Tartu Ülikool
Loodus- ja täppisteaduste valdkond
Matemaatika ja statistika instituut

Hendrik Vija

Greeni-Tao teoreem

Matemaatika

Bakalaureusetöö (9 EAP)

Juhendaja: PhD Lauri Tart

Tartu 2023

GREENI-TAO TEOREEM

Bakalaureusetöö

Hendrik Vija

Lühikokkuvõte

Greeni-Tao teoreem, 21. sajandi tuntumaid tulemusi arvuteooria valdkonnas, ütleb, et algarvude hulgas leidub kuitahes pikki mittekonstantseid aritmeetilisi jadasid. Esitame Greeni-Tao teoreemi mõnevõrra moodsama tõestuse, mis on originaaltõestusest lihtsam, lühem ja vähem tehniline.

CERCS teaduseriala: P120 Arvuteooria, väljateooria, algebraline geomeetria, algebra, rühmateooria.

Märksõnad: Aritmeetiline progressioon, algarvud, analüütiline arvuteooria, graafiteooria, normeeritud ruumid.

GREEN-TAO THEOREM

Bachelor thesis

Hendrik Vija

Abstract

The Green-Tao theorem, one of the more famous results of 21st century number theory, states that the prime numbers contain arbitrarily long non-constant arithmetic progressions. We present a simpler, shorter and less technical modern proof of the theorem.

CERCS research specialisation: P120 Number theory, field theory, algebraic geometry, algebra, group theory.

Key Words: Arithmetic progression, prime numbers, analytic number theory, graph theory, normed spaces.

Sisukord

Sissejuhatus

Selles töös esitame tõestuse Greeni-Tao teoreemile. Teoreemi sõnastus on järgmine:

Teoreem (Green-Tao)**.** *Algarvude hulgas leidub kuitahes pikki mittekonstantseid aritmeetilisi jadasid.*

Teoreemi tõestasid esmakordselt briti matemaatik Ben Green ja austraalia matemaatik Terence Tao 2004. aastal. Lihtsa sõnastuse ja tabava väite tõttu oli tegemist tuntud hüpoteesiga juba ammu enne seda, kui see tõestati. Seetõttu kogus uudis teoreemi tõestusest rohkelt tähelepanu ning paljud matemaatikud on põhjalikult uurinud tõestuse kõiki osi. Tänu sellele tähelepanule on erinevatel matemaatikutel õnnestunud esmase tõestuse leidmisest saati lihtsustada kõiki tõestuse põhilisi osi. See töö on täielikult referatiivne ja põhineb iiri matemaatiku David Conloni ning ameerika matemaatikute Jacob Foxi ja Yufei Zhao 2014. aasta artiklil, mis on kokkuvõte selleks ajaks leitud lihtsustustest, sealhulgas autorite enda omadest. (Conlon, Fox ja Zhao, 2014)

Täpsustame siinkohal, et Greeni-Tao teoreemi sõnastuses esinev fraas „leidub kuitahes pikki mittekonstantseid aritmeetilisi jadasid“ ei tähenda lõpmata pikkade jadade olemasolu, vaid seda, et iga $k \in \mathbb{N}$ korral leiduvad algarvud $a_1, \ldots, a_k$, nii et $a_2 - a_1 = a_3 - a_2 = \ldots = a_k - a_{k-1} \neq 0$. Edaspidi nimetame sellist k-elemendilist aritmeetilist jada *k-jadaks*. On kerge näha, et iga k-jada sisaldab endas ühtlasi ka l-jadasid iga $l < k$ jaoks.

Juhtudel $k = 1, 2$ on teoreemi väide ilmne, sest mistahes kaks arvu moodustavad aritmeetilise jada. Ka juhul $k = 3$ leiame kergesti, et algarvud $3, 5, 7$ moodustavad sobiva jada, kuid siis võime küsida, kas 3-jadasid on lõpmata palju. Vastus sellele küsimusele pole enam ilmselge.

Teoreemi tõestus põhineb Szemerédi teoreemil:

Teoreem (Szemerédi)**.** *Igas naturaalarvude hulgas $A \subseteq \mathbb{N}$, mille ülemine asümptootiline tihedus on positiivne, leidub iga $k \in \mathbb{N}$ korral k-jadasid.*

Peatükis 1 sõnastame vajalikud definitsioonid ning näeme, et Szemerédi teoreemi ei saa rakendada otse algarvude hulgale. Sellest hoolimata võtame töös eelduseks Szemerédi teoreemi kehtimise ning tõestame selle põhjal sarnase väite ka teatud hulkade jaoks, mille ülemine asümptootiline tihedus on 0, sealhulgas algarvude hulga jaoks. Seda väidet hakkame nimetama *suhteliseks Szemerédi teoreemiks*. Erijuhul $k = 3$ nimetame tavalist ja suhtelist Szemerédi teoreemi vastavalt tavaliseks ja suhteliseks Rothi teoreemiks.

Selle töö peatükis 1 tutvume lähemalt matemaatiliste tööriistadega, mida meil läheb vaja Greeni-Tao teoreemi tõestamiseks. Peatükis 2 tõestame Rothi teoreemi graafiteooria meetoditega. See tõestus sisaldab graafikonstruktsiooni, mis viib meid suhtelise Szemerédi teoreemi sõnastuseni, esmalt peatükis 3 juhul $k = 3$ ning siis peatükis 4 ka üldjuhul. Peatükkides 5 ja 6 tõestame vastavalt tiheda mudeli teoreemi ja loenduslemma, mis on suhtelise Szemerédi teoreemi tõestuse põhiosad ning peatükis 7 tõestame nende põhjal suhtelise Szemerédi teoreemi. Peatükkides 8 ja 9 konstrueerime algarvude hulgale tiheda mudeli ning näitame, et see tõesti rahuldab suhtelise Szemerédi teoreemi sõnastuses esitatud tingimusi. Sellega saab Greeni-Tao teoreem tõestatud.

Szemerédi teoreem, mida töös eeldame, on väga keeruline tulemus ning kõik selle teadaolevad tõestused on mahukad ning kasutavad omakorda põhjalikke teadmisi teistest valdkondadest, näiteks Fourier' analüüsist või hüpergraafide teooriast. Töös eeldame veel Dirichlet' teoreemi arvuteooria valdkonnast, Ruzsa-Szemerédi teoreemi graafiteooriast ning mitmeid teoreeme matemaatilise analüüsi erinevatest alamvaldkondadest. Neist esimesed kaks on samuti üpriski mahuka ja keerulise tõestusega tulemused.

Autori panus töösse on töö põhiallika arutluskäikude põhjalikum lahtikirjutamine, näiteks märkus 1.3 ja lemmad 1.5, 7.1, 8.3 on olulisemad väited, mis on allikas esitatud ilma põhjenduseta.

1 Töö teoreetiline taust

Szemerédi teoreemi sõnastus kasutab naturaalarvude hulga alamhulga ülemise asümptootilise tiheduse mõistet. Hiljem vajame ka suhtelise ülemise asümptootilise tiheduse mõistet.

Definitsioon 1.1 (Ülemine asümptootiline tihedus)**.** Hulga $A \subseteq \mathbb{N}$ *ülemiseks asümptootiliseks tiheduseks* nimetame piirväärtust

$$\limsup_{N\to\infty} \frac{|A \cap [N]|}{N}, \text{ kus } [N] := \{1, 2, \ldots, N\}.$$

Definitsioon 1.2 (Suhteline ülemine asümptootiline tihedus)**.** Olgu $A \subseteq S \subseteq \mathbb{N}$. Hulga A *suhteliseks ülemiseks asümptootiliseks tiheduseks hulgas* S nimetame piirväärtust

$$\limsup_{N\to\infty} \frac{|A \cap [N]|}{|S \cap [N]|}.$$

Märkus 1.3. Szemerédi teoreemist järeldub tegelikult mitte ainult ühe k-jada, vaid lõpmata paljude k-jadade olemasolu iga $k \in \mathbb{N}$ korral. Oletades vastuväiteliselt, et hulga $A \subseteq \mathbb{N}$ ülemine asümptootiline tihedus on $\varepsilon > 0$, kuid mingi k korral on k-jadasid lõplik arv m, saame valida iga $i = 1, \ldots, m$ korral i. k-jadast välja ühe elemendi $a_i \in \mathbb{N}$. Siis hulgas $A' := A \setminus \{a_1, \ldots, a_m\}$ ei ole ühtegi k-jada.

Kuid teisalt saame näidata, et ka hulga A' ülemine asümptootiline tihedus on ε. Kuna ühelt poolt $A' \subseteq A$, kuid teisalt on hulgas A' vaid m elementi vähem kui hulgas A, siis iga $N \in \mathbb{N}$ korral

$$|A \cap [N]| \geqslant \left|A' \cap [N]\right| \geqslant |A \cap [N]| - m.$$

Siit saame järeldada, et ühelt poolt

$$\varepsilon = \limsup_{N\to\infty} \frac{|A \cap [N]|}{N} \geqslant \limsup_{N\to\infty} \frac{|A' \cap [N]|}{N},$$

samas kui teiselt poolt

$$\limsup_{N\to\infty}\frac{|A'\cap[N]|}{N}\geqslant\limsup_{N\to\infty}\frac{|A\cap[N]|-m}{N}=\limsup_{N\to\infty}\frac{|A\cap[N]|}{N}-\limsup_{N\to\infty}\frac{m}{N}=\varepsilon.$$

See tähendab, et hulga A' ülemine asümptootiline tihedus on ε, nagu soovitud. Järelikult Szemerédi teoreemi põhjal leidub hulgas A' k-jadasid, mis annab vastuolu. Seega on hulgas A k-jadasid lõpmata palju, mida tahtsimegi näidata. Peaaegu identne argument toimib ka suhtelise Szemerédi teoreemi jaoks.

Kuna algarvude hulga ülemine asümptootiline tihedus on algarvude jaotuse teoreemi tõttu $\limsup\limits_{N\to\infty}\frac{1}{\ln N}=0$, ei järeldu Szemerédi teoreemist otse Greeni-Tao teoreem. Võrdluseks suhteline Szemerédi teoreem, mille sõnastame peatükis 4, kehtib juhul, kui uuritav hulk $A\subseteq\mathbb{N}$ on positiivse suhtelise ülemise asümptootilise tihedusega hulgas $S\subseteq\mathbb{N}$, mis rahuldab teatud *pseudojuhuslikkuse* tingimusi. Pseudojuhuslikkuse all mõtleme seda, et hulga S poolt teatud reegli järgi konstrueeritud graaf rahuldab teatud tingimusi, mida rahuldaks juhuslikult konstrueeritud sama servade arvuga graaf. Täpsemalt tutvume kasutatava pseudojuhuslikkuse tingimusega peatükkides 3 ja 4.

Et jõuda suhtelise Szemerédi teoreemi sõnastuseni, uurime kõigepealt Szemerédi teoreemi erijuhtu $k=3$ ehk Rothi teoreemi. Rothi teoreemi sõnastamiseks paneme tähele, et Szemerédi teoreem on samaväärne väitega, et kui naturaalarvude hulgas $A\subseteq\mathbb{N}$ ei leidu mingi $k\in\mathbb{N}$ korral ühtegi k-jada, siis hulga A ülemine asümptootiline tihedus on 0 ehk $\lim\limits_{N\to\infty}\frac{|A\cap[N]|}{N}=0$. Peatükis 2 uurime, kui suur saab olla hulk $A\subseteq[N]$, kui temas ei leidu ühtegi k-jada. Selle väljendamiseks kasutame järgnevat tähistust:

Definitsioon 1.4. Olgu $f,g:\mathbb{R}\to\mathbb{R}$.

1. Ütleme, et $f(x)=o(g(x))$, kui

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=0.$$

2. Ütleme, et $f(x) = O(g(x))$, kui leiduvad sellised reaalarvud x_0 ja $M > 0$, et iga $x > x_0$ korral $|f(x)| \leqslant M \cdot |g(x)|$, teisisõnu kui

$$\limsup_{x \to \infty} \left| \frac{f(x)}{g(x)} \right| < \infty.$$

Selle töö vältel uurime funktsioonide asümptootilist käitumist peaaegu alati N suhtes, vastasel juhul on vastav muutuja lisatud o või O juurde alaindeksina.

Et sõnastada ja tõestada Rothi teoreemi ning suhtelist Szemerédi teoreemi, hakkame uurima $[N]$ alamhulkade asemel hoopis Abeli rühma $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$ alamhulki. Need on aritmeetiliste jadade mõttes peaaegu identsed: iga aritmeetiline jada hulgas $[N]$ on ilmselgelt ka aritmeetiline jada rühmas $\mathbb{Z}_N$.

Vastupidine küll ilma lisaeeldusteta ei kehti, näiteks on $\overline{-1}, \overline{0}, \overline{1}$ 3-jada rühmas $\mathbb{Z}_N$, kuid vastavad hulga $[N]$ elemendid $N-1, N, 1$ ei moodusta 3-jada. Sellest probleemist möödapääsemiseks kasutame Greeni-Tao teoreemi tõestuses järgnevat lemmat:

Lemma 1.5. *Olgu $N \in \mathbb{N}$ ning moodustagu arvude $a_1, \ldots, a_k \in \mathbb{N} \cap [N/2, N)$ jäägiklassid k-jada rühmas $\mathbb{Z}_N$. Siis ka arvud $a_1, \ldots, a_k$ ise moodustavad k-jada.*

Tõestus. Kui $\overline{a_1}, \ldots, \overline{a_k}$ moodustavad k-jada rühmas $\mathbb{Z}_N$, siis kehtib $\overline{a_2 - a_1} = \ldots = \overline{a_k - a_{k-1}} \neq \overline{0}$. Järelikult arvudest $a_2 - a_1, \ldots, a_k - a_{k-1}$ iga kahe vahe on arvu N kordne. Kuid paneme tähele, et kuna iga $i = 1, \ldots, k$ korral $a_i \in \mathbb{N} \cap [N/2, N)$, siis $a_2 - a_1, \ldots, a_k - a_{k-1} \in \mathbb{Z} \cap (-N/2, N/2)$, mistõttu ühegi kahe vahe neist ei saa olla nullist erinev arvu N kordne, teisisõnu on nad kõik võrdsed. Samuti, kuna nad ei kuulu jäägiklassi $\overline{0}$, ei ole nad võrdsed nulliga. Seega $a_2 - a_1 = \ldots a_k - a_{k-1} \neq 0$ ehk arvud $a_1, \ldots, a_k$ moodustavad k-jada, nagu soovitud. $\square$

Nüüd oleme valmis sõnastama Rothi teoreemi.

Teoreem 1.6 (Roth). *Olgu $A \subseteq \mathbb{Z}_N$ selline, et temas ei leidu 3-jadasid. Siis $|A| = o(N)$.*

Suhtelises Rothi ja Szemerédi teoreemis on meil vaja hulga $A \subseteq \mathbb{Z}_N$ asemel hoopis rühma $\mathbb{Z}_N$ elementidel määratud funktsiooni f, mille väärtused on mittenegatiivsed reaalarvud. See funktsionaalne käsitlus on hulgateoreetilise käsitluse üldistus, sest mistahes hulga $A \subseteq \mathbb{Z}_N$ võime samastada tema karakteristliku funktsiooniga 1_A, mille väärtus on 1 hulga A elementidel ning 0 mujal, kuid funktsionaalne käsitlus lubab meil kasutada ka teistsuguseid, *kaalutud* funktsioone. Sellisel juhul vajame aga analoogi hulga A asümptootilise tiheduse mõistele. Selleks kasutame funktsiooni keskväärtuse mõistet:

Definitsioon 1.7 (Funktsiooni keskväärtus)**.** Olgu $f : X \to \mathbb{R}$, kusjuures olgu $|X|$ lõplik. Funktsiooni f *keskväärtuseks* nimetame suurust

$$\mathbb{E}_{x \in X} f = |X|^{-1} \sum_{x \in X} f(x).$$

Keskväärtust saame leida ka siis, kui f on mitme muutuja funktsioon, sel juhul on X muutujate määramispiirkondade otsekorrutis. Kui määramispiirkond on kontekstist selge, jätame selle välja kirjutamata.

Kasutame ka tähistust, kus jätame funktsioonide argumendid kirjutamata. Näiteks kirjutades sama määramispiirkonnaga X funktsioonide f, ν jaoks $f \leqslant \nu$, mõtleme, et $f(x) \leqslant \nu(x)$ iga $x \in X$ korral. Samuti mõtleme 1 all funktsiooni, mille kõik väärtused on ühed ning funktsioonide korrutamise all mõtleme alati nende punktiviisilist korrutamist, näiteks $f\nu(x) = f(x) \cdot \nu(x)$.

2 Rothi teoreemi tõestus

Selles peatükis tõestame Rothi teoreemi. Selleks eeldame, et hulgas $A \subseteq \mathbb{Z}_N$ ei leidu 3-jadasid ning näitame, et $|A| = o(N)$. Selleks kasutame järgmist tuntud teoreemi ekstremaalsest graafiteooriast:

Teoreem 2.1 (Ruzsa-Szemerédi)**.** *Olgu G graaf, milles on n tippu ning iga serv osaleb maksimaalselt ühes kolmnurgas. Siis graafis G on $o(n^2)$ serva.*

Rothi teoreemi tõestus. Olgu G_A kolmealuseline graaf alustega $X = Y = Z = \mathbb{Z}_N$, kus servad graafi aluste vahel on defineeritud järgnevalt:

- Tippude paari $(x, y) \in X \times Y$ vahel on serv parajasti siis, kui $2x + y \in A$;
- Tippude paari $(x, z) \in X \times Z$ vahel on serv parajasti siis, kui $x - z \in A$;
- Tippude paari $(y, z) \in Y \times Z$ vahel on serv parajasti siis, kui $-y - 2z \in A$.

Graafi on kujutatud joonisel 1. See ja kõik järgnevad joonised pärinevad töö põhiallikast, kuid neid on tõlgitud ja kohandatud.

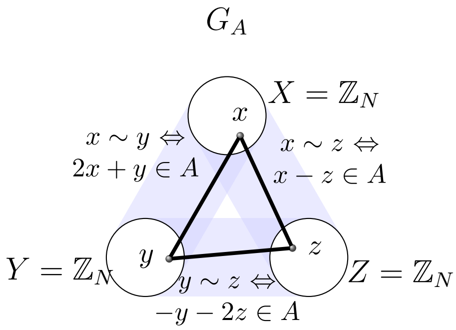
Tippude $(x, y, z) \in X \times Y \times Z$ vahelised servad moodustavad kolmnurga parajasti siis, kui rühma $\mathbb{Z}_N$ kõik kolm elementi

$$-y - 2z, \quad x - z, \quad 2x + y$$

kuuluvad hulka A. Need elemendid moodustavad aritmeetilise jada, mille vahe on $x + y + z$. See tähendab, et kolmnurk graafis G_A tippudega x, y, z vastab aritmeetilisele jadale $-y - 2z, x - z, 2x + y$ vahega $x + y + z$ rühmas $\mathbb{Z}_N$.

Rothi teoreemi eelduste põhjal ei ole rühmas $\mathbb{Z}_N$ 3-jadasid. Seega ainsad aritmeetilised jadad rühmas $\mathbb{Z}_N$ on need, kus jada vahe on $x + y + z = 0$.

Fikseerides $x \in X, y \in Y$, näeme, et on täpselt üks võimalus valida $z \in Z$, nii et $x + y + z = 0$, see on $z = -x - y$. See tähendab, et iga serv $(x, y) \in X \times Y$ kuulub



Joonis 1: Rothi teoreemi tõestuses kasutatav kolmealuseline graaf G_A.

täpselt ühte kolmnurka. Analoogselt kuulub ka iga serv $(x, z) \in X \times Z$ ning iga serv $(y, z) \in Y \times Z$ täpselt ühte kolmnurka.

Ruzsa-Szemerédi teoreem (Teoreem 2.1) ütleb, et sellisel juhul on graafi G_A servade arv $o(n^2)$, kus n on graafi tippude arv. Kuna graafis G_A on $3|\mathbb{Z}_N| = 3N$ tippu, on servade arv seega $o((3N)^2) = o(9N^2) = o(N^2)$.

Kuid teisalt paneme tähele, et iga $x \in X$ ja $a \in A$ jaoks leidub täpselt üks $y \in Y$, nii et $2x + y = a \in A$. See on $y = a - 2x$. See tähendab, et aluste X ja Y vahel leidub täpselt $|A| \cdot |X| = |A|N$ serva. Analoogselt näeme, et ka aluste X ja Z ning aluste Y ja Z vahel leidub täpselt $|A|N$ serva. Kokku on graafis G_A seega $3|A|N$ serva.

Võttes need kaks tulemust kokku, saame et $3|A|N = o(N^2)$ ehk $|A| = o(N)$, mida tahtsimegi tõestada. $\square$

Kasutatud graafikonstruktsiooni abil õnnestus meil hulga $A \subseteq \mathbb{Z}_N$ liikmete aritmeetiliseks jadaks olemist kirjeldada lineaarvormide kaudu, millest igaüks kasutab täpselt kahe muutujat kolmest, mis omakorda võimaldab meil kasutada graafiteoreetilisi meetodeid. Peatükkides 3 ja 4 näeme, et sarnane (hüper)graafikonstruktsioon aitab meid ka nii Rothi kui Szemerédi teoreemide suhteliste versioonide sõnastamisel ja tõestamisel.

3 Suhteline Rothi teoreem

Selles peatükis sõnastame suhtelise Rothi teoreemi.

Rothi teoreemis oli vaja 3-jadade olemasoluks seda, et rühma $\mathbb{Z}_N$ alamhulk A oleks positiivse ülemise asümptootilise tihedusega. Suhtelises Rothi teoreemis ei pea hulk A olema positiivse ülemise asümptootilise tihedusega. See-eest peab leiduma rühma $\mathbb{Z}_N$ alamhulk S, nii et

1. $A \subseteq S \subseteq \mathbb{Z}_N$,

2. A on positiivse suhtelise ülemise asümptootilise tihedusega hulgas S,

3. S rahuldab teatud *pseudojuhuslikkuse* tingimusi.

Edaspidi kasutame rühma $\mathbb{Z}_N$ alamhulkade $A \subseteq S$ asemel rühma $\mathbb{Z}_N$ elementidel määratud mittenegatiivsete reaalarvuliste väärtustega funktsioone. Olgu hulgale A vastav funktsioon f ning hulgale S vastav funktsioon ν. Eelnevalt sõnastatud tingimused 1 ja 2 sõnastame ümber järgnevalt:

1. $f \leqslant \nu$, teisisõnu iga $x \in \mathbb{Z}_N$ korral $f(x) \leqslant \nu(x)$,

2. Funktsioonide f ja ν keskväärtused rahuldavad mingi $\delta > 0$ jaoks tingimusi $\mathbb{E}\nu = 1 + o(1), \mathbb{E}f \geqslant \delta$.

Märkus 3.1. Motiveeriva näitena selle definitsiooni jaoks käsitleme olukorda, kus $A \subseteq S \subseteq \mathbb{Z}_N$ korral defineerime ν ja f vastavalt hulkade S ja A karakteristlike funktsioonide kordsetena:

$$\nu(x) = \frac{N}{|S|}1_S(x), \quad f(x) = \nu(x)1_A(x) = \frac{N}{|S|}1_A(x).$$

Sel juhul

$$\mathbb{E}\nu = \frac{1}{N}\sum_{x\in S}\frac{N}{|S|} = 1, \quad \mathbb{E}f = \frac{1}{N}\sum_{x\in A}\frac{N}{|S|} = \frac{|A|}{|S|}.$$

Siis paneme tähele, et tingimus $\mathbb{E}f \geqslant \delta$ kehtib mingi $\delta > 0$ ja iga $N \in \mathbb{N}$ korral parajasti siis, kui alati $|A| \geqslant \delta|S|$ ehk parajasti siis, kui hulk A on positiivse suhtelise ülemise tihedusega hulgas S.

Märkus 3.2. Et tingimus $\mathbb{E}\nu = 1 + o(1)$ omaks tähendust, peab funktsioon $\nu : \mathbb{Z}_N \to [0, \infty)$ olema defineeritud lõpmata paljude naturaalarvude N jaoks. Seega on õigem öelda, et ν on tegelikult hoopis funktsioonide jada $(\nu_N)_{N \in \mathbb{N}}$.

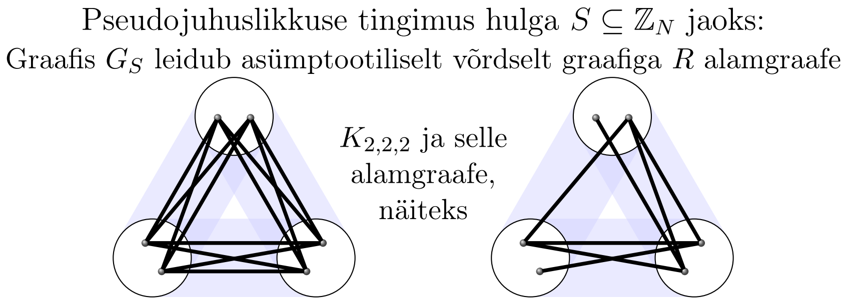
Viimane samm suhtelise Rothi teoreemi sõnastamiseks on defineerida *pseudojuhuslikkuse* tingimus, mida hulk S (ehk funktsionaalses käsitluses funktsioon ν) rahuldama peab. Esmalt sõnastame selle hulga S jaoks, seejärel asendame selle funktsiooniga ν.

Olgu $p := \frac{|S|}{N}$. Defineerime kolmealuselise graafi G_S analoogselt graafiga G_A eelmises peatükis. Nagu eelmises peatükis paneme tähele, et iga $x \in X$ ja $a \in S$ korral leidub täpselt üks $y \in Y$, nii et $2x + y = a$. See tähendab, et valides juhuslikult paari $(x, y) \in X \times Y$, on tõenäosus täpselt p, et graafis G_S on nende tippude vahel serv. Analoogselt on ka iga serva $(x, z) \in X \times Z$ ja $(y, z) \in Y \times Z$ leidumise tõenäosus graafis G_S täpselt p.

Selleks, et kontrollida graafi G_S *pseudojuhuslikkust*, võrdleme teda juhuslikult genereeritud kolmealuselise graafiga R, kus aluste vahel on iga tipu tõmbamise tõenäosus p. Täpsemalt võrdleme neid selle poolest, kas neis esineb sama alamgraafina sama palju graafi $K_{2,2,2}$ ning kõiki graafi $K_{2,2,2}$ alamgraafe. Graaf $K_{2,2,2}$ on täielik kolmealuseline graaf, mille iga alus koosneb kahest tipust. Graaf $K_{2,2,2}$ ning üks selle alamgraaf on toodud joonisel 2.

Et graafil $K_{2,2,2}$ on 6 tippu ja 12 serva, on graafis R selliste alamgraafide arv $(1 + o(1))N^6p^{12}$, sest iga tipu valikuks on (asümptootiliselt) N võimalust ning iga serva leidumise tõenäosus graafis R on p. Järelikult tahame, et ka graafis G_S oleks $(1 + o(1))N^6p^{12}$ alamgraafi $K_{2,2,2}$.

Graafi $K_{2,2,2}$ mistahes alamgraafi H jaoks, millel on $v(H) \leqslant 6$ tippu ja $e(H) \leqslant$



Joonis 2: Pseudojuhuslikkuse tingimus suhtelises Rothi teoreemis.

12 serva, näeme analoogselt, et tahame, et graafis G_S oleks $(1 + o(1))N^{v(H)}p^{e(H)}$ alamgraafi H.

Nüüd läheme üle hulgateoreetiliselt käsitluselt funktsionaalsele ehk hulgalt S üle funktsioonile ν. Selleks defineerime kaalutud kolmealuselise graafi ν, kus servade kaalud on määratud funktsiooni ν väärtuste põhjal järgmiste seostega:

- $\nu(x, y) = \nu(2x + y)$, kui $(x, y) \in X \times Y$;
- $\nu(x, z) = \nu(x - z)$, kui $(x, z) \in X \times Z$;
- $\nu(y, z) = \nu(-y - 2z)$, kui $(y, z) \in Y \times Z$.

Kirjutame eelnevad tingimused ümber keskväärtuste abil. Loomulik valik funktsiooni ν jaoks on $\nu(x) = \frac{N}{|S|}1_S(x)$. Sellisel juhul erinevad servade kaalud graafides G_S ja ν vaid konstantse kordaja poolest ning graafi tipud $x, x' \in X, y, y' \in Y, z, z' \in Z$ moodustavad alamgraafi $K_{2,2,2}$ parajasti siis, kui nende vahel leiduvad kõik servad

$$(x, y), (x', y), (x, y'), (x, z), (x', z), (x, z'), (x', z'), (y, z), (y', z), (y, z'), (y', z').$$

Selleks peavad kõik nendele servadele vastavad lineaarvormide väärtused kuuluma hulka S ehk funktsiooni ν väärtus nende lineaarvormide väärtustel olema $\frac{N}{|S|}$. Seega leidub graafis $(1+o(1))N^6p^{12}$ alamgraafi $K_{2,2,2}$ parajasti siis, kui

$$\begin{aligned}\mathbb{E}[\nu(y,z)\nu(y',z)\nu(y,z')\nu(y',z')\nu(x,z)\nu(x',z)\nu(x,z')\nu(x',z')\\ \cdot\nu(x,y)\nu(x',y)\nu(x,y')\nu(x',y')] =\\ = (N^6)^{-1}\cdot(1+o(1))N^6p^{12}\cdot\left(\frac{N}{|S|}\right)^{12} = 1+o(1). \quad (1)\end{aligned}$$

Analoogselt näeme, et selleks, et graafis G_S leiduks graafi $K_{2,2,2}$ mingit alamgraafi H $(1+o(1))N^{v(H)}p^{e(H)}$ tükki, peab võrdus (1) kehtima ka siis, kui eemaldada üks või rohkem selle vasaku poole teguritest. Seda pseudojuhuslikkuse tingimust kasutamegi suhtelises Rothi teoreemis. Sõnastame ta nii graafide kui funktsioonide jaoks.

Definitsioon 3.3 (3-lineaarvormide tingimus)**.** Ütleme, et kaalutud kolmealuseline graaf ν, mille alused on X, Y ja Z, rahuldab 3-*lineaarvormide tingimust*, kui

$$\begin{aligned}\mathbb{E}_{x,x'\in X,\ y,y'\in Y,\ z,z'\in Z}[\nu(y,z)\nu(y',z)\nu(y,z')\nu(y',z')\nu(x,z)\nu(x',z)\nu(x,z')\nu(x',z')\\ \cdot\nu(x,y)\nu(x',y)\nu(x,y')\nu(x',y')] = 1+o(1) \quad (2)\end{aligned}$$

ning võrdus (2) jääb kehtima ka ühe või enama teguri eemaldamisel keskväärtusest.

Sarnaselt ütleme, et funktsioon $\nu\colon \mathbb{Z}_N\to[0,\infty)$ rahuldab 3-*lineaarvormide tingimust*, kui

$$\begin{aligned}\mathbb{E}_{x,x',y,y',z,z'\in\mathbb{Z}_N}[\nu(-y-2z)\nu(-y'-2z)\nu(-y-2z')\nu(-y'-2z')\nu(x-z)\nu(x'-z)\\ \cdot\nu(x-z')\nu(x'-z')\nu(2x+y)\nu(2x'+y)\nu(2x+y')\nu(2x'+y')] = 1+o(1) \quad (3)\end{aligned}$$

ning võrdus (3) jääb kehtima ka ühe või enama teguri eemaldamisel keskväärtusest.

Rothi teoreem väidab 3-jada olemasolu rühmas $\mathbb{Z}_N$ ehk et leiduvad $x\in\mathbb{Z}_N, \overline{0}\neq d\in\mathbb{Z}_N$, nii et kõik kolm arvu $x, x+d, x+2d$ kuuluvad hulka A. See tähendab, et $f(x), f(x+d), f(x+2d)$ peavad olema positiivsed ehk $f(x)f(x+d)f(x+2d)>0$.

Suhteline Rothi teoreem väidab, et mingi konstandi $c > 0$ jaoks kehtib

$$\mathbb{E}[f(x)f(x+d)f(x+2d)] \geq c - o(1),$$

sest siis on piisavalt suure $N \in \mathbb{N}$ korral võrrandi (3) vasakul poolel olev keskväärtus positiivne ehk peavad leiduma $x, d \in \mathbb{Z}_N$, nii et $f(x)f(x+d)f(x+2d) > 0$:

Teoreem 3.4 (Suhteline Rothi teoreem)**.** *Olgu* $\nu\colon \mathbb{Z}_N \to [0, \infty)$ *funktsioon, mis rahuldab* 3*-lineaarvormide tingimust. Siis iga* $\delta > 0$ *korral leidub* $c = c(\delta) > 0$*, nii et iga funktsiooni* $f\colon \mathbb{Z}_N \to [0,\infty)$ *jaoks, mis rahuldab tingimusi* $0 \leqslant f \leqslant \nu$ *ja* $\mathbb{E}f \geq \delta$*, kehtib*

$$\mathbb{E}_{x,d\in\mathbb{Z}_N}[f(x)f(x+d)f(x+2d)] \geq c - o(1).$$

4 Suhteline Szemerédi teoreem

Selles peatükis sõnastame suhtelise Szemerédi teoreemi.

Suhtelise Rothi teoreemi (ehk suhtelise Szemerédi teoreemi juhu $k = 3$) sõnastamiseks konstrueerisime kaalutud kolmealuselise graafi, kus iga kahe erinevatesse alustesse kuuluva tipu vahelise serva kaal sõltub funktsiooni $\nu \colon \mathbb{Z}_N \to [0, \infty)$ väärtusest teatud lineaarvormil, mis sõltub täpselt nendele kahele tipule vastavatest rühma $\mathbb{Z}_N$ elementidest.

Suhtelise Szemerédi teoreemi üldjuhul konstrueerime k-aluselise hüpergraafi ν alustega $X_1 = \ldots = X_k = \mathbb{Z}_N$, kus iga $k-1$ erinevatesse alustesse kuuluva tipu vahelise serva kaal sõltub funktsiooni ν väärtusest ühel järgmistest lineaarvormidest (selguse eesmärgil kirjutame välja ka kordajad 0 ja 1):

- $0x_1 - 1x_2 - \ldots - (k-1)x_k$,
- $1x_1 + 0x_2 - \ldots - (k-2)x_k$,
- $\ldots$,
- $(k-1)x_2 + (k-2)x_2 + \ldots + 0x_k$.

Teisisõnu, iga $j = 1, \ldots, k$ korral sõltub j. lineaarvorm kõigist muutujatest peale x_j ning lineaarvormid moodustavad aritmeetilise jada vahega $x_1 + \ldots + x_k$. Kompaktsemalt kirjutades on graafi ν servade kaalud määratud iga $j = 1, \ldots, k$ jaoks lineaarvormidega

$$\nu(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k) = \nu\left(\sum_{i=1}^{k}(j-i)x_i\right).$$

Suhtelises Rothi teoreemis kasutatud 3-lineaarvormide tingimus seisnes selles, et kolmealuselises kaalutud graafis ν leiduks asümptootiliselt sama palju alamgraafi $K_{2,2,2}$ ning selle alamgraafe kui juhuslikult genereeritud sama servade tihedusega kolmealuselises graafis.

Üldjuhul uurime, kui palju on hüpergraafis ν alamhüpergraafi $K_k^{(k-1)}$. See on k-aluseline hüpergraaf, mille kõik servad ühendavad $k-1$ tippu, mis kuuluvad erinevatesse alustesse. Täpsemalt koosneb graafi $K_k^{(k-1)}$ iga alus kahest tipust ning kõikvõimalikud servad on tõmmatud.

Et graafi ν tippudest $x_1^{(0)}, x_1^{(1)} \in X_1, \ldots, x_k^{(0)}, x_k^{(1)} \in X_k$ ja nendevahelistest servadest moodustuv alamgraaf oleks $K_k^{(k-1)}$, peab kehtima järgmine tingimus: alati, kui valida

1. alustest $X_1, \ldots, X_k$ välja $k-1$ tükki (see on samaväärne sellega, et valida $j \in \{1, \ldots, k\}$ ning siis kõik alused peale X_j),
2. iga väljavalitud aluse X_i jaoks, kumba selle aluse elementi ($x_i^{(0)}$ või $x_i^{(1)}$) serv sisaldab,

peab tekkinud serv olema ka hüpergraafis ν. Seega saame järgneva tingimuse:

Definitsioon 4.1 (k-lineaarvormide tingimus)**.** Ütleme, et funktsioon $\nu : \mathbb{Z}_N \to [0, \infty)$ rahuldab *k-lineaarvormide tingimust*, kui kehtib võrdus

$$\mathbb{E}_{x_1^{(0)}, x_1^{(1)}, \ldots, x_k^{(0)}, x_k^{(1)} \in \mathbb{Z}_N} \Big[\prod_{j=1}^{k} \prod_{\omega \in \{0,1\}^{[k] \setminus \{j\}}} \nu \Big(\sum_{i=1}^{k} (j-i) x_i^{(\omega_i)} \Big)^{n_{j,\omega}} \Big] = 1 + o(1) \tag{4}$$

kõigi astendajate kombinatsioonide $n_{j,\omega} \in \{0, 1\}$ jaoks.

Me soovime hulgas A leida k-jadasid ehk siis elemente $x, d \in \mathbb{Z}_N, d \neq \overline{0}$, nii et elemendid $x, x+d, \ldots, x+(k-1)d$ kuuluksid kõik hulka A. Teisisõnu soovime, et $f(x)f(x+d)\ldots f(x+(k-1)d) > 0$.

Teoreem 4.2 (Suhteline Szemerédi teoreem)**.** *Olgu $k \geq 3$ naturaalarv ning rahuldagu funktsioon $\nu \colon \mathbb{Z}_N \to [0, \infty)$ k-lineaarvormide tingimust. Siis leidub iga $\delta > 0$ jaoks $c = c(k, \delta) > 0$, nii et iga funktsiooni $f \colon \mathbb{Z}_N \to [0, \infty)$ korral, mis rahuldab tingimusi $0 \leqslant f \leqslant \nu$ ja $\mathbb{E}f \geq \delta$, kehtib võrratus*

$$\mathbb{E}_{x,d \in \mathbb{Z}_N} [f(x)f(x+d)f(x+2d) \cdots f(x+(k-1)d)] \geq c - o(1). \tag{5}$$

Märkus 4.3. Me kasutame suhtelise Szemerédi teoreemi tõestamiseks Szemerédi teoreemi. Kuna läksime hulgateoreetiliselt käsitluselt üle funktsionaalsele, on meil vaja ka Szemerédi teoreem sõnastada funktsioonide keeles. Seega kasutame kaalutud Szemerédi teoreemi:

Teoreem 4.4 (Kaalutud Szemerédi teoreem)**.** *Iga $k \geq 3$ ja $\delta > 0$ korral leidub $c = c(k, \delta) > 0$, nii et iga funktsioon $f\colon \mathbb{Z}_N \to [0,1]$, mille korral $\mathbb{E}f \geq \delta$, rahuldab tingimust*

$$\mathbb{E}_{x,d \in \mathbb{Z}_N}[f(x)f(x+d)f(x+2d)\cdots f(x+(k-1)d)] \geq c - o(1). \tag{6}$$

Järgmises kolmes peatükis tõestame suhtelise Szemerédi teoreemi. Esitame siinkohal tõestuse skeemi.

Peatükis 5 tõestame *tiheda mudeli teoreemi*, mis ütleb, et leidub funktsioon $\tilde{f}\colon \mathbb{Z}_N \to [0,1]$, mis on teatud *lõikenormi* poolest väga lähedane funktsioonile f.

Peatükis 6 tõestame *loenduslemma*, mis ütleb, et kui funktsioonid f ja $\tilde{f}$ on lõikenormi poolest lähedased, siis nad tekitavad hulgas $\mathbb{Z}_N$ asümptootiliselt võrdselt k-jadasid, teisisõnu et

$$\mathbb{E}_{x,d}[f(x)f(x+d)\cdots f(x+(k-1)d)] = \mathbb{E}_{x,d}[\tilde{f}(x)\tilde{f}(x+d)\cdots \tilde{f}(x+(k-1)d)] - o(1).$$

Viimaks, kuna kaalutud Szemerédi teoreemi (Teoreemi 4.4) põhjal on viimase võrrandi parem pool mingi konstandi $c > 0$ jaoks suurem-võrdne avaldisest $c - o(1)$, järeldub sama ka vasaku poole kohta, mis tõestab suhtelise Szemerédi teoreemi.

5 Tiheda mudeli teoreem

Selles peatükis tõestame *tiheda mudeli teoreemi*, mis ütleb, et funktsiooni $f : \mathbb{Z}_N \to [0, \infty), \mathbb{E}f \geqslant \delta$ jaoks leidub funktsioon $\tilde{f} \colon \mathbb{Z}_N \to [0,1]$, mis on teatud *lõikenormi* poolest väga sarnane funktsioonile f, samuti rahuldab ta tingimust $\mathbb{E}\tilde{f} \geq \delta - o(1)$.

Lõikenormi mõiste pärineb graafiteooriast: kui kaalutud kahealuselise graafi g (alustega X, Y) servade kaalud on määratud funktsiooniga $g : X \times Y \to \mathbb{R}$, siis nii graafi kui funktsiooni g *lõikenorm* $\|g\|_\square$ on defineeritud seosega

$$\|g\|_\square := \sup_{A \subseteq X, B \subseteq Y} \left| \mathbb{E}_{x \in X, y \in Y} [g(x,y) 1_A(x) 1_B(y)] \right|. \tag{7}$$

Üldistame lõikenormi mõistet kaalutud k-aluselistele hüpergraafidele. Selleks võtame kasutusele uue tähistuse:

Olgu $X_1, \ldots, X_k$ hulgad ja $x_1 \in X_1, \ldots, x_k \in X_k$. Siis tähistame iga $i = 1, \ldots, k$ jaoks $X_{-i} := \prod_{j \in [k] \setminus \{i\}} X_j$ ja $x_{-i} := (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k) \in X_{-i}$.

Nüüd defineerime üldistatud lõikenormi nii hüpergraafide kui ka hulgal $\mathbb{Z}_N$ määratud funktsioonide jaoks:

Definitsioon 5.1. Olgu g kaalutud r-aluseline hüpergraaf alustega $X_1, \ldots, X_r$, kus iga serv ühendab r tippu. Siis hüpergraafi g lõikenorm $\|g\|_{\square,r}$ on määratud seosega

$$\|g\|_{\square,r} := \sup_{A_1, \ldots, A_r \in \mathbb{Z}_N^{r-1}} \left| \mathbb{E}_{x_1 \in X_1, \ldots, x_r \in X_r} [g(x_1, \ldots, x_r) 1_{A_1}(x_{-1}) 1_{A_2}(x_{-2}) \cdots 1_{A_r}(x_{-r})] \right| \tag{8}$$

ning kaalutud k-aluselises hüpergraafis g alustega $X_1, \ldots, X_k$, kus iga serv ühendab $k-1$ tippu, võtame $\|g\|_{\square,k-1} = \max\{\|g_{-1}\|_{\square,k-1}, \ldots, \|g_{-k}\|_{\square,k-1}\}$.

Olgu $f \colon \mathbb{Z}_N \to \mathbb{R}$ ning olgu $r \geqslant 2$ naturaalarv. Siis funktsiooni f lõikenorm $\|f\|_{\square,r}$

on määratud seosega

$$\|f\|_{\square,r} := \sup_{A_1,\ldots,A_r \subseteq \mathbb{Z}_N^{r-1}} \left|\mathbb{E}_{x_1,\ldots,x_r\in\mathbb{Z}_N}[f(x_1+\cdots+x_r)1_{A_1}(x_{-1})\cdots 1_{A_r}(x_{-r})]\right|. \quad (9)$$

Edasises kasutame funktsioonide $\mathbb{Z}_N \to \mathbb{R}$ vektorruumil määratud skalaarkorrutist

$$\langle f, \varphi\rangle = \mathbb{E}_x[f(x)\varphi(x)].$$

Teisendame võrrandis (9) esineva keskväärtuse sellisele kujule. Selleks kasutame funktsioonide *konvolutsiooni*. Defineerime kahe funktsiooni konvolutsiooni seosega

$$h_1 * h_2(z) := \mathbb{E}_x[h_1(x)h_2(z-x)].$$

Juhul $r = 2$ saame muutujavahetuse $x + y = z$ abil kirjutada

$$\mathbb{E}_{x,y}[f(x+y)1_A(x)1_B(y)] = \mathbb{E}_{x,z}[f(z)1_A(x)1_B(z-x)] = \langle f, 1_A * 1_B\rangle.$$

Olgu Φ_2 selliste funktsioonide $\varphi : \mathbb{Z}_N \to \mathbb{R}$ hulk, mis esituvad mingite $A, B \subseteq \mathbb{Z}_N$ jaoks konvolutsioonina $1_A * 1_B$. Siis

$$\|f\|_{\square,2} = \sup_{A,B\subseteq\mathbb{Z}_N} |\langle f, 1_A * 1_B\rangle| = \sup_{\varphi\in\Phi_2} |\langle f,\varphi\rangle|.$$

Et minna juhult $r = 2$ üle üldjuhule, kirjutame kahe funktsiooni konvolutsiooni definitsiooni ümber kujul

$$h_1 * h_2(z) = \mathbb{E}_x[h_1(x)h_2(z-x)] = \mathbb{E}_{\substack{x,y\in\mathbb{Z}_N \\ x+y=z}}[h_1(x)h_2(y)].$$

Analoogselt defineerime r funktsiooni $h_1, \ldots, h_r \colon \mathbb{Z}_N^{r-1} \to \mathbb{R}$ jaoks nende *üldistatud konvolutsiooni* $(h_1, \ldots, h_r)^* \colon \mathbb{Z}_N \to \mathbb{R}$ seosega

$$(h_1, \ldots, h_r)^*(x) = \mathbb{E}_{\substack{y_1,\ldots,y_r\in\mathbb{Z}_N \\ y_1+\cdots+y_r=x}}[h_1(y_{-1})h_2(y_{-2})\cdots h_r(y_{-r})].$$

Olgu Φ_r selliste funktsioonide $\varphi : \mathbb{Z}_N \to \mathbb{R}$ hulk, mis esituvad mingite $A_1, \ldots, A_r \subseteq \mathbb{Z}_N^{r-1}$ korral üldistatud konvolutsioonina $(1_{A_1}, 1_{A_2}, \ldots, 1_{A_r})^*$. Siis

$$\|f\|_{\square,r} = \sup_{A_1,\ldots,A_r \subseteq \mathbb{Z}_N^{r-1}} |\langle f, (1_{A_1}, \ldots, 1_{A_r})^* \rangle| = \sup_{\varphi \in \Phi_r} |\langle f, \varphi \rangle| .$$

Järgmine lemma on üks hulga Φ_r põhiomadusi.

Lemma 5.2. *Hulk Φ_r on kinnine funktsioonide korrutamise suhtes.*

Tõestus. Piisab näidata, et kui $\varphi = (1_{A_1}, \cdots, 1_{A_r})^*$ ja $\varphi' = (1_{B_1}, \ldots, 1_{B_r})^*$, kus $A_1, \ldots, A_r, B_1, \ldots, B_r \subseteq \mathbb{Z}_N^{r-1}$, siis $\varphi\varphi' \in \Phi_r$. Kirjutame iga vektori $y = (y_1, \ldots, y_r) \in \mathbb{Z}_N^r$ jaoks $\Sigma y = y_1 + \cdots + y_r$. Siis iga $x \in \mathbb{Z}_N$ korral

$$\begin{aligned}
\varphi(x)\varphi'(x) &= \mathbb{E}_{\substack{y,y' \in \mathbb{Z}_N^r \\ \Sigma y = \Sigma y' = x}} [1_{A_1}(y_{-1}) 1_{B_1}(y'_{-1}) \cdots 1_{A_r}(y_{-r}) 1_{B_r}(y'_{-r})] \\
&= \mathbb{E}_{\substack{y,z \in \mathbb{Z}_N^r \\ \Sigma y = x, \Sigma z = 0}} [1_{A_1}(y_{-1}) 1_{B_1}(y_{-1} + z_{-1}) \cdots 1_{A_r}(y_{-r}) 1_{B_r}(y_{-r} + z_{-r})] \\
&= \mathbb{E}_{\substack{y,z \in \mathbb{Z}_N^r \\ \Sigma y = x, \Sigma z = 0}} [1_{A_1 \cap (B_1 - z_{-1})}(y_{-1}) \cdots 1_{A_r \cap (B_r - z_{-r})}(y_{-r})] \\
&= \mathbb{E}_{\substack{z \in \mathbb{Z}_N^r \\ \Sigma z = 0}} [(1_{A_1 \cap (B_1 - z_{-1})}, \ldots, 1_{A_r \cap (B_r - z_{-r})})^*(x)].
\end{aligned}$$

Oleme funktsiooni $\varphi\varphi'$ esitanud üldistatud konvolutsioonina, seega $\varphi\varphi' \in \Phi_r$. $\square$

Edasises fikseerime arvu r ning kirjutame $\|\cdot\|_{\square,r}$ asemel $\|\cdot\|$ ning Φ_r asemel Φ.

Defineerime normi $\|\cdot\|$ duaalse normi $\|\psi\|^* := \sup_{\|f\| \leqslant 1} \langle f, \psi \rangle$. Tähistades $f' = \frac{f}{\|f\|}$, kehtib $\|f'\| = 1$, seega saame normi $\|\cdot\|^*$ definitsioonist $|\langle f, \psi \rangle| = \|f\| \, |\langle f', \psi \rangle| \leqslant \|f\| \, \|\psi\|^*$.

Näitame, et (kinnine) ühikkera B normi $\|\cdot\|^*$ suhtes on hulga $X := \Phi \cup -\Phi$ *kumer kate* ehk hulk $\mathrm{ch}(X) := \{tx_1 + (1-t)x_2 \mid t \in [0,1], x_1, x_2 \in X\}$.

Hulga X iga element kuulub ühikkerasse, sest oletades vastuväiteliselt, et mingi $\varphi \in \Phi, \|f\| \leqslant 1$ korral $|\langle f, \varphi \rangle| > 1$, saame, et $1 \geqslant \|f\| = \sup_{\varphi' \in \Phi} |\langle f, \varphi' \rangle| \geqslant$

$|\langle f, \varphi\rangle| > 1$, vastuolu. Analoogselt ka hulga $-\Phi$ iga element kuulub ühikkerasse ning kolmnurga võrratusest normi $\|\cdot\|^*$ jaoks saame, et ühikkerasse kuuluvad ka lineaarkombinatsioonid $tx_1 + (1-t)x_2, t \in [0,1], x_1, x_2 \in X$. Seega $\mathrm{ch}(X) \subseteq B$.

Tõestamaks, et $B \subseteq \mathrm{ch}(X)$, oletame vastuväiteliselt, et $\psi \in B$, aga $\psi \notin \mathrm{ch}(X)$. Järgmiseks kasutame analüüsis tuntud *teoreemi eraldavast hüpertasandist* (Boyd ja Vandenberghe, 2004):

Teoreem 5.3 (Teoreem eraldavast hüpertasandist)**.** *Olgu $B \subseteq \mathbb{R}^n$ kumer ja kinnine ning olgu $a \in \mathbb{R}^n, a \notin B$. Siis leidub vektor $v \in \mathbb{R}^n$, nii et $\langle a, v\rangle > 1$ ja $\langle b, v\rangle \leqslant 1$ iga $b \in B$ korral.*

Kuna vektorid ruumis $\mathbb{R}^n$ on samastatavad funktsioonidega $\mathbb{Z}_N \to \mathbb{R}$ ning skalaarkorrutised ruumides erinevad vaid konstantse kordaja poolest, siis teoreemi 5.3 põhjal leidub funktsioon f, nii et $|\langle f, \varphi\rangle| \leqslant 1$ iga $\varphi \in X$ jaoks ning $\langle f, \psi\rangle > 1$. Kuid esimesest võrratusest järeldub, et $\|f\| \leqslant 1$ ning seega teisest, et $\|\psi\|^* > 1$, vastuolu eeldusega $\psi \in B$. Seega $B = \mathrm{ch}(X)$. Kuna lemma 5.2 põhjal on Φ ja seega ka $X = \Phi \cup -\Phi$ kinnised funktsioonide korrutamise suhtes, siis ka $B = \mathrm{ch}(X)$ jaoks saame, et kui $x = a\varphi_1 + (1-a)\varphi_2, y = b\varphi_3 + (1-b)\varphi_4$, kus $a, b \in [0,1], \varphi_i \in X, i \in \{1,2,3,4\}$, siis

$$xy = ab\varphi_1\varphi_3 + (1-a)b\varphi_2\varphi_3 + a(1-b)\varphi_1\varphi_4 + (1-a)(1-b)\varphi_2\varphi_4.$$

Kuna $\varphi_i\varphi_j \in X \subseteq B$, kus $i, j \in \{1,2,3,4\}$, siis $\|\varphi_i\varphi_j\|^* \leqslant 1$, seega kolmnurga võrratuse põhjal $\|xy\|^* \leqslant ab + (1-a)b + a(1-b) + (1-a)(1-b) = 1$ ehk $xy \in B$. Seega iga $\varphi, \psi \colon \mathbb{Z}_N \to \mathbb{R}$ korral $\|(\varphi/\|\varphi\|^*)(\psi/\|\psi\|^*)\|^* \leqslant 1$ ehk

$$\|\varphi\psi\|^* \leqslant \|\varphi\|^* \|\psi\|^* . \tag{10}$$

Näitame, et $\|\cdot\| \leqslant \|\cdot\|_1$ ja $\|\cdot\|_\infty \leqslant \|\cdot\|^*$, kus $\|f\|_1 := \mathbb{E}|f(x)|$ ning $\|\psi\|_\infty := \max|\psi(x)|$. Neist võrratustest esimene kehtib, sest kasutades asjaolu, et $\varphi(x) \leqslant 1$

iga $\varphi \in \Phi, x \in \mathbb{Z}_N$ korral, saame

$$\|f\| = \sup_{\varphi\in\Phi} |\langle f, \varphi\rangle| = \sup_{\varphi\in\Phi} |\mathbb{E}_x f(x)\varphi(x)| \leqslant \mathbb{E}_x|f(x)| = \|f\|_1 .$$

Teise võrratuse jaoks olgu x' selline, et ψ saavutab seal oma maksimumi ning defineerime $f(x) := \begin{cases} N, & \text{kui } x = x', \\ 0, & \text{muidu.} \end{cases}$. Siis $|\langle f, \varphi\rangle| = |\mathbb{E}_x f(x)\varphi(x)| = \left|\frac{1}{N} f(x')\varphi(x')\right| = |\varphi(x')|$, mistõttu $\|f\| \leqslant 1$ ja

$$\|\psi\|^* \geq |\langle f, \psi\rangle| = \left|\psi(x')\right| = \|\psi\|_\infty .$$

Nüüd sõnastame ja tõestame *tiheda mudeli teoreemi.*

Teoreem 5.4 (Tiheda mudeli teoreem)**.** *Iga $\varepsilon > 0$ korral leidub $\varepsilon' > 0$, nii et alati kui $\nu\colon \mathbb{Z}_N \to [0,\infty)$ ja $f\colon \mathbb{Z}_N \to [0,\infty)$ rahuldavad tingimusi $\|\nu - 1\|_{\square,r} \leqslant \varepsilon'$ ja $f \leqslant \nu$, leidub funktsioon $\tilde{f}\colon \mathbb{Z}_N \to [0,1]$, nii et $\|f - \tilde{f}\|_{\square,r} \leqslant \varepsilon$.*

Tõestus. Olgu üldisust kitsendamata $\epsilon \leqslant \frac{1}{10}$. Piisab näidata, et mingi $\varepsilon' > 0$ ja kõigi teoreemi eeldusi rahuldavate ν, f korral leidub $\tilde{f}\colon \mathbb{Z}_N \to [0, 1 + \frac{\varepsilon}{2}]$, nii et $\|f - \tilde{f}\| \leqslant \frac{\varepsilon}{2}$, sest siis rahuldaks $\tilde{f}' : \mathbb{Z}_N \to [0,1], \tilde{f}'(x) = \tilde{f}(x)(1 + \frac{\varepsilon}{2})^{-1}$ tingimust $\|f - \tilde{f}'\| \leqslant \|f - \tilde{f}\| + \|\tilde{f} - \tilde{f}'\| \leqslant \frac{\varepsilon}{2} + \|\tilde{f} - \tilde{f}'\|_1 \leqslant \varepsilon$, nagu soovitud.

Oletame vastuväiteliselt, et sellist $\tilde{f}$ ei leidu. Siis olgu

$$K_1 := \{\tilde{f}\colon \mathbb{Z}_N \to [0, 1 + \frac{\varepsilon}{2}]\} \quad \text{ja} \quad K_2 := \{h\colon \mathbb{Z}_N \to \mathbb{R} \mid \|h\| \leqslant \frac{\varepsilon}{2}\}.$$

Siis K_1, K_2 on kumerad ja kinnised. Eelduse põhjal $f \notin K_1 + K_2 = \{\tilde{f} + h : \tilde{f} \in K_1, h \in K_2\}$. Kuid kuna ka $K_1 + K_2$ on kumer ja kinnine, leidub teoreemi 5.3 põhjal $\psi\colon \mathbb{Z}_N \to \mathbb{R}$, nii et

(a) $\langle f, \psi\rangle > 1$,

(b) $\langle g, \psi\rangle \leqslant 1$ iga $g \in K_1 + K_2$ korral.

Võttes $g = (1+\frac{\varepsilon}{2})1_{\psi>0} \in K_1 \subseteq K_1+K_2$, defineerime $x_+ := \max\{0, x\}$ ja $\psi_+(x) := \psi(x)_+$, seega saame $\langle 1, \psi_+\rangle = \left\langle \frac{g}{1+\frac{\varepsilon}{2}}, \psi_+ \right\rangle \leqslant (1+\frac{\varepsilon}{2})^{-1}$.

Võttes aga $g \in K_2 \subseteq K_1 + K_2$, saame, et iga $\|g\| \leqslant \frac{\varepsilon}{2}$ korral $\langle g, \psi\rangle \leqslant 1$, seega iga $\|g\| \leqslant 1$ korral $\langle g, \psi\rangle \leqslant \frac{2}{\varepsilon}$. Järelikult $\|\psi\|_\infty \leqslant \|\psi\|^* \leqslant \frac{2}{\varepsilon}$.

Järgmiseks kasutame tuntud Weierstrassi teoreemi funktsioonide lähendatavusest polünoomidega:

Teoreem 5.5 (Weierstrass)**.** *Olgu $f : [a, b] \to \mathbb{R}$ pidev. Siis iga $\varepsilon > 0$ korral leidub polünoom $P \in \mathbb{R}[X]$, nii et iga $x \in [a, b]$ korral $|f(x) - P(x)| < \varepsilon$.*

Teoreemi 5.5 põhjal leidub $P \in \mathbb{R}[X]$, nii et $|P(x) - x_+| \leqslant \frac{\varepsilon}{8}$ iga $x \in [-\frac{2}{\varepsilon}, \frac{2}{\varepsilon}]$ korral. Olgu $P(x) = p_d x^d + \cdots + p_1 x + p_0$ ja $R = |p_d| \left(\frac{2}{\varepsilon}\right)^d + \cdots + |p_1| \left(\frac{2}{\varepsilon}\right) + |p_0|$.

Olgu $P\psi(x) := P(\psi(x))$. Kolmnurga võrratuse, võrrandi (10) ja omaduse $\|\psi\|^* \leqslant \frac{2}{\varepsilon}$ põhjal saame

$$\|P\psi\|^* \leqslant \sum_{i=0}^{d} |p_i|\, \|\psi^i\|^* \leqslant \sum_{i=0}^{d} |p_i|\, (\|\psi\|^*)^i \leqslant \sum_{i=0}^{d} |p_i| \left(\frac{2}{\varepsilon}\right)^i = R.$$

Kuna $\|\nu - 1\| \leqslant \varepsilon'$, siis omaduse $|\langle f, \psi\rangle| \leqslant \|f\|\, \|\psi\|^*$ põhjal

$$|\langle \nu - 1, P\psi\rangle| \leqslant \|\nu - 1\|\, \|P\psi\|^* \leqslant \varepsilon' R.$$

Kuna $\|\psi\|_\infty \leqslant \frac{2}{\varepsilon}$, siis polünoomi P definitsiooni põhjal $\|P\psi - \psi_+\|_\infty \leqslant \frac{\varepsilon}{8}$.

Järelikult, kuna $\langle 1, f\rangle = \mathbb{E}[f(x)] = \|f\|_1 \leqslant \|f\|_\infty$, siis

$$\langle \nu, P\psi\rangle \leqslant \langle 1, P\psi\rangle + \varepsilon' R \leqslant \langle 1, \psi_+\rangle + \frac{\varepsilon}{8} + \varepsilon' R \leqslant \left(1 + \frac{\varepsilon}{2}\right)^{-1} + \frac{\varepsilon}{8} + \varepsilon' R.$$

Tõestame, et $\|1\|^* = 1$. Ühelt poolt $\|1\|^* \geqslant \|1\|_\infty = 1$, kuid teisalt, kuna $1 \in \Phi$, siis $\|f\| = \sup_{\varphi\in\Phi} |\langle f, \varphi\rangle| \geqslant \langle f, 1\rangle$, mistõttu $\|1\|^* = \sup_{\|f\|\leqslant 1} \langle f, 1\rangle \leqslant 1$.

Omadusest $|\langle f,\psi\rangle| \leqslant \|f\|\,\|\psi\|^*$ koos kolmnurga võrratusega ja omadusega $\|1\|^* = 1$ saame $\|\nu\|_1 = \langle \nu, 1\rangle \leqslant \|\nu - 1\| + 1 \leqslant 1 + \varepsilon'$. Järelikult

$$\langle f,\psi\rangle \leqslant \langle f,\psi_+\rangle \leqslant \langle \nu,\psi_+\rangle \leqslant \langle \nu, P\psi\rangle + \|\nu\|_1 \|P\psi - \psi_+\|_\infty$$

$$\leqslant \left(1 + \frac{\varepsilon}{2}\right)^{-1} + \frac{\varepsilon}{8} + \varepsilon' R + \left(1 + \varepsilon'\right)\frac{\varepsilon}{8}.$$

Kuna $\varepsilon \leqslant \frac{1}{10}$, siis piisavalt väikese ε' korral on viimase avaldise väärtus maksimaalselt 1, sest $\varepsilon' = 0$ korral on avaldise väärtus $\left(1 + \frac{\varepsilon}{2}\right)^{-1} + \frac{\varepsilon}{4}$, mis on iga $0 < \varepsilon \leqslant \frac{1}{10}$ korral väiksem kui 1 ning avaldis on ε' suhtes pidev. See annab vastuolu tingimusega (a), mis tõestab tiheda mudeli teoreemi.

□

6 Loenduslemma

Selles peatükis sõnastame ja tõestame *loenduslemma*. Alustame loenduslemma sõnastamisest ja tõestamisest juhul $k = 3$, sest see käsitleb hüpergraafide asemel tavalisi graafe, mille korral on lihtsam näha argumentide taga olevat intuitsiooni. Peatüki lõpus näeme, et tõestus laieneb üldjuhule ilma uusi ideid lisamata.

Loenduslemma väidab, et kui kaalutud (hüper)graafid on lähedased lõikenormi poolest, siis neis on sarnane kogus aritmeetilisi jadasid. Esmalt tõestame loenduslemma lihtsama juhu, kus mõlema graafi kaalud on tõkestatud.

Lause 6.1 (Loenduslemma tõkestatud funktsioonide jaoks)**.** *Olgu g ja $\tilde{g}$ kaalutud kolmealuselised graafid alustega X, Y, Z ja kaaludega hulgast $[0,1]$. Kui $\|g - \tilde{g}\|_\square \leqslant \varepsilon$, siis*

$$|\mathbb{E}_{x\in X, y\in Y, z\in Z}[g(x,y)g(x,z)g(y,z) - \tilde{g}(x,y)\tilde{g}(x,z)\tilde{g}(y,z)]| \leqslant 3\varepsilon.$$

Tõestus. Tõestame, et iga $a\colon X \to [0,1]$ ja $b\colon Y \to [0,1]$ korral

$$|\mathbb{E}_{x\in X, y\in Y}[(g(x,y) - \tilde{g}(x,y))a(x)b(y)]| \leqslant \varepsilon \tag{11}$$

Kuna keskväärtus avaldises (11) on iga $a(x)$ ja $b(y)$ väärtuse suhtes lineaarne, omandab ta ekstreemumid siis, kui funktsioonide a, b kõik väärtused kuuluvad hulka $\{0,1\}$ ehk a, b on mingite hulkade karakteristlikud funktsioonid. Kuid karakteristlike funktsioonide jaoks saame lõikenormi definitsioonist, et (11) kehtib, seega kehtib (11) ka kõigi a, b korral. Fikseerides ajutiselt z, kasutades (11) ning leides seejärel keskväärtuse üle z, saame näha, et

$$|\mathbb{E}_{x\in X, y\in Y, z\in Z}[g(x,y)g(x,z)g(y,z) - \tilde{g}(x,y)g(x,z)g(y,z)]| =$$

$$|\mathbb{E}_{z\in Z}[\mathbb{E}_{x\in X, y\in Y}[g(x,y)g(x,z)g(y,z) - \tilde{g}(x,y)g(x,z)g(y,z)]]| \leqslant \varepsilon.$$

Analoogselt ka

$$|\mathbb{E}[\tilde{g}(x,y)g(x,z)g(y,z) - \tilde{g}(x,y)\tilde{g}(x,z)g(y,z)]| \leqslant \varepsilon$$

ja

$$|\mathbb{E}[\tilde{g}(x,y)\tilde{g}(x,z)g(y,z) - \tilde{g}(x,y)\tilde{g}(x,z)\tilde{g}(y,z)]| \leqslant \varepsilon.$$

Soovitud väite saame liites võrrandid kokku ja rakendades kolmnurga võrratust. $\square$

Edasises kasutame korduvalt tuntud Cauchy-Schwarzi võrratust keskväärtuste jaoks:

Teoreem 6.2 (Cauchy-Schwarzi võrratus)**.** *Olgu $X, Y \subseteq \mathbb{R}$ lõplikud hulgad ja $f : X \to \mathbb{R}, g : Y \to \mathbb{R}$. Siis*

$$\mathbb{E}_{x\in X}[f(x)^2]\mathbb{E}_{y\in Y}[g(y)^2] \geqslant [\mathbb{E}_{x\in X, y\in Y} f(x)g(y)]^2.$$

Erijuhul $X = Y, f = g$ võtab Cauchy-Schwarzi võrratus kuju $\mathbb{E}_{x\in X}[f(x)^2] \geqslant [\mathbb{E}_{x\in X} f(x)]^2 = \mathbb{E}_{x,x'\in X} f(x)f(x')$, mida kasutame mitmes järgnevas lemmas.

Lemma 6.3. *Olgu $\nu, g, \tilde{g}$ kaalutud kolmealuselised graafid alustega X, Y, Z. Rahuldagu ν 3-lineaarvormide tingimust ning kehtigu $0 \leqslant g \leqslant \nu$ ja $0 \leqslant \tilde{g} \leqslant 1$. Siis*

$$\mathbb{E}_{x\in X, y\in Y, z,z'\in Z}[(\nu(x,y)-1)g(x,z)g(x,z')g(y,z)g(y,z')] = o(1)$$

ning tingimus jääb kehtima, kui ühes või rohkemas teguritest funktsioon g asendada funktsiooniga $\tilde{g}$.

Tõestus. Esmalt tõestame juhu, kus igas liikmes on funktsioon g. Cauchy-Schwarzi võrratusest saame

$$
\begin{aligned}
&\left|\mathbb{E}_{x,y,z,z'}[(\nu(x,y)-1)g(x,z)g(x,z')g(y,z)g(y,z')]\right|^2\\
&= (\mathbb{E}_{y,z,z'}[\mathbb{E}_x[(\nu(x,y)-1)g(x,z)g(x,z')]g(y,z)g(y,z')])^2\\
&\leqslant \mathbb{E}_{y,z,z'}[(\mathbb{E}_x[(\nu(x,y)-1)g(x,z)g(x,z')])^2 g(y,z)g(y,z')]\ \mathbb{E}_{y,z,z'}[g(y,z)g(y,z')]\\
&\leqslant \mathbb{E}_{y,z,z'}[(\mathbb{E}_x[(\nu(x,y)-1)g(x,z)g(x,z')])^2 \nu(y,z)\nu(y,z')]\ \mathbb{E}_{y,z,z'}[\nu(y,z)\nu(y,z')].
\end{aligned}
$$

3-lineaarvormide tingimuse põhjal on teine liige maksimaalselt $1+o(1)$, seega piisab uurida vaid esimest tegurit. Rakendades taas Cauchy-Schwarzi võrratust, saame

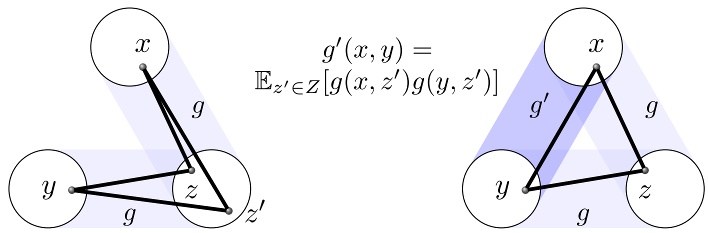
$$
\begin{aligned}
&\left|\mathbb{E}_{y,z,z'}[(\mathbb{E}_x[(\nu(x,y)-1)g(x,z)g(x,z')])^2\nu(y,z)\nu(y,z')]\right|^2\\
&= \left|\mathbb{E}_{x,x',y,z,z'}[(\nu(x,y)-1)(\nu(x',y)-1)g(x,z)g(x,z')g(x',z)g(x',z')\nu(y,z)\nu(y,z')]\right|^2\\
&= \left|\mathbb{E}_{x,x',z,z'}[\mathbb{E}_y[(\nu(x,y)-1)(\nu(x',y)-1)\nu(y,z)\nu(y,z')]g(x,z)g(x,z')g(x',z)g(x',z')]\right|^2\\
&\leqslant \mathbb{E}_{x,x',z,z'}[(\mathbb{E}_y[(\nu(x,y)-1)(\nu(x',y)-1)\nu(y,z)\nu(y,z')])^2 g(x,z)g(x,z')g(x',z)g(x',z')]\\
&\qquad \cdot \mathbb{E}_{x,x',z,z'}[g(x,z)g(x,z')g(x',z)g(x',z')]\\
&\leqslant \mathbb{E}_{x,x',z,z'}[(\mathbb{E}_y[(\nu(x,y)-1)(\nu(x',y)-1)\nu(y,z)\nu(y,z')])^2 \nu(x,z)\nu(x,z')\nu(x',z)\nu(x',z')]\\
&\qquad \cdot \mathbb{E}_{x,x',z,z'}[\nu(x,z)\nu(x,z')\nu(x',z)\nu(x',z')].
\end{aligned}
$$

3-lineaarvormide tingimuse põhjal on teine liige $1+o(1)$. Esimeses liikmes saame

$$
\begin{aligned}
&\mathbb{E}_{x,x',z,z'}[(\mathbb{E}_y[(\nu(x,y)-1)(\nu(x',y)-1)\nu(y,z)\nu(y,z')])^2 \nu(x,z)\nu(x,z')\nu(x',z)\nu(x',z')]\\
&= \mathbb{E}_{x,x',y,y',z,z'}[(\nu(x,y)-1)(\nu(x',y)-1)(\nu(x,y')-1)(\nu(x',y')-1)\nu(y,z)\nu(y,z')\\
&\quad \cdot \nu(y',z)\nu(y',z')\nu(x,z)\nu(x,z')\nu(x',z)\nu(x',z')],
\end{aligned}
$$

kus pärast sulgude avamist on 3-lineaarvormide tingimuse põhjal kõik 16 liiget $1+o(1)$. Neist miinusmärgiga on need, kus täpselt 1 või 3 sulust on valitud -1, seega on neid kokku 8. Järelikult taanduvad ühed välja ja uuritav liige on $o(1)$.

Kui mõnes liikmes on g asemel $\tilde{g}$, siis tõestus kulgeb identselt, kuid hindame funkt-



$$\mathbb{E}[g(x,z)g(x,z')g(y,z)g(y,z')] = \mathbb{E}[g'(x,y)g(x,z)g(y,z)]$$

Joonis 3: Loenduslemma juhu $k = 3$ tõestuses kasutatud graafikonstruktsioon.

siooni $\tilde{g}$ ülalt mitte funktsiooniga ν, vaid funktsiooniga 1. $\square$

Lemmas 6.3 uurisime neljast servast koosnevate tsüklite (edaspidi 4-*tsüklite*) hulka graafis: $\mathbb{E}_{x,y,z,z'}[g(x,z)g(x,z')g(y,z)g(y,z')]$. Defineerime abigraafi $g'\colon X \times Y \to [0,\infty)$ seosega $g'(x,y) := \mathbb{E}_{z'}[g(x,z')g(y,z')]$. Joonisel 3 on kujutatud graafi g'.

Osutub, et funktsioon g' käitub väga sarnaselt tõkestatud funktsiooniga, mille kohta oleme loenduslemma juba tõestanud (lause 6.1). Teisisõnu oleme asendanud kolmealuselises graafis g ühe komplekti servi g_{XY} sobivama komplekti servadega g'_{XY}. Defineerime analoogselt ka servad g'_{YZ} ja g'_{XZ}.

Sõnastame ja tõestame nüüd loenduslemma juhul $k = 3$:

Teoreem 6.4 (Loenduslemma juhul $k = 3$)**.** *Olgu $\nu, g, \tilde{g}$ kaalutud kolmealuselised graafid alustega X, Y, Z. Rahuldagu ν 3-lineaarvormide tingimust ning kehtigu $0 \leqslant g \leqslant \nu$ ja $0 \leqslant \tilde{g} \leqslant 1$. Kui $\|g - \tilde{g}\|_{\square} = o(1)$, siis*

$$|\mathbb{E}_{x\in X, y\in Y, z\in Z}[g(x,y)g(x,z)g(y,z) - \tilde{g}(x,y)\tilde{g}(x,z)\tilde{g}(y,z)]| = o(1). \qquad (12)$$

Tõestus. Tõestame väite matemaatilise induktsiooniga graafi ν selliste komponentide $\nu_{XY}, \nu_{XZ}, \nu_{YZ}$ arvu l suhtes, mis ei ole võrdsed funktsiooniga 1. Kui $l = 0$ ehk $\nu = 1$, järeldub lausest 6.1 soovitud väide. Kui $l > 0$, eeldame väite kehtimist $l - 1$ jaoks, samuti olgu (vajadusel tähiseid permuteerides) $\nu_{XY} \neq 1$. Defineerime abigraafid $\nu', g', \tilde{g}' \colon X \times Y \to [0, \infty)$ seostega

$$\begin{aligned}\nu'(x,y) &:= \mathbb{E}_z[\nu(x,z)\nu(y,z)],\\ g'(x,y) &:= \mathbb{E}_z[g(x,z)g(y,z)],\\ \tilde{g}'(x,y) &:= \mathbb{E}_z[\tilde{g}(x,z)\tilde{g}(y,z)].\end{aligned}$$

Olgugi et ν' ja g' väärtused ei ole ülalt tõkestatud ühega, tõkestame neid „jõuga“ ning näitame, et selle mõju on piisavalt väike. Defineerime $g'_{\wedge 1} := \min\{g', 1\}$ ja $\nu'_{\wedge 1} := \min\{\nu', 1\}$. Esmalt näeme, et

$$\mathbb{E}[g(x,y)g(x,z)g(y,z) - \tilde{g}(x,y)\tilde{g}(x,z)\tilde{g}(y,z)] = \mathbb{E}[gg' - \tilde{g}\tilde{g}']$$

$$= \mathbb{E}[g(g' - \tilde{g}')] + \mathbb{E}[(g - \tilde{g})\tilde{g}'], \tag{13}$$

kus esimene keskväärtus on leitud üle $x \in X, y \in Y, z \in Z$ ning ülejäänud üle $x \in X, y \in Y$ (ka ülejäänud peatükis võtame keskväärtusi üle $x \in X, y \in Y$, kui ei ole märgitud teisiti). Võrrandi (13) parema poole teine liige on $\mathbb{E}[(g(x,y) - \tilde{g}(x,y))\tilde{g}(x,z)\tilde{g}(y,z)]$, mille absoluutväärtus on analoogselt võrrandi (11) tõestamiseks kasutatud arutluskäiguga maksimaalne, kui $\tilde{g}$ rolli viimases kahes teguris võtta mingid karakteristlikud funktsioonid. Seega on see liige maksimaalselt $\|g - \tilde{g}\|_{\square} = o(1)$.

Jääb hinnata võrrandi (13) parema poole esimest liiget. Cauchy-Schwarzi võrratusest saame

$$(\mathbb{E}[g(g' - \tilde{g}')])^2 \leqslant \mathbb{E}[g(g' - \tilde{g}')^2]\ \mathbb{E}[g] \leqslant \mathbb{E}[\nu(g' - \tilde{g}')^2]\ \mathbb{E}[\nu]$$

$$= \mathbb{E}_{x,y}[\nu(x,y)(\mathbb{E}_z[g(x,z)g(y,z) - \tilde{g}(x,z)\tilde{g}(y,z)])^2]\ \mathbb{E}_{x,y}[\nu(x,y)],$$

kus 3-lineaarvormide tingimusest $\mathbb{E}_{x,y}[\nu(x,y)] = 1 + o(1)$. Esimeses teguris saame

$$\mathbb{E}_{x,y}[\nu(x,y)(\mathbb{E}_z[g(x,z)g(y,z) - \tilde{g}(x,z)\tilde{g}(y,z)])^2]$$
$$= \mathbb{E}_{x,y,z,z'}[\nu(x,y)(g(x,z)g(y,z) - \tilde{g}(x,z)\tilde{g}(y,z))(g(x,z')g(y,z') - \tilde{g}(x,z')\tilde{g}(y,z'))],$$

kus avades sulud ja rakendades korduvalt lemmat 6.3 näeme, et see erineb vaid $o(1)$ võrra avaldisest

$$\begin{aligned}\mathbb{E}_{x,y,z,z'}&[(g(x,z)g(y,z) - \tilde{g}(x,z)\tilde{g}(y,z))(g(x,z')g(y,z') - \tilde{g}(x,z')\tilde{g}(y,z'))] \\ &= \mathbb{E}_{x,y}[(\mathbb{E}_z[g(x,z)g(y,z) - \tilde{g}(x,z)\tilde{g}(y,z)])^2] = \mathbb{E}[(g' - \tilde{g}')^2]. \quad (14)\end{aligned}$$

3-lineaarvormide tingimusest saame, et $\mathbb{E}[\nu'] = 1 + o(1)$ ja $\mathbb{E}[\nu'^2] = 1 + o(1)$, seega Cauchy-Schwarzi võrratusest saame

$$(\mathbb{E}[|\nu' - 1|])^2 \leqslant \mathbb{E}[(\nu' - 1)^2] = o(1). \quad (15)$$

Tahame näidata, et võrrandi (14) parem pool on $o(1)$. Selleks paneme tähele, et

$$\mathbb{E}[(g' - \tilde{g}')^2] = \mathbb{E}[(g' - \tilde{g}')(g' - g'_{\wedge 1})] + \mathbb{E}[(g' - \tilde{g}')(g'_{\wedge 1} - \tilde{g}')]. \quad (16)$$

Kuna $0 \leqslant g' \leqslant \nu'$, siis

$$0 \leqslant g' - g'_{\wedge 1} = \max\{g' - 1, 0\} \leq \max\{\nu' - 1, 0\} \leqslant |\nu' - 1|. \quad (17)$$

Võrranditest (15) ja (17) ning tingimustest $0 \leqslant g \leqslant \nu$ ja $0 \leqslant \tilde{g} \leqslant 1$ saame, et võrrandi (16) parema poole esimese liikme absoluutväärtus on maksimaalselt

$$\mathbb{E}[(\nu' + 1)|\nu' - 1|] = \mathbb{E}[(\nu' - 1)|\nu' - 1|] + 2\mathbb{E}[|\nu' - 1|] = o(1).$$

Järgmiseks tõestame, et

$$\|g'_{\wedge 1} - \tilde{g}'\|_{\square} = o(1). \quad (18)$$

Selleks paneme tähele, et iga $A \subseteq X, B \subseteq Y$ korral kehtib

$$\mathbb{E}_{x,y}[(g'_{\wedge 1} - \tilde{g}')(x,y)1_A(x)1_B(y)] = \mathbb{E}[(g'_{\wedge 1} - \tilde{g}')1_{A\times B}]$$

$$= \mathbb{E}[(g'_{\wedge 1} - g')1_{A\times B}] + \mathbb{E}[(g' - \tilde{g}')1_{A\times B}].$$

Võrrandite (17) ja (15) põhjal on esimese liikme absoluutväärtus maksimaalselt $\mathbb{E}[|\nu' - 1|] = o(1)$. Teise liikme saame ümber kirjutada kujul

$$\mathbb{E}_{x,y,z}[1_{A\times B}(x,y)g(x,z)g(y,z) - 1_{A\times B}(x,y)\tilde{g}(x,z)\tilde{g}(y,z)].$$

See on aga (12), kus oleme asendanud $\nu_{XY}, g_{XY}, \tilde{g}_{XY}$ vastavalt funktsioonidega 1, $1_{A\times B}$, $1_{A\times B}$. Seega oleme vähendanud arvu l ühe võrra ehk uuritav liige on $o(1)$ induktsiooni eelduse põhjal. See tõestab võrduse (18).

Võrrandi (16) parema poole teise liikme saame lahti kirjutada kujul

$$\mathbb{E}[(g' - \tilde{g}')(g'_{\wedge 1} - \tilde{g}')] = \mathbb{E}[g'g'_{\wedge 1}] - \mathbb{E}[g'\tilde{g}'] - \mathbb{E}[\tilde{g}'g'_{\wedge 1}] + \mathbb{E}[\tilde{g}'^2]. \tag{19}$$

Tõestame, et kõik keskväärtused (19) paremal poolel on $\mathbb{E}[(\tilde{g}')^2] + o(1)$. Tõepoolest,

$$\mathbb{E}[g'g'_{\wedge 1}] - \mathbb{E}[(\tilde{g}')^2] = \mathbb{E}_{x,y,z}[g'_{\wedge 1}(x,y)g(x,z)g(y,z) - \tilde{g}'(x,y)\tilde{g}(x,z)\tilde{g}(y,z)].$$

See on aga (12), kus oleme asendanud $\nu_{XY}, g_{XY}, \tilde{g}_{XY}$ funktsioonidega $1, g'_{\wedge 1}, \tilde{g}'$ ning mis jällegi kehtib induktsiooni eelduse põhjal. Analoogselt on ka teised keskväärtused võrrandi (19) paremal pool $\mathbb{E}[(\tilde{g}')^2] + o(1)$. Seega võrrandi (19) parem pool on $o(1)$, mis tõestab soovitud väite. □

Nüüd sõnastame ja tõestame loenduslemma üldjuhu.

Teoreem 6.5 (Loenduslemma)**.** *Olgu $\nu, g, \tilde{g}$ kaalutud k-aluselised hüpergraafid (kus iga serv ühendab $k-1$ tippu) alustega $X_1, \ldots, X_k$. Rahuldagu ν k-lineaarvormide* *tingimust ning kehtigu* $0 \leqslant g \leqslant \nu$ *ja* $0 \leqslant \tilde{g} \leqslant 1$. *Kui* $\|g - \tilde{g}\|_{\square} = o(1)$, *siis*

$$|\mathbb{E}[g(x_{-1})g(x_{-2}) \cdots g(x_{-k}) - \tilde{g}(x_{-1})\tilde{g}(x_{-2}) \cdots \tilde{g}(x_{-k})]| = o(1). \tag{20}$$

Tõestus. Tõestus on pea täielikult analoogne juhu $k = 3$ tõestusega. Tõestame väite matemaatilise induktsiooniga graafi ν selliste komponentide $\nu_{-1}, \ldots, \nu_{-k}$ arvu l suhtes, mis ei ole võrdsed funktsiooniga 1. Kui $l = 0$ ehk $\nu = 1$, järeldub lause 6.1 otsesest üldistusest soovitud väide. Kui $l > 0$, eeldame väite kehtimist $l - 1$ jaoks, samuti olgu (vajadusel tähiseid permuteerides) $\nu_{-1} \neq 1$.

Ka lemmal 6.3 leidub otsene üldistus, mille tõestamiseks peame lihtsalt rohkem kordi analoogselt rakendama Cauchy-Schwarzi võrratust. Järgmiseks defineerime abigraafid $\nu', g', \tilde{g}' \colon X_{-1} \to [0, \infty)$ seostega

$$\begin{aligned}
\nu'(x_{-1}) &= \mathbb{E}_{x_1 \in X_1}[\nu(x_{-2}) \cdots \nu(x_{-k})], \\
g'(x_{-1}) &= \mathbb{E}_{x_1 \in X_1}[g(x_{-2}) \cdots g(x_{-k})], \\
\tilde{g}'(x_{-1}) &= \mathbb{E}_{x_1 \in X_1}[\tilde{g}(x_{-2}) \cdots \tilde{g}(x_{-k})].
\end{aligned}$$

Siis taandame analoogselt juhu $k = 3$ tõestusega võrduse (20) tõestuse võrduse (20) kujule, kus oleme asendanud funktsiooni ν_{-1} funktsiooniga 1, mis järeldub seega induktsiooni eeldusest. Seega oleme loenduslemma tõestanud. □

7 Suhtelise Szemerédi teoreemi tõestus

Selles peatükis tõestame tiheda mudeli teoreemi ja loenduslemma abil suhtelise Szemerédi teoreemi.

Lemma 7.1. *Rahuldagu ν k-lineaarvormide tingimust. Siis $\|\nu - 1\|_{\square,k-1} = o(1)$.*

Tõestus. Olgu üldisust kitsendamata $\|\nu - 1\|_{\square,k-1} = \|\nu_{-k} - 1\|_{\square,k-1}$. Siis lõikenormi definitsiooni (8) põhjal piisab näidata, et iga $A_1, \ldots, A_{k-1} \subseteq \mathbb{Z}_N^{k-2}$ korral

$$\mathbb{E}_{x_1,\ldots,x_{k-1}\in\mathbb{Z}_N}[(\nu(x_1,\ldots,x_{k-1}) - 1)1_{A_1}(x_{-1})\cdots 1_{A_{k-1}}(x_{-(k-1)})] = o(1).$$

Rakendades $k-1$ korda Cauchy-Schwarzi võrratust ning kasutades võrratust $1_A \leqslant 1$ saame, et

$$\begin{aligned}
&\left|\mathbb{E}_{x_1,\ldots,x_{k-1}}[(\nu(x_1,\ldots,x_{k-1}) - 1)1_{A_1}(x_{-1})\cdots 1_{A_{k-1}}(x_{-(k-1)})]\right|^{2^{k-1}} \\
&= \Big|\mathbb{E}_{x_1,\ldots,x_{k-2}}[1_{A_{k-1}}(x_{-(k-1)}) \\
&\quad \cdot \mathbb{E}_{x_{k-1}}[(\nu(x_1,\ldots,x_{k-1}) - 1)1_{A_1}(x_{-1})\cdots 1_{A_{k-2}}(x_{-(k-2)})]]\Big|^{2^{k-1}} \\
&\leqslant \Big|\mathbb{E}_{x_1,\ldots,x_{k-2}}[1_{A_{k-1}}(x_{-(k-1)}) \\
&\quad \cdot \mathbb{E}_{x_{k-1}^{(0)},x_{k-1}^{(1)}}[(\nu(x_1,\ldots,x_{k-1}^{(0)}) - 1)1_{A_1}(x_{-1})\cdots 1_{A_{k-2}}(x_{-(k-2)})] \\
&\quad \cdot (\nu(x_1,\ldots,x_{k-1}^{(1)}) - 1)1_{A_1}(x_{-1})\cdots 1_{A_{k-2}}(x_{-(k-2)})]\Big|^{2^{k-2}} \\
&\leqslant \ldots \\
&\leqslant \mathbb{E}_{x_1^{(0)},x_1^{(1)},\ldots,x_{k-1}^{(0)},x_{k-1}^{(1)}}\Big[\prod_{j=1}^{k}\prod_{\omega\in\{0,1\}^{[k-1]}}\left(\nu\left(x_1^{(\omega_1)},\ldots,x_{k-1}^{(\omega_{k-1})}\right) - 1\right)\Big]
\end{aligned}$$

Avades viimases avaldises avaldises sulud näeme, et iga liige on k-lineaarvormide tingimuse tõttu $1 + o(1)$, kusjuures ühed koonduvad, sest täpselt pooled liikmetest (ehk $k \cdot 2^{k-2}$ liiget) on miinusmärgiga. Seega on uuritav avaldis $o(1)$, nagu soovitud. $\square$

Suhtelise Szemerédi teoreemi tõestus. Lemma 7.1 põhjal $\|\nu - 1\|_{\square,k-1} = o(1)$. Seega tiheda mudeli teoreemi põhjal leidub (funktsioonide jada) $\tilde{f}\colon \mathbb{Z}_N \to [0,1]$, nii et $\left\|f - \tilde{f}\right\|_{\square,k-1} = o(1)$.

Defineerime $j = 1, \ldots, k$ korral lineaarvormid $\psi_j\colon X_{-j} \to \mathbb{Z}_N$ seosega

$$\psi_j(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_k) := \sum_{i \in [k] \setminus \{j\}} (j - i)x_i.$$

Konstrueerime kaalutud k-aluselised hüpergraafid $\nu, g, \tilde{g}$ seostega

$$\nu_{-j}(x_{-j}) := \nu(\psi_j(x_{-j})), \quad g_{-j}(x_{-j}) := f(\psi_j(x_{-j})), \quad \tilde{g}_{-j}(x_{-j}) := \tilde{f}(\psi_j(x_{-j})).$$

Soovime tõestada, et

$$\|\nu_{-j} - 1\|_{\square} = \|\nu - 1\|_{\square,k-1} \tag{21}$$

ja

$$\|g_{-j} - \tilde{g}_{-j}\|_{\square} = \|f - \tilde{f}\|_{\square,k-1}, \tag{22}$$

kus mõlemas võrrandis iga j korral vasak pool viitab hüpergraafi lõikenormile (8) ning parem pool hulgal $\mathbb{Z}_N$ määratud funktsioonide lõikenormile (9). Teeme argumendi näitlikustamiseks esmalt võrrandi (22) tõestuse läbi juhul $k = j = 4$. Võrrandi vasak pool on

$$\sup_{A_1, A_2, A_3 \subseteq \mathbb{Z}_N^2} \left|\mathbb{E}[(f - \tilde{f})(3x_1 + 2x_2 + x_3)1_{A_1}(x_2, x_3)1_{A_2}(x_1, x_3)1_{A_3}(x_1, x_2)]\right|, \tag{23}$$

ja parem pool on

$$\sup_{B_1, B_2, B_3 \subseteq \mathbb{Z}_N^2} \left|\mathbb{E}[(f - \tilde{f})(x_1 + x_2 + x_3)1_{B_1}(x_2, x_3)1_{B_2}(x_1, x_3)1_{B_3}(x_1, x_2)]\right|. \tag{24}$$

Eeldades, et arvud N ja $(k-1)!$ on ühistegurita (selliseid arve on lõpmata palju, millest meile piisab), saame muuta avaldised võrdseks muutujavahetustega $3x_1 \leftrightarrow x_1$ ja $2x_2 \leftrightarrow x_2$, mis ei mõjuta supreemumit, sest teineteisele vastavad ka hulgad

B_1 ja $A_1 := \{(2x_2, x_3) \mid (x_2, x_3) \in B_1\}$.

Valides arvude k ja j kohale midagi muud, muutuvad ainult r väärtus lõikenormis ning lineaarvorm, mis on võrrandis (23) funktsiooni $f - \tilde{f}$ argumendiks, kuid avaldised erinevad endiselt vaid muutujavahetuste poolest, mida saame alati teha, sest kõigi lineaarvormide kõik kordajad on täisarvud vahemikus $[-(k-1),(k-1)]$. Tõestades võrrandit (21), asendub vaid $(f - \tilde{f})$ funktsioniga $\nu - 1$, muu jääb identseks. Seega oleme tõestanud võrrandid (21) ja (22).

Nüüd järeldub võrrandist (22), et $\|g - \tilde{g}\|_{\square} = \|f - \tilde{f}\|_{\square,k-1} = o(1)$. Siis järeldub loenduslemmast (Teoreem 6.5), et

$$\mathbb{E}_{x_1,\ldots,x_k \in \mathbb{Z}_N^k}[g_{-1}(x_{-1}) \cdots g_{-k}(x_{-k})] = \mathbb{E}_{x_1,\ldots,x_k \in \mathbb{Z}_N^k}[\tilde{g}_{-1}(x_{-1}) \cdots \tilde{g}_{-k}(x_{-k})] + o(1). \tag{25}$$

Kasutades g definitsiooni ja asendades $x = \psi_1(x_{-1})$ ja $d = x_1 + \cdots + x_k$, näeme, et eelmise võrrandi vasak pool võtab kuju

$$\mathbb{E}_{x_1,\ldots,x_k \in \mathbb{Z}_N^k}[f(\psi_1(x_{-1})) \cdots f(\psi_k(x_{-k}))] = \mathbb{E}_{x,d \in \mathbb{Z}_N}[f(x)f(x+d) \cdots f(x+(k-1)d)].$$

Teisendades analoogselt ka võrrandi (25) paremat poolt, võtab ta kuju

$$\mathbb{E}_{x,d \in \mathbb{Z}_N}[f(x)f(x+d) \cdots f(x+(k-1)d)] = \mathbb{E}_{x,d \in \mathbb{Z}_N}[\tilde{f}(x)\tilde{f}(x+d) \cdots \tilde{f}(x+(k-1)d)] + o(1),$$

mis on kaalutud Szemerédi teoreemi (Teoreem 4.4) põhjal $c(k, \delta) - o(1)$, mida tahtsimegi tõestada. $\square$

8 Algarvude hulga tihe mudel

Selles peatükis kasutame suhtelist Szemerédi teoreemi, et tõestada Greeni-Tao teoreem.

Selleks kasutame von Mangoldti funktsiooni:

Definitsioon 8.1 (von Mangoldti funktsioon)**.** Von Mangoldti funktsioon $\Lambda : \mathbb{N} \to \mathbb{R}$ on defineeritud seosega

$$\Lambda(n) = \begin{cases} \log p, & \text{kui } n = p^k, \text{ kus } p \in \mathbb{P}, k \in \mathbb{N}, \\ 0, & \text{muidu.} \end{cases}$$

Siin ja edaspidi mõtleme $\log(n)$ all alati naturaallogaritmi.

Meil on vaja funktsioon ν defineerida nii, et ta rahuldaks k-lineaarvormide tingimust ehk oleks piisavalt „juhuslik“. Funktsiooni $\Lambda(n)$ väärtused jaotuvad aga ebaühtlaselt arvu n erinevate jäägiklasside vahel erinevate moodulite järgi. Näiteks mooduli 2 järgi näeme, et ainult üks algarv ($p = 2$) on paaris ning kõik ülejäänud algarvud on paaritud.

Selliste kalduvuste eemaldamiseks kasutame *W-meetodit*. Rahuldagu $w = w(N)$ tingimust $\lim_{N\to\infty} w = \infty$, kuid kasvagu w väga aeglaselt N suhtes. Siis saame vältida funktsiooni väärtuste ebaühtlast jaotust moodulite $p \leqslant w$ järgi, võttes $W := \prod_{p \leqslant w, p \in \mathbb{P}} p$ ning defineerides *modifitseeritud von Mangoldti funktsiooni* kujul

$$\widetilde{\Lambda}(n) := \begin{cases} \frac{\varphi(W)}{W} \log(Wn + 1) & \text{kui } Wn + 1 \in \mathbb{P}, \\ 0 & \text{muidu.} \end{cases}$$

Näitame, et kehtib tingimus

$$\sum_{n \leqslant N} \widetilde{\Lambda}(n) = (1 + o(1))N. \tag{26}$$

Selleks kasutame Dirichlet' teoreemi:

Teoreem 8.2 (Dirichlet' teoreem)**.** *Olgu $m \in \mathbb{N}$. m jäägiklassi hulgast modulo m leidub $\varphi(m)$ jäägiklassis lõpmata palju algarve. Algarvud on jaotatud nende vahel ühtlaselt: iga jäägiklassi $\overline{a}$ jaoks, kus leidub lõpmata palju algarve, kehtib*

$$\sum_{n \leqslant N, n \in \mathbb{P}, n \in \overline{a}} \Lambda(n) = (1 + o(1)) \frac{N}{\varphi(m)}.$$

Summeerides funktsiooni $\widetilde{\Lambda}$ üle arvude $1 \leqslant n \leqslant N$, „katame ära" kõik algarvud vahemikus $[1, WN + 1]$, mis kuuluvad modulo W jäägiklassi $\overline{1}$. Seega

$$\sum_{n \leqslant N} \widetilde{\Lambda}(n) = \frac{\varphi(W)}{W} \sum_{n \leqslant WN+1, n \in \mathbb{P}, n \in \overline{1}} \Lambda(n) = \frac{\varphi(W)}{W}(1+o(1)) \frac{WN+1}{\varphi(W)} = (1+o(1))N,$$

nagu soovitud.

Hiljem vajame veel järgmist lemmat funktsiooni $\widetilde{\Lambda}$ kohta:

Lemma 8.3. *Iga naturaalarvu $k \geqslant 3$ korral* $\sum\limits_{\frac{N}{2} \leqslant n < N} \widetilde{\Lambda}(n)^k = o(N^2)$.

Tõestus. Piisavalt suure N jaoks saame, et ühelt poolt

$$\frac{1}{2} \log(WN) = \log(\sqrt{WN}) < \log(\frac{WN}{2} + 1) \leqslant \log(Wn + 1)$$

ning teiselt poolt

$$\log(Wn + 1) \leqslant \log(W(N-1) + 1) < \log(WN),$$

seega tähistades $C = \frac{\varphi(W)}{W}$, kehtib $\frac{C}{2} \log(WN) < \widetilde{\Lambda}(n) < C \log(WN)$ iga $\frac{N}{2} \leqslant n < N$ jaoks, mille korral $\widetilde{\Lambda}(n) \neq 0$.

Võrrandi (26) põhjal piisavalt suure N korral $\sum\limits_{\frac{N}{2} \leqslant n < N} \widetilde{\Lambda}(n) \leqslant (1 + o(1))N < 2N.$, mis koos võrratusega $\widetilde{\Lambda}(n) > \frac{C}{2} \log(WN)$ tähendab, et funktsioonil $\widetilde{\Lambda}$ on maksimaalselt $\frac{4N}{C \log(WN)}$ nullist erinevat väärtust. Et igaüks neist on maksimaalselt

$C\log(WN)$, saame et

$$\sum_{\frac{N}{2}\leqslant n<N} \widetilde{\Lambda}(n)^k \leqslant \frac{4N}{C\log(WN)} \cdot (C\log(WN))^k = 4N\,(C\log(WN))^{k-1}.$$

Kui W kasvab piisavalt aeglaselt, siis $\log(WN)^{k-1} = o(N)$, sest kirjutades $N = e^m, W \leqslant e^m$ saame $\log(WN)^{k-1} = (2m)^{k-1}$, mis on m suhtes polünoom, seega ta on $o(e^m) = o(N)$.

Sel juhul aga

$$\sum_{\frac{N}{2}\leqslant n<N} \widetilde{\Lambda}(n)^k = (2C^{k-1})N\log(WN)^{k-1} = o(N^2),$$

nagu soovitud. $\square$

Samuti kasutame *Möbiuse funktsiooni* $\mu(n)$, mille väärtus on $(-1)^l$, kui n on $l \geqslant 0$ paarikaupa erineva algarvu korrutis (nimetame selliseid arve n *ruuduvabadeks*) ning 0 muidu. Funktsioon μ on seotud von Mangoldti funktsiooniga Λ seosega

$$\Lambda(n) = \sum_{d|n} \mu(d)\log(n/d).$$

Selle põhjal konstrueerime veel ühe alternatiivkuju von Mangoldti funktsiooni jaoks:

Definitsioon 8.4. Olgu $R > 0$ ning olgu $\chi\colon \mathbb{R} \to [0,1]$ pidevalt diferentseeruv, nii et $\chi(x) > 0$ parajasti lõigus $x \in [1,-1]$. Defineerime funktsiooni $\Lambda_{\chi,R}(n)$ seosega

$$\Lambda_{\chi,R}(n) := \log R \sum_{d|n} \mu(d)\chi\left(\frac{\log d}{\log R}\right).$$

Lause 8.5. *Olgu pidevalt diferentseeruva $\chi\colon \mathbb{R} \to [0,1]$ korral $\chi(x) > 0$ parajasti lõigus $x \in [1,-1]$. Olgu $m,t \in \mathbb{N}$. Olgu $\psi_1,\ldots,\psi_m\colon \mathbb{Z}^t \to \mathbb{Z}$ lineaarkujutused, millest ükski ei ole teise kordne. Olgu $R(N) = o(N^{1/(10m)})$ ning $w(N)$ piisavalt aeglaselt kasvavad funktsioonid. Olgu $W := \prod_{p\leqslant w} p$. Olgu $\theta_i, i = 1,\ldots,m$ sellised* *lineaarkujutused, et* $\theta_i(x) := W\psi_i(x) + 1$ *iga* $i = 1, \dots, m$ *ja* $x \in \mathbb{Z}^t$ *korral. Olgu* $B := \prod_{i=1}^t I_i$*, kus iga hulk* I_i *koosneb vähemalt* R^{10m} *järjestikusest täisarvust. Siis*

$$\mathbb{E}_{\mathbf{x}\in B}[\Lambda_{\chi,R}(\theta_1(\mathbf{x}))^2 \cdots \Lambda_{\chi,R}(\theta_m(\mathbf{x}))^2] = (1+o(1))\left(\frac{Wc_\chi \log R}{\varphi(W)}\right)^m, \tag{27}$$

kus

$$c_\chi := \int_0^\infty \left|\chi'(x)\right|^2 \, dx.$$

Tõestus. Tõestuse suure mahu ja keeruliste suurushinnangute tõttu tõestame lause eraldi järgmises peatükis. □

Lause 8.6. *Olgu pidevalt diferentseeruva* $\chi\colon \mathbb{R} \to [0,1]$ *jaoks* $\chi(0) = 1$ *ning* $\chi(x) > 0$ *parajasti siis, kui* $x \in [1, -1]$*. Olgu* $k \geq 3$ *ja* $R := N^{k^{-1}2^{-k-3}}$*. Kasvagu* $w(N)$ *piisavalt aeglaselt ning olgu* $W := \prod_{p \leqslant w} p$*. Siis defineerides* $\nu\colon \mathbb{Z}_N \to [0, \infty)$ *seosega*

$$\nu(n) := \begin{cases} \frac{\varphi(W)}{W}\frac{\Lambda_{\chi,R}(Wn+1)^2}{c_\chi \log R} & \text{kui } \frac{N}{2} \leqslant n < N, \\ 1 & \text{muidu,} \end{cases} \tag{28}$$

rahuldab ν k*-lineaarvormide tingimust.*

Tõestus. Meil on vaja näidata, et kehtib võrdus

$$\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_N^t}[\nu(\psi_1(\mathbf{x}))\cdots\nu(\psi_m(\mathbf{x}))] = 1 + o(1), \tag{29}$$

kui $\psi_1, \dots, \psi_m$, $m \leqslant k2^{k-1}$ on erinevad lineaarvormid k-lineaarvormide tingimuses.

Olgu $Q = Q(N)$ aeglaselt kasvav funktsioon. Jaotame hulga $\mathbb{Z}_N$ Q peaaegu võrdseks vahemikuks ning t-dimensionaalse kuubi $\mathbb{Z}_N^t$ (edaspidi nimetame t-dimensionaalseid risttahukaid *kastideks*) Q^t kastiks järgnevalt:

$$B_{u_1,\dots,u_t} = \prod_{j=1}^t \left(\left[u_j\frac{N}{Q}, (u_j+1)\frac{N}{Q}\right) \cap \mathbb{Z}_N\right) \subseteq \mathbb{Z}_N^t, \qquad u_1, \dots, u_t \in \mathbb{Z}_Q.$$

Kuna kastide suurused on peaaegu võrdsed, on võrrandi (29) vasak pool

$$\mathbb{E}_{u_1,\dots,u_t\in\mathbb{Z}_Q}[\mathbb{E}_{\mathbf{x}\in B_{u_1,\dots,u_t}}[\nu(\psi_1(\mathbf{x}))\cdots\nu(\psi_m(\mathbf{x}))]]+o(1).$$

Nimetame kasti $B_{u_1,\dots,u_t}$ *heaks*, kui iga $j=1,\dots,m$ korral $\{\psi_j(\mathbf{x}):\mathbf{x}\in B_{u_1,\dots,u_t}\}$ on kas hulga $[\frac{N}{2},N)\subseteq\mathbb{Z}_N$ alamhulk või ei oma sellega ühtegi ühist elementi. Vastasel juhul ütleme, et kast on *halb*. Eeldame, et Q kasvab N suhtes piisavalt aeglaselt, et $\frac{N}{Q}\geq R^{10m}=o(N)$. Siis näeme lause 8.5 ja funktsiooni ν definitsiooni põhjal, et heade kastide jaoks kehtib võrdus

$$\mathbb{E}_{\mathbf{x}\in B_{u_1,\dots,u_t}}[\nu(\psi_1(\mathbf{x}))\cdots\nu(\psi_m(\mathbf{x}))]=1+o(1).$$

Halbade kastide jaoks kasutame võrratust $\nu(n)\leqslant 1+\frac{\varphi(W)}{W}\frac{\Lambda_{\chi,R}(Wn+1)^2}{c_\chi\log R}$. Kasutades seda võrratust keskväärtuse igas liikmes ning rakendades igale liikmele ka lauset 8.5, saame

$$\mathbb{E}_{\mathbf{x}\in B_{u_1,\dots,u_t}}[\nu(\psi_1(\mathbf{x}))\cdots\nu(\psi_m(\mathbf{x}))]\leqslant 2^m+o(1)=O(1).$$

Et näidata, et heade ja halbade kastide kokkupanemisel on keskväärtus $1+o(1)$, on meil vaja näidata, et halbade kastide osakaal kõigist kastidest on $o(1)$.

Olgu $B_{u_1,\dots,u_t}$ halb kast. Siis mingi lineaarvormi ψ_i, $i\in\{1,\dots,t\}$ korral omab kasti kujutis elemente nii hulgas $[\frac{N}{2},N)$ kui ka väljaspool seda. Kuna ψ_i on pidev kõigi muutujate suhtes, peab seega leiduma $\mathbf{x}\in\prod_{j=1}^t[u_j\frac{N}{Q},(u_j+1)\frac{N}{Q})\subseteq(\mathbb{R}/N\mathbb{Z})^t$, nii et $\psi_i(\mathbf{x})=\overline{0}$ või $\overline{\frac{N}{2}}$.

Defineerides $\mathbf{y}:=\frac{Q}{N}\mathbf{x}$, näeme et $\mathbf{y}\in\prod_{j=1}^t[u_j,u_j+1)\subseteq(\mathbb{R}/Q\mathbb{Z})^t$ rahuldab kas võrdust $\psi_i(\mathbf{y})=\overline{0}$ või $\overline{\frac{Q}{2}}$. Siis, kuna $\mathbf{y}$ kõik koordinaadid erinevad vastavatest koordinaatidest $(u_1,\dots,u_t)$ maksimaalselt 1 võrra ning ψ_i kordajad ei sõltu arvust N, peab $\psi_i(u_1,\dots,u_t)$ olema kas $\overline{O(1)}$ või $\overline{\frac{Q}{2}+O(1)}$.

Kuna ψ_i on lineaarvorm, mis ei ole samaselt null, siis maksimaalne osakaal koordinaatide korteeže $(u_1, \dots, u_t) \in \mathbb{Z}_Q^t$, mis seda tingimust rahuldavad, on $O(\frac{1}{Q})$. Seda sellepärast, et fikseerides kõik koordinaadid peale ühe (näiteks sellise, mille kordaja lineaarvormis on ± 1), on viimase koordinaadi jaoks Q valikut, millest $O(1)$ annavad sobiva lineaarvormi väärtuse.

Võttes ühendi üle $i = 1, \dots, t$, näeme, et halbade kastide osakaal on maksimaalselt $t \cdot O(\frac{1}{Q}) = o(1)$, kuna $t \in \mathbb{N}$ on fikseeritud ning Q valiku tõttu $\lim_{N\to\infty} \frac{1}{Q} = 0$. $\square$

Järgmise lause abil näitame peagi, et lauses 8.6 konstrueeritud funktsioon ν tõesti rahuldab tingimust $\nu \geqslant f$.

Lause 8.7. *Iga $k \geq 3$ jaoks leidub $\delta_k > 0$, nii et piisavalt suure N korral leidub funktsioon $\nu\colon \mathbb{Z}_N \to [0, \infty)$, mis rahuldab k-lineaarvormide tingimust ning tingimust $\nu(n) \geq \delta_k \widetilde{\Lambda}(n)$ iga $\frac{N}{2} \leqslant n < N$ korral.*

Tõestus. Olgu $\delta_k = k^{-1} 2^{-k-4} c_\chi^{-1}$ ning olgu ν defineeritud lause 8.6 põhjal. Piisab näidata, et piisavalt suure N korral kehtib $\delta_k \widetilde{\Lambda}(n) \leqslant \nu(n)$ iga $\frac{N}{2} \leqslant n < N$ jaoks.

Funktsiooni $\widetilde{\Lambda}(n)$ definitsiooni põhjal piisab käsitleda juhtu, kus $Wn+1$ on algarv, sest mujal on $\widetilde{\Lambda}(n) = 0$. Kui w kasvab piisavalt aeglaselt, siis piisavalt suure N korral

$$\log R = k^{-1} 2^{-k-3} \log N \geq k^{-1} 2^{-k-4} \log(WN + 1) = c_\chi \delta_k \log(WN + 1).$$

Kui $Wn + 1 > R$ on algarv, kehtib $\Lambda_{\chi,R}(Wn + 1) = \log R$, seega

$$\delta_k \widetilde{\Lambda}(n) = \delta_k \frac{\varphi(W)}{W} \log(Wn + 1) \leqslant \delta_k \frac{\varphi(W)}{W} \log(WN + 1) \leqslant \frac{\varphi(W)}{W} \frac{\log R}{c_\chi} = \nu(n),$$

nagu soovitud. $\square$

Nüüd saame ära tõestada Greeni-Tao teoreemi:

Greeni-Tao teoreemi tõestus. Olgu $f\colon \mathbb{Z}_N \to [0,\infty)$ defineeritud seosega

$$f(n) = \begin{cases} \delta_k \widetilde{\Lambda}(n) & \text{kui } \frac{N}{2} \leqslant n < N, \\ 0 & \text{muidu,} \end{cases}$$

kus δ_k on lausest 8.7 pärinev konstant. Võrrandi (26) põhjal

$$\sum_{1\leqslant n<N} f(n) = \sum_{\frac{N}{2}\leqslant n<N} \delta_k \widetilde{\Lambda}(n) = \sum_{1\leqslant n<N} \delta_k \widetilde{\Lambda}(n) - \sum_{1\leqslant n<\frac{N}{2}} \delta_k \widetilde{\Lambda}(n) =$$

$$= (1+o(1))\delta_k N - (1+o(1))\delta_k \frac{N}{2} = (1/2 + o(1))\delta_k N.$$

See tähendab, et piisavalt suure N korral $\mathbb{E}f \geq \frac{\delta_k}{3}$. Kuna lause 8.7 põhjal $0 \leqslant f \leqslant \nu$ ning ν rahuldab k-lineaarvormide tingimust, järeldub suhtelisest Szemerédi teoreemist $\delta = \delta_k/3$ jaoks , et

$$\mathbb{E}[f(x)f(x+d)\cdots f(x+(k-1)d)] \geq c(k, \frac{\delta_k}{3}) - o(1). \tag{30}$$

Lemmast 8.3 näeme, et tingimust $d = 0$ rahuldavate liikmete mõju keskväärtusele on $\frac{1}{N^2}\delta_k o(N^2) = o(1)$. Seega piisavalt suure N korral leiduvad $d \neq 0$ ja $x \in [\frac{N}{2}, N)$, nii et $f(x)f(x+d)\cdots f(x+(k-1)d) > 0$. Kuna funktsioon f omandab positiivseid väärtusi ainult vahemikus $[\frac{N}{2}, N)$, järeldub lemmast 1.5, et arvud $x, x+d, \ldots, x+(k-1)d$ moodustavad k-jada mitte ainult rühmas $\mathbb{Z}_N$, vaid ka hulgas $\mathbb{Z}$. Seega arvud $(x+jd)W+1$, kus $j = 0, \ldots, k-1$, on algarvudest koosnev k-jada, mida tahtsimegi tõestada. $\square$

9 Suurushinnangud

Meil on jäänud tõestada lause 8.5. Selleks on vaja mitmeid suurushinnanguid. Et avaldised ei läheks liiga keeruliseks, teeme esmalt tõestuse põhiosa läbi vaid „pealiikme“ jaoks ning näitame peatüki teises pooles, et ülejäänud liikmed (nö „vealiikmed“) tulemust ei mõjuta.

Kirjutades võrrandi (27) vasakus pooles igal pool välja funktsiooni $\Lambda_{\chi,R}$ definitsiooni, saame avaldise

$$(\log R)^{2m} \sum_{d_1,d_1',\ldots,d_m,d_m' \in \mathbb{N}} \left(\prod_{j=1}^{m} \mu(d_j)\chi\left(\frac{\log d_j}{\log R}\right) \mu(d_j')\chi\left(\frac{\log d_j'}{\log R}\right) \right) \mathbb{E}[1_{d_j,d_j' | \theta_j(\mathbf{x})\ \forall j}]. \tag{31}$$

Kuna $\mu(d) = 0$, kui d jagub mingi algarvu ruuduga, käsitleme ainult liikmeid, kus arvud $d_1, d_1', \ldots, d_m, d_m'$ on ruuduvabad. Kuna funktsiooni χ positiivsed väärtused asuvad lõigus $[-1, 1]$, siis võime eeldada, et $d_1, d_1', \ldots, d_m, d_m' \leqslant R$. Olgu $D := \text{VÜK}(d_1, d_1', \ldots, d_m, d_m')$. Kuna kasti B laius igas mõõtmes on vähemalt R^{10m}, võime käsitleda tema suurimat võimalikku alamkasti $B' \subseteq B$, mille laius igas mõõtmes jagub arvuga $D \leqslant R^{2m}$. Siis saame B' jagada omakorda alamkastideks mõõtmetega $D \times D \times \ldots \times D$, millest igaühes on keskväärtus sama, mis keskväärtus üle $\mathbf{x} \in \mathbb{Z}_D^t$. Samuti vektorite $\mathbf{x} \in B \backslash B'$ osakaal kogu kastis B on väiksem kui $m \cdot \frac{R^{2m}}{R^{10m}} = O(R^{-8m})$, seega

$$\mathbb{E}_{\mathbf{x} \in B}[1_{d_j,d_j' | \theta_j(\mathbf{x})\, \forall j}] = \mathbb{E}_{\mathbf{x} \in \mathbb{Z}_D^t}[1_{d_j,d_j' | \theta_j(\mathbf{x})\, \forall j}] + O(R^{-8m}).$$

Seega, kuna arvude $d_1, d_1', \ldots, d_m, d_m'$ jaoks on maksimaalselt R^{2m} valikut, võime hinnata avaldist (31) kui

$$(\log R)^{2m} \sum_{d_1,d_1',\dots,d_m,d_m'\in\mathbb{N}} \left(\prod_{j=1}^{m} \mu(d_j)\chi\left(\frac{\log d_j}{\log R}\right) \mu(d_j')\chi\left(\frac{\log d_j'}{\log R}\right)\right)$$
$$\cdot\, \mathbb{E}_{\mathbf{x}\in\mathbb{Z}_D^t}[1_{d_j,d_j'|\theta_j(\mathbf{x})\ \forall j}], \quad (32)$$

kus viga on maksimaalselt $O(R^{-6m}(\log R)^{2m})$.

Soovime siin ümber kirjutada avaldisi $\chi\left(\frac{\log d}{\log R}\right)$. Selleks kasutame Fourier' pööret ning selle tuntud omadusi (Stein ja Shakarchi, 2011):

Definitsioon 9.1 (Fourier' pööre)**.** Funktsiooni $f : \mathbb{R} \to \mathbb{R}$ *Fourier' pöördeks* $\hat{f} : \mathbb{R} \to \mathbb{C}$ nimetame funktsiooni

$$\hat{f}(\xi) := \int_{\mathbb{R}} f(x)e^{-2\pi i\xi x}\,dx.$$

Lause 9.2 (Fourier' pöörde omadused)**.** *Olgu $f : \mathbb{R} \to \mathbb{R}$ ja $\hat{f}$ tema Fourier' pööre.*

1. *Funktsioon f avaldub tema Fourier' pöörde $\hat{f}$ kaudu võrrandiga*

$$f(x) = \int_{\mathbb{R}} \hat{f}(\xi)e^{-ix\xi}\,d\xi.$$

2. *Kui funktsioon f on pidevalt diferentseeruv ning omab nullist erinevaid väärtusi vaid lõplikus vahemikus, siis funktsiooni $e^x f(x)$ Fourier' pööre rahuldab tingimust $(\widehat{e^x f})(\xi) = O_\xi((1+|\xi|)^{-A})$ iga $A > 0$ korral.*

3. *Kui funktsioon f on pidevalt diferentseeruv ning omab nullist erinevaid väärtusi vaid lõplikus vahemikus, siis funktsiooni $f'(x)$ Fourier' pööre on $\hat{f}'(\xi) = 2\pi i\xi\hat{f}(\xi)$.*

Olgu h funktsiooni $e^x\chi(x)$ Fourier' pööre. Siis lause 9.2.1 põhjal

$$e^x\chi(x) = \int_{\mathbb{R}} h(\xi)e^{-ix\xi}\,d\xi.$$

Asendades sinna $x = \frac{\log d}{\log R}$ ning võrrandit algebraliselt teisendades saame

$$\chi\left(\frac{\log d}{\log R}\right) = \int_{\mathbb{R}} d^{-\frac{1+i\xi}{\log R}} h(\xi)\, d\xi.$$

Et kasutada seda integraali avaldises (32), piirame teda lõigule $I = [-(\log R)^{\frac{1}{2}}, (\log R)^{\frac{1}{2}}]$. Kuna χ on pidevalt diferentseeruv ning omab nullist erinevaid väärtusi vaid vahemikus $[-1, 1]$, siis lause 9.2.2 põhjal näeme, et $h(\xi) = O_\xi((1+|\xi|)^{-A})$ iga $A > 0$ korral. Seega viga integreerimispiirkonna muutmisest on ($A > 1$ korral) maksimaalselt

$$2\int_{(\log R)^{\frac{1}{2}}}^{\infty} \left| d^{-\frac{1+i\xi}{\log R}} h(\xi) \right| d\xi \leqslant O\left(d^{-\frac{1}{\log R}} \int_{(\log R)^{\frac{1}{2}}}^{\infty} (1+|\xi|)^{-A}\, d\xi \right) = O(d^{-\frac{1}{\log R}} (\log R)^{-\frac{A-1}{2}}),$$

kus võime asendada $\frac{A-1}{2} \to A$, kuna väide kehtib iga $A > 0$ korral. Seega

$$\chi\left(\frac{\log d}{\log R}\right) = \int_I d^{-\frac{1+i\xi}{\log R}} h(\xi)\, d\xi + O(d^{-\frac{1}{\log R}} (\log R)^{-A}). \tag{33}$$

Siis $\chi(\log d/\log R) = O(d^{-\frac{1}{\log R}})$, sest vealiikme jaoks järeldub hinnang sellest, et $(\log R)^{-A} = o(1)$ ning integraali jaoks sellest, et

$$\int_I d^{-\frac{1+i\xi}{\log R}} h(\xi)\, d\xi \leqslant d^{-\frac{1}{\log R}} \int_I h(\xi)\, d\xi = d^{-\frac{1}{\log R}} O(1).$$

Teeme iga $j = 1, \ldots, m$ korral muutujavahetused

$$z_j := \frac{1+i\xi_j}{\log R} \qquad \text{ja} \qquad z'_j := \frac{1+i\xi'_j}{\log R}.$$

Võrrandi (33) põhjal saame nüüd

$$\prod_{j=1}^{m} \chi\left(\frac{\log d_j}{\log R}\right)\chi\left(\frac{\log d_j'}{\log R}\right) = \int_I \cdots \int_I \prod_{j=1}^{m} d_j^{-z_j} d_j'^{-z_j'} h(\xi_j)h(\xi_j')\, d\xi_j d\xi_j' + O\left((\log R)^{-A} \prod_{j=1}^{m} (d_j d_j')^{-\frac{1}{\log R}}\right). \quad (34)$$

Nüüd hindame (34) põhjal avaldist (32) (vealiikmeid uurime hiljem):

$$(\log R)^{2m} \int_I \cdots \int_I \sum_{d_1, d_1', \ldots, d_m, d_m' \in \mathbb{N}} \mathbb{E}_{\mathbf{x} \in \mathbb{Z}_D^t} [1_{d_j, d_j' | \theta_j(\mathbf{x})\ \forall j}] \cdot \prod_{j=1}^{m} \mu(d_j) d_j^{-z_j} \mu(d_j') d_j'^{-z_j'} h(\xi_j) h(\xi_j')\, d\xi_j d\xi_j'. \quad (35)$$

Me võisime siin ära vahetada summa ja integraali, sest summas on vaid lõplik arv nullist erinevaid liikmeid. Jagades nüüd arvud $d_1, d_1', \ldots, d_m, d_m'$ avaldises (35) algarvude korrutisteks, saame

$$(35) = (\log R)^{2m} \int_I \cdots \int_I \prod_p E_p(\xi) \cdot \prod_{j=1}^{m} h(\xi_j)h(\xi_j')\, d\xi_j d\xi_j', \quad (36)$$

kus $\xi = (\xi_1, \xi_1', \ldots, \xi_m, \xi_m') \in I^{2m}$ ning

$$E_p(\xi) := \sum_{d_1, d_1', \ldots, d_m, d_m' \in \{1, p\}} \mathbb{E}_{\mathbf{x} \in \mathbb{Z}_p^t} [1_{d_j, d_j' | \theta_j(\mathbf{x})\ \forall j}] \prod_{j=1}^{m} \mu(d_j) d_j^{-z_j} \mu(d_j') d_j'^{-z_j'}.$$

Kui $p \leqslant w$, siis W definitsiooni põhjal iga j korral $p \nmid \theta_j(\mathbf{x}) = W\psi_j(\mathbf{x}) + 1$, seega $E_p(\xi) = 1$. Kui $p > w$, siis keskväärtus on 1, kui kõik arvud d_j, d_j' on ühed, keskväärtus on $\frac{1}{p}$, kui $d_j d_j' = 1$ iga j korral peale täpselt ühe (sest siis täpselt $\frac{1}{p}$ vektoritest $\mathbf{x} \in \mathbb{Z}_p^t$ rahuldavad vajalikku jaguvust), ning ülejäänud juhtudel on keskväärtus maksimaalselt $\frac{1}{p^2}$ (sest maksimaalselt $\frac{1}{p^2}$ vektoritest $\mathbf{x} \in \mathbb{Z}_p^t$ rahuldavad vajalikke jaguvusi) ning korrutise väärtus $O(1)$, kusjuures vastavaid liidetavaid summas on $< 2^{2m} = O(1)$.

Järelikult juhul $p > w$

$$E_p(\xi) = 1 - p^{-1}\sum_{j=1}^{m}(p^{-z_j} + p^{-z'_j} - p^{-z_j - z'_j}) + O(p^{-2}) = (1 + O(p^{-2}))E'_p(\xi),$$

kus iga algarvu p korral

$$E'_p(\xi) := \prod_{j=1}^{m} \frac{(1 - p^{-1-z_j})(1 - p^{-1-z'_j})}{1 - p^{-1-z_j-z'_j}}.$$

Kuid siis

$$\prod_p E_p(\xi) = \prod_{p>w}(1 + O(p^{-2}))E'_p(\xi) = (1 + O(w^{-1}))\left(\prod_{p\leqslant w} E'_p(\xi)\right)^{-1}\prod_p E'_p(\xi). \tag{37}$$

Selle võrrandi paremal poolel esineb korrutis $\prod_p E'_p(\xi)$.

Asendades siin sisse $E'_p(\xi)$ definitsiooni ning vahetades ära korrutiste järjekorra, näeme, et meil tekib hulganisti korrutisi kujul $\prod_p (1 - p^{-s})^{-1}$, kus $s \in \mathbb{R}$. See on aga Riemanni dzeetafunktsioon

$$\zeta(s) := \prod_p (1 - p^{-s})^{-1}.$$

Me kasutame Riemanni dzeetafunktsiooni järgnevat omadust:

Lemma 9.3. *Piirkonnas* $\operatorname{Re} s > 1$ *rahuldab Riemanni dzeetafunktsioon* ζ *tingimust*

$$\zeta(s) = (s-1)^{-1} + O(1),$$

eeldusel, et $s - 1 = O(1)$.

Lemma 9.3 ütleb meile, et

$$\prod_p E_p'(\xi) = \prod_{j=1}^{m} \frac{\zeta(1+z_j+z_j')}{\zeta(1+z_j)\zeta(1+z_j')} \approx \prod_{j=1}^{m} \frac{z_j z_j'}{z_j+z_j'}, \tag{38}$$

kus $\approx$ tähendab erinemist $o(1)$ võrra. (Siin kasutame arvudena $\xi_1, \xi_1', \ldots \xi_m, \xi_m' \in I$ arve $|z_1|, |z_1'|, \ldots, |z_m|, |z_m'| = O((\log R)^{-\frac{1}{2}})$.)

Algarvude $p \leqslant w$ jaoks saame kasutada hinnangut $E_p'(\xi) \approx (1-p^{-1})^m$, kuna $z_j, z_j' = o(1)$ nende definitsiooni põhjal.

Seega

$$\prod_{p \leqslant w} E_p'(\xi) \approx \prod_{p \leqslant w} (1-p^{-1})^m = \left(\prod_{p \leqslant w} \frac{p-1}{p} \right)^m = \left(\frac{\varphi(W)}{W} \right)^m. \tag{39}$$

Asendades võrrandid (37), (38) ja (39) võrrandisse (36), näeme, et (vealiiget uurime hiljem)

$$(36) \approx (\log R)^{2m} \left(\frac{W}{\varphi(W)} \right)^m \int_I \cdots \int_I \prod_{j=1}^{m} \frac{z_j z_j'}{z_j+z_j'} h(\xi_j)h(\xi_j')\, d\xi_j d\xi_j'. \tag{40}$$

Nüüd on veel vaja vaid hinnata integraali

$$\int_I \int_I \frac{z_j z_j'}{z_j+z_j'} h(\xi_j)h(\xi_j')\, d\xi_j d\xi_j' = \frac{1}{\log R} \int_I \int_I \frac{(1+i\xi_j)(1+i\xi_j')}{2+i(\xi_j+\xi_j')} h(\xi_j)h(\xi_j')\, d\xi_j d\xi_j'.$$

Me tahame siin asendada integreerimispiirkonna $I = [-(\log R)^{1/2}, (\log R)^{1/2}]$ reaalteljega $\mathbb{R}$. Kuna suure $|\xi|$ korral hääbub h seose $h(\xi) = O_\xi((1+|\xi|)^{-A})$ järgi, siis analoogselt võrrandi (33) tõestusega näeme, et integreerimispiirkonna vahetus mõjutab integraali väärtust vaid $O(\log^{-A} R)$ võrra iga $A > 0$ korral.

Siis tahame näidata, et

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \frac{(1+i\xi)(1+i\xi')}{2+i(\xi+\xi')} h(\xi)h(\xi')\, d\xi d\xi' = \int_0^\infty |\chi'(x)|^2\, dx = c_\chi. \tag{41}$$

Kasutades võrdust

$$\frac{1}{2+i(\xi+\xi')} = \int_0^\infty e^{-(1+i\xi)x} e^{-(1+i\xi')x}\, dx,$$

saame võrrandi (41) vasaku poole ümber kirjutada kujul

$$\int_0^\infty \left(\int_{\mathbb{R}} h(\xi)(1+i\xi)e^{-(1+i\xi)x}\, d\xi \right)^2 dx.$$

Sulgudes olev avaldis on lausete 9.2.1 ja 9.2.3 põhjal $-\chi'(x)$, seega tingimus (41) on tõestatud. Nüüd, asendades võrrandi (41) võrrandisse (40), olemegi tõestanud lause 8.5, nagu soovitud.

Nüüd hindame täpsemalt vealiikmeid varasemas arutluskäigus, et veenduda selle korrektsuses.

Hinnang võrrandis (35). Meil on vaja hinnata avaldiste (35) ja (32) vahet ehk avaldise (34) vealiikme mõju avaldisele (32). Võttes igast liikmest absoluutväärtuse ja kasutades asjaolu, et $\mu(d) \neq 0$ parajasti siis, kui d on ruuduvaba, saame hinnangu

$$\begin{aligned}
&(\log R)^{2m} \sum_{d_1,d_1',\ldots,d_m,d_m' \in \mathbb{N}} \left(O(D^{-\frac{1}{\log R}}(\log R)^{-A}) \right) \mathbb{E}_{\mathbf{x}\in\mathbb{Z}_D^t}[1_{d_j,d_j'|\theta_j(\mathbf{x})\ \forall j}] \\
&\leqslant (\log R)^{O(1)-A} \sum_{\substack{d_1,d_1',\ldots,d_m,d_m' \\ \text{ruuduvabad}}} \mathbb{E}_{\mathbf{x}\in\mathbb{Z}_D^t}[1_{d_j,d_j'|\theta_j(\mathbf{x})\ \forall j}](d_1d_1'\cdots d_md_m')^{-\frac{1}{\log R}} \\
&= (\log R)^{O(1)-A} \prod_p \sum_{d_1,d_1',\ldots,d_m,d_m' \in \{1,p\}} \mathbb{E}_{\mathbf{x}\in\mathbb{Z}_p^t}[1_{d_j,d_j'|\theta_j(\mathbf{x})\ \forall j}](d_1d_1'\cdots d_md_m')^{-\frac{1}{\log R}}.
\end{aligned}$$

Keskväärtus $\mathbb{E}_{\mathbf{x}\in\mathbb{Z}_p^t}[1_{d_j,d_j'|\theta_j(\mathbf{x})\ \forall j}]$ on 1, kui arvud d_i, d_i' on kõik ühed, ning muidu maksimaalselt $\frac{1}{p}$. Seega saame uuritavat avaldist ülalt piirata avaldisega

$$\begin{aligned}
&(\log R)^{O(1)-A}\prod_p\left(1+p^{-1}\sum_{\substack{d_1,d_1',\dots,d_m,d_m'\in\{1,p\}\\ d_1d_1'\dots d_m'\neq 1}}(d_1d_1'\cdots d_md_m')^{-\frac{1}{\log R}}\right)\\
&=(\log R)^{O(1)-A}\prod_p\left(1+p^{-1}((p^{-\frac{1}{\log R}}+1)^{2m}-1)\right)\\
&\leqslant(\log R)^{O(1)-A}\prod_p\left(1-p^{-1-\frac{1}{\log R}}\right)^{-O(1)}\\
&=(\log R)^{O(1)-A}\zeta(1+\frac{1}{\log R})^{O(1)}.
\end{aligned}$$

Seega avaldiste (35) ja (32) vahe on $O((\log R)^{O(1)-A})$, mis on piisavalt suure A korral piisavalt väike.

Hinnang võrrandis (38). Kuna $|\xi_j|,\left|\xi_j'\right|\leqslant(\log R)^{\frac{1}{2}}$, siis järelikult $|z_j|,\left|z_j'\right|=O((\log R)^{\frac{1}{2}})$. Seega

$$\begin{aligned}
\prod_{j=1}^m\frac{\zeta(1+z_j+z_j')}{\zeta(1+z_j)\zeta(1+z_j')}&=\prod_{j=1}^m\frac{((z_j+z_j')^{-1}+O(1))}{(z_j^{-1}+O(1))(z_j'^{-1}+O(1))}\\
&=(1+O((\log R)^{-\frac{1}{2}}))\prod_{j=1}^m\frac{z_jz_j'}{z_j+z_j'}.
\end{aligned}\tag{42}$$

Hinnang võrrandis (39). Kuna algarvude $p\leqslant w$ jaoks $|z|\log p=O(1)$, siis

$$1-p^{-1-z}=1-p^{-1}e^{-z\log p}=1-p^{-1}(1+O(|z|\log p))=(1-p^{-1})(1+O(|z|\,p^{-1}\log p)).$$

Järelikult kõigi $p\leqslant w$ ja $\xi_1,\xi_1',\dots,\xi_m,\xi_m'\in I$ korral

$$E_p'(\xi)=\left(1+O\left(\frac{\log p}{p\,(\log R)^{\frac{1}{2}}}\right)\right)(1-p^{-1})^m$$

ning seega ka

$$\prod_{p\leqslant w} E'_p(\xi) = \left(1 + O\left(\frac{w}{(\log R)^{\frac{1}{2}}}\right)\right) \prod_{p\leqslant w}(1-p^{-1})^m. \tag{43}$$

Hinnang võrrandis (40). Kasutades võrrandeid (37), (42) ja (43), näeme, et võrrandi (40) poolte jagatis on $1 + O(\frac{1}{w} + \frac{w}{(\log R)^{\frac{1}{2}}}) = 1 + o(1)$, eeldusel, et w kasvab N suhtes piisavalt aeglaselt.

Kasutatud allikad

Boyd, Stephen P ja Lieven Vandenberghe (2004). *Convex optimization.* Cambridge University Press, lk. 49.

Conlon, David, Jacob Fox ja Yufei Zhao (2014). „The Green-Tao theorem: an exposition“. *EMS Surveys in Mathematical Sciences* 1.2, lk. 257–291.

Stein, Elias M ja Rami Shakarchi (2011). *Fourier analysis: an introduction.* Köide 1. Princeton University Press, lk. 134–137.

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Hendrik Vija,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose Greeni-Tao teoreem, mille juhendaja on Lauri Tart, reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada digitaalarhiivi DSpace kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi DSpace kaudu Creative Commonsi litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaaloman-di ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Hendrik Vija

9. mai 2023