ulomkov opazimo, če si definiramo $s_1(z) = \frac{az+1}{z} = a + \frac{1}{z}$ in $s_2(z) = \frac{bz+1}{z} = b + \frac{1}{z}$; tedaj je namreč

$$s(z) = s_1(s_2(z)) = a + \frac{1}{b + \frac{1}{z}}$$

končen verižni ulomek in hkrati Möbiusova transformacija. Transformacije bomo natančno definirali v poglavju 3 in jih uporabili v dokazu glavnega izreka tega članka, za začetek pa si natančneje oglejmo verižne ulomke.

2 Verižni ulomki

Osredotočimo se na enostavne verižne ulomke. Enostaven verižni ulomek ima vse $a_i = 1$. Vpeljimo tudi krajši zapis, ki je v končni obliki enak

$$[b_0, b_1, \dots, b_n] = b_0 + \frac{1}{b_1 + \frac{1}{\dots + \frac{1}{b_n}}},$$

kjer je b_0 celo število, b_1, b_2, \dots pa naravna števila; enostaven verižni ulomek je nato enak

$$[b_0, b_1, b_2, \dots] = b_0 + \frac{1}{b_1 + \frac{1}{b_2 + \dots}} = \lim_{n \rightarrow \infty} [b_0, b_1, \dots, b_n]. \quad (2)$$

Limita v (2) vedno obstaja, saj približki $[b_0, b_1, \dots, b_n]$ verižnega ulomka strogo naraščajo za sode n , so v tem primeru navzgor omejeni z $[b_0, b_1]$ in posledično konvergirajo. Za lihe n ti približki strogo padajo, so omejeni z $[b_0]$ in zato prav tako konvergirajo. Ker je razlika med zaporednima približkoma obratno sorazmerna z n^2 , sta limiti enaki, posledično pa limita (2) obstaja.

Za verižni ulomek $[b_0, b_1, b_2, \dots]$ rečemo, da je *periodičen* s periodo k , če je $b_n = b_{n+k}$ za vsa naravna števila n . To zapišemo $\overline{[b_0, \dots, b_{k-1}]}$. Za ulomek rečemo, da je *sčasoma periodičen*, če je $b_n = b_{n+k}$ za vse dovolj velike n . Opazimo, da v primeru, ko je $[b_0, b_1, b_2, \dots]$ periodičen s periodo k , velja $b_0 = b_k \geq 1$. Člen b_0 je torej naravno število, vrednost periodičnega verižnega ulomka pa tako vedno večja od 1.

2.1 Kvadratna iracionalna števila

Definicija 1. Realno število x je **kvadratno iracionalno število** (kvadratni iracional), če je iracionalno število in ničla kvadratnega polinoma P s celoštevilskimi koeficienti.

Naj bo x kvadratni iracional. Tedaj je ničla celoštevilskega kvadratnega polinoma in je zato oblike $x = \frac{a+b\sqrt{c}}{d}$, kjer so a, b, c in d cela števila, izmed katerih b, c in d ne smejo biti enaki nič, $c > 0$ pa ni popolni kvadrat. Ko vstavimo x v kvadratni celoštevilski polinom P , katerega ničla je, vidimo da je polinom do množenja s skalarjem enolično določen. Druga ničla polinoma P je *algebraična konjugirana vrednost* x , ki jo označimo z $x^* = \frac{a-b\sqrt{c}}{d}$.

Za zapis realnih števil z enostavnimi verižnimi ulomki velja, da lahko vsako racionalno število zapišemo kot končen verižni ulomek, vsakemu iracionalnemu številu pa pripada enolično določen verižni ulomek oblike (2). O verižnih ulomkih kvadratnih iracionalov lahko povemo še več, za iracionalno število $x = [b_0, b_1, b_2, \dots]$ namreč veljata naslednji lastnosti.

Trditev 1. *Verižni ulomek $[b_0, b_1, b_2, \dots]$ je sčasoma periodičen natanko tedaj, ko je x kvadratni iracional.*

Trditev 2. *Verižni ulomek $[b_0, b_1, b_2, \dots]$ je periodičen natanko tedaj, ko je x kvadratni iracional, katerega algebraična konjugirana vrednost x^* leži na intervalu $(-1, 0)$.*

Prvo ekvivalenco sta dokazala Euler, ki je pokazal, da sčasoma periodičen ulomek predstavlja kvadratni iracional, in Lagrange, ki je dokazal obrat. Drugo trditev pa je pokazal Galois.

Osrednji namen tega članka je pokazati, kako lahko s pomočjo Möbiusovih transformacij na verižnih ulomkih dokažemo naslednji Galois-ev izrek.

Izrek (Galois-ev izrek). *Za $x = [b_0, \dots, b_{k-1}]$ velja $[b_{k-1}, \dots, b_0] = -\frac{1}{x^*}$.*

Za konec poglavja si oglejmo zgled, ki pokaže veljavnost Galois-evega izreka na ulomku s periodo dolžine 2.

Zgled 1. *Naj $a, b \in \mathbb{N}$ in $\alpha = [a, b]$. S substitucijo dobimo $\alpha = a + 1/(b + 1/\alpha)$. Torej je α negibna točka $s(z) = a + 1/(b + 1/z)$. Za negibni točki funkcije s velja, da sta rešitvi enačbe*

$$bz^2 - abz - a = 0. \quad (3)$$

To sta torej α in α^ . Ker iz Vietovih formul sledi $\alpha\alpha^* = -\frac{a}{b} < 0$, velja $\alpha > 0 > \alpha^*$.*

Definirajmo še $\beta = [b, a]$. Enak premislek nas pripelje do ugotovitve, da sta β in β^ rešitvi $az^2 - abz - b = 0$ ter da velja $\beta > 0 > \beta^*$. Če na zadnji enačbi uporabimo transformacijo $w = -1/z$, dobimo enačbo (3). Torej za rešitve enačbe (3) velja $\{\alpha, \alpha^*\} = \{-1/\beta, -1/\beta^*\}$ in zato $\beta = -1/\alpha^*$, kar bi nam povedal tudi Galois-ev izrek.*

3 Širši pogled

Pred dokazom izreka, si bomo v tem poglavju ogledali lastnosti kompleksnih funkcij in ravnine, natančneje pa si bomo ogledali tudi Möbiusove transformacije.

3.1 Kompleksna ravnina

Ko kompleksni ravnini \mathbb{C} dodamo novo točko ∞ , s tem tvorimo *razširjeno kompleksno ravnino*, ki jo označimo $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$.

Definicija 2. Funkcija g z domeno \mathbb{C}_∞ je **Möbiusova transformacija**, če jo lahko zapišemo v obliki (1), kjer so a, b, c in d kompleksna števila za katera velja $ad - bc \neq 0$.

Če je $c \neq 0$, potem v (1) velja $g(\infty) = \frac{a}{c}$ in $g(-\frac{d}{c}) = \infty$, sicer je $g(\infty) = \infty$.

Vsaka Möbiusova transformacija g je bijekcija \mathbb{C}_∞ , saj ima inverz funkcije oblike (1) enak $g^{-1}(z) = \frac{-dz+b}{cz-a}$. Vidimo, da je g^{-1} Möbiusova transformacija, kratek račun pa nam utemelji, da to velja tudi za kompozitum dveh transformacij. Množica vseh Möbiusovih transformacij je torej grupa. Pri njihovem komponiranju si lahko pomagamo z množenjem matrik, ki kot člene vsebujejo koeficiente funkcije. Da dobimo na ustreznih mestih enake koeficiente, nam utemeljita naslednji enakosti

$$\frac{a_1 \frac{a_2 z + b_2}{c_2 z + d_2} + b_1}{c_1 \frac{a_2 z + b_2}{c_2 z + d_2} + d_1} = \frac{(a_1 a_2 + b_1 c_2)z + (a_1 b_2 + b_1 d_2)}{(c_1 a_2 + d_1 c_2)z + (c_1 b_2 + d_1 d_2)},$$

$$\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \cdot \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} = \begin{bmatrix} a_1 a_2 + b_1 c_2 & a_1 b_2 + b_1 d_2 \\ c_1 a_2 + d_1 c_2 & c_1 b_2 + d_1 d_2 \end{bmatrix}.$$

Označimo *razširjeno realno os* kot $\mathbb{R}_\infty = \mathbb{R} \cup \infty$. V primeru, ko so vsi koeficienti Möbiusove transformacije g realna števila, g ohranja \mathbb{R}_∞ . Tedaj velja

$$\operatorname{Im}[g(z)] = \frac{(ad - bc)\operatorname{Im}[z]}{|cz + d|^2},$$

kar pomeni, da g ohranja zgornjo kompleksno polravnino $\mathbb{H} = \{x + iy ; x, y \in \mathbb{R}, y > 0\}$ natanko tedaj, ko velja $ad - bc > 0$. Primer za to je funkcija $h(z) = -\frac{1}{z}$. Nasprotno pa se v primeru, ko $ad - bc = -1$, kompleksni polravnini ravno zamenjata, kot je to pri $k(z) = \frac{1}{z}$.

Omenimo še, kako \mathbb{C}_∞ opremimo z metriko. Stereografska projekcija je znan homeomorfizem med \mathbb{C} in enotsko sfero \mathbb{S} brez ene točke v \mathbb{R}^3 . To

projekcijo lahko razširimo do homeomorfizma med \mathbb{C}_∞ in celotno sfero \mathbb{S} ter nato prenesemo Evklidsko metriko iz \mathbb{S} v metriko χ na \mathbb{C}_∞ . Za metrični prostor $(\mathbb{C}_\infty, \chi)$ je nato vsaka Möbiusova transformacija g homeomorfizem prostora \mathbb{C}_∞ samega vase.

V primeru zgornje polravnine \mathbb{H} pa ob vpeljavi norme $||z|| = |z|/y$, kjer je $|z|$ absolutna vrednost kompleksnega števila z in $y = \text{Im}[z]$, dobimo Poincaré-jev model polravnine, ki je eden izmed standardnih modelov hiperbolične ravnine. Tu so Möbiusove transformacije, ki ohranjajo \mathbb{H} ($ad - bc > 0$), ravno vse izometrije \mathbb{H} . Meja prostora ustreza \mathbb{R}_∞ .

3.2 Modularna grupa

Definicija 3. Modularna grupa Γ je grupa vseh Möbiusovih transformacij oblike (1) s celoštevilskimi koeficienti a, b, c in d , za katere velja $ad - bc = 1$.

Kot smo že omenili, elementi Γ na \mathbb{H} delujejo kot izometrije hiperbolične metrike, njihovo delovanje na \mathbb{R}_∞ pa je tesno povezano s teorijo verižnih ulomkov. V grupi namreč med drugim leži tudi funkcija $s(z) = a+1/(b+1/z)$.

Posebno zanimive so *loksodromične izometrije* \mathbb{H} . To so Möbiusove transformacije, ki ohranjajo \mathbb{H} in imajo dve različni negibni točki. Primer takšne funkcije je $z \mapsto 2z$, katere negibni točki sta 0 in ∞ . Ob njihovi obravnavi pridemo do pomembne ugotovitve glede kvadratnih iracionalov.

Trditev 3. *Realno število x je kvadratno iracionalno število natanko tedaj ko je negibna točka nekega loksodromičnega elementa g modularne grupe Γ . Tedaj je algebraična konjugirana vrednost x^* druga negibna točka g .*

Tudi te trditve ne bomo dokazovali, uporabili pa jo bomo pri dokazovanju izreka. V ta namen si oglejmo še eno pomembno lastnost. Če je g loksodromična funkcija z negibnima točkama u in v , potem je ena izmed njih, recimo u , *privlačna negibna točka*, druga (v tem primeru v) pa je *odbojna negibna točka*. To pomeni da ob večkratni aplikaciji funkcije g na elementu $z \neq v$ v limiti $n \rightarrow \infty$ velja $g^n(z) = g(g(\dots(g(z)))) \rightarrow u$. V že omenjenem primeru $z \mapsto 2z$ je ∞ privlačna negibna točka, 0 pa je odbojna. Praviloma velja, da je negibna točka w poljubne funkcije f privlačna oziroma odbojna, če velja zaporedoma $|f'(w)| < 1$ oziroma $|f'(w)| > 1$.

4 Dokaz izreka

S pridobljenim znanjem bomo v tem poglavju dokazali Galois-ev izrek. Dokaz temelji na naslednji lemi, ki posploši pomen algebraične konjugirane vrednosti števila, saj b_i v lemi niso nujno cela števila. Za lažji zapis bomo v lemi namesto komponiranja uporabljali množenje, kot na primer $s_1 s_2(z) = s_1(s_2(z))$.

Lema 1. *Za funkcije s_i , $i \in \{1, \dots, k\}$ oblike $s : z \mapsto b + 1/z$, kjer je $b \geq 1$, ima končni kompozitum $S = s_1 \cdots s_k$ privlačno negibno točko $\zeta \in (1, \infty)$ in odbojno negibno točko $\tilde{\zeta} \in (-1, 0)$.*

Dokaz. Naj bo $S = s_1 \cdots s_k$, za $s_i(z) = b_i + 1/z$ in $b_i \geq 1$. Vsak s_i slika interval $[1, \infty)$ v omejen zaprt interval $I \subset (1, \infty)$, torej naredi kompozitum S enako. Ker velja

$$|s_i(z_1) - s_i(z_2)| < |z_1 - z_2|,$$

za vse i , lahko tudi to posplošimo na cel S . Po Banachovem skrčitvenem načelu iz teh lastnosti sledi obstoj negibne točke $\zeta \in (1, \infty)$ kompozituma S . Za določitev vrste negibne točke si pomagamo z velikostjo odvoda. Ker velja $|s'_i(z)| < 1$ na celem $(1, \infty)$, lahko preko verižnega pravila za odvajanje kompozituma sklepamo $|S'(\zeta)| < 1$. To pomeni, da je ζ privlačna negibna točka kompozituma S .

Definirajmo še $\tilde{S} = s_k \cdots s_1$. Analogni postopek nas pripelje do ugotovitve, da ima \tilde{S} privlačno negibno točko $\tilde{\zeta} \in (-1, 0)$. Naj bo $\sigma(z) = -1/z$. Tedaj velja

$$\delta s_i(z) = -\frac{1}{b + \frac{1}{z}} = \frac{1}{-\frac{1}{z} - b} = s_i^{-1} \delta(z),$$

za vse z in i . Posledično velja tudi $\delta S = \tilde{S}^{-1} \delta$. Ko v zadnjo enačbo vstavimo $\delta(\tilde{\zeta})$ vidimo, da je to negibna točka kompozituma S . Ker $\delta(\tilde{\zeta}) \in (-1, 0)$ je to od ζ različna negibna točka in je zato odbojna. \square

Združimo zdaj vse v dokazu Galois-evega izreka, ki ga zaradi preglednosti še enkrat zapišimo.

Izrek (Galois-ev izrek). *Za $x = \overline{[b_0, \dots, b_{k-1}]}$ velja $\overline{[b_{k-1}, \dots, b_0]} = -\frac{1}{x^*}$.*

Dokaz. Naj bodo $s_i = b_i + 1/z$, $i = 0, 1, 2, \dots$ funkcije z $b_i \geq 1$ za vse i . Po definiciji velja

$$[b_0, b_1, b_2, \dots] = \lim_{n \rightarrow \infty} s_0 \cdots s_n(\infty),$$

kjer smo za argument v kompozitumu vstavili ∞ , saj je to limitna točka argumentov. Predpostavimo da je zaporedje b_0, b_1, b_2, \dots periodično s periodo

k in označimo $S = s_0 \cdots s_{k-1}$. Privlačno negibno točko S označimo z $\zeta > 1$. Naj bo

$$K = \{\infty, s_0(\infty), s_0 s_1(\infty), \dots, s_0 \cdots s_{k-2}(\infty)\}.$$

S pomočjo leme sklepamo, da je $K \subset (1, \infty)$ in ker leži odbojna negibna točka S na intervalu $(-1, 0)$ velja $S^n(z) \rightarrow \zeta$ za vse $z \in K$. To bi lahko ekvivalentno povedali kot $s_0 \cdots s_n(\infty) \rightarrow \zeta$ ko $n \rightarrow \infty$. Torej je

$$\overline{[b_0, \dots, b_{k-1}]} = [b_0, \dots, b_{k-1}] = \zeta.$$

Ker je ζ enak periodičnemu verižnemu ulomku, je kvadratni iracional in po trditvi 3 velja, da je $S \in \Gamma$. Iz te trditve sledi tudi, da je druga (odbojna) negibna točka S enaka ζ^* .

Kot v lemi zdaj obrnimo periodo zaporedja na b_{k-1}, \dots, b_0 in označimo privlačno negibno točko kompozituma $\tilde{S} = s_{k-1} \cdots s_0$ z $\tilde{\zeta}$. Po enakem premisleku kot zgoraj velja

$$\overline{[b_{k-1}, \dots, b_0]} = \tilde{\zeta}.$$

Zdaj še enkrat uporabimo lemo pri sklepu, da je tudi $-1/\tilde{\zeta}$ odbojna negibna točka S . Ker ima S natanko eno odbojno odbojno negibno točko, torej velja $-1/\tilde{\zeta} = \zeta^*$. To nas pripelje do željenega rezultata

$$\overline{[b_{k-1}, \dots, b_0]} = -\frac{1}{\zeta^*}.$$

□

Angleško-slovenski slovar strokovnih izrazov

algebraic conjugate	algebraična konjugirana vrednost
Banach fixed-point theorem	Banachovo skrčitveno načelo
continued fraction	verižni ulomek
eventually periodic	časoma periodičen
extended complex plane	razširjena kompleksna ravnina
extended real axis	razširjena realna os
fixed point	negibna točka
loxodromic isometries	loksodromične izometrije
modular group	modularna grupa
Möbius map	Möbiusova transformacija
periodic	periodičen
quadratic irrational	kvadratno iracionalno število

Literatura

- [1] A. F. Beardon, *Möbius Maps and Periodic Continued Fractions*, Mathematics Magazine **88** (2015) 272–277.
- [2] *Hyperbolic geometry*, v: Wikipedia, The Free Encyclopedia, [ogled 27. 2. 2020], dostopno na https://en.wikipedia.org/wiki/Hyperbolic_geometry.
- [3] *Poincaré half-plane model*, v: Wikipedia, The Free Encyclopedia, [ogled 27. 2. 2020], dostopno na https://en.wikipedia.org/wiki/Poincar%C3%A9_half-plane_model.