POLYGON
VII. kötet 1. szám 1997.

Zárójeles megjegyzések*

CSÁKÁNY BÉLA

A zárójelekről a matematika oktatása során kevés szó esik. Használatukat, amely gyakran magától értetődő, néha a hagyomány, máskor a tudatos célszerűség diktálja. Mindnyájan érezzük, hogy a zárójel nem csupán írásjel, hanem matematikai fogalom is. Mégis, az utóbbiak családjában a zárójelnek Hamupipőke szerepe jut. Ez a dolgozat segíteni kívánja az olvasót, hogy derék királyfi módjára fedezze fel: Hamupipőke szép is, jó is — a zárójelekről is lehet érdekes és tartalmas matematikai megállapításokat tenni. A magasabb analízist vagy algebrát igénylő meggondolásokat — vannak ilyenek! — elkerüljük, ezért csak ízelítőkre szorítkozhatunk. Befejezésként fel fogjuk hívni az Olvasó figyelmét egy, zárójelekkel kapcsolatos, elemi eszközökkel is megközelíthető, még nyitott kérdéskörre.

1. Kellenek-e zárójelek? Az algebra, tartalmát tekintve, *a műveletekkel ellátott halmazok vizsgálata.* Ugyanakkor az algebra, mint tevékenység, *a szimbólumok kezelésének gyakorlata.* Az utóbbi definíciót Garrett Birkhoff (1911—1996), a kiemelkedő amerikai algebrista adta. A *szimbólum* szó jelentése a magyar nyelvben: jelkép. A Birkhoff által használt angol *symbol* szó jelentése ennél általánosabb: bármi, ami másvalami ábrázolására szolgál, vagy képének tekinthető; így pl. bármely betű vagy más szokásos *jel*, amelyet dolgok, mennyiségek, műveletek, függvények stb. megnevezése, megadása céljából használunk. Példaként tekintsünk három egyszerű, különböző nehézségű algebrai tevékenységet:

(I) Tízes számrendszerben megadott természetes számok összeadása:

$$123 + 98 = 221.$$

(II) Idegen ciklusok szorzataként felírt permutációk szorzása:

$$(12)(345) \cdot (13)(245) = (14235).$$

(III) Annak a bizonyítása, hogy négyzetes mátrixnak legfeljebb egy inverze van.

* Ez a cikk az MKM KF 402/96 sz. kutatási program keretében készült. Köszönet illeti Führerné Nagy Györgyit, a soproni Erdészeti és Faipari Egyetem docensét a cikk témájáért, valamint Mark Krusemeyert, a Carleton College (Northfield, Minnesota) tanárát a cikk címének ötletéért.

Az első esetben számjegyekből, azaz bizonyos számok jeleiből álló sorozatokat alakítunk át ugyanilyen sorozatokká. A második példában a permutálandó halmaz elemeinek jeleiből és zárójelekből álló sorozatokat alakítunk át megint csak ugyanilyen jelekből képezett sorozatokká. A harmadik példában azt kell megmutatnunk, hogy az $AA' = A'A = AA'' = A''A = E$ feltevésből következik az $A' = A''$ állítás; tehát mind a feltevés, mind az állítás megadható jelek sorozatával. A bizonyítás pedig abban áll, hogy A'-t alkalmasan kezelve A''-vé alakítjuk át: $A' = A'E = A'(AA'') = (A'A)A'' = EA'' = A''$. Mindhárom esetben úgy „csinálunk algebrát", hogy jeleket kezelünk valamilyen szabályok által megengedett módon. Igaz, algebrai bizonyításokban a jelsorozatok (formulák, egyenletek stb.) mellett gyakran sok a szöveg, de ez nem szükségszerű: csupán a könnyebb megértést szolgálja.

A második és harmadik példában *zárójeleket* is kezelünk; az átalakítások során ezek helye és száma megváltozik. Eszerint az algebra többek között a zárójelek kezelésének mestersége is. A zárójelek azonban példáinkban lényegesen különböznek a többi jelektől, amelyek számot, halmazelemet, mátrixot, műveletet jelölnek, míg a (és) jelek — a *kezdőzárójel* és a *végzárójel* — láthatóan csak azt mondják meg, mely másfajta jeleket kell az átalakítás során összetartozónak tekintenünk. A zárójelek tehát más jelek kezeléséhez használatos segédjelek. Az világos, hogy halmazok elemein végzendő műveletek vizsgálatához — ami az algebra tartalma — szükség van a halmazok, az elemek és a műveletek jeleire. De kellenek-e további segédjelek, vagy csak megszokásból, esetleg a kényelem kedvéért használjuk ezeket?

Figyeljük meg, hogy a (II) és a (III) példában a zárójelek funkciója lényegesen különböző. Szerepük a (II) esetben egyszerűen az *elhatárolás*. Írhatnánk ezt is: 12 345 · 13 245 = 14235; a zárójelek csupán arra szolgálnak, hogy pl. az 1, 2 elemek összetartozását és a 3, 4, 5 elemektől való különállását szembeötlővé tegyék. Ez a szerep hasonló ahhoz, amelyet a zárójelek a nem matematikai szövegekben játszanak: elkülönítik a szöveg egy részét, amely pl. valamilyen szemléltetést vagy magyarázatot tartalmaz (ezt teszik ebben a mondatban is). A (III) bizonyításban azonban a zárójelek nemcsak elválasztják az ottani matematikai objektumokat (a mátrixokat), hanem *a műveletek elvégzésének sorrendjét is kijelölik*. Valóban, az $A'(AA'')$ és az $(A'A)A''$ mátrixszorzatok tényezői ugyanazok, e tényezők sorrendje is megegyezik, de az elsőnél az AA'' szorzást, a másodiknál az $A'A$ szorzást kell először elvégeznünk.

Mindez azonban még mindig nem győz meg bennünket a zárójelek szükségességéről. Hiszen (III)-hoz éppen azt használtuk fel, hogy a mátrixok szorzása asszociatív, ami azt jelenti, hogy három mátrix szorzata — amelyet természetesen csak két lépésben képezhetünk, mert a mátrixszorzás definíciója *két* mátrixra vonatkozik — csupán e mátrixok sorrendjétől függ, de nem függ az elvégzendő két

szorzás sorrendjétől, azaz nem függ a szorzások sorrendjét kijelölő zárójelezéstől. Ha pedig ez így van, akkor a zárójeleket ebben az esetben is elhagyhatnánk, hiszen most is csak a figyelemfelkeltés a szerepük, csak most nem matematikai objektumok (mátrixok) különállására, hanem éppen a szorzás asszociatív voltára hívják fel a figyelmet.

Bármennyire is próbáljuk azonban kisebbíteni a zárójelek jelentőségét, bármennyire világos is, hogy gyakran csak gondolkozásunk kényelmét szolgálják, az bizonyos, hogy $(1+2)\cdot 3 \neq 1+2\cdot 3$, mutatva, hogy a zárójelek elhagyása valamely számot megadó jelsorozatból megváltoztathatja azt a számot. Ezért a címben feltett kérdésre a válasz csak *igen* lehet.

Ez az érv ugyan megcáfolhatatlan, de a következőkben megmutatjuk, hogy érvényessége csak egy esetleges — bár megszokottsága miatt természetesnek látszó — megállapodáson múlik: azon, hogy a közönséges műveletek jelét nem ama számok (mátrixok, vektorok stb.) elé írjuk, amelyeken a műveletet éppen végezzük, nem is utánuk, hanem *közéjük*: $5+8$, $5-8$, $5\cdot 8$, $5/8$ és így tovább. Az egyváltozós függvények esetén a jelölés módja fölöttébb változatos: x^2, $\cos x$, e^x stb. Ezért kialakult az a gyakorlat, hogy közelebbről meg nem határozott egyváltozós függvény értékét jelölve a függvény jelét mindig a változó vagy annak értéke elé írjuk: $f(x)$, $g(5)$ stb. Itt is használunk ugyan zárójelet, ám ezt egyes konkrét esetekben el is hagyhatjuk: $\cos\pi$ helyett nem írunk $\cos(\pi)$-t. Három- vagy többváltozós függvényt jelölve hasonlóképpen járunk el: $f(x,y,z)$, $f(3,5,8)$. Ezt a gyakorlatot követhetjük kétváltozós függvényeknél is, de olyan kétváltozós függvény esetén, amely *művelet*, rendszerint ragaszkodunk a hagyományos jelöléshez, olyannyira, hogy pl. egy közelebbről meg nem határozott műveletet alkalmazva az 5 és 8 számokra, az eredményt a kínálkozó $f(5,8)$ helyett $5\circ 8$ alakban írjuk fel. Itt a karika a meghatározatlan művelet jele, hasonlóan ahhoz, ahogy pl. n gyakran jelöl meghatározatlan természetes számot.

A megszokott kétváltozós műveletek jelölése tehát eltér az általános függvényjelöléstől. Az utóbbit alkalmazva $a-b$ helyett $-(a,b)$-t vagy éppen $-ab$-t kellene írnunk. Ennek azonban több hátránya lenne a szokásos jelöléshez képest, amely megtakarítja a zárójeleket és kizárja a többértelműséget; ha $5-8$ helyett $-(5,8)$-at, vagy -58-at írnánk, aligha a -3 számra gondolnánk. E példát szemügyre véve felvetődik a kérdés: nem szűnnek-e meg ezek a nehézségek, ha nem konkrét számokkal, hanem csupa egyetlen betűből álló jelekkel dolgozunk? Lássunk egy példát! Jelentsen f és g egyváltozós valós függvényt, a és b pedig valós számot. Jelenthet-e egy valós számot az

$$f-agb$$

jelsorozat, s ha igen, megfejthető-e egyértelműen, melyiket? Mivel g egyváltozós függvény jele, és az egyetlen b jel áll utána, gb-nek tulajdoníthatunk jelentést: azt a számot, amelyet közönségesen $g(b)$-vel jelölünk. Mivel a kivonás kétváltozós, és két, külön-külön jelentéssel bíró dolog áll a jele után, a és gb, azért $-agb$-nek is van jelentése: az, amit szokás szerint $a - g(b)$ jelöl. Végül, hasonló meggondolással, $f - agb$ jelentése a szokásos jelöléssel: $f(a - g(b))$. Azt is figyeljük meg, hogy pl. a második lépésben nem volt más választásunk: az agb jelsorozat nem bontható értelmesen ag-re és b-re, mert ag-nek nincs jelentése. Hasonlóképpen az első és a harmadik lépésben is csak egy lehetőségünk volt. Ezzel azt is beláttuk, hogy az $f - agb$ jelsorozat nemcsak értelmes, hanem egyértelmű is.

Ugyancsak van jelentése $-afgb$-nek is, mégpedig $a - f(g(b))$, de $af - gb$-nek már nincs. Általában, ha egy jelsorozat szám — és nem függvény — jelével kezdődik, akkor nincs jelentése, kivéve persze, ha ebből az egyetlen jelből áll.

Előbbi gondolatmenetünk sokkal általánosabban is alkalmazható. Jelentsenek $f_1, f_2, \ldots$ függvényeket, $n_1, n_2, \ldots$ pedig természetes számokat. Legyen f_1 n_1-változós, f_2 n_2-változós függvény és így tovább. Tegyük fel, hogy mindegyik f_i függvény művelet ugyanazon az M halmazon (azaz az M-ből képezett, n_i elemű — szokásos szóhasználattal n_i *hosszúságú* — sorozatok halmazának M-be való leképezése). Az M halmaz tetszőleges elemeiből és az indexes f jelekből képezett véges hosszúságú sorozatokat nevezük *szavaknak*. Előfordulhat, hogy egy szó M-nek egy elemét jelenti. Legyen pl. f_1 az összeadás, f_2 a cos függvény, M a valós számok halmaza; ekkor az $f_1 3 f_2 \pi$ szó éppúgy megfejthető, mint az imént $f - agb$. Valóban, f_1 kétváltozós, az őt követő 3 jel (azaz 1 hosszúságú szó) és az $f_2\pi$ szó megfelelően a 3 illetve a $\cos\pi = -1$ valós számot jelenti; ezért $f_1 3 f_2 \pi$ jelentése $3 + (-1) = 2$. Az olyan szót, amely M valamelyik elemét jelenti, *kifejezésnek* nevezzük, azt az elemet pedig, amelyet jelent, e kifejezés *értékének*. Kifejezésünk felírásakor nem használtunk zárójelet. Erre nincs is szükség, mert

(*) *bármely szónak legfeljebb egy értéke van.*

Tekintsünk ugyanis egy tetszőleges s szót; tegyük fel, hogy s kifejezés. Az s szó hosszúsága szerinti teljes indukcióval megmutatjuk, hogy s-nek egyetlen értéke van. Ha az s szó egyetlen jelből áll, az szükségképpen egy $a \in M$ elem jele, másként s-nek nem lenne értéke. Így viszont s értéke éppen a. Legyen ezután s hosszúsága nagyobb 1-nél, és tegyük fel, hogy az s-nél rövidebb szavaknak legfeljebb egy értéke van. Mivel s kifejezés, azért valamely f_i függvényjellel kezdődik, amelyet n_i számú kifejezés követ:

$$s = f_i k_1 \ldots k_{n_i}. \tag{1}$$

Az s kifejezés értékét úgy kapjuk meg, hogy az f_i műveletet elvégezzük a $k_1 \ldots k_{n_i}$ kifejezések értékein. Elég tehát azt belátnunk, hogy s nem írható fel (1)-től különböző $f_i l_1 \ldots l_{n_i}$ alakban, ahol $l_1, \ldots$ ugyancsak kifejezések. Tegyük fel, hogy s-nek van ilyen alakja is, és legyen r a legkisebb olyan természetes szám, hogy $k_r \neq l_r$.

Akkor k_r és l_r közül az egyik, mondjuk k_r, a másiknak néhány — de nem az összes — jeléből, mégpedig az l_r valamelyik jelétől balra elhelyezkedő összes jelből áll. Ilyenkor azt mondjuk, hogy a k_r szó az l_r szónak *kezdőszelete* (míg a kimaradó jelek l_r-nek *végszeletét* alkotják; ezek a fogalmak egyébként bármilyen véges jelsorozatra alkalmazhatók). Feladatunkat ezzel annak a ténynek a belátására vezettük vissza, hogy kifejezés kezdőszelete nem lehet kifejezés.

Tegyük fel, hogy ez nem igaz; akkor van *legrövidebb* olyan s kifejezés is, amelynek alkalmas t kezdőszelete is kifejezés. Ez az s nem állhat M egyetlen eleméből, így — hogy ne kelljen újabb jelölést bevezetnünk — őt is felvehetjük az (1) alakban. Ekkor $t = f_i m_1 \ldots m_{n_i}$, és mivel $t \neq s$, létezik legkisebb olyan q természetes szám, hogy $k_q \neq m_q$. Mint gondolatmenetünk első felében, most is látjuk, hogy a k_q és m_q kifejezések egyike kezdőszelete a másiknak. Ám s hosszabb k_q és m_q bármelyikénél, találtunk tehát a legrövidebbnél is rövidebb olyan kifejezést, amelynek egy kezdőszelete is kifejezés. Ebből az ellentmondásból következik, hogy állításunk, s vele (*) is, igaz.

Az elvégzett meggondolás azt mutatja, hogy bizonyos, eléggé természetes és csak a műveletek eredményének jelölésére vonatkozó megállapodás mellett a műveletek elvégzése sorrendjének kijelöléséhez a zárójelek valóban fölöslegesek. Abban kell csupán megállapodnunk, hogy minden művelet jelét mindig azoknak az elemeknek vagy elemet jelölő kifejezéseknek az elejére írjuk, amelyeken a műveletet el kell végezni. Ezt a jelölési módot Jan Łukasiewicz lengyel logikus vezette be, ezért rendszerint *lengyel prefix jelölésnek* nevezzük. A „prefix" szó arra utal, hogy a műveletjelek állnak elől. Mindezt hasonló módon végiggondolhatnánk arra az esetre is, amikor a műveletjeleket az elemek jelei után írjuk — az a „postfix" jelölés. Ilyen jelölések alkalmazásakor természetes követelmény, hogy legyen előre megadva a műveletek (jeleinek) változószáma, továbbá az elemek jelei egyértelműen váljanak el egymástól. Az utóbbi követelményt azonban nehezen teljesíthetjük, ha számokkal dolgozunk, amelyeket számjegyek sorozataival kell megadnunk: $+121$ jelenthetné $12 + 1$-et és $1 + 21$-et is. Ha a többértelműséget el akarjuk kerülni, akkor szétválasztó jeleket kell használnunk (pl. zárójeleket...). Kis elemszámú halmazok elemeit viszont jelölhetjük úgy, hogy szétválasztó jelekre ne legyen szükség. Ha pl. $M = \{\uparrow, \downarrow\}$ az „igaz" és „hamis" logikai értékek halmaza, $f_1, f_2, \ldots$ pedig logikai függvények, akkor a lengyel prefix jelölés elvi nehézség nélkül használható. Annak, hogy még ilyenkor is ritkán használjuk, gyakorlati oka van: kisiskolás korunktól hozzászoktunk a függvények változatos jelöléséhez, többek között ahhoz is, hogy a kétváltozós és a nem kétváltozós műveleteket másképp jelöljük, nem is beszélve a hatványfüggvény (pl. x^2) vagy akár az osztás (pl. $\frac{1}{2}$) rendhagyó jelöléséről. Mire megismerjük azt a tényt, hogy a műveleteket lehetne egységes módon és zárójelmentesen is írni, már késő: akkorra úgy megbarátkozunk a jelölések sokféleségével, hogy eszünkbe sem jut megválni tőle. Nem utolsó szempont az sem, hogy bármilyen matematikai szöveg írásakor (zárthelyi dolgozattól a tudományos monográfiáig) a potenciális olvasóval szembeni elemi szakmai szolidaritás azt kívánja, hogy szövegünket, képleteinket tegyük minél könnyebben érthetővé. Ezért a lengyel jelölést csak egyes kutatási és didaktikai meggondolásokban alkalmazzuk.

Az a tény, hogy alkalmas jelölési rendszert használva zárójelekre nincs szükség, semmiképpen sem törvényesíti az olyan, elsőéves hallgatók által olykor írásban elkövetett merényleteket, mint pl.[1]

$$z = \cos -\pi - {}^{-}k\pi/2 + i \sin -\pi - {}^{-}k\pi/2.$$

Mint didaktikai ötlet, a negatív számokkal való ismerkedéskor átmeneti használatra elfogadható az ilyen jelölés:

$$-3 - {}^{-}4 = {}^{+}1$$

(lásd [1], 15. oldal), de az senkitől sem várható el, hogy helyesen értelmezze a fentebbihez hasonló többértelmű vagy értelmetlen képleteket.

A zárójelektől tehát a hétköznapi matematikai gyakorlatban nem kell félnünk és nem lehet megválnunk. Ugyanakkor, takarékoskodni a zárójelekkel — „oly cél, minőt óhajthat a kegyes". A gyakorlatban ezt úgy érjük el, hogy rögzítjük a műveletek elvégzésének, valamint a gyakran használt függvények alkalmazásának sorrendjét, s csak akkor használunk zárójeleket, ha ettől a sorrendtől eltérünk, de a zárójeleken belül ekkor is a rögzített sorrendet vesszük figyelembe. Ilyen sorrend rögzítését precedencia- vagy prioritási szabálynak nevezzük. Bolygónk iskoláiban a következő precedenciaszabályt[2] tanítjuk:

1. Hatványozás
2. Szorzás
3. Függvényértékképzés
4. Összeadás, kivonás

Ennek alapján világos, hogy mit jelent $\cos xy + z^k$. Megfelelő zárójelezésekkel felírhatjuk az összes olyan négyváltozós függvényt, amelyek ugyanezeket a függvény- és műveletjeleket[3], valamint ugyanezeket a változójeleket tartalmazzák, ugyanilyen sorrendben felírva, csak éppen a cos függvényt és a három műveletet kell más sorrendben alkalmazni:

$$\cos(xy + z^k), (\cos x)y + z^k, (\cos x)(y + z^k), \cos x(y + z^k), \ldots$$

[1] Führerné Nagy Györgyi közlése.

[2] Elképzelhető olyan kultúra, amelyben az összeadást végzik el először, például azért, mert azt a legkönnyebb; tehát $a(b + c)$ helyett $ab + c$-t, utóbbi helyett meg $(ab) + c$-t írnak.

[3] A hatványozást itt kétváltozós műveletnek tekintjük a z és k változókkal.

(Hány van belőlük, és miért éppen ezt a néhányat írtuk fel?)

Elvárnánk, hogy az osztásnak is legyen egyértelműen meghatározott helye a precedencialistán, de ez nem így van. Ha az osztást jelző törtvonalat ferdén írjuk, akkor az osztás helye az összeadás előtt van; ha viszont vízszintes törtvonalat használunk, akkor az osztás az utolsó.

Eddig *különböző műveletek* sorrendjéről volt szó. Ha ugyanazt a műveletet kell alkalmaznunk többször, akkor a következő módon takarékoskodhatunk a zárójelekkel. Megállapodunk abban, hogy a műveleteket balról jobbra haladva végezzük, tehát pl. $(a \circ b) \circ c$ helyett egyszerűen $a \circ b \circ c$-t írunk; csak akkor írunk ki zárójelet, ha ettől a szabálytól el kell térnünk. Példa: $(a \circ b) \circ ((c \circ d) \circ e)$ helyett elegendő $a \circ b \circ (c \circ d \circ e)$. A közönséges összeadás asszociatív: többtagú összeget fölösleges zárójelezni; mindamellett gépiesen alkalmazzuk a „balról-jobbra" szabályt. Ugyanígy végezzük a kivonást is (amely nem asszociatív):

$$a - b - c = (a - b) - c \neq a - (b - c).$$

Az összeadás és a kivonás holtversenyben állnak a precedencialistán. Ezért a „váltakozó előjelű összeget" (amelyben tulajdonképpen kivonás és összeadás váltakozik) ugyancsak a „balról-jobbra" szabállyal számítjuk ki:

$$a - b + c = (a - b) + c \neq a - (b + c).$$

A „balról-jobbra" szabályt azonban nem használjuk a hatványozásra: 3^{3^3} nem a $(3^3)^3$ hatványt jelenti — amelyet inkább $3^{3\cdot 3}$ alakban írunk fel —, hanem a $3^{(3^3)}$ hatványt. A három számjegyből felírható legnagyobb szám (régi iskolai feladat!) tehát 9^{9^9}.

2. Valódi szorzatok és szabályos zárójelsorozatok. Jelölje $a \circ b$ az a és b valós számok számtani közepét. Ez a $\circ$ művelet nem asszociatív: $(1 \circ 0) \circ 0 = \frac{1}{4} \neq$ $\neq 1 \circ (0 \circ 0) = \frac{1}{2}$. Eszerint $(a \circ b) \circ c$ és $a \circ (b \circ c)$ különböző *háromváltozós műveletek* a valós számok halmazán. Hasonlóképpen

$$(2) \quad ((a \circ b) \circ c) \circ d, \quad (a \circ (b \circ c)) \circ d, \quad (a \circ b) \circ (c \circ d), \quad a \circ ((b \circ c) \circ d), \quad a \circ (b \circ (c \circ d))$$

négyváltozós műveletek. Alkalmas helyettesítésekkel beláthatjuk, hogy ezek a műveletek is mind különbözők.

Felmerül a kérdés: ilyen úton hány különböző n-változós műveletet származtathatunk a $\circ$ műveletből? A módszer láthatóan abban áll, hogy egy $a \circ b \circ c \circ \ldots$ véges jelsorozatban zárójeleket helyezünk el *értelmes módon*, azaz úgy, hogy ezáltal megadjuk a $\circ$ jellel jelölt kétváltozós műveletnek az $a, b, c, \ldots$ elemeken való elvégzése egy lehetséges — bár nem feltétlenül egyértelműen meghatározott[4] — sorrendjét, s eközben feleslegesen nem használunk zárójelet.

A felmerült kérdést a következő részekre bontjuk: az $a \circ b \circ c \circ \ldots$ véges sorozat értelmes módon végzett zárójelezései

(a) hogyan ismerhetők fel,
(b) hogyan írhatók fel,
(c) hányan vannak,

végül pedig

(d) csupa különböző műveletet határoznak-e meg?

Vegyük észre, hogy az első három kérdés vizsgálatához szükségtelen a (2) jelsorozatban különböző $a, b, c, \ldots$ jeleket használnunk: pl. $(a \circ (b \circ c)) \circ d$ helyett tekinthetjük az $(a \circ (a \circ a)) \circ a$ jelsorozatot — ez a zárójelezés értelmességén nem változtat. Egyszerűség kedvéért az éppen vizsgált kétváltozós műveletet mindig szorzásnak mondjuk és így is jelöljük. Így (2) helyett ezt írhatjuk:

(2′) $$((aa)a)a, \quad (a(aa))a, \quad (aa)(aa), \quad a((aa)a), \quad a(a(aa)).$$

Az olyan jelsorozatot, amelyben n számú a jel és — bármilyen elrendezésben — zárójelek állnak, *n tényezős szorzatnak* nevezzük. Ebben az értelemben szorzat tehát $))a(a)()$ és aaa is [!]. Az olyan szorzatot, amelyben a zárójelek értelmes módon vannak elhelyezve, *valódi szorzatnak*, a nem valódi szorzatot pedig *álszorzatnak* nevezzük.

(a) Négy-öt tényezős szorzatról ránézésre eldönthető, hogy jól zárójelezett-e. Ám pl. a

(3) $$((a((a((aa)(aa)))a)(aa)))a$$

szorzatnál a ránézés nem segít. Minden legalább háromtényezős szorzatra használható azonban egy olyan algoritmus, amelynek lényege, hogy megpróbáljuk kiszámítani a szorzatot, s ha valamely ponton ehhez túl sok vagy túl kevés a zárójel, akkor megállapítjuk, hogy a szorzat nem valódi. Ezt az eljárást *természetes algoritmusnak* nevezzük. A természetes algoritmus a következő lépésekből áll:

1. Megkeressük szorzatunkban azokat a zárójelpárokat, amelyek egy olyan kezdő- és egy olyan végzárójelből állnak, hogy ezek között egyetlen zárójel sincs.

Ezeket *elsőrendű pároknak* nevezzük. Először mindig azokat a szorzásokat kell elvégezni, amelyeket elsőrendű zárójelpárok jelölnek ki.

2. Ha szorzatunkban nincs elsőrendű pár, akkor benne kezdőzárójel nem előz meg végzárójelet, így a zárójelek semmiféle szorzások sorrendjét nem jelölhetik ki.

[4] Gondoljunk arra, hogy az $(a \circ b) \circ (c \circ d)$ négyváltozós művelet elvégzésekor mindegy, hogy $a \circ b$-t vagy $c \circ d$-t számítjuk ki először.

Ha van olyan elsőrendű pár, amelyben egyetlen a áll, úgy ez a pár fölösleges. Ha van olyan elsőrendű pár, amelyben legalább három a áll, akkor ezen a-k szorzásának sorrendjét a zárójelezés nem jelöli ki. A szorzat egyik esetben sem valódi.

3. Ha minden elsőrendű párban két a áll, akkor a szorzat elsőrendű párok által határolt, (aa) alakú részletszorzatainak mindegyikét helyettesítsük egyetlen a-val.

Pl. (3)-ból ekkor a

(3_1) $$((a((a(aa))a)a))a$$

szorzatot kapjuk.

4. A kapott szorzatra alkalmazzuk ismét az 1. és a 2. lépést.

Most az 1. lépéssel talált zárójelpárokat *másodrendűnek* nevezzük.

5. Ha a 2. lépés mutatja, hogy kapott szorzatunk nem valódi, akkor eredeti szorzatunk sem valódi.
6. Ha az ötödik lépés nem dönti el, hogy a szorzat valódi-e, alkalmazzuk ismét sorban a 3., 4., 5. és 6. lépést.

Eközben megtaláljuk a harmadrendű stb. zárójelpárokat. Mivel a 3. lépés csökkenti a szorzat tényezőinek a számát, a természetes algoritmus vagy megmutatja, hogy szorzatunk nem valódi, vagy kéttényezős szorzatot szolgáltat. Eredeti szorzatunk akkor és csak akkor valódi, ha ez a kéttényezős szorzat is valódi, ami pontosan akkor igaz, ha aa alakú.

Algoritmusunk (3)-ból (3_1) után a

(3_2) $$((a((aa)a)a))a$$
(3_3) $$((a(aa)a))a$$
(3_4) $$((aaa))a$$

szorzatokat származtatja. Mivel (3_4) nem valódi, (3) sem az.

Tekintsük zárójelek egy (véges) sorozatát, pl. a $(\,) \,) \, (\, (\,) \,)$ sorozatot. Ha egy szorzatban ez a zárójelsorozat áll (az elemek jeleinek elhagyása után ez marad vissza), a szorzat nem lehet valódi, hiszen sorozatunkban három kezdőzárójel és négy végzárójel van, pedig a valódi szorzatokban a zárójelek párosával szerepelnek. Hagyjuk el az utolsó zárójelet; most már egyenlő a kezdő- és a végzárójelek száma, az így kapott zárójelsorozat mégsem állhat valódi szorzatban. Vegyük ugyanis a 3 hosszúságú kezdőszeletét: $(\,) \,)$. Tegyük fel, hogy ezek egy szorzat első három zárójelét képezik. Alkalmazzuk erre a szorzatra a természetes algoritmust. A 3. lépés mindig egy vagy több, szomszédos kezdőzárójelből és végzárójelből álló párt tüntet el, ezért az utolsó lépésben olyan kéttényezős szorzat marad, amelyben az első zárójel végzárójel, ezért tehát ez a szorzat — s vele együtt az eredeti szorzat — álszorzat. Láthatóan ez azon múlott, hogy a zárójelsorozatban volt olyan kezdőszelet, amelyben több a végzárójel, mint a kezdőzárójel. Ezek a megfigyelések a következő definíciót sugallják. Nevezzünk egy zárójelsorozatot *szabályosnak*, ha benne a kezdő- és a végzárójelek száma egyenlő, és minden kezdőszeletében legalább annyi kezdőzárójel van, mint végzárójel. Ezután megfigyelésünket — amellyel az álszorzatok gyakran egyszerűen leleplezhetők — így mondhatjuk ki: valódi szorzatban a zárójelek sorozata szabályos. Ennek a megfordítása sajnálatos módon nem igaz: aaa és $((aa))$ egyike sem valódi, pedig zárójelsorozatuk (az üres sorozat is!) szabályos.

(b) Minden valódi n tényezős szorzat $n-2$ zárójelpárt tartalmaz, mert a természetes algoritmus végén belőle az aa szorzat marad, és valahányszor az algoritmus során a 3. lépést alkalmazva egy tényező eltűnik, mindannyiszor egy zárójelpár is eltűnik. Tekintsünk egy S valódi n tényezős szorzatot (legyen $n > 2$), és vegyük a benne levő legmagasabb rendű zárójelpárokat. Ezek a természetes algoritmus során jelennek meg; számuk 1 vagy 2. Kettőnél több nem lehet belőlük, mert akkor az algoritmus befejező lépésében elhagyva e zárójelpárokat, kettőnél több a jelből álló sorozatot kapnánk, ellentmondásban azzal, hogy S valódi.

Eszerint S szükségképpen vagy $a(P)$, vagy $(Q)(R)$, vagy $(P)a$ alakú, ahol csak a legmagasabb rendű zárójelpárokat tüntettük fel. Itt P, Q és R maguk is valódi szorzatok, mert belőlük az S-re alkalmazott természetes algoritmus utolsó előtti lépésében aa keletkezik, különben az algoritmus azt mutatná, hogy S nem valódi. Azt is látjuk, hogy P $n-1$ tényezős, Q és R pedig k és l tényezősek, ahol $k, l \geq 2$, és $k + l = n$.

Ennek alapján bármely n-re fel tudjuk írni az összes n tényezős, jól zárójelezett szorzatokat. A továbbiakban tekintsük az egyetlen jelből álló a sorozatot is egytényezős valódi szorzatnak. Láttuk, hogy aa az egyetlen kéttényezős valódi szorzat. Ha $n \geq 2$, és már ismerjük az összes, n-nél kevesebb tényezős valódi szorzatot, akkor minden n tényezős valódi szorzatot felírhatunk, mégpedig egyetlen módon, $a(P)$, $(Q)(R)$ vagy $(P)a$ alakban, ahol P $n-1$ tényezős, Q és R pedig legalább kéttényezősek, és tényezőik számának összege n. Figyeljük meg, hogy pontosan ilyen módon írtuk fel (2′)-ben az összes négytényezős valódi szorzatot!

3. A valódi szorzatok száma. A (c) kérdésre adandó választ egy matematikatörténeti érdekességű gondolatmenettel sejtjük meg, majd elemi kombinatorikai — úgyszólván eszköztelen — módszerrel igazoljuk.

Az előző meggondolás segítségével rekurzív összefüggést tudunk felírni az n tényezős valódi szorzatok számára. Jelöljük ezt a számot Z_n-nel. Mint láttuk, $Z_1 = Z_2 = 1$, továbbá $n > 2$-re

$$Z_n = Z_1 Z_{n-1} + Z_2 Z_{n-2} + \ldots + Z_{n-1} Z_1, \tag{4}$$

mivel $Z_1 Z_{n-1}$ az összes, különböző $a(P)$ alakú n tényezős valódi szorzatok száma, $Z_2 Z_{n-2}$ az összes különböző, $(Q)(R)$ alakú n tényezős valódi szorzatok közül azoknak a száma, amelyekben Q kéttényezős (s ezért R $n-2$ tényezős) és így tovább.

Innen $Z_3 = 1 \cdot 1 + 1 \cdot 1 = 2$, $Z_4 = 1 \cdot 2 + 1 \cdot 1 + 2 \cdot 1 = 5$, és hasonlóképpen $Z_5 = 14$, $Z_6 = 42$ stb.

Az $1, 2, 5, 14, 42, \ldots$ számsorozatot már Euler említette egyik, 1760-ban Szentpétervárott megjelent munkájában. Ebben barátjának, a Pozsonyból Göttingába került magyar származású orvos-fizikus Segner Jánosnak — a róla elnevezett öntözőkerék feltalálójának — az ötletét kidolgozva, a szabályos n-szög átlókkal történő háromszögekre bontásainak számát határozta meg. Később, 1838-ban és 1839-ben Eugène Catalan öt cikket szentelt e számok vizsgálatának. Innen ered elnevezésük: *Catalan-számok*. Kortársával, a Fibonacci-számok képletének névadójával, Jacques Binet-vel együtt, a Moivre-tól eredő generátorfüggvény-módszert alkalmazva egyszerű képletet vezettek le a Catalan-számokra. Gondolatmenetük a következő:

Tegyük fel, hogy létezik olyan $g(x)$ valós függvény, amely hatványsorba fejthető, és hatványsorának együtthatói éppen a $Z_0, Z_1, Z_2, \ldots Z_n, \ldots$ számok:

$$g(x) = Z_0 + Z_1 x + Z_2 x^2 + \ldots + Z_n x^n + \ldots \tag{5}$$

Az ilyen függvényt a $Z_0, Z_1, Z_2, \ldots Z_n, \ldots$ számsorozat — esetünkben a kézenfekvő $Z_0 = 0$ megállapodással megtoldott Catalan-számok — *generátorfüggvényének* nevezzük. Felhasználva a (4) összefüggést, nyerjük:

$$\begin{aligned}(g(x))^2 &= Z_1 Z_1 x^2 + (Z_1 Z_2 + Z_2 Z_1) x^3 + (Z_1 Z_3 + Z_2 Z_2 + Z_3 Z_1) x^4 + \ldots = \\ &= Z_2 x^2 + Z_3 x^3 + Z_4 x^4 + \ldots = g(x) - x.\end{aligned}$$

Másodfokú egyenletet kaptunk $g(x)$-re, amelynek megoldása:

$$g(x) = \frac{1 - \sqrt{1 - 4x}}{2}.$$

Mivel $g(0) = 0$, a négyzetgyökmennyiség negatív értékét kellett vennünk. A $g(x)$ függvényt most binomiális sorba fejthetjük:

$$\begin{aligned}
g(x) &= \frac{1}{2} - \frac{1}{2}(1+(-4x))^{\frac{1}{2}} \\
&= \frac{1}{2} - \frac{1}{2}\left(1+\frac{1}{2}(-4x)+\frac{\frac{1}{2}(-\frac{1}{2})}{1\cdot 2}(-4x)^2+\ldots+\frac{\frac{1}{2}(-\frac{1}{2})\ldots(-\frac{2n-3}{2})}{1\cdot 2\ldots n}(-4x)^n+\ldots\right) \\
&= x + x^2 + \ldots + (-1)(-1)^{n-1}(-1)^n\frac{4^n}{2^{n+1}}\frac{1\cdot 3\ldots(2n-3)}{1\cdot 2\ldots n}x^n + \ldots
\end{aligned}$$

Átalakítás után x^n együtthatója:

$$2^{n-1}\frac{1\cdot 3\ldots(2n-3)}{1\cdot 2\ldots n} = \frac{(2n-2)!}{n!(n-1)!} = \frac{1}{2(2n-1)}\binom{2n}{n}.$$

Ezért minden n-re

(6)
$$Z_n = \frac{1}{2(2n-1)}\binom{2n}{n}.$$

Az $n+1$-edik Catalan-szám (6)-ból

(7)
$$Z_{n+1} = \frac{1}{n+1}\binom{2n}{n}.$$

Minthogy itt a jobb oldal kicsit egyszerűbb, szokás Z_{n+1}-et nevezni az n-edik Catalan-számnak, és Z_{n+1} helyett C_n-t írni. A (6) és (7) képletek mutatják, hogy minden n-re

$$Z_{n+1} = \frac{2(2n-1)}{n+1}Z_n,$$

ami (4)-nél sokkal egyszerűbb rekurzív összefüggés, emellett azt is látjuk belőle, hogy a szomszédos Catalan-számok hányadosa 4-hez tart.

A klasszikus gondolatmenettel kapcsolatban két kifogás is felvetődhet. Egyrészt nem tudjuk, létezik-e a $g(x)$ generátorfüggvény. Ez az aggály gyakran felmerül, amikor rekurzív módon megadott számsorozatra generátorfüggvény-módszerrel találunk képletet. Valóban nem bizonyítottuk be, hogy (6) következik (4)-ből és a $Z_1 = Z_2 = 1$ feltételekből, de a (6) összefüggést — aminek megsejtéséhez a fenti meggondolás vezetett el — most már teljes indukcióval is igazolhatnánk. Ehelyett azonban (7)-re mutatunk be olyan bizonyítást, amely válaszol a lehetséges másik kifogásra is: a fenti gondolatmenet titokzatos, nem derül ki belőle, hogyan kerül a számelméletben nemegyszer felhasznált $\binom{2n}{n}$ binomiális együttható a kép(let)be.

A következő bizonyítás Dov Tamari 1962-ben közzétett gondolatmenetét követi. Jelölje W_n az n kezdő- és n végzárójelből felírható szabályos zárójelsorozatok számát. Mivel ezekből a zárójelekből összesen $\binom{2n}{n}$ különböző sorozat képezhető, a belőlük felírható *nem szabályos* zárójelsorozatok száma $\binom{2n}{n} - W_n$. A (7) képlet igazolásához elegendő a következő két tényt belátnunk:

A. Minden n-re $Z_{n+1} = W_n$.

B. Az n kezdő- és n végzárójelből felírható, nem szabályos sorozatok száma $\binom{2n}{n-1}$.

Mivel minden n-re $\binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1}\binom{2n}{n}$, az A és B állításokból valóban következik (7).

Az A állítás bizonyításához kölcsönösen egyértelmű megfeleltetést adunk meg az $n+1$ tényezős valódi szorzatok halmaza és az n zárójelpárból képezhető szabályos sorozatok halmaza között. Hogy eközben a valódi szorzatokban álló és a szabályos sorozatokat alkotó zárójeleket egymástól megkülönböztessük, az utóbbiakat szögletes zárójelként fogjuk írni. Tekintsünk egy $n+1$ tényezős valódi szorzatot; jelölje ezt S. Az S-ben levő zárójelek n számú elvégzendő szorzást jelölnek ki. Minden ilyen szorzás első tényezőjét tegyük szögletes zárójelek közé. Ezután az így előálló jelsorozatból hagyjuk el az eredetileg benne levő jeleket, tehát az elemek jeleit és a kerek zárójeleket. Feleltessük meg az S szorzatnak a visszamaradó zárójelsorozatot, amely — képzési módja miatt — n zárójelpárból álló szabályos sorozat. Példák:

$$\begin{aligned}
(aa)(aa) &\mapsto [([a]a)]([a]a) \mapsto [\,[\,]\,]\,[\,] \\
a((a(aa))a) &\mapsto [a]([([a]([a]a))]a) \mapsto [\,]\,[\,[\,]\,[\,]\,]
\end{aligned}$$

Ezt a megfeleltetést jelöljük τ-val. Formálisan τ a következő szabály alakjában írható fel:

$$\tau : \begin{cases} a \mapsto \emptyset \\ aa \mapsto [\,] \\ a(P) \mapsto [\,]\,\tau P \\ (Q)(R) \mapsto [\tau Q]\,\tau R \\ (P)a \mapsto [\tau P] \end{cases}$$

Az összes valódi szorzat felírási módja — amelyet a (b) kérdésre válaszolva ismertünk meg — mutatja, hogy ez a szabály minden S valódi szorzatra egyértelműen megadja a hozzárendelt τS szabályos zárójelsorozatot.

Belátjuk, hogy τ-nak létezik inverz leképezése, azaz olyan σ leképezés, amely minden n párból képezett Γ szabályos zárójelsorozathoz egy $\sigma\Gamma$ $n+1$ tényezős valódi szorzatot rendel úgy, hogy mindig $\tau\,\sigma\,\Gamma = \Gamma$, és minden S valódi szorzatra $\sigma\,\tau\,S = S$. A σ leképezést is algoritmussal írhatjuk le, amely megmondja, hogy adott Γ szabályos sorozatba hogyan kell (kerek) zárójeleket és elemek jeleit beírni, hogy ezután az eredetileg benne levő szögletes zárójeleket elhagyva a kívánt $\sigma\Gamma$ valódi szorzatot kapjuk. Ezt az algoritmust ugyancsak formális szabályba foglaljuk.

Előkészítésül vegyük észre, hogy ha egy Λ szabályos zárójelsorozatnak egyetlen kezdőszelete sem szabályos, akkor Λ szükségképpen $[\Gamma]$ alakú, ahol Γ is szabályos; különben pedig van legrövidebb szabályos kezdőszelete, és ekkor *ez* lesz $[\Gamma]$ alakú, ahol Γ szabályos, ráadásul a $[\Gamma]$-hoz tartozó Δ végszelet is szabályos lesz. Eszerint minden nem üres Λ szabályos zárójelsorozat vagy $[\Gamma]$, vagy $[\Gamma]\,\Delta$ alakú, ahol Γ is, Δ is szabályos, és mindkettő rövidebb, mint Λ. Ezek alapján felírhatjuk a σ leképezést definiáló szabályt:

$$\sigma:\begin{cases} \emptyset \mapsto a \\ [\,] \mapsto aa \\ [\,]\,\Delta \mapsto a(\sigma\Delta) \\ [\,\Gamma\,]\,\Delta \mapsto (\sigma\Gamma)(\sigma\Delta) \\ [\,\Gamma\,] \mapsto (\sigma\Gamma)a \end{cases}$$

minden nem üres Γ és Δ zárójelsorozatra. Próbáljuk ki a fenti példákon, hogy τ és σ valóban „visszacsinálják” egymást! Ez általánosan is belátható a szorzatok tényezőszáma, ill. a sorozatok hosszúsága szerinti teljes indukcióval, a szabályokban felsorolt eseteket külön-külön megvizsgálva, például így:

$$(Q)(R) \overset{\tau}{\mapsto} [\tau\,Q]\,\tau R \overset{\sigma}{\mapsto} (\sigma\,\tau\,Q)(\sigma\,\tau\,R) = (Q)(R),$$
$$[\Gamma] \overset{\sigma}{\mapsto} (\sigma\,\Gamma)a \overset{\tau}{\mapsto} [\tau\,\sigma\,\Gamma] = [\Gamma].$$

A befejező egyenlőségek indukciófeltevésből következnek, mert pl. Γ rövidebb, mint $[\Gamma]$. Ezzel az A állítást beláttuk.

Ahhoz, hogy a B állítást igazoljuk, vegyük észre, hogy $n+1$ kezdőzárójelből és $n-1$ végzárójelből $\binom{2n}{n-1}$ sorozat képezhető. Elegendő tehát kölcsönösen egyértelmű megfeleltetést találnunk az *ilyen* sorozatok halmaza (jelölje $H_{n+1,n-1}$) és az n zárójelpárból felírható *nem szabályos* sorozatok halmaza (jelölje I_{2n}) között.

Az összes valódi szorzat felírása érdekében bevezettük a szorzat elsőrendű, másodrendű stb. zárójelpárjait. Ezt hasonlóan megtehetjük zárójelsorozatokra is: az egymás utáni kezdőzárójel-végzárójel párok elsőrendűek. Nevezzük az ilyen párok törlését *redukálásnak*. A redukálás után megmaradó sorozatban elsőrendűvé váló párok lesznek az eredeti sorozat másodrendű párjai és így tovább. Ha egy sorozat szabályos, akkor az a sorozat is szabályos lesz, amely belőle redukálással keletkezik (miért?). Ezért a redukálást folytatva a szabályos sorozat mindig teljesen elfogy. Ha viszont egy n zárójelpárból álló sorozat nem szabályos, akkor

van benne olyan végzárójel, amelyet legfeljebb annyi kezdőzárójel előz meg, mint végzárójel, s az ilyen végzárójelek közül a balról legelső sohasem eshet ki a redukálások során. Nem szabályos jelsorozatból tehát nem juthatunk üres sorozathoz egymás utáni redukálásokkal. Tovább nem redukálható sorozathoz azonban el kell jutnunk, s az csak $)\ldots)(\ldots($ alakú — röviden: $)^k(^k$ alakú — lehet, azaz benne egy vagy több végzárójel alkot kezdőszeletet, amelyet ugyanannyi kezdőzárójel követ. Az így kapott sorozatot nevezzük a kiindulási sorozat *maradványának*. Ezt az eljárást elvégezhetjük olyan zárójelsorozatokon is, amelyekben a kezdő- és a végzárójelek száma nem egyenlő. Speciálisan, ha egy zárójelsorozatban kettővel több kezdőzárójel van, mint végzárójel, akkor ez maradványára is igaz marad, tehát a maradvány $)^{k-2}(^k$ $(k \geq 2)$ alakú lesz.

Cseréljük ki minden n zárójelpárból álló nem szabályos sorozatban a maradvány *utolsó* végzárójelét kezdőzárójelre. Ez az átalakítás I_{2n} leképezése $H_{n+1,n-1}$-be. Ennek a leképezésnek van inverze, mégpedig: minden $n+1$ kezdő- és $n-1$ végzárójelet tartalmazó sorozat maradványa *első* kezdőzárójelének végzárójelre változtatása. Egy-egy példa a két leképezésre:

$$[\,[\,]\,[\,]\,]\,\underline{]}\,\underline{]}\,[\,]\,[\,[\,]\,[\,]\,]\,[\,]\,\underline{\underline{]}}\,\underline{[}\,\underline{[}\,\underline{[}\,[\,] \longmapsto [\,[\,]\,[\,]\,]\,\underline{]}\,\underline{]}\,[\,]\,[\,[\,]\,[\,]\,]\,[\,]\,\underline{\underline{[}}\,\underline{[}\,\underline{[}\,\underline{[}\,[\,]$$
$$\underline{]}\,[\,[\,]\,]\,\underline{\underline{]}}\,\underline{[}\,\underline{[} \longleftarrow \underline{]}\,[\,[\,]\,]\,\underline{\underline{[}}\,\underline{[}\,\underline{[}$$

Az aláhúzott zárójelek alkotják a maradványt; a kétszer aláhúzott zárójeleket kell megváltoztatni. Nyertük, hogy I_{2n}-nek létezik kölcsönösen egyértelmű leképezése $H_{n+1,n-1}$-re, és éppen ezt kellett bizonyítanunk. Ezzel igazoltuk a (7) képletet; beláttuk tehát, hogy a rögzített tényezőszámú valódi szorzatok számát a Catalan-számok adják meg.

Tamari a fenti elemi bizonyítást absztrakt algebrai nyelvezetet használva írta meg [4] cikkében, ezért a kombinatorikusok elsiklottak felette. Richard Guy, a kombinatorika egyik „nagy öregje”, néhány éve még nyilvánosan kételkedett abban, hogy közvetlen kapcsolat is található a valódi zárójelezések és a szabályos zárójelsorozatok között [2]. Ezen felbuzdulva Mark Krusemeyer a múlt évben újra felfedezte Tamari módszerét az A állítás bizonyítására [3].

4. Az értelmes zárójelezéssel nyerhető műveletek száma. Térjünk vissza az $a \circ b \circ c \circ \ldots$ jelsorozathoz, és írjuk a jobban kezelhető

$$a_1 \circ a_2 \circ \ldots \circ a_n \tag{8}$$

alakban. Válaszolunk a (d) kérdésre: ha $\circ$ a számtani közép jele, hány *különböző* n-változós műveletet kaphatunk (8)-ból értelmes zárójelezésekkel? Világos, hogy *legfeljebb* annyit, ahány értelmes zárójelezés van, ezek száma pedig az n-tényezős valódi szorzatok száma: a Z_n Catalan-szám. Megmutatjuk, hogy különböző értelmes zárójelezések különböző műveleteket szolgáltatnak. Ebből következik, hogy a (d) kérdésre a választ *pontosan* a Z_n Catalan-számok adják.

Figyeljük meg, hogy a (8) sorozatból nyerhető bármely művelet visszavezethető az alapműveletekre, mégpedig bármelyik ilyen művelet eredménye az $a_1, a_2, \dots a_n$ számokon

$$\frac{1}{2^{k_1+1}} a_1 + \frac{1}{2^{k_2+1}} a_2 + \dots + \frac{1}{2^{k_n+1}} a_n, \tag{9}$$

ahol minden egyes k_i azt mondja meg, hogy a_i az adott zárójelezésnél hány zárójelpár belsejében van. Valóban, minden a_i eggyel többször csöppen bele számtani közép képzésébe, és így eggyel többször feleződik, mint ahány zárójelpárban van. Tegyük fel most már, hogy (8) két értelmes zárójelezése különböző. Akkor lesz olyan i, hogy a_i e két zárójelezésnél különböző számú zárójelpárba kerül. Ezt n szerinti teljes indukcióval igazolhatjuk, felhasználva, hogy (8) bármilyen értelmes zárójelezés után vagy $a_1 \circ (a_2 \circ \dots \circ a_n)$, vagy $(a_1 \circ \dots \circ a_k) \circ (a_{k+1} \circ \dots \circ a_n)$, vagy pedig $(a_1 \circ \dots \circ a_{n-1}) \circ a_n$ alakú, ahogyan azt a (b) kérdésre válaszolva észrevettük.

Ekkor a kétféle zárójelezéssel kapott műveletet (9) alakba írva át azt tapasztaljuk, hogy bennük a_i együtthatója különböző. De akkor a két művelet eredménye is különbözni fog azon a szám-n-esen, amelyben az i-edik elem 1, a többi elem 0. Ezzel beláttuk, hogy a két művelet különböző, és éppen erre volt szükségünk.

Tekintsünk most egy másik műveletet a valós számokon: a közönséges szorzást. Az $a_1 a_2 \dots a_n$ sorozatból természetesen ugyancsak Z_n különböző módon alkothatunk valódi szorzatot, de ezek mindegyike *ugyanaz* a művelet lesz, hiszen a közönséges szorzás asszociatív, s ezért n tényezős szorzatot akármilyen értelmes módon is zárójelezünk, az eredmény mindig ugyanaz a szám lesz. *Ezt* a tényt szoktuk úgy megfogalmazni, hogy asszociatív műveletre érvényes az általános asszociativitás. Ha tehát a (d) kérdést a számtani közép képzésének művelete helyett a valós számok szorzására vonatkozóan tesszük fel, akkor a válasz: e műveletből minden n-re *egy* n-változós műveletet kaphatunk értelmes zárójelezéssel.

Ha ezután egy tetszőleges halmazon egy tetszőleges kétváltozós műveletet tekintünk, ahhoz is képezhetjük azoknak a természetes számoknak a sorozatát, amelyek megmondják, hogy az adott műveletből hány különböző n-változós művelet származtatható értelmes zárójelezésekkel. Ezt a sorozatot nevezzük az adott művelet *asszociatív spektrumának.* A számtani közép asszociatív spektruma tehát a Catalan-számok $Z_1, Z_2, Z_3, \dots$ sorozata, míg a szorzásé (és az összeadásé is) az $1, 1, 1, \dots$ sorozat. Az asszociatív spektrum a műveletnek az asszociatívságtól való távolságát jellemzi; minél nagyobbak az elemei, a művelet annál messzebb van attól, hogy asszociatív legyen.

Nevezzünk a rövidség kedvéért *spektrumnak* minden olyan számsorozatot, amely valamilyen művelet asszociatív spektruma. Néhány egyszerű észrevétel:

1. Minden spektrum első két eleme 1, és ha a harmadik is 1, akkor minden további elem 1.
2. Ha $c_1, c_2, \dots$ spektrum, akkor minden n-re $1 \le c_n \le Z_n$.
3. Ha $c_1, c_2, \dots$ spektrum, akkor minden n-re

$$c_n \le c_1 c_{n-1} + c_2 c_{n-2} + \dots + c_{n-1} c_1.$$

A spektrumokról nagyon keveset tudunk. Olvasónk azonban siker reményében próbálhat válaszolni a következő kérdésekre:

Spektrum-e az $1, 1, 2, 4, \dots, 2^n, \dots$ sorozat?

Létezik-e olyan $c_1, c_2, \dots$ spektrum és olyan i, hogy $c_i > c_{i+1}$?

IRODALOM

[1] Eglesz Istvánné, Kovács Csongorné, Sz. Földvári Vera, *Matematika* (*általános iskola 5.*), Tankönyvkiadó, 1979.

[2] Guy, R. K., The second strong law of small numbers, *Mathematics Magazine*, **63** (1990), 3–20.

[3] Krusemeyer, M., A parenthetical note (to a paper of Guy), *Mathematics Magazine*, **69** (1996), 257–260.

[4] Tamari, D., The algebra of bracketings and their enumeration, *Nieuw Archief voor Wiskunde* (3), **10** (1962), 131–146.

Csákány Béla, JATE Bolyai Intézet