Адитивна комбинаторика и Ротова теорема о трочланим аритметичким прогресијама

***др Горан Ђанковић*, Београд**

1. Увод

У овом чланку ћемо приказати доказ следеће теореме, коју је доказао Клаус Рот[1] 1953. године.

Ротова теорема. *Постоји нека позитивна константа C таква да било који подскуп A природних бројева у интервалу $[1, N]$, кардиналности*

$$|A| \geqslant C\frac{N}{\log\log N}$$

садржи нетривијалну аритметичку прогресију дужине 3.

Овде читалац свакако треба да замисли да је N велик број, или да $N \to \infty$, јер се ионако тиме тежина проблема још повећава. Вредност константе C је апсолутна (неки одређен позитиван број), али није битно колика је она тачно (што ће бити јасније после дискусије која следи, због присуства фактора $(\log\log N)^{-1}$).

Ова теорема је један од најранијих и најкласичнијих резултата области математике која се данас зове *адитивна комбинаторика*. Проблем је очигледно комбинаторне природе: A је *произвољан* подскуп интервала $[1, N]$ и питамо се да ли у таквом подскупу можемо наћи неку *адитивну структуру*, у овом случају 3 елемента $a_1, a_2, a_3 \in A$, тако да је $a_3 - a_2 = a_2 - a_1$ (овај услов је *адитиван*, па отуда и комбинаторна питања везана за овакве услове припадају именованој области).

Како размишљати о оваквом проблему? Најпре, потребна нам је величина која ће мерити *тежину* проблема, а то је у нашем случају $\delta := \frac{|A|}{N}$. Вредност $\delta \in (0, 1)$ можемо посматрати као *густину* подскупа A у интервалу $[1, N]$, или као *вероватноћу* да случајно изабрани елемент n тог интервала припада и подскупу A.

Наш циљ је да нађемо „парче" адитивне структуре у подскупу A. Интуитивно је јасно да ако је густина δ већа, то ће нам налажење овакве подструктуре бити лакше. Нпр. ако је $\delta = 0{,}7$, можемо се лако уверити да A мора садржати бар једну аритметичку прогресију дужине 3. На пример, можемо претпоставити да $3 \mid N$ и поделити интервал $[1, N]$ на трочлане аритметичке прогресије $\{1, \frac{N}{3}+1, \frac{2N}{3}+1\}$, $\{2, \frac{N}{3}+2, \frac{2N}{3}+2\}$, $\{3, \frac{N}{3}+3, \frac{2N}{3}+3\}, \ldots$. Ако би у свакој од њих бар по један елемент био ван A, густина подскупа A не би била већа од $2/3$. Дакле, бар једна од наведених трочланих аритметичких прогресија мора припадати скупу A. Међутим, ако подскуп A постаје све ређи и ређи, јасно је да проблем постаје све тежи и тежи. У Ротовој теореми, густина је $\delta = \frac{C}{\log\log N}$. Како овде N може да буде произвољно велико, ова густина може да буде и мања од нпр. 0,000000000000001. Дакле, подскуп A може да буде изузетно *редак* и при том потпуно произвољан, а Ротова теорема тврди да

[1] Klaus Roth (1925–), британски математичар немачког порекла, добитник Филдсове медаље 1958. године

ћемо чак и тада успети да пронађемо адитивну правилност у њему (у овом случају 3−аритметичку прогресију).

Формулација питања заиста звучи елементарно и заиста, многи средњошколци би могли размишљати о њему. Покушаћемо да и доказ представимо у „елементарном духу" и од читаоца се очекује једино познавање тригонометријског облика комплексног броја и основних тригонометријских функција. Међутим, иако ће све у доказу бити *коначно* и *дискретно*, доказ[2] (филозофски и суштински) припада области математике која се зове *Фуријеова* (или општије *хармонијска*) *анализа*. Дакле, пробаћемо да разумемо комбинаторни проблем, анализом. То је и други циљ овог чланка – упознавање средњошколаца са хармонијском анализом, веома важном и моћном математичком дисциплином. Са друге стране, идеје и сама структура доказа су изузетно дубоки, али и употребљиви и у многим другим контекстима.

2. ДИСКРЕТНА ФУРИЈЕОВА АНАЛИЗА

Неке је N произвољан природан број. Означимо са $\mathbb{Z}_N = \{0, 1, 2, \ldots, N-2, N-1\}$ потпун систем остатака по модулу N. Нека је $f : \mathbb{Z}_N \to \mathbb{C}$ произвољна комплексна функција на скупу $\mathbb{Z}_N$. Наравно, ако је читаоцу лакше, f можемо видети и као $f_1 + if_2$, где су f_1 и f_2 две реалне функције, а i имагинарна јединица. Додатно, како је скуп $\mathbb{Z}_N$ коначан, функција на њему је само једна N−торка комплексних бројева: $(f(0), f(1), \ldots, f(N-1))$.

За $\theta \in \mathbb{R}$ уведимо следећу ознаку:

$$e(\theta) := e^{2\pi i\theta} = \cos(2\pi\theta) + i\sin(2\pi\theta).$$

Дакле, $e(\theta)$ се налази на јединичном кругу у комплексној равни и кад θ пролази интервал $[0, 1)$, комплексни бројеви $e(\theta)$ пролазе цео јединични круг. Наравно, због периодичности функција cos и sin, и функција $e(\theta)$ је периодична, са периодом 1. Такође, препуштамо читаоцу да се увери да за све $\theta_1, \theta_2 \in \mathbb{R}$ важи следећа фундаментална адициона формула:

$$e(\theta_1 + \theta_2) = e(\theta_1) \cdot e(\theta_2). \tag{1}$$

Приметимо још да за комплексно коњуговање важи: $\overline{e(\theta)} = e(-\theta)$.

Сада можемо дефинисати нову функцију $\widehat{f} : \mathbb{Z}_N \to \mathbb{C}$ на скупу остатака по модулу N, која ће играти кључну улогу у доказу Ротове теореме:

$$\widehat{f}(k) = \sum_{n \in \mathbb{Z}_N} f(n) e\left(-\frac{kn}{N}\right), \quad \text{за } k \in \mathbb{Z}_N. \tag{2}$$

Приметите да због 1−периодичности функције $e(\theta)$, ова дефиниција не зависи од избора представника класа остатака по модулу N у сумацији на десној страни. Функцију $\widehat{f}$ зовемо *дискретна Фуријеова трансформација* полазне функције $f : \mathbb{Z}_N \to \mathbb{C}$.

[2]Ову варијанту доказа је дао британски математичар Гауерс (Timothy Gowers, 1963 –), добитник Филдсове медаље 1998.

Ортогоналност

Приметимо да важи и следећи фундаментални идентитет: за $N \in \mathbb{N}$ и $a \in \mathbb{Z}$ је

$$\frac{1}{N} \sum_{k \in \mathbb{Z}_N} e\left(\frac{ak}{N}\right) = \begin{cases} 1, & \text{ако } N \mid a, \\ 0, & \text{ако } N \nmid a. \end{cases} \tag{3}$$

Ако $N \mid a$, сваки од ak/N је цео број, па је и $e(ak/N) = 1$, за све $k \in \mathbb{Z}_N$ и идентитет је тривијалан у овом случају.

Ако $N \nmid a$, приметимо да је онда $e(a/N) = \cos(2\pi a/N) + i \sin(2\pi a/N) \neq 1$. Али, користећи (1), имамо да је

$$e\left(\frac{a}{N}\right) \sum_{k \in \mathbb{Z}_N} e\left(\frac{ak}{N}\right) = \sum_{k \in \mathbb{Z}_N} e\left(\frac{a(k+1)}{N}\right) = \sum_{k \in \mathbb{Z}_N} e\left(\frac{ak}{N}\right),$$

где последња једнакост следи јер ако k пролази потпуним системом остатака по модулу N, онда и вредности $k+1$ пролазе потпуним системом остатака по модулу N (ово је кључно!). Дакле, добијамо да је

$$\left(e\left(\frac{a}{N}\right) - 1\right) \cdot \sum_{k \in \mathbb{Z}_N} e\left(\frac{ak}{N}\right) = 0,$$

одакле следи идентитет (3) и у овом случају.

Формула Фуријеове инверзије

За било које $m \in \mathbb{Z}_N$, замењујући (2), налазимо да је

$$\begin{aligned} \frac{1}{N} \sum_{k \in \mathbb{Z}_N} \widehat{f}(k) e\left(\frac{km}{N}\right) &= \frac{1}{N} \sum_{k \in \mathbb{Z}_N} e\left(\frac{km}{N}\right) \sum_{n \in \mathbb{Z}_N} f(n) e\left(-\frac{kn}{N}\right) \\ &= \sum_{n \in \mathbb{Z}_N} f(n) \frac{1}{N} \sum_{k \in \mathbb{Z}_N} e\left(\frac{k(m-n)}{N}\right) \\ &= f(m), \end{aligned}$$

где смо последњу једнакост добили применом ортогоналности (3). Дакле, добили смо да за произвољну функцију f на $\mathbb{Z}_N$ важи:

$$f(m) = \frac{1}{N} \sum_{k \in \mathbb{Z}_N} \widehat{f}(k) e\left(\frac{km}{N}\right) = \frac{1}{N} \sum_{k \in \mathbb{Z}_N} \widehat{f}(k) \left[\cos\left(\frac{2\pi km}{N}\right) + i \sin\left(\frac{2\pi km}{N}\right)\right].$$

Ова фундаментална формула изражава вредност полазне функције, преко вредности њене Фуријеове трансформације $\widehat{f}$ и зове се *формула Фуријеове инверзије*. Она нам каже да функције $m \mapsto e(km/N)$, за $k = 0, 1, 2, \ldots, N-1$ представљају *основне функције* преко којих можемо изразити све остале функције f на скупу $\mathbb{Z}_N$. Коефицијенти $\widehat{f}(k)$ у овој суми су управо вредности Фуријеове трансформације и зато их некад зовемо и *Фуријеови коефицијенти* функције f. Дакле, интуитивно, Фуријеов коефицијент $\widehat{f}(k)$

мери на који начин и колико основна функција $e(km/N)$ утиче на понашање и изглед оригиналне функције $m \mapsto f(m)$.

Како $e(km/N)$ само означава збир косинуса и синуса, читалац свакако може да замисли да нам формула Фуријеове инверзије казује да се *свака* функција f на $\mathbb{Z}_N$ може разложити на збир *таласа* (сваки cos и сваки sin је по један талас). Фуријеови коефицијенти $\widehat{f}(k)$ онда одређују *амплитуде* појединих компонентних таласа, док параметар k одређује *фреквенцију* таласа $m \mapsto \cos(2\pi km/N)$ и $m \mapsto \sin(2\pi km/N)$.

Парсевалова формула

Биће нам потребан и следећи идентитет, који важи за све функције $f : \mathbb{Z}_N \to \mathbb{C}$:

$$\sum_{k\in\mathbb{Z}_N} |\widehat{f}(k)|^2 = N \sum_{n\in\mathbb{Z}_N} |f(n)|^2. \tag{4}$$

Доказ поново добијамо заменом дефиниције (2) у сабирке на левој страни, променом редоследа сумације и применом формуле ортогоналности (3):

$$\begin{aligned}\sum_{k\in\mathbb{Z}_N} |\widehat{f}(k)|^2 &= \sum_{k\in\mathbb{Z}_N} \widehat{f}(k)\overline{\widehat{f}(k)} = \sum_{k\in\mathbb{Z}_N}\sum_{n\in\mathbb{Z}_N} f(n)e\left(-\frac{kn}{N}\right) \sum_{m\in\mathbb{Z}_N} \overline{f(m)}\,\overline{e\left(-\frac{km}{N}\right)} \\ &= \sum_{n\in\mathbb{Z}_N}\sum_{m\in\mathbb{Z}_N} f(n)\overline{f(m)} \cdot \sum_{k\in\mathbb{Z}_N} e\left(\frac{k(m-n)}{N}\right) = N\sum_{n\in\mathbb{Z}_N} f(n)\overline{f(n)},\end{aligned}$$

јер због (3) „преживе" само производи $f(n)\overline{f(m)}$, за које $n \equiv m \pmod N$, тј. за које су n и m представници *исте* класе остатака по модулу N.

Функције индикатори

Нека је A произвољан подскуп целих бројева у интервалу $[0, N)$, који онда можемо видети и као подскуп скупа $\mathbb{Z}_N$ (за сваки коначан скуп природних бројева A, постоји неки довољно велик природан број N који ово задовољава). Овом скупу ћемо придружити његову *индикатор–функцију* $A : \mathbb{Z}_N \to \mathbb{C}$, која је дефинисана са

$$A(n) = \begin{cases} 1, & \text{ако је } n \in A, \\ 0, & \text{ако } n \notin A. \end{cases}$$

Такође, уводимо и *избалансирану индикатор–функцију* $f_A : \mathbb{Z}_N \to \mathbb{C}$, дефинисану са $f_A(n) = A(n) - |A|/N$. Њена средња вредност на $\mathbb{Z}_N$ је једнака 0.

Приметимо на овом месту да је вредност $0-$фреквенције индикатор–функције $A(n)$ подскупа A једнака

$$\widehat{A}(0) = \sum_{n\in\mathbb{Z}_N} A(n)e(0) = |A|.$$

И уопште, вредност $\widehat{f}(0)$ Фуријеове трансформације $\widehat{f}$ произвољне функције f у тачки 0 је једнака $\widehat{f}(0) = \sum_{n\in\mathbb{Z}_N} f(n)$. Ако је f ненегативна реална функција, интуитивно, $\widehat{f}(0)$ представља укупну „масу" функције f. Зато је и $\widehat{f_A}(0) = 0$.

3. Дирихлеова теорема о диофантским апроксимацијама

Дирихлеова[3] **теорема.** *Нека је α произвољан реалан број и Q произвољан природан број. Онда постоји рационалан број a/q, за који је* НЗД$(a,q)=1$ *и именилац $q \leqslant Q$, такав да важи*

$$\left|\alpha - \frac{a}{q}\right| < \frac{1}{qQ}.$$

Дакле, Дирихлеова теорема нам каже да сваки реалан број α можемо *апроксимирати* неким рационалним бројем, при чему можемо истовремено контролисати величину имениоца тог рационалног броја и квалитет апроксимације једним параметром Q, као у теореми.

Доказ. Доказ је лагана примена *Дирихлеовог принципа* („о зечевима и кавезима"). Поделимо интервал $[0,1)$ на Q дисјунктних интервала $\left[\frac{j}{Q}, \frac{j+1}{Q}\right)$, $j = 0,1,2,\ldots,Q-1$. Посматрајмо $Q+1$ тачака

$$k\alpha - [k\alpha], \qquad k = 0,1,2,\ldots,Q,$$

које се све налазе у $[0,1)$. Дакле имамо $Q+1$ тачака („зечева") и Q подинтервала („кавеза"), што значи да се бар две тачке, рецимо $k_1\alpha - [k_1\alpha]$ и $k_2\alpha - [k_2\alpha]$, $k_1 < k_2$, налазе у истом подинтервалу. Дакле,

$$|k_2\alpha - [k_2\alpha] - (k_1\alpha - [k_1\alpha])| < \frac{1}{Q},$$

па тврђење следи ако означимо $k_2 - k_1$ са q, а $[k_2\alpha] - [k_1\alpha]$ са a. □

4. Доказ Ротове теореме

Нека је, дакле, $A \subseteq [0, N-1]$, при чему ћемо претпоставити још и да је N непарно (довољан нам је било који интервал који садржи скуп A, па можемо и ово додатно захтевати). Означимо $\delta = \frac{|A|}{N}$.

Скуп A садржи подскуп свих својих парних и подскуп свих својих непарних елемената. Означимо са B онај од та два подскупа који је већи (ако су исте кардиналности, онда било који од њих). Нека су $A, B : \mathbb{Z}_N \to \mathbb{C}$ одговарајуће индикатор–функције ових подскупова и $\widehat{A}, \widehat{B} : \mathbb{Z}_N \to \mathbb{C}$ њихове Фуријеове трансформације. Пођимо од следеће величине:

$$V = \frac{1}{N} \sum_{k \in \mathbb{Z}_N} \widehat{B}(k)^2 \widehat{A}(-2k).$$

[3] Peter Gustav Lejeune Dirichlet (1805–1859), немачки математичар, пионир и оснивач аналитичке теорије бројева

Ако заменимо дефиницију (2), променом редоследа сумирања добијамо да је

$$V = \frac{1}{N} \sum_{k \in \mathbb{Z}_N} \sum_{x \in \mathbb{Z}_N} B(x) e\left(-\frac{kx}{N}\right) \sum_{y \in \mathbb{Z}_N} B(y) e\left(-\frac{ky}{N}\right) \sum_{z \in \mathbb{Z}_N} A(z) e\left(\frac{2kz}{N}\right)$$
$$= \sum_{x \in \mathbb{Z}_N} B(x) \sum_{y \in \mathbb{Z}_N} B(y) \sum_{z \in \mathbb{Z}_N} A(z) \cdot \frac{1}{N} \sum_{k \in \mathbb{Z}_N} e\left(\frac{k(2z - x - y)}{N}\right).$$

Према (3), последњи фактор је 1 ако и само ако је $x + y - 2z$ дељиво са N, па закључујемо да величина V у ствари броји елементе следећег скупа:

$$\{(x, y, z) \in B \times B \times A \mid x + y \equiv 2z \pmod{N}\}.$$

Како смо (згодно) претпоставили да је N *непарно*, једначине $x + y - 2z = \pm N$ немају решења, јер су $x, y \in B$, тј. исте парности. Како су сви $x, y, z \in [0, N-1]$, конгруенција $x + y \equiv 2z \pmod{N}$ се своди на *једнакост* $x + y = 2z$. Дакле добили смо да је

$$\frac{1}{N} \sum_{k \in \mathbb{Z}_N} \widehat{B}(k)^2 \widehat{A}(-2k) = |\{(x, y, z) \in B \times B \times A \mid x + y = 2z\}|.$$

Једначина $x+y = 2z$ има тривијална решења $x = y = z \in B$ и њих има $|B|$. Међутим, ако је $x + y = 2z$ и нпр. $y > x$, онда је $y - z = z - x$, тј. x, z, y је једна *нетривијална* трочлана аритметичка прогресија у скупу A. Дакле, број нетривијалних трочланих аритметичких прогресија у скупу A, за које су $x, y \in B$, је једнак

$$\frac{1}{N} \sum_{k \in \mathbb{Z}_N} \widehat{B}(k)^2 \widehat{A}(-2k) - |B|.$$

То нису *све* трочлане аритметичке прогресије које се могу наћи у скупу A, али свеједно, ми само желимо да нађемо *једну* такву $3-$прогресију, па је довољно да покажемо да је горња вредност *позитивна*. Издвајањем $0-$фреквенције, тј. члана $k = 0$, налазимо да је последња величина једнака

$$\frac{|B|^2|A|}{N} - |B| + \frac{1}{N} \sum_{\substack{k \in \mathbb{Z}_N \\ k \not\equiv 0 \pmod{N}}} \widehat{B}(k)^2 \widehat{A}(-2k). \tag{5}$$

Зашто смо ово урадили? Вредности $\widehat{A}(0) = |A|$ и $\widehat{B}(0) = |B|$ су *„велике"*, док су вредности $\widehat{A}(k), \widehat{B}(k)$, за све остале k *„мале"*. Нпр.

$$\widehat{B}(k) = \sum_{n \in \mathbb{Z}_N} B(n) e\left(-\frac{kn}{N}\right) = \sum_{n \in \mathbb{Z}_N} B(n) \cos\left(\frac{2\pi kn}{N}\right) - i \sum_{n \in \mathbb{Z}_N} B(n) \sin\left(\frac{2\pi kn}{N}\right),$$

а како је B ненегативна функција, а cos и sin *осцилују*, очекујемо велико *„поништење"* у обе суме на десној страни (јер очекујемо подједнак број позитивних и негативних сабирака у њима). Да ли је то заиста тако, за сада нас не брине, издвајањем $0-$фреквенције

смо интуитивно издвојили *„главни члан"* (највећи допринос) наше суме и то је трик који је најчешће од помоћи.

Дошли смо до кључног места у доказу. Вредност (5) ћемо проучити раздвајањем два случаја.

Први случај. Сви Фуријеови коефицијенти $\widehat{A}(k)$ индикатор–функције A за $k \not\equiv 0$ су „мали". Прецизније, претпоставимо да за све $k \in \mathbb{Z}_N$, $k \not\equiv 0$ важи неједнакост (којом квантификујемо и прецизирамо придев „мали" за потребе доказа):

$$|\widehat{A}(k)| \leqslant \frac{\delta^2 N}{4}.$$

Како је N непарно, за $k \not\equiv 0$ је свакако и $-2k \not\equiv 0$. Сада користећи неједнакост троугла $|z_1 + z_2| \leqslant |z_1| + |z_2|$, за комплексне бројеве z_1, z_2, добијамо да је

$$\begin{aligned}\left|\frac{1}{N}\sum_{\substack{k\in\mathbb{Z}_N\\ k\not\equiv 0}} \widehat{B}(k)^2\widehat{A}(-2k)\right| &\leqslant \frac{1}{N}\sum_{\substack{k\in\mathbb{Z}_N\\ k\not\equiv 0}} |\widehat{B}(k)|^2|\widehat{A}(-2k)| \\ &\leqslant \frac{\delta^2}{4}\sum_{\substack{k\in\mathbb{Z}_N\\ k\not\equiv 0}} |\widehat{B}(k)|^2 = \frac{\delta^2}{4}\left(\sum_{k\in\mathbb{Z}_N} |\widehat{B}(k)|^2 - |\widehat{B}(0)|^2\right) \\ &= \frac{\delta^2}{4}\left(N\sum_{n\in\mathbb{Z}_N} |B(n)|^2 - |B|^2\right) = \frac{\delta^2}{4}\left(N|B| - |B|^2\right)\end{aligned}$$

где смо на преласку у трећи ред искористили Парсевалову формулу (4), као и да је $\widehat{B}(0) = |B|$.

Сада је величина (5) већа или једнака од

$$\begin{aligned}\frac{|B|^2|A|}{N} - |B| - \frac{\delta^2}{4}\left(N|B| - |B|^2\right) &\geqslant |B|\left(\delta|B| - 1 - \frac{\delta^2}{4}N\right) \\ &\geqslant \frac{|A|}{2}\left(\frac{\delta|A|}{2} - 1 - \frac{\delta|A|}{4}\right) = \frac{|A|}{2}\left(\frac{\delta|A|}{4} - 1\right),\end{aligned}$$

где смо искористили да је $|B| \geqslant |A|/2$. Вредност на десној страни је *строго позитивна* и за много мање δ него што је у тврђењу теореме (где је $\delta \geqslant C(\log\log N)^{-1}$). Дакле, у овом случају смо нашли аритметичке прогресије дужине 3 и то, много њих. Доказали смо да скуп A садржи бар једну 3−аритметичку прогресију, тако што смо оценили одоздо број 3−аритметичких прогресија унутар скупа A и доказали да је та доња граница позитивна! Приметите колико је тој позитивности допринео „велики" Фуријеов коефицијент 0−фреквенције.

Наравно у овом случају смо успели, јер смо и претпоставили да у (5) једино 0−фреквенција игра улогу, док је утицај осталих фреквенција (за $k \not\equiv 0$) занемарљив. Међутим, шта ако ипак још нека друга фреквенција значајније утиче на „звук" који производи скуп A? Долазимо до другог (тежег) случаја.

Други случај. Постоји неко $k \in \mathbb{Z}_N$, $k \not\equiv 0$, такво да је

$$|\widehat{A}(k)| > \frac{\delta^2 N}{4}.$$

Како је $\widehat{A}(k) = \sum_{n\in\mathbb{Z}_N} A(n)e(-kn/N)$ и $\sum_{n\in\mathbb{Z}_N} e(-kn/N) = 0$ на основу (3), јер $-k \not\equiv 0 \pmod N$, из наше претпоставке у овом случају следи да је и

$$\left|\sum_{n\in\mathbb{Z}_N} (A(n)-\delta)e\left(-\frac{kn}{N}\right)\right| > \frac{\delta^2 N}{4}. \tag{6}$$

„Сецкање" интервала. Уведимо *први параметар* Q, $1 \leqslant Q \leqslant N$, који је за сада слободан, али ћемо га ускоро погодно одабрати. Дирихлеова теорема о диофантским апроксимацијама нам каже да свакако можемо наћи неки разломак a/q такав да је $q \leqslant Q$, НЗД$(a,q) = 1$ и да је

$$\left|\frac{k}{N} - \frac{a}{q}\right| \leqslant \frac{1}{qQ}. \tag{7}$$

Поделимо интервал $[0, N-1]$ на аритметичке прогресије по модулу q. Свака од ових q прогресија има највише $\frac{N}{q}+1$ чланова. Сада уведимо и *други параметар* M: сваку од ових q прогресија додатно поделимо на по M подинтервала (почетних $M-1$ интервала исте кардиналности и у последњем M−том остатак). Дакле, интервал $[0, N-1]$ смо поделили на укупно qM подскупова, од којих сваки има највише $(N/qM)+1$ елемената и сваки је подниз узастопних елемената неке аритметичке прогресије модуло q.

Шта је идеја овог „сецкања"? Посматрајући грубо, комплексни бројеви (наше „основне функције") $e(-kn/N)$ којима множимо у (6) су *скоро константни* на сваком од оваквих подскупова. Нека је нпр. $J = \{n_1, n_1+q, n_1+2q, \ldots, n_1+(|J|-1)q\}$ један од таквих подскупова. Због (7) можемо написати $\frac{k}{N} = \frac{a}{q} + \theta$, где знамо да је $|\theta| \leqslant \frac{1}{qQ}$, па је

$$e\left(-\frac{kn}{N}\right) = e\left(-\left(\frac{a}{q}+\theta\right)n\right) = e\left(-\frac{an}{q}\right)e(-\theta n).$$

Фактор $e(-an/q)$ је константан (модула 1) за све $n \in J$, јер је по конструкцији поделе, цео подскуп J садржан у *истом* аритметичкој прогресији по модулу q. Дакле на J варира само други фактор $e(-\theta n)$, али не превише. Први n_1 и последњи члан $n_1+(|J|-1)q$ било ког подскупа J се разликују за највише $q\left(\frac{N}{qM}+1-1\right) = \frac{N}{M}$, па користећи познате неједнакости $|\sin x| \leqslant |x|$ и $1-\cos x \leqslant \frac{x^2}{2}$ добијамо да је

$$\begin{aligned}\left|\sum_{n\in J}(A(n)-\delta)e\left(-\frac{kn}{N}\right)\right| &= \left|\sum_{n\in J}(A(n)-\delta)e\left(-\frac{an}{q}\right)e(-\theta n)\right| \\ &= \left|\sum_{j=0}^{|J|-1}(A(n_1+jq)-\delta)e\left(-\frac{a(n_1+jq)}{q}\right)e(-\theta(n_1+jq))\right| \\ &= \left|e\left(-\frac{an_1}{q}\right)e(-\theta n_1)\right| \cdot \left|\sum_{j=0}^{|J|-1}(A(n_1+jq)-\delta)e(-jq\theta)\right|\end{aligned}$$

$$= \left| \sum_{j=0}^{|J|-1} (A(n_1+jq)-\delta)\left(1+\cos(2\pi jq\theta)-1-i\sin(2\pi jq\theta)\right) \right|$$
$$\leqslant \left| \sum_{j=0}^{|J|-1} (A(n_1+jq)-\delta) \right| + 2|J| \cdot 2\pi \frac{N}{M}|\theta|$$
$$\leqslant \left| \sum_{j=0}^{|J|-1} (A(n_1+jq)-\delta) \right| + 4\pi \frac{|J|N}{MqQ}.$$

Из неједнакости (6), после примене неједнакости троугла за подсуме које одговарају нашој партицији интервала $[0, N-1] = J_1 \cup J_2 \cup \cdots \cup J_{qM}$, сада имамо

$$\frac{\delta^2 N}{4} < \left| \sum_{n\in\mathbb{Z}_N} (A(n)-\delta) e\left(-\frac{kn}{N}\right) \right| \leqslant \sum_J \left| \sum_{n\in J} (A(n)-\delta) e\left(-\frac{kn}{N}\right) \right|$$
$$\leqslant \sum_J \left(\left| \sum_{n\in J} (A(n)-\delta) \right| + 4\pi \frac{|J|N}{MqQ} \right) = \sum_J \left| \sum_{n\in J} (A(n)-\delta) \right| + 4\pi \frac{N^2}{MqQ}.$$

На овом месту изаберимо наше параметре:

$$Q = \sqrt{N}; \quad q \leqslant Q \text{ из Дирихлеове теореме}; \quad M = 32\pi \frac{\sqrt{N}}{q\delta^2}.$$

За овај избор параметара последња неједнакост постаје

$$\frac{\delta^2 N}{8} < \sum_J \left| \sum_{n\in J} (A(n)-\delta) \right|.$$

Додатно, знамо да је и $\sum_J \sum_{n\in J} (A(n)-\delta) = \sum_{n\in\mathbb{Z}_N} (A(n)-\delta) = 0$. Сабирањем претходне неједнакости и ове једнакости, и како има највише qM подсупова J чије одговарајуће суме $\sum_{n\in J}(A(n)-\delta)$ су позитивне, закључујемо да постоји *бар један* подскуп J за који важи неједнакост

$$\sum_{n\in J} (A(n)-\delta) > \frac{\delta^2 N}{16qM}, \quad \text{тј.} \quad \sum_{n\in J} A(n) > \delta|J| + \frac{\delta^2 N}{16qM}.$$

Како је $|J| \leqslant \frac{N}{qM} + 1$, добијамо даље и неједнакост

$$\frac{1}{|J|} \sum_{n\in J} A(n) > \delta + \frac{\delta^2}{16} - \frac{1}{|J|} > \delta + \frac{\delta^2}{32}.$$

Међутим, израз на левој страни је тачно број елемената нашег скупа A који се налазе у подскупу J подељен укупним бројем елемената у J. Дакле на левој страни се налази *'релативна густина'* скупа A у подскупу J. Густина скупа A у целом интервалу

$[0, N-1]$ је била δ, а добијена неједнакост нам каже да је релативна густина A у J *већа*, бар $\delta + \frac{\delta^2}{32}$. Дакле овом прецизном анализом смо нашли један подскуп J у коме се густина скупа A повећала! Ово је за нас добро, јер интуитивно, ако се густина повећава, већа је и наша шанса да пронађемо аритметичке прогресије дужине 3 у датом подскупу. Сада ћемо се послужити (још) једним триком. Подскуп J није био произвољан, већ је парче аритметичке прогресије по модулу q. Зато тај подскуп J једном транслацијом и дилатацијом фактором $1/q$ можемо пресликати на интервал $[0, |J|-1]$ који је садржан у интервалу

$$\left[0, \frac{N}{qM}\right] = \left[0, \frac{\delta^2}{32\pi}\sqrt{N}\right].$$

Међутим, транслација и дилатација ће чувати 3−аритметичке прогресије! Ако их је било у J биће их и у скалираном интервалу и обрнуто! Дакле, овим триком смо ситуацију „подскуп A у интервалу $[0, N-1]$ густине δ", заменили ситуацијом „подскуп (који одговара неким елементима скупа A) у интервалу $\left[0, \frac{\delta^2}{32\pi}\sqrt{N}\right]$ густине бар $\delta + \frac{\delta^2}{32}$". Проблем смо свели на исти, али сада тражимо 3−аритметичке прогресије у скупу веће густине.

Овај процес сада можемо итерирати. Ако га поновимо још једном, густина ће се повећати на $\delta + \frac{\delta^2}{32} + \frac{1}{32}(\delta + \frac{\delta^2}{32})^2 > \delta + \frac{\delta^2}{16}$ итд. Ако га поновимо $\lceil \frac{32}{\delta} \rceil$ пута, густина ће се дуплирати (биће бар 2δ). После још $\lceil \frac{16}{\delta} \rceil$ итерација, густина ће се поново дуплирати ($\geqslant 2^2\delta$) итд. После највише $\lceil \log_2 \frac{0,7}{\delta} \rceil$ оваквих дуплирања густине, тј. после највише

$$\frac{32}{\delta} + \frac{16}{\delta} + \frac{8}{\delta} + \frac{4}{\delta} + \frac{2}{\delta} + \frac{1}{\delta} + \frac{1}{2\delta} + \frac{1}{4\delta} + \frac{1}{8\delta} + \cdots + \log_2 \frac{0,7}{\delta} \leqslant \left\lfloor \frac{100}{\delta} \right\rfloor$$

описаних итерација, можемо постићи да је релативна густина скупа A у крајњем издвојеном подскупу (који се низом транслација и дилатација може свести на неки интервал) бар 0,7. Још у Уводу смо видели да онда лако можемо наћи трочлану аритметичку прогресију у овом парчету, чија сва 3 елемента припадају скупу A.

Паралелно, после прве итерације, нова дужина интервала је $\frac{\delta^2}{32\pi}\sqrt{N}$, што је веће од нпр. $N^{1/3}$, у опсегу за параметар δ који је много шири и од онога што се тврди у теореми. После $\lfloor \frac{100}{\delta} \rfloor$ итерација, дужина новодобијеног интервала (тј. угњежденог парчета неке аритметичке прогресије) ће бити бар $N^{(1/3)^{100/\delta}}$. Наравно, да бисмо на крају добили прави интервал (на коме можемо да применимо тривијални аргумент из Увода, а не нпр. интервал од само 2 елемента), желимо да наш крајњи интервал има дужину бар нпр. 1000. Дакле услов да овај итеративни аргумент „пролази" је да буде

$$N^{(1/3)^{100/\delta}} \geqslant 1000,$$

одакле после два логаритмовања, налазимо да је неопходно да важи

$$\delta \geqslant \frac{C}{\log\log N},$$

за неку позитивну константу C. Ово је управо услов Ротове теореме, који гарантује да ћемо описаним итеративним процесом успети да нађемо аритметичку прогресију дужине 3. □

5. Интерпретација доказа и поглед унапред

Структура доказа је типична за много аргумената у адитивној комбинаторици и инспирисала је многе наредне спектакуларне резултате. Прва кључна идеја је бројање трочланих аритметичких прогресија уз помоћ Фуријеове анализе (видети израз (5)). Тиме је комбинаторни проблем детектован аналитички, а затим и „нападнут" моћним аналитичким методама.

Друга кључна идеја је подела на два случаја, тј. *дихотомија* у понашању скупа A унутар интервала $[1, N]$. Први случај је да су сви ненула Фуријеови коефицијенти индикатор–функције $A(n)$ скупа A „мали". То можемо интерпретирати на следећи начин: скуп A изгледа као да је потпуно случајно и униформно распоређен унутар интервала $[1, N]$. Ова униформност у расподели нам је омогућила да врло лако (из функције која броји 3−прогресије) добијемо постојање аритметичких прогресија дужине 3.

У другом случају, постоји бар један Фуријеов коефицијент $\widehat{A}(k)$, за неко $k \neq 0$, који је „велик". Интерпретација овога је да скуп A није униформно распоређен унутар интервала $[1, N]$, већ има *склоност* ка одређеној аритметичкој прогресији (то је прогресија $a \pmod q$ у доказу) унутар интервала $[1, N]$. На тој аритметичкој прогресији се повећава вероватноћа да ћемо наћи елементе скупа A. Свака итерација у доказу је прелазак на нову аритметичку подпрогресију претходно издвојене аритметичке прогресије унутар интервала $[1, N]$. После сваке итерације, вероватноћа да нађемо елементе скупа A на новој подпрогресији се *повећава*.

Дакле, у првом случају је A скоро равномерно распоређен, и та равномерност иде у нашу корист, док ако A није равномерно распоређен, онда ипак може да се нађе нека аритметичка подпрогресија (која има адитивну структуру сличну интервалу, тј. то је само нека дилатација краћег интервала) на којој сада A има већу густину, па нам у овом случају ово повећање густине иде у прилог!

Приметите да је описана дихотомија *контролисана* величином Фуријеових коефицијената, тј. да је контекст Фуријеове анализе допринео и квалитативном разумевању комбинаторне ситуације.

Следећи чувени резултат адитивне комбинаторике (и једна од најдубљих и најчувенијих теорема комбинаторике уопште) је Семередијева торема.

Семередијева[4] теорема. *За дати природан број k и дато реално $\delta > 0$, постоји неко $N(k, \delta) \in \mathbb{N}$ такво да за све $N \geqslant N(k, \delta)$ и за произвољан подскуп $A \subseteq [1, N]$ за који је $|A| \geqslant \delta N$, унутар скупа A можемо пронаћи неку аритметичку прогресију дужине k.*

Можда многи читаоци све до сада нису били претерано импресионирани овим теоремама. Међутим, ево одмах *спектакла*! Шта ако је A подскуп *простих бројева* у интервалу $[1, N]$? Да ли у скупу простих бројева можемо да нађемо k−точлане аритметичке прогресије? Ово питање одједном поред комбинаторног, добија и аритметички карактер, тј. може се сада видети и као питање у *теорији бројева*! Читалац свакако зна да су нека (многа) питања о простим бројевима међу најтежим отвореним (недоказаним) проблемима читаве математике! Одједном, цела описана машинерија не делује узалуд!

[4] Endre Szemerédi (1940–), мађарски математичар, добитник Абелове награде 2012.

Да ли Ротову теорему можемо да искористимо да нађемо аритметичке прогресије дужине 3, чија су сва 3 члана прости бројеви? Нажалост не! На основу тзв. *теореме о простим бројевима*, у интервалу $[1, N]$ се налази око $\frac{N}{\log N}$ простих бројева, па је густина подскупа простих бројева $\frac{1}{\log N}$, што је много мање од $\frac{C}{\log\log N}$, а што је неопходно за Ротову теорему. Дакле, овај адитивни проблем у скупу простих бројева је још (много) тежи. Међутим, ван дер Корпут[5] је још 1939. године доказао да прости бројеви садрже бесконачно много трочланих аритметичких прогресија.

Ван дер Корпутова теорема. *У скупу A простих бројева из интервала $[1, N]$ постоји бар $C\frac{N^2}{(\log N)^3}$ различитих нетривијалних аритметичких прогресија дужине 3, где је $C > 0$ нека апсолутна константа.*

Ван дер Корпутов доказ је, дакле, старији од Ротове теореме и користи сасвим другачију технику – ослања се на тзв. *метод Виноградова* из *аналитичке теорије бројева*, који је специјалнији јер се односи баш на просте бројеве, а не на произвољне подскупове $A \subseteq [1, N]$.

Одмах је постављено и питање да ли прости бројеви садрже и бесконачно много аритметичких прогресија дужине k, за било које $k \geqslant 4$. На пример, $5, 11, 17, 23, 29$ је једна аритметичка прогресија простих бројева дужине 5, док је $199, 409, 619, 829, 1039, 1249, 1459, 1669, 1879, 2089$ једна аритметичка прогресија у скупу простих бројева дужине 10. Дакле, питање је да ли има оваквих прогресија произвољно велике дужине k и да ли за свако фиксирано k, има бесконачно много различитих $k-$точланих аритметичких прогресија у простим бројевима?

Ово питање је остало јако дуго отворено, јер ван дер Корпутов метод није могао да се уопшти на случај $k \geqslant 4$, док су прости бројеви сувише ретки (малобројни) да би могла да се примени Семередијева теорема. Али 2004. је доказан следећи резултат.

Тао[6]–Гринова[7] теорема. *За свако $k \geqslant 4$, скуп простих бројева садржи бесконачно много различитих аритметичких прогресија дужине k.*

Тао–Гринов доказ је ингениозна комбинација неколико елемената: Семередијеве теореме, адитивне комбинаторике, хармонијске анализе, ергодичке теорије и аналитичке теорије бројева. Многе идеје доказа се ослањају на рад и идеје неколико великих математичара 20. века, а груба структура доказа се такође, као и доказ Ротове теореме, ослања на *дихотомију* између „случајности и уноформности" са једне и „постојања додатне структуре" са друге стране. У даљем раду, Тао и Грин су дали и асимптотску формулу за укупан број $k-$аритметичких прогресија у простим бројевима у интервалу $[1, N]$, кад $N \to \infty$. За ово је био потребан још много софистицираније математички апарат и развој „*Фуријеове анализе вишег реда*" (Гауерс је први приметио, да „линеарна" Фуријеова анализа која је била довољна за ван дер Корпутов доказ ($k = 3$) није довољна и за $k = 4$ и да нам је за тај случај потребна „квадратна" Фуријеова анализа).

[5] Johannes van der Corput (1890–1975), холандски математичар
[6] Terence Tao (1975–), аустралијски математичар, добитник Филдсове медаље 2006.
[7] Ben Green (1977–), британски математичар

Статијата прв пат е објавена во списанието ТАНГЕНТА на ДМ на Србија во 2013/14 година