Matematikai mozaik

Matematikai mozaik

www.typotex.hu

Tartalom

ELŐSZÓ.....vii
TALÁLJA KI, MELYIK SZÁMRA GONDOLTAM!..... 9
MATEMATIKA A SAKKTÁBLÁN..... 17
1 MILLIMÉTER = 1000 KILOMÉTER..... 23
1.JEGYZETEK.....45
TUD-E ÖN FEJBEN ÖTÖDIK GYÖKÖT VONNI?..... 49
1. 1. NÉHÁNY SZÁMOLÁSI KÖNNYÍTÉS..... 49
2. 2. ÖSSZEADUNK ÉS KIVONUNK EGYSZERRE..... 50
3.3. SZÁMKITALÁLÁS.....51
4. 4. A 11-ES PRÓBA..... 52
5. 5. ÖTÖDIK GYÖKÖT VONUNK FEJBEN..... 53
6. 6. AZ ÖTÖDIK HATVÁNY UTOLSÓ SZÁMJEGYE..... 53
7. 7. VÉDEKEZZÜNK A CSAPDÁK ELLEN!..... 54
8. 8. TÖKÉLETESÍTJÜK TUDOMÁNYUNKAT..... 54
9. 9. SZÁMOLÁSI EGYSZERŰSÍTÉSEK MAGYARÁZATA..... 55
10. 10. SZORZÁS RÉSZLETSZORZATOK NÉLKÜL..... 57
11. 11. OSZTÁS MELLÉKSZÁMÍTÁSOK NÉLKÜL..... 59
12. 12. NEGATÍV SZÁMJEGYEK SEGÍTSÉGÜL VÉTELE..... 60
PÓK ÉS LÉGY.....63
A HÉRÓN-FÉLE HÁROMSZÖKEGRŐL.....77
A KÖNIGSBERGI HIDAK, A KILENC ÖSVÉNY ÉS MÁS GRÁFELMÉLETI PROBLÉMÁK..... 103
A GALTON-DESZKA.....121
1.BEVEZETÉS.....121
2. A GALTON-DESZKA LEÍRÁSA..... 121
3. EGY JÁTÉK A GALTON-DESZKÁVAL. A GALTON-DESZKA ÉS A BINOMIÁLIS ELOSZLÁS.....123
4. A GALTON-DESZKA ÉS A POISSON-ELOSZLÁS..... 132
5. VÉLETLEN SZÁMOK ELŐÁLLÍTÁSA A GALTON-DESZKA SEGÍTSÉGÉVEL..... 139
6. A GALTON-DESZKA ÉS A MARKOV-LÁNCOK..... 139
7. EGYENLETES ELOSZLÁS ELŐÁLLÍTÁSA GALTON-DESZKÁVAL..... 142
PARKETTÁK A GEOMETRIA SZEMSZÖGÉBŐL..... 147
ÉRDEKES SZÁMOK.....165
HÁNY SZÍN KELL A TÉRKÉP SZÍNEZÉSÉHEZ?..... 179
1. 1. A NÉGYSZÍN PROBLÉMA..... 179
2. 2. A TÉRKÉP LÉNYEGTELEN MÓDOSÍTÁSA. EULER TÉTELE..... 180
3. 3. A PÁRONKÉNT SZOMSZÉDOS ORSZÁGOK MAXIMÁLIS SZÁMA..... 182
4. 4. AZ ÖTSZÍN PROBLÉMA..... 185
5. 5. TETSZŐLEGES TÉRKÉPEK SZÍNEZÉSE..... 192
6. 6. KÉT GÖMB SZÍNEZÉSI PROBLÉMÁJA..... 192
7. 7. A MÖBIUS-SZALAG ÉS A TÓRUSZ SZÍNEZÉSE..... 193
8. 8. TETSZŐLEGES FELÜLETRE RAJZOLT TÉRKÉPEK SZÍNEZÉSE..... 195
9. 9. A NÉGYSZÍN TÉTEL ÁTFOGALMAZÁSAI..... 196
10.10. FELADATOK.....197
11.11. MEGOLDÁSOK.....198
A SZERENCSEJÁTÉKOK ÉS A VALÓSZÍNŰSÉGSZÁMÍTÁS..... 201
1.1. BEVEZETÉS.....201
2.2. A KÁRTYAKEVERÉSRŐL..... 201
3. 3. A KÁRTYÁK ELOSZLÁSÁRA VONATKOZÓ FELADATOK..... 206

4.4. JÁTÉKSTRATÉGIÁK 212
5.5. EGY MATEMATIKUS HARCA A JÁTÉKKASZINÓK ELLEN 226
BŰVÖSNÉGYZETEK 229
CSALAFINTA FELÜLETEK 251
1. BEVEZETÉS 251
2. A MÖBIUS-SZALAG 251
3. A FELÜLET FOGALMA 257
4. A FELÜLETEK OSZTÁLYOZÁSA 267
5. EGYOLDALÚ ZÁRT FELÜLETEK 270
EGY IGAZÁN CSUDÁLATOS BIZONYÍTÁS 283
A BARKOCHBA JÁTÉK ÉS AZ INFORMÁCIÓELMÉLET 301
1.1. BEVEZETÉS 301
2.2. A BARKOCHBA JÁTÉK ÉS AZ INFORMÁCIÓ MÉRÉSE 301
3.3. INFORMÁCIÓ ÉS BIZONYTALANSÁG 307
4.4. A SHANNON-FÉLE FORMULA 312
5.5. A BARKOCHBA JÁTÉK ÉS A KÓDOLÁS 317
PRÍMKÉPLETEK 319

ELŐSZÓ

helyett egy kis történelem, már ami ezzel a könyvvel kapcsolatos:

1969-ben a budapesti Gondolat Kiadó tizenhat dolgozatból álló cikkgyűjteményt jelentetett meg Matematikai érdekességek címmel. A cikkek a Tudományos Ismeretterjesztő Társulat József Attila Szabadegyetemén rendezett Matematika mindenkinek című sorozat legkiemelkedőbb előadásainak anyagát ölelték fel. Az előadások célja a matematika népszerűsítése volt, témáik szélesebb érdeklődésre tartottak számot. A cikkek nagy része közismert kérdésekkel foglalkozott: számkitalálós játékokkal, a sakk, a totó-lottó, a Barkochba és különféle szerencsejátékok matematikai problémáival, de ezenkívül is számos érdekes témával találkozhatott a könyv olvasója; ezeknek segítségével könnyű, játékos, szórakoztató formában egészíthette ki és fejleszthette tovább matematikai ismereteit.

1977-ben az Urania-Verlag (Leipzig–Jena–Berlin) németre fordíttatta és Mathematisches Mosaik címmel kiadta a könyvet. Még ugyanennek az évnek folyamán megjelent a volt NDK-ban a könyv második kiadása, sőt ugyancsak 1977-ben az Aulis Verlag Deubner & Co KG Köln az NSZK-ban is közrebocsátotta a könyvnek ezt a változatát.

1980-ban a szófiai Tehnika kiadó magyarról bolgár nyelvre fordíttatta és Matematicseszka mozajka címmel jelentette meg a könyvet.

1987-ben a varsói Wiedza Powszechna a német nyelvű változatot lengyelre ültette át és ezt adta ki Mozaika matematyczna címmel.

Közben Magyarországon meglehetősen rövid idő alatt elfogyott a megjelent 4200 példány, és az újabb olvasók hiába szerettek volna hozzájutni a könyvhöz, ez a fáradozásuk sikertelennek bizonyult, mert bár ismételt javaslatok érkeztek a Gondolat Kiadó akkori igazhatójához a könyv új kiadását sürgetve, ő nem vette figyelembe ezeket.

1997-ben a Typotex Kiadó pályázatot nyújtott be a Művelődési és Közoktatási Minisztériumba. Ebben anyagi támogatást kért a Matematikai érdekességek című könyv új, átdolgozott változatának megjelentetéséhez. Az illetékesek 1998 márciusában közölték, hogy a Typotex Kiadó megnyerte a pályázatot. Ezek után kezdődött meg az új kiadás előkészítése. A következőkben szeretnénk röviden ismertetni, hogy lényegében mi a különbség az eredeti és az átdolgozott kiadás között.

Kezdjük a címmel! Az idegennyelvű fordítások címének hatására az új kiadás címe: Matematikai mozaik.

Az eredeti kiadásnak két cikke maradt ki az átdolgozottból, mégpedig A totó-lottójáték matematikai problémái című, mert ez lényeges átdolgozást igényelne, amire a szerző: Révész Pál – más irányú elfoglaltságai miatt – nem tudott vállalkozni. Továbbá a szerző: Lánczi Ivánné kívánságára a Szórakozás számrendszerekkel című cikk.

Az átdolgozott kiadás új cikkei: Vadkerty Tibor – Hódi Endre: A Hérón-féle háromszögekről. Ez elsőként A Matematika Tanítása című szakmódszertani folyóiratban jelent meg 1979-ben, és még abban az esztendőben élénk visszhangot váltott ki a hazai szakmai körökben, de még 1996-ban is hivatkozott rá egy tekintélyes külföldi szakfolyóirat.

Az olvasók bizonyára nagy érdeklődéssel olvassák majd Rónyai Lajos: Egy igazán csudálatos bizonyítás című cikkét, amelyben a szerző a matematika egyik legizgatóbb problémájának megoldására irányuló erőfeszítéseket követi nyomon a probléma felvetésétől kezdve, egészen napjainkban történt megoldásáig.

Az új kötet utolsó cikke Laczkovich Miklóstól származik. Címe: Prímképletek. Ez eredetileg a Középiskolai Matematikai Lapokban jelent meg. A cikkben az illusztris szerző olyan, régi számelméleti problémákat tárgyal terjedelmes matematikai apparátussal, amelyek megoldására irányuló törekvések még napjainkban is vezettek új eredményekre. A cikk meglehetősen magas színvonalú, egyes – fontos eredményeket közlő – részletei itt jelentek meg először magyar nyelven.

Valamennyi – régi és új – cikk szövegét átolvastuk, az észrevett sajtóhibákat kijavítottuk (és reméljük, hogy nem vétettünk újakat). Ezenkívül elvégeztük a szükségessé vált korszerűsítéseket mind szakmai, mind helyesírási, mind stiláris szempontból.

Reméljük, hogy az új, átdolgozott kiadás ugyanannyi örömet okoz majd a könyv olvasóinak, mint az eredeti okozott annak idején.

Budapest, 1998 Karácsonyán	Hódi Endre

TALÁLJA KI, MELYIK SZÁMRA GONDOLTAM!

(SZÁMOK KITALÁLÁSÁRA VONATKOZÓ TRÉFÁS FELADATOK MATEMATIKAI HÁTTERE)

VARGA TAMÁS

A gondolatolvasásban csak babonás emberek hisznek. Valamely információnak mindig anyagi úton kell eljutnia a „gondolatolvasóhoz", amiből az következtetni tud. A látszólagos titokzatosságot az okozza, hogy a megfigyelő vagy az *információ útját* nem tudja nyomon követni, vagy a következtetést nem tudja levonni az információból a gondolt névre, számra stb. Matematikáról lévén szó – ha szórakoztató formában is –, nem pedig bűvészetről, ezért csak olyan számkitalálós trükkökről beszélünk, amelyekben az információ útja körül nincs semmiféle ügyeskedés, zsonglőrködés, csupán az okozhat meglepetést, hogy a kapott adatokból hogyan lehet kitalálni a gondolt számot.

Ilyen értelemben beszélhetünk gondolt számok és más adatok *kitalálásáról*, vagy – ha tetszik – most már félreértés veszélye nélkül mondhatjuk: gondolatolvasásról. Ennek a gondolatolvasásnak egyik-másik trükkjét az iskolásgyerekek is megtanulják már az általános iskola nyolcadik osztályában.

„Gondoltam egy számot" – mondja a tanár. A gyerekek tudomásul veszik, és leírják a füzetükbe:

x.

„Hozzáadtam 4-et." A gyerekek így folytatják:

$x+4$.

„Az eredményt megszoroztam 5-tel, és így 50-et kaptam":

$(x+4)\cdot 5=50$.

Innen persze könnyű kitalálni, visszafelé okoskodva, hogy amit a tanár 5-tel megszorzott, az 10 volt, s amihez 4-et adva 10-et kapott, az 6 volt, és ez éppen a gondolt szám. Az adott esetben szinte még a lejegyzés – az *egyenlet* felírása – is felesleges, hiszen ezt a két műveletet könnyű fejben tartani és az eredményből kiindulva visszapergetni. De ha több művelet van, és nem ilyen egyszerűek a számok, akkor már jól jön az egyenlet felírása.

Vannak olyan esetek is, amikor a visszafelé okoskodás ebben az egyszerű formában nem alkalmazható, más módszerekre van szükség. Ha például az előbb 5-tel való szorzás után így folytatta volna a tanár: „…az eredményből elvettem a gondolt szám kétszeresét, és így 50-et kaptam", akkor a lejegyzés ugyan nem nehéz:

$(x+4)\cdot 5-2x=50$,
de a gondolt szám kitalálásához tudni kell, hogy az

$(x+4)$
-nek az 5-szöröse ugyanannyi, mint az

x
-nek az 5-szöröse meg a 4-nek az 5-szöröse, vagyis

$5x+20$
:

$(5x+20)-2x=50$,

és ebből az

$(5x+20)$
-ból úgy lehet elvenni a

$2x$
-et, hogy az

$5x$
-ből elvesszük, a 20-hoz pedig nem nyúlunk:

$3x+20=50.$
Innen már könnyű kitalálni visszafelé okoskodva, hogy a gondolt szám ebben az esetben 10.

Lássunk most egy olyan változatot, amit az iskolákban nem nagyon tanítanak. Először is változtassuk meg a szereposztást! Valaki gondol egy számot, a másik (a „gondolatolvasó”) megmondja neki, hogy milyen műveleteket végezzen a számmal, aztán megkérdezi az eredményt, és ennek alapján kitalálja a gondolt számot.

Legyen az előírás például a következő: „Gondoljon egy számot! Akármekkora pozitív egész szám lehet. Szorozza meg 2-vel! Adjon az eredményhez 4-et! Amit kapott, azt szorozza meg 5-tel! Vonjon ki ebből 7-et! Mi az eredmény? 173? Akkor 16-ra gondolt”. Nem kell számolóművésznek lenni ahhoz, hogy valaki egy másodpercnél rövidebb idő alatt kitalálja az eredményül kapott számból, a 173-ból, a gondolt számot, a 16-ot. Még az egyenletek megoldásához sem kell értenie. Nincs szüksége arra, hogy elvégezze az eredményként kapott számmal az előírt négy művelet fordítottját. Csak azt kell tennie, hogy elhagyja az eredmény utolsó jegyét, és a megmaradó számból kivon 1-et. Az előírt műveletek ugyanis ezt „teszik”

x
-szel:

$(2x+4)\square 5-7,$
de ez ugyanannyi, mint

$10x+20-7,$
vagyis

$10x+13,$
vagyis

$10(x+1)+3.$
Ennek a számnak az utolsó jegye mindig 3, ha

x
pozitív egész szám, hiszen ez a szám 10-nek egy többszörösénél 3-mal nagyobb. Az utolsó jegynek, a 3-nak az elhagyása azt jelenti, hogy a számból kivonunk 3-at, és az eredményt elosztjuk 10-zel. Akkor pedig a

$10(x+1)+3$
-ból

$x+1$
lesz. Ebből már csak ki kell vonni egyet, hogy a gondolt szám maradjon. Az utolsó jegyet elhagyni és a megmaradó számból 1-et kivonni a leggyengébb fejszámolónak is egy rövid pillanat műve. Gondolhatott a másik akár 999-re is, csak neki volt hosszadalmas a számolás, amíg 10003-ig eljutott, nekünk ugyanolyan könnyű a gondolt szám megállapítása ebben az esetben is:

$1000-1=999$
.

Fokozhatjuk a hatást, ha még egy gondolt számot – de most már egyjegyűt! – hozzáadatunk az előbbi műveletsorozat eredményéhez. Ha ezt a számot

y
-nal jelöljük, akkor

$10(x+1)+3+y$
lesz a műveletsorozat eredménye, és ebből *mindkét gondolt számot* kitalálhatjuk. Kivonunk belőle 3-at, lesz

$10(x+1)+y$,
és ennek az utolsó jegye

y
(ha

y
csakugyan egyjegyű!), a megmaradó résznél 1-gyel kisebb szám pedig megint

x
.

Bonyolíthatjuk tovább is a trükköt. „Szorozza meg az eredményt 25-tel, vonjon ki belőle 74-et, ezt még szorozza meg 4-gyel, adjon az eredményhez 97-et és még egy számot, ami most kétjegyű is lehet.” Ha ebből 1101-et kivonunk, akkor az utolsó két jegy, a megelőző jegy és a még megmaradó jegyek megadják az utoljára kigondolt kétjegyű, az előtte gondolt egyjegyű és az először gondolt akárhányjegyű számot. Számoljunk csak utána! Ugyanezt a feladatot variálva, kitalálhatjuk egyszerre három kockadobás eredményét (ekkor a számok eleve egyjegyűek) vagy valakinek a születési évét, hónapját, napját is. Ha bízunk abban, hogy nem tévedünk évtizedet, akkor az utóbbi esetben a születési év utolsó jegye lehet a másodszorra megkérdezett egyjegyű szám, a hónap sorszámát kérdezhetjük például először, a napot utoljára. Ha azonban nem vagyunk biztosak az évtizedben, akkor módosítanunk kell a trükköt. Aki az alapelvet megértette, annak már ez sem lesz nehéz.

$\{[10(x+1)+3+y]\cdot 25-74\}\cdot 4+97+z= \quad =(250x+250+75+25y-74)\cdot 4+97+z= \quad =1000x+100y+1004+97+z= =1000x+100y+z+1101$

Talán az előbbieknél is jobban meglepi az olvasót a következő egyszerű trükk:

„Gondoljon egy számot! Adjon hozzá 6-ot! Szorozza meg az egészet 2-vel! Adjon hozzá még 4-et! Vegye a kapott számnak a felét! Vonja ki belőle a gondolt számot! Biztos benne, hogy jól számolt? Akkor 8 maradt.”

Hogyan? Hát bármiféle információ nélkül kitalálhatom azt a számot, amelyet a másik ember hosszas fejszámolgatás után éppen fejben tart, amikor még csak azt sem tudom, hogy milyen számból indult ki? Lehetséges volna mégis a gondolatolvasás? Szó sincs róla, egyszerű matematikai ügyeskedés az egész. Olyan műveletsorozatot végeztettünk a másikkal, amelyben 8-nak *kellett* maradnia. Lássuk csak:

$((x+6)\cdot 2+4/2)-x$,
ez a műveletsorozat végeredménye. Hozzuk egyszerűbb alakra! A számláló mindkét tagját eloszthatjuk 2-vel;

$x+6+2-x$,
erről pedig már mindenki látja, hogy 8, akármi is az

x
.

Egy kis algebrával – sőt már egy kis józan ésszel is – akárhány olyan műveletsorozatot kigondolhatunk, amely a kiindulástól függetlenül mindig ugyanarra az eredményre vezet. Az alábbi trükköt könnyű megjegyezni, annál nehezebb megmagyarázni. Gondoljon valaki egy háromjegyű számot, amelyben a szélső jegyek különbsége legalább 2. (Pl. 365 megfelel céljainknak, de 465 vagy 565 nem.) Fordítsa meg a gondolt szám jegyeinek sorrendjét, és vegye a két szám különbségét![1] Fordítsa meg a különbségben is a jegyek sorrendjét, és adja össze ezt a két számot! Az eredmény mindig 1089.

Végezzük el a mondott műveleteket néhány olyan számon, amelyek követelményeinknek megfelelnek:

563	400	913	801	
–365	–004	–319	–108	
198	396	594	693	
+891	+693	+495	+396	
1089	1089	1089	1089	

Persze ez a négy eset még nem bizonyítja, hogy sok száz esetben is 1089 lesz az eredmény. Próbáljuk megkeresni a törvényszerűséget, amely ezt biztosítja.

Először is észrevehetjük, hogy a kivonással kapott szám (a közbülső eredmény) középső jegye példáinkban mindig 9, két szélső jegyének pedig az összege 9. Véletlen ez, vagy máskor is mindig így van?

Nagyobb számból vonunk ki kisebbet, tehát az első jegy mindenesetre nagyobb fönt a kisebbítendőben, mint lent a kivonandóban (feltevésünk szerint nem lehet egyenlő). De akkor az utolsó jegy éppen megfordítva, fönt kisebb, lent nagyobb. Kivonáskor tehát „marad 1", amit átviszünk a tízesekhez. Ott egyébként egyenlő számok vannak, és így 0 lenne a különbség, de az 1 átvitele miatt 9 lesz, és a százasokhoz megint átviszünk 1-et. Ezért a közbülső eredmény első jegye 1-gyel kisebb lesz, mint eredetileg a szélső jegyek különbsége. Azt már sejtjük, mire volt jó kikötni, hogy a szélső jegyek különbsége legalább 2 legyen: így a közbülső eredmény első jegye legalább 1, vagyis háromjegyű számot kapunk. Mindjárt látni fogjuk, hogy ez miért fontos, de egyelőre azt akarjuk megérteni, hogy a közbülső eredmény szélső jegyeinek miért 9 az összege. Akármekkora is az eredeti szám szélső jegyeinek különbsége, a közbülső eredmény utolsó jegye ezt 10-re egészíti ki (ha 2 a különbség, akkor 8 lesz az utolsó jegy; ha 3, akkor 7 lesz stb.). A százasok kivonásakor viszont 1-gyel kevesebbet kapunk, mint amennyi a szélső jegyek különbsége volt. A közbülső eredmény első és utolsó jegye tehát nem 10-re, hanem 9-re egészítik ki egymást.

Most még azt kell végiggondolnunk, hogy ha a közbülső eredmény jegyeinek sorrendjét megfordítjuk, és a két számot összeadjuk, miért kapunk mindig 1089-et. Számoljunk csak utána: az utolsó jegyek összege 9, az első jegyeké is annyi, de az 9 százast jelent, a középső jegyek összege 180, összesen

$9+900+180=1089$.

Ha nem kötöttük volna ki, hogy a szélső jegyek különbsége legalább 2, akkor a közbülső eredmény kétjegyű is lehetett volna (ti. 99), sőt egyjegyű is (ti. 0), és ha ezekben a számokban megfordítjuk a jegyek sorrendjét, és összeadjuk a két számot, akkor nem 1089 lesz az eredmény, hanem 198 és 0. (Az első esetben egy fogással elérhetjük, hogy 1089 legyen a végeredmény: azt mondjuk, hogy háromjegyűnek tekintjük a 99-et is, vagyis 099-nek írjuk; ekkor

$099+990=1089$

. A második esetben még ez sem segít.)

Sokféle számkitaláló trükk alapul a kilences maradékok tételén, amelyet így fogalmazhatunk: „Ha egy természetes számot elosztunk 9-cel, akkor ugyanazt a maradékot kapjuk, mint ha a szám jegyeinek összegét

[1]Különbségen itt és alább mindig pozitív különbséget értünk (a nagyobb számból vonjuk ki a kisebbet).

osztottuk volna el 9-cel.” Ez más szóval azt jelenti, hogy bármely többjegyű szám 9-nek egy többszörösével különbözik a saját számjegyeinek összegétől. (Ezt például háromjegyű számokra így láthatjuk be:

$(100x+10y+z)$
-nek és jegyei összegének, vagyis

$(x+y+z)$
-nek a különbsége

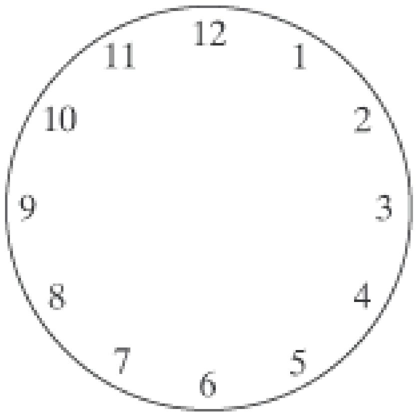
$99x+9y$
, és ez csakugyan többszöröse 9-nek, pontosan

$(11x+y)$
-szorosa. Kétjegyű vagy háromnál többjegyű számok esetében is egészen hasonló a bizonyítás.) Ha tehát egy ismeretlen számot megszoroz valaki 9-cel, a szorzatnak kihúzza az első jegyét, és megmondja a többi jegy összegét, akkor ebből már kitalálhatjuk a kihúzott első jegyet: csak ki kell egészítenünk a kapott számot 9 legközelebbi, nála nagyobb többszörösére. Például

$555\times 9=4995$
, utolsó három jegyének összege 23, és ezt éppen 4 egészíti ki 9 következő nagyobb többszörösére, 27-re. (Ha nem az első jegyet húzatjuk át, akkor ki kell kötnünk, hogy az áthúzott jegy ne 0 legyen, különben a trükk nem mindig sikerül. Aki megértette a fentieket, az könnyen rájöhet, miért.)

Persze nemcsak 9-cel szorozhatjuk meg a gondolt számot, hanem 9-nek bármely többszörösével is; sőt 9-cel osztható számhoz jut a partnerünk akkor is, ha például kivonja a gondolt számból a szám jegyeinek összegét vagy a jegyei megfordításával kapott számot stb. Amint így vagy úgy elértük, hogy 9-cel osztható száma legyen, a trükköt ugyanúgy folytathatjuk tovább, mint az előbb.

Lássunk most egy egészen más természetű, de nem kevésbé meglepő számkitalálós játékot! Rajzoljuk le egy óra számlapját, és kérjünk meg valakit, hogy gondoljon egy egész számot 1 és 12 között (a határokat is beleértve)! Többször egymás után koppantani fogunk ceruzánkkal a számlapon, neki minden koppantásra eggyel tovább kell számolnia. Ha például 8-at gondolt, akkor az első koppantáskor már 9-re kell gondolnia, a második koppantáskor 10-re, és így tovább. Amikor eléri a 20-at, akkor ceruzánk a számlapnak éppen arra a számára fog mutatni, amire eredetileg gondolt.



1. ábra

Hogyan koppantsunk, hogy csakugyan így legyen? Az első 7 koppantás mehet össze-vissza, hiszen a 20-at akkor sem éri el a hetedikre, ha történetesen a lehető legnagyobb számra, 12-re gondolt. Nyolcadszorra azonban elérheti a 20-at, tudniillik éppen abban az esetben, ha 12-re gondolt; koppantsunk tehát nyolcadiknak a 12-re. Ha a 11-re gondolt, akkor kilencedik koppantásnál éri el a 20-at; koppantsunk tehát kilencediknek a 11-re. És így tovább, a tizedik koppantásnak a 10-re, a tizenegyediknek a 9-re, …, a tizenkilencediknek az 1-re kell esnie, és akkor akármelyik számra gondolt is a másik, éppen akkor mutatunk rá a gondolt számra, amikor ő a 20-hoz ér a számlálásban. Mi persze nem vesszük észre, mikor történik ez meg, de ő mégis azt hiszi, hogy kitaláltuk a gondolt számot, hiszen pontosan az ő számát érte a koppantás, amikor ő a 20-hoz ért. Hatásosabb a trükk, ha megkérjük, hogy csak 19-ig számoljon fejben, a 20-at már mondja ki hangosan. Akkor persze mi sem folytatjuk tovább a kopogást, hanem diadalmasan ránézünk, és így teljes lesz az illúziója természet feletti képességeinkről – amíg meg nem magyarázzuk neki a fortélyt.

A számkitalálásoknak egy érdekes fajtája a találgatósdi. Ha például valaki 1 és 100 között (a határokat is beleértve) gondol egy egész számot, hányadik próbálkozásra tudnánk kitalálni, melyik számra gondolt? Ha nagy szerencsénk van, esetleg elsőre eltaláljuk, de lehet olyan pechünk is, hogy csak a századikra hibázunk rá, *feltéve, hogy az illető csak arról informál bennünket minden egyes találgatásunk után, hogy eltaláltuk-e a gondolt számot vagy nem*. Ha azonban csak egy kicsit is jobban informál minket; tudniillik megmondja, hogy a mi számunk sok-e vagy kevés, akkor a hetedikre mindenképpen (esetleg hamarabb is) kitalálhatjuk a számot. Persze nem úgy, hogy vaktában tapogatózunk, hanem szisztematikus felderítő munkával. Egy példán könnyebb megmagyarázni, hogyan.

Gondoltam egy számot 1-től 100-ig.

50

sok

25

sok

12

sok

6

kevés

9

kevés

11

sok

Akkor 10.

Eltaláltad.

Látjuk, hogy az elv: a még szóba jöhető számköz felezése (vagy pontosan, vagy közelítőleg). Két felezés után az eredeti számköz negyedéhez, három felezés után a nyolcadához, …, hét felezés után a 128-adához jutunk. (Persze ha a számköz hossza páratlan számmal fejezhető ki, akkor meg kell elégednünk közelítő felezéssel és így ez a szám kissé módosulhat.) Ez azt mutatja, hogy nemcsak 1-től 100-ig, hanem 1-től 128-ig is kitalálhatjuk a gondolt számot hét próbálgatással; valójában csak 127-ig, mert a hetedik felezéssel a 126-tól 128-ig terjedő számközt felezzük, 127-et mondunk, és ha a felelet még ekkor is „kevés”, akkor a „128” már a nyolcadik próbálgatás lesz. (Persze az utolsó próbálgatásra már nem illik a „próbálgatás” név, hiszen akkor már biztosra megyünk, mint az előbb is a 10-nél.)

Hasonló módszerrel nemcsak számokat találhatunk ki, hanem dátumokat is (az év bármely napját legrosszabb esetben kilencedikre), kezdőbetűket, szavakat stb.

Utoljára hagytam az úgynevezett „bűvészkártyával” való számkitalálás kérdését.

Hat papírlapon – ezek a bűvészkártyák – pozitív egész számok vannak, amelyek nagyság szerint következnek egymás után 59-ig vagy 60-ig. Nem folyamatosan, sok szám kimarad közben, az egyik kártyán ezek, a másikon azok, valami bonyolult rendszer szerint. Gondolni kell egy számot 1-től 60-ig, és a „gondolatolvasónak” odaadni mindazokat a kártyákat, amelyeken megtalálható a gondolt szám. Ennek alapján ő pillanatok alatt megmondja, hogy mi a gondolt szám. A többi kártyát még csak meg sem nézi.

Könnyű dolga van: csak össze kell adnia a neki átadott kártyák első számait, ezek összege lesz a gondolt szám. Ha pl. a 16-tal, 4-gyel és 8-cal kezdődő kártyákat adják a kezébe, akkor a gondolt szám

$16+4+8=28$

. Ezt a titkot a használati utasítás is elárulja; azt azonban nem, hogy *min alapul* a kártyáknak ez a csodálatos tulajdonsága. Megpróbálunk most feleletet adni erre.

Nézzük csak meg a kártyákon levő számokat; nem fedezünk-e fel rajtuk valami szabályosságot, ami útbaigazít.

Az első (*A* jelű) kártyán a páratlan számokat találjuk 1-től 59-ig. A többi kártyán nem látható ilyen egyszerű szabályosság. Feltűnő azonban, hogy első számaik – éppen azok a bizonyos első számok, amelyeket a használati utasítás szerint össze kell adni – 2, 4, 8, 16, 32; vagyis mindegyik kétszer akkora, mint az *előző*. A matematikusok úgy nevezik ezeket a számokat, hogy „2 hatványai”. Például

$2=2^1$

,

$4=2^2=2\cdot 2$

,

$8=2^3=2\cdot 2\cdot 2$

stb. Az első kártya legelső száma 1; ez is 2 hatványai közé tartozik, 2-nek a 0-adik hatványa (hogy miért, az most mellékes).

A „gondolatolvasó” tehát, akármi volt is a gondolt szám, mindig 2 hatványai közül ad össze valahányat, és a használati utasítás szerint ezeknek a hatványoknak az összege lesz a gondolt szám. Úgy látszik, bármely számot fel lehet írni (vagy ahogy mondani szokás, elő lehet állítani) 2 hatványainak összegeként, mégpedig úgy, hogy minden hatvány legfeljebb egyszer fordul elő. Próbáljuk csak ki például a 13-at:

$13=8+4+1$

Nézzük meg az 55-öt is:

$55=32+16+4+2+1$

De vajon biztos, hogy *ez mindig* sikerül? Könnyű belátni, hogy igen, nem is csak 60-ig, hanem azon túl is akármeddig, a következő eljárással:

Az illető számból először levonjuk 2-nek azt a legnagyobb hatványát, amely még nem nagyobb nála (pl. az 55-nél a 32 még nem nagyobb, de a 64 már nagyobb;

$55-32=23$

). Az biztos, hogy így 2-nek az illető hatványánál (itt 32-nél) kisebb számot kapunk, mert különben nem ez lett volna a legnagyobb hatvány, amit elvehetünk. A megmaradó számból megint levonjuk 2-nek a lehető legnagyobb hatványát, és így tovább, míg az eredmény 0 nem lesz. A levont számok összege maga az eredeti szám.

Milyen rendszer alapján írta hát fel a bűvészkártya kitalálója az egyes kártyákra a számokat?

Először is külön egy nagy lapra felírt minden számot 2 hatványainak összegeként:

1 =1 2 =2 3 =2+1 4 =4 5 =4+1 6 =4+2 7 =4+2+1 stb.
Azután ez első kártyára a lista alapján ráírta azokat a számokat, amelyeknek a felbontásához előfordult az 1:

1,3,5,7 stb.
(A páratlan számok felbontásában mindig előfordul, a párosakéban soha. Ugyanis az 1-en kívül 2-nek minden hatványa páros, és így a hatványok összege is páros, ha az 1 nem szerepel a felbontásban; de ha szerepel, akkor páratlanná teszi az összeget.)

Egy másik kártyára azokat a számokat írta fel, amelyeknek a felbontásában előfordult a 2:

2,3,6,7 stb.
Egy harmadikra azokat, amelyeknek a felbontásában előfordult a 4 és így tovább.

Ezeknek a kártyáknak az *első száma* éppen 1, 2, 4 stb., hiszen ezek azok a legkisebb számok, amelyeknek a felbontásában előfordul az 1, ill. a 2, ill. 4 stb.

Milyen kártyákra kerül így rá pl. a 13? Azokra, amelyeknek az első száma szerepel a 13 felbontásában (

8+4+1
-ben): a 8-cal, a 4-gyel és az 1-gyel kezdődő kártyákra; a többire nem. Aki tehát a 13-at gondolja, az éppen ezeken a kártyákon találja meg a gondolt számot, és így a „gondolatolvasónak” csakugyan nincs más dolga, mint összeadni a kártyák kezdőszámait.

Miért éppen 60-ig szerepelnek a kártyákon a számok?

Ha a számkitaláló megállt volna 7-nél (mint mi az előbb a táblázat készítésében), akkor három kártyára lett volna szüksége – az 1-gyel, 2-vel és a 4-gyel kezdődőkre –, hiszen 1-től 7-ig minden szám felbontásában 2-nek csak ezek a hatványai szerepelnek. A 8-hoz azonban már új kártyára lett volna szüksége, hiszen 8-nak a „felbontása” maga a 8, 2-nek a következő hatványa. Ezzel az új kártyával 15-ig lehet elmenni; minden szám 2 hatványaira való felbontásában csak 1, 2, 4 és 8 szerepel. A 16-hoz már egy ötödik kártya kell (ami éppen 16-tal kezdődik), ezzel el lehet menni 32-ig, 32-höz megint egy új kártya szükséges, ezzel el lehet jutni 63-ig. A „bűvös kártya” készítője csak azért hagyta ki a 61-et, 62-t és a 63-at, hogy kerek számig jusson. Hét kártyával már 100-on túl, pontosan 127-ig lehet biztosan kitalálni egy pozitív egész számot.

És most jusson eszünkbe, hogy az előbb kiderítettük: 7 próbálgatással éppen 127-ig lehet biztosan kitalálni egy pozitív egész számot. Ez persze nem véletlen egyezés. Mindkettő azzal kapcsolatos, hogy

127=27–1
. Ha az olvasó elgondolkozik azon, ami a két trükkben közös, ezen keresztül egy kis betekintést kaphat a matematika egy érdekes új ágának, az *információelméletnek* a gondolatvilágába.

MATEMATIKA A SAKKTÁBLÁN

RÉVÉSZ PÁL

A matematikai gondolkozás és a sakkjátékban szükséges gondolkodás hasonlósága valószínűleg mindenki számára nyilvánvaló. Mindkettő világos logikai készséget és nagyfokú kombináló képességet igényel. Ez a hasonlóság felveti a kérdést, hogy lehet-e a sakkjátékban matematikai módszereket használni, lehetséges-e megtalálni a legjobb stratégiát, vagy például sakkrejtvények megoldásában használhatók-e matematikai módszerek.

Különösen reményt keltő ebből a szempontból a matematika egy viszonylag új ágának, a *játékelméletnek* a kialakulása. Ez azt a feladatot tűzte maga elé, hogy általános módszereket adjon annak eldöntésére, hogy egyes játékokban vajon létezik-e valamelyik játszó fél számára nyerő stratégia, illetve hogyan tudják az egyes játékosok megtalálni a legjobb stratégiát. (Megemlítjük, hogy a játékelmélet haszna nem elsősorban az egyes játékok analízisében, hanem a matematikai statisztikában, a matematikának az alkalmazások szempontjából egyik legfontosabb fejezetében mutatkozik meg.) A játékelmélet eredményeiből eddig nem sikerült érdemleges következtetést levonni a sakkjátékra vonatkozóan. (Még arra a két kérdésre sem ismeretes a válasz, hogy létezik-e valamelyik félnek nyerő stratégiája, vagy hogyha egyik fél sem csinál hibát, akkor a játék eredménye döntetlen lesz-e?)

Ennek elsősorban az az oka, hogy a sakkjátékban igen nagyszámú kombinációt kellene áttekinteni. Ilyen nagyszámú eset áttekintése még a legmodernebb elektronikus számítógépek segítségével is lehetetlen. Megoldható feladat olyan elektronikus számítógép konstruálása, amely alkalmas 2 vagy 3 lépéses sakkfeladványok megoldására. Olyan sakkozó gép is konstruálható, amely két lépésben nem néz el tisztet, de természetesen egy ilyen gép kikaphat jobb sakkozótól.

A sakkjátéknál fellépő lehetőségek számát jellemzi, hogy mivel az első alkalommal mindkét játékos lehetséges lépéseinek száma 20, így miután mindkét játékos egy lépést tett, a lehetséges helyzetek száma 400. Amint az ellenőrizhető, a világos második lépése utáni lehetséges helyzetek száma 5206. (Itt az esetek összeszámlálásakor nem vettük külön esetként figyelembe azt, hogy egy állás több módon is létrejöhet. Így pl. az e4–e5–Hc3 és a Hc3–e5–e4 eseteket azonosaknak tekintjük. Ha ezeket különbözőknek tekintjük, az 5206-nak közel kétszeresét kapjuk.)

Valójában azt mondhatjuk, hogy a matematika szerepe a sakkjátékban igen minimális. (Más kérdés az, hogy a matematikával való foglalkozás útján szerzett kombinatív készség a sakkjátékban igen jól használható.) Érdekesebb a fordított kérdés: milyen szerepe van a sakknak a matematikában, pontosabban: milyen matematikai problémák vetődnek fel a sakkjátékkal kapcsolatban. Most ezzel az utóbbi kérdéssel foglalkozunk.

A monda szerint már a sakkjáték felfedezője is felvetett egy matematikai problémát az őt megjutalmazni kívánó uralkodónak. Ugyanis jutalmát búzaszemekben kérte, mégpedig úgy, hogy tegyen az uralkodó a sakktábla első mezőjére egy búzaszemet, a következőre kettőt, az azután következőre négyet és így tovább, mindig kétszer annyi szemet, mint amennyi az előzőn volt, végig a 64 mezőn. A történet szerint az uralkodó felháborodott, hogy a feltaláló ilyen „csekélységet” kér jutalmul, de amikor hozzákezdett a számoláshoz, kiderült, hogy a kérést nem tudja teljesíteni, ennyi búza az egész világon nem létezik.

Próbáljuk meg a számítási munkát mi elvégezni az uralkodó helyett! A sakktábla egyes mezőire jutó búzaszemek száma 1, 2, 4, 8, 16, …,

263

. Így a szükséges búzaszemek száma:

1+2+4+8+16+…+263.

Az összeg kiszámításához csak azt kell észrevennünk, hogy a sakktábla első

k
kockájára lerakott búzaszemek számának összege eggyel kevesebb, mint a

$(k+1)$
-edik mezőre lerakott búzaszemek száma, azaz

$1+2+4+8+16+\ldots+2k-1=2k-1$.
Így

$1+2+4+8+16+\ldots+263=264-1$.
Megemlítjük, hogy

264
egy húszjegyű szám.

Egy másik, a sakktáblával (de még nem a sakkfigurákkal) kapcsolatos feladat *Neumann Jánostól* származik. Tegyük fel, hogy van egy sakktáblánk és egy olyan dominókockánk, amely a sakktábla két, egymás melletti mezőjét képes lefedni. Vágjuk ki a sakktábla két ellentétes sarkán lévő mezőket (pl. a1-et és h8-at), kérdés, a megmaradó tábla lefedhető-e dominókkal úgy, hogy minden mező egyszer és csak egyszer legyen lefedve.

Ha valaki megpróbálkozik a feladat elvégzésével, azt fogja tapasztalni, hogy 30 dominó elhelyezése után két, nem egymás melletti lefedetlen mező marad, amelyeket már egy 31-edik dominóval nem tud lefedni. Természetesen kérdés, hogy ez a próbálkozó ügyetlensége miatt van-e, vagy a feladat tényleg megoldhatatlan. A következő szellemes ötlettel belátható, hogy a feladatot valóban lehetetlen megoldani. Minden egyes dominó a sakktáblának egy fekete és egy fehér mezőjét fogja lefedni, viszont a két kivágott kocka mindegyike fehér (vagy mindegyike fekete). Így a feladat nyilvánvalóan megoldhatatlan. Ugyanígy megoldhatatlan a dominókkal való lefedés, ha bármely két olyan mezőt vágunk ki, amely egyszínű. Valamivel nehezebb belátni, hogy ha két különböző színű kockát vágunk ki a sakktábláról bárhonnan, akkor a lefedés megoldható.

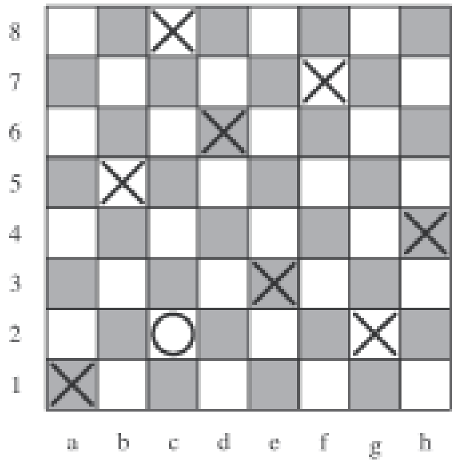
A sakkfigurák közül talán legegyszerűbb a bástya lépéseinek áttekintése. Foglalkozzunk először a következő kérdéssel: hogyan lehet egy sakktáblára bástyákat elhelyezni úgy, hogy közülük bármely kettő ne üthesse egymást, azaz hogy bármely két bástya különböző sorban és különböző oszlopban legyen. Nyilvánvaló, hogy nyolc bástya elhelyezhető ily módon a sakktáblán (pl. úgy, hogy a tábla valamelyik átlójára tesszük a bástyákat), de nyolcnál több már nem. Kérdezzük, hogy hány különböző módon lehet nyolc bástyát elhelyezni a sakktáblán úgy, hogy ne legyen kettő, amelyik üti egymást. A feladat megoldása meglepően nagy szám:

40 320
. A következő módon lehet ezt belátni: ha nyolc bástyát elhelyezünk a sakktáblán úgy, hogy nincs kettő, amelyik üti egymást, akkor minden sorban lesz egy és csak egy bástya. Az első sor bármely helyére (tehát 8 hely mindegyikére) tehetjük az első bástyát. A második bástyát elhelyezzük a második sorba, vigyázva arra, hogy ne ugyanabba az oszlopba kerüljön, ahol az első bástya van; így a második bástyát 7 különböző helyre tehetjük (bárhova is tettük az elsőt). Ugyanígy látható, hogy a harmadik bástyát a harmadik sorba hatféleképpen lehet elhelyezni. Az eljárást folytatva adódik, hogy a 8 bástyát

8□7□6□5□4□3□2□1=40 320
-féleképpen lehet elhelyezni.

Meg kell jegyeznünk, hogy itt különböző elhelyezéseknek tekintjük pl. az a1 b2 c3 d4 e5 f6 g7 h8 és az a8 b7 c6 d5 e4 f3 g2 h1 elhelyezéseket, vagyis az olyan elhelyezéseket, amelyek egymásból a tábla elforgatásával, illetve tükrözéssel nyerhetők. Nehezebb az a kérdés, hogy hány lényegesen különböző elhelyezése létezik a nyolc bástyának, tehát olyan elhelyezése, amelyek között nincs kettő, mely forgatással, illetve tükrözéssel lenne nyerhető egymásból.

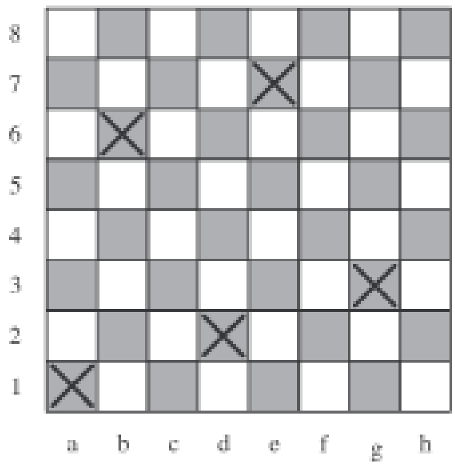
A királynővel kapcsolatos hasonló kérdésnél még az sem eleve világos, hogy hány királynő helyezhető el a sakktáblán úgy, hogy ne legyen kettő, amelyik üti egymást. Itt is belátható, hogy nyolcnál több királynő nem helyezhető el, de ennyi elhelyezése sem könnyű feladat. Hogy ez egyáltalán lehetséges, mutatja a következő példa: a1 b5 c8 d6 e3 f7 g2 h4 (lásd 1. ábra).



1. ábra

Könnyen ellenőrizhető, hogy a királynőket így elhelyezve valóban nincs kettő, amelyik ütné egymást. (Megjegyezzük, hogy a nem lényegesen különböző elhelyezések száma 92.) Nyilvánvaló, hogy ha ezekre a helyekre bástyákat vagy futókat rakunk, akkor a bástyáknak, illetve a futóknak olyan elhelyezését kapjuk, ahol nincs kettő, amely ütné egymást. Ha például ezekre a helyekre futókat rakunk, akkor elfér a táblára még egy futó (pl. c 2-re) (lásd 1. ábra), amely a már lerakott futók egyikét sem üti. Figyeljük meg azt is ezen a példán, hogy a nyolc vezér közül pontosan négy áll világos és négy sötét mezőn. Ellenőrizhető, hogy mind a 92 lehetséges elhelyezés rendelkezik az említett tulajdonsággal, azaz pontosan négy királynő áll világos mezőn.

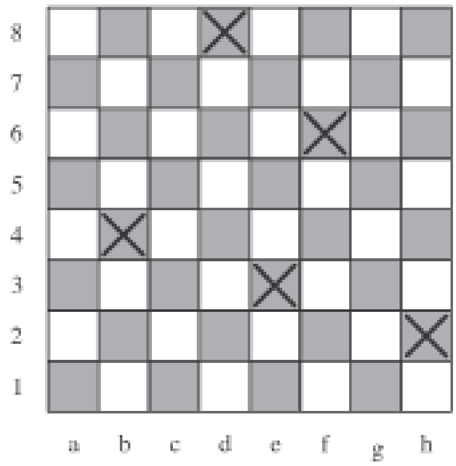
Felvetődik a kérdés: hány királynő helyezhető el sötét mezőre úgy, hogy ne legyen kettő, amelyik üti egymást. Példánk mutatja, hogy négy bizonyosan. Kérdés, hogy négynél több elhelyezhető-e? Könnyen látható, hogy öt királynő elhelyezhető sötét mezőkön úgy, hogy ne üssék egymást, például a következő módon: a1 b6 d2 e7 g3 (2. ábra).



2. ábra

Nehezebb annak igazolása, hogy hat királynő elhelyezése már nem oldható meg.

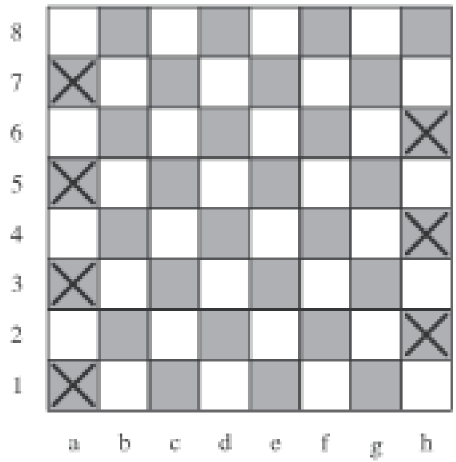
Ezeknek a feladatoknak lényegében fordítottja a következő kérdés: mi a királynőknek az a minimális száma, amely elegendő ahhoz, hogy az egész táblát „lefogja”. Pontosabban, hány királynőt kell leraknunk a sakktáblára, ha azt akarjuk elérni, hogy minden mező a királynők legalább egyike által támadva legyen. Ezt a feladatot már 5 királynő el tudja látni, például a következő elhelyezésben: b4 d8 e3 f6 h2 (3. ábra).



3. ábra

(Ebben a példában a királynők mindegyike sötét mezőn áll.) Ellenőrizhető, hogy négy királynő nem elegendő az említett feladat végrehajtására. De ha csak azt tekintjük feladatunknak, hogy a királynőkkel a tábla világos mezőit lefedjük, akkor már elég négy királynő.

Futókkal kapcsolatban is felvethető, hogy hány futó helyezhető el a sakktáblán úgy, hogy ne legyen kettő, amelyik üti egymást. Elegendő kizárólag sötét futókkal foglalkoznunk. Könnyen látható, hogy legfeljebb 7 sötét futó helyezhető el így; ugyanis az (a7 b8), (a5 b6 c7 d8), (a3 b4 c5 d6 e7 f8), (a1 b2 c3 d4 e5 f6 g7 h8), (c1 d2 e3 f4 g5 h6), (e1 f2 g3 h4) és (g1 h2) „átlók” mindegyikén legfeljebb egy futó állhat. Hogy 7 futó valóban elhelyezhető, mutatja a következő példa: (a1 a3 a5 a7 h2 h4 h6) (4. ábra).



4. ábra

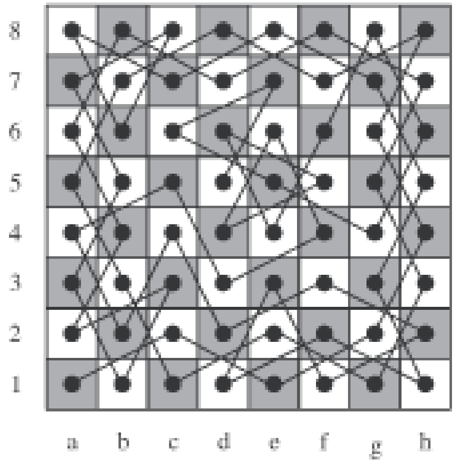
Összeszámolható az összes lehetséges elhelyezés száma is. Az a7 és b8 mezők egyikére (és csak egyikére) kell tennünk egy futót; ha a7-re tettük a futót, akkor h2-re is kell tennünk; ha h8-ra tettük, akkor g1-re is kell tennünk. Továbbá az a5 és d8 mezők egyikére kell tennünk futót. Aszerint, amint a5-re vagy d8-ra tettünk futót, futót kell tennünk h4-re, illetve e1-re. Hasonlóan az a3 és f8 mezők egyikére kell tennünk futót és ennek megfelelően h6-ra, illetve c1-re. Végül egy futót kell elhelyeznünk az a1 és h8 mezők egyikén. Ez a gondolatmenet mutatja, hogy a 7 világos futót

2□2□2□2=24=16

-féleképpen lehet a sakktáblán elhelyezni úgy, hogy ne legyen kettő, amelyik üti egymást. Fenti gondolatmenet világosan mutatja, hogy minden lehetséges elhelyezés olyan, hogy mind a 7 futó a tábla szélén áll.

Természetesen huszárokkal kapcsolatosan is felvethetők ilyen típusú kérdések. Nem nehéz arra válaszolni, hogy hány huszár helyezhető el a sakktáblán úgy, hogy ne legyen kettő, amelyik üti egymást, vagy hogy mi az a minimális számú huszár, amely elegendő ahhoz, hogy a tábla minden mezője a huszárok legalább egyike által támadva legyen.

Huszárokkal kapcsolatban nem teljesen evidens az sem, hogy egy huszárral a tábla bármely mezőjéről el lehet jutni a tábla bármely más mezőjére. Ez természetesen igaz, sőt igaz az is, hogy a sakktábla bejárható egyetlen huszárral úgy, hogy minden mezőt egyszer és csak egyszer érintünk. Például jó a következő út: a1 – c2 – e1 – g2 – h4 – g6 – h8 – f7 – d8 – b7 – a5 – b3 – c1 – e2 – g1 – h3 – g5 – h7 – f8 – d7 – b8 – a6 – b4 – a2 – c3 – b1 – a3 – b5 – a7 – c8 – b6 – a8 – c7 – e8 – g7 – h5 – g3 – h1 – f2 – d1 – e3 – f1 – h2 – f3 – d2 – c4 – b2 – a4 – c5 – d3 – f4 – e6 – d4 – f5 – d6 – e4 – f6 – g8 – h6 – g4 – e5 – c6 – e7 – d5 (5. ábra).

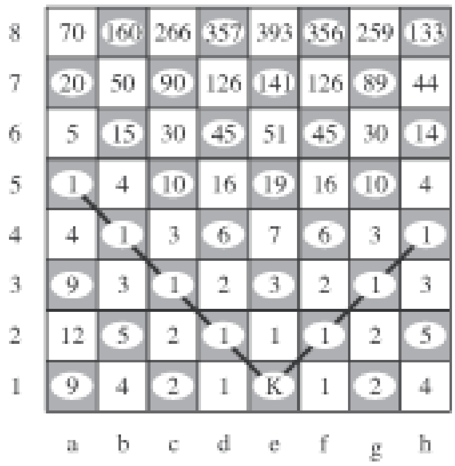


5. ábra

Természetesen egy sakkozót az a kérdés is érdekel, hogy egy adott helyről hány lépésben tud egy huszárral egy másik adott helyre eljutni. Ellenőrizhető, hogy bármely helyről bármely helyre legfeljebb hat lépésben el lehet jutni. (Nyilvánvaló, hogy pl. a1-ről h8-ra hatnál kevesebb lépésben nem lehet eljutni.)

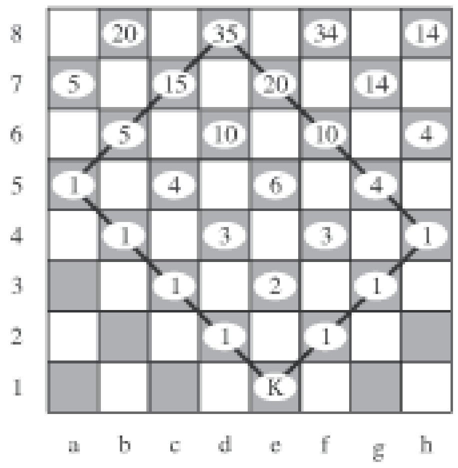
A király szerepe főleg a végjátékban érdekes. Fontos kérdés például, hogy egy gyalogot a király „utol tud-e érni”. Ismeretes, hogy a „bemenő” gyalogot az ellenfél királya akkor és csak akkor tudja utolérni, ha a király benne van a gyalog négyzetében, vagyis abban a négyzetben, amelynek egyik csúcsa a gyalog és egyik éle az első, illetve nyolcadik vonal aszerint, amint a bemenő gyalog világos vagy sötét.

Kiszámítható, hogy egy adott helyről egy másik adott helyre hány lépésben tud eljutni a király. Érdekesebb az a kérdés, hogy egy adott helyről (pl. a kezdeti helyzetből) hányféleképpen tud minimális számú lépéssel eljutni egy másik adott helyre. Erre a kérdésre ad választ a 6. ábra.



6. ábra

Megjegyezzük, hogy ennek a táblázatnak képzési módja nagyon hasonlít az ún. Pascal-háromszög képzési módjához; ha a királynak csak átló-irányú lépéseket engednénk meg, a táblázat maga is hasonló lenne a Pascal-háromszöghöz. (Pontosan a Pascal-háromszöget kapnánk végtelen sakktábla használata esetén.) Az így adódó táblázatot a 7. ábra tartalmazza, a szaggatott vonalon belüli számok pontosan megegyeznek a Pascal-háromszög megfelelő elemeivel.



7. ábra

1 MILLIMÉTER = 1000 KILOMÉTER

(HIBÁS BIZONYÍTÁSOK)

VARGA TAMÁS

1. példa. Bebizonyítom, hogy

(2□6/6□5)=(2/5)
. Egyszerűsítek 6-tal:

(2□□6/□6□5)=(2/5)
Az állítás igaz. A bizonyítás helyes.

2. példa. Bebizonyítom, hogy

(26/65)=(2/5)
. Egyszerűsítek 6-tal:

(2□6/□65)=(2/5)
Az állítás igaz. A bizonyítás azonban hibás. Huszonhat hatvanötöd csakugyan egyenlő kétötöddel, de ez nem következik abból, hogy a két 6-os számjegy törlése után

(2/5)
alakú kifejezést kapunk. Véletlenül jó eredményt kapunk

(26=2□13, 65=5□13)
, de a módszer, mint a következő példa mutatja, nem általános érvényű.

3. példa. Bebizonyítom, hogy

(27/75)=(2/5)
. Egyszerűsítek 7-tel, ahogy az előbb 6-tal:

(2□7/□75)=(2/5).
Az előbbi módszer most nyilvánvalóan tévútra vezetett.

75=5□15
, de 27 nem egyenlő

2□15
-tel; 27 egyetlen egész szám kétszeresével sem egyenlő (1. jegyzet).

Az állítás téves. A bizonyítás hibás. Most olyan példa volna soron, amelyben az állítás téves, a bizonyítás mégis helyes. Ilyen példát azonban nem tudunk, mert ilyen példa nem létezik. Hibás bizonyításból adódhat olyan eredmény is, ami igaz, olyan is, ami nem; de helyes bizonyítás nem vezethet téves eredményre, hacsak a kiindulópont is nem téves. Az alább következő példák nagyrészt abba a két típusba sorolhatók, amelyeket a 2. és 3. példa szemléltet; de a bennük található hiba nem lesz mindig olyan átlátszóan egyszerű, mint itt. Igaz, az olvasó ismereteitől is függ, hogy azonnal látja-e a hibát, vagy sem. Van, akit könnyű becsapni; van, akit nehéz. Becsapós bizonyításokkal éppen azért érdemes megismerkednie az olvasónak, hogy legközelebb már ne lehessen olyan könnyen félrevezetni. Jó bizonyításokat is azért mutatunk néhol a hibásak mellett, hogy különbséget tudjon tenni köztük. Az tudja meg, mit jelent valamit bebizonyítani, aki látja ezt a különbséget.

4. példa. Figyeljük meg ezt a két kivonást:

987654321 9+8+7+6+5+4+3+2+1 123456789□ 1+2+3+4+5+6+7+8+9□ 864197532 8+6+4+1+9+7+5+3+2

Az első kétségtelenül helyes. A másodikat is ugyanúgy végeztük el, nem törődve a

+
jelekkel. Talán csak nem okoznak zavart? Próbáljuk ki fönn is, középütt is, lenn is, mi a számok összege:

9+8+7+6+5+4+3+2+1 =45 1+2+3+4+5+6+7+8+9 =45, 8+6+4+1+9+7+5+3+2 =45.
Érthető, hogy mind a három eredmény ugyanaz, hiszen ugyanazokat a számokat adtuk össze, csak a sorrendjük volt más-más. De akkor a kivonással azt „bizonyítottuk be”, hogy

45–45=45
. Úgy látszik, összegeket mégsem lehet ugyanúgy kivonni, mint többjegyű számokat (2. jegyzet).

5. példa. Melyik több, 6 kilogramm vagy 6 tonna? A következő bizonyítás – minden tapasztalat ellenére – mintha azt mutatná, hogy egyenlők:

2 kg=2000 g 3 kg=3000 g□ 6 kg=6 000 000 g=6000 kg=6 tonna,
hiszen egyenlőket egyenlőkkel szorozva egyenlőket kapunk.

Az egyenlőség körül nincs is hiba, csak a szorzás körül. Ott sem a számokban, csak a nevükben. Kilogrammszor kilogramm a mindennapi életben nem szokott előfordulni, de a fizikában van értelme, és akkor ez éppúgy négyzetkilogramm, ahogy kilométerszer kilométer is négyzetkilométer, nem pedig kilométer. Például

2 km□3 km=6 km2
. Ha így számítjuk:

2000 m□3000 m=6 000 000 m2
, ugyanazt az eredményt kapjuk, hiszen minden négyzetkilométerben egymillió négyzetméter van. Ugyanígy 1 négyzetkilogramm is egymillió négyzetgrammal egyenlő, nem ezerrel.

6. példa. Egy embernek két szülője, négy nagyszülője, nyolc dédszülője van, és így tovább. Ha egy nemzedékkel följebb megyünk, mindig kétszer annyi; tíz generációval előbb több, mint ezerannyi (mert

2□2□2□2□2□2□2□2□2□2=210=1024
), még tíz generációval előbb ennek is több, mint az ezerszerese, vagyis a mostaninak több, mint milliószorosa, 3000 billiónál is több. …Pedig húsz generáció még csak a középkorba visz vissza. Az okoskodás nyilvánvalóan hibás. Hol van benne a hiba?

Két hiba is van. Egyik sem matematikai természetű, hanem abból származik, hogy a matematikai modell nem jól illeszkedik a valósághoz. „Kicsiben” még az is elképzelhető, hogy jól illeszkedik: például, ha valaki egyetlen gyermeke szüleinek, szülei is egykék, azok szülei is, és így tovább néhány nemzedéken át. Akkor csakugyan felfelé menve egy-egy következő nemzedékhez kétszer annyian tartoznak, lefelé menve viszont feleannyian. Ez a kihaló családfa-típus, illetve annak egy egyszerű fajtája. Az előbbi gondolatmenet egyik hibája, hogy ebből a ritka típusból akart egyetemesen érvényes modellt faragni. De van egy más hibája is: egy ember

n
generációval előbbi felmenőinek száma nem lehet ugyan több

2*n*
-nél, de kevesebb lehet. Ha például valakinek a szülei első unokatestvérek (ami nem is olyan ritkaság), akkor a szüleinek a négy-négy nagyszülője közül kettő közös, neki tehát nem nyolc dédszülője van, csak hat. Hasonló „őscsökkenés” lép fel akkor is, ha másodunokatestvérek vagy még távolabbi rokonok kötnek házasságot.

7. példa. Mint ismeretes,

$2\square 2$
néha 5. Ezt be is lehet bizonyítani, nem is egy módon. Néhány „bizonyítás” következik.

Először induljunk ki egy olyan állításból, amit senki sem vonhat kétségbe:

$a2-a2=a2-a2$

A bal oldalon emeljük ki a közös

a
tényezőt, a jobb oldalt pedig úgy bontsuk tényezőkre, mint két négyzet különbségét:

$a(a-a)=(a+a)(a-a)$

Az

$a-a$
tényezővel végigosztva:

$a=2a$

Mindkét oldalt

$3a$
-val növelve:

$4a=5a$

Most még végigoszthatunk

a
-val és a 4-et

$2\square 2$
alakban írhatjuk:

$2\square 2=5.$

8. példa. Jelentsen

b
és

c
bármely két számot (például

$b=4$
,

$c=1$
). Jelöljük az összegüket

a
-val:

$b+c=a.$

(Ha

$b=4$
és

$c=1$
, akkor

$a=5$
.) Szorozzuk meg mindkét oldalt

$(a–b)$
-vel:

$a \cdot b+a \cdot c–b \cdot b–b \cdot c=a \cdot a–a \cdot b.$

Vonjuk ki mindkét oldalból

$a \cdot c$
-t:

$a \cdot b–b \cdot b–b \cdot c=a \cdot a–a \cdot b–a \cdot c.$

Észrevehetjük, hogy a bal oldalon is van egy közös tényező (

b
), a jobb oldalon is van egy (

a
). Emeljük ki ezeket:

$b(a–b–c)=a(a–b–c).$

Osszuk végig

$(a–b–c)$
-vel! Azt kapjuk, hogy

$b=a.$

A fent említett esetben, amikor

$b=4$
, vagyis

$2 \cdot 2$
és

$a=5$
, azt kapjuk:

$2 \cdot 2=5.$

Hol van a hiba

$2 \cdot 2=5$
-nek ezekben a bizonyításaiban? Természetesen ott, hogy – alig burkolt formában – 0-val osztottunk. Először az

$a–a$

volt egyenlő 0-val, másodszor

$(a–b–c)$
(hiszen

$a=b+c$
). Ha pedig 0-val osztunk, akár burkoltan, akár nyíltan, akkor minden képtelenségre el lehetünk készülve. (Gondoljunk arra, hogy

$0\cdot 1=0\cdot 2$
és mégsem igaz az, hogy

$1=2$
.)

9. példa. Lássuk csak kendőzetlenül is a 0-val való osztást! Induljunk ki abból, hogy

$0=0.$

Ez eddig igaz is. Ezt így is írhatjuk:

$0\cdot 2\cdot 2=0\cdot 5,$
ami még mindig igaz, hiszen

$0\cdot 2\cdot 2$
is,

$0\cdot 5$
is egyenlő 0-val. De ha ezzel a 0-val végigosztunk, azt kapjuk:

$2\cdot 2=5.$

Persze azt is írhattuk volna, hogy

$0\cdot 1=0\cdot 1\ 000\ 000\ 000$
és így

$1=1\ 000\ 000\ 000,$
például

1 mm=1 000 000 000 mm=1 000 000 m=1000 km,
amint a címben állítottuk. Erre még más csalfa bizonyítást is adunk majd. Most azonban lássuk a

$2\cdot 2=5$
-nek egy kevésbé átlátszó bizonyítását.

10. példa. Kétségtelenül igaz, hogy

$16–36=25–45,$
hiszen mindkét oldalon

$–20$
szerepel, más-más formában. A bal oldalon 16 teljes négyzet (4-nek a négyzete),

$–36$
pedig 4-nek és

$–(9/2)$

-nek a kétszeres szorzata. Ha még hozzáadjuk

$-\frac{9}{2}$
-nek a négyzetét, két tag négyzete lesz belőle; emlékezzünk csak vissza az

$a^2+2ab+b^2\equiv(a+b)^2$
, vagy ha tetszik, az

$a^2-2ab+b^2\equiv(a-b)^2$
azonosságra. A jobb oldalról is szinte ugyanezt mondhatjuk: 25 teljes négyzet (5-nek a négyzete),

-45
pedig 5-nek és

$-\frac{9}{2}$
-nek a kétszeres szorzata, és így

$-\frac{9}{2}$
négyzetének hozzáadása után itt is két tag négyzetét kapjuk.

Szerencsénk volt: mindkét oldalt ugyanannak a tagnak a hozzáadásával alakíthatjuk két tag négyzetévé, és ezután az egyenlőség érvényben marad:

$16-36+\left(-\frac{9}{2}\right)^2=25-45+\left(-\frac{9}{2}\right)^2,$
vagyis

$\left(4-\frac{9}{2}\right)^2=\left(5-\frac{9}{2}\right)^2,$
vagy másképp

$2\cdot 2-\frac{9}{2}=5-\frac{9}{2},$
és így,

$\frac{9}{2}$
-et adva mindkét oldalhoz

$2\cdot 2=5.$

A hibát könnyen megtalálja az, aki sorról sorra utánaszámol; mi az, ami még igaz, mi az, ami már nem.

$16-36+\left(-\frac{9}{2}\right)^2=-20+\frac{81}{4}=\frac{1}{4}$
és ugyanennyi

$25-45+\left(-\frac{9}{2}\right)^2$
is.

$\left(4-\frac{9}{2}\right)^2$
és

$\left(5-\frac{9}{2}\right)^2$
is

$\frac{1}{4}$
; az első mint

$\left(-\frac{1}{2}\right)^2$
, vagyis

$(-(1/2))(-(1/2))$
, a második mint

$((1/2))2$
. De

$4-(9/2)$
már nem ugyanannyi, mint

$5-(9/2)$
: az előbbi

$-(1/2)$
, az utóbbi 1-gyel több,

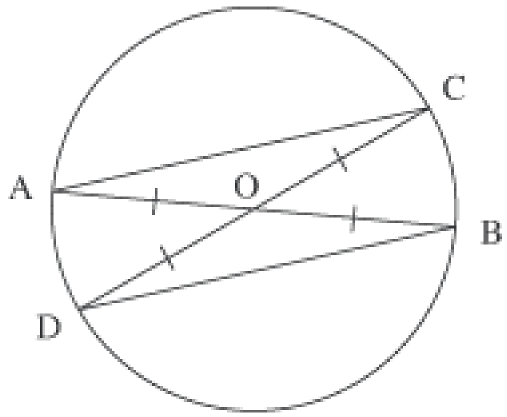
$(1/2)$
. A hiba tehát a jogtalan négyzetgyökvonásban volt. Két szám négyzete lehet úgy is egyenlő, hogy maguk a számok nem egyenlők: egyenlő abszolút értékű pozitív és negatív számoknak mindig egyenlő a négyzete, pedig pozitív és negatív számok nem lehetnek egyenlők egymással.

11. példa. Húzzunk meg egy körben egy átmérőt (az 1. ábrán

AB
), és a végpontjaiból két egymással párhuzamos húrt (

AC
és

BD
)! Bebizonyítjuk, hogy ezek a húrok egyenlő hosszúságúak.



1. ábra

Az

AB
átmérő felezőpontja a kör középpontja,

O
. Kössük ezt össze

C
-vel és

D
-vel:

OA=*OB*=*OC*=*OD*,
mert mind a kör sugarai.

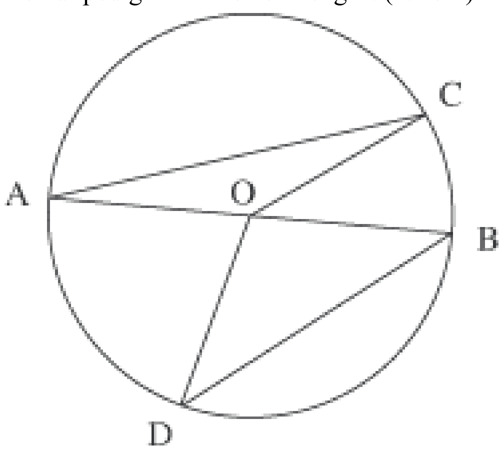
AOC□=*BOD*□
mint csúcsszögek. Tehát két oldal és a közbezárt szög egyezése folytán

AOC□□*BOD*□,
és így ezeknek a háromszögeknek a harmadik oldala is megegyezik:

AC=*BD*.

Éppen ezt kellett bebizonyítanunk.

Helyesebben, kellett volna, de nem bizonyítottuk be. Az igaz ugyan, hogy a mondott módon keletkező húrok egyenlők, de a fenti bizonyítás hibás. A hiba nyilvánvaló abból, hogy nem használtunk ki egy lényeges feltételt: azt, hogy a két húr párhuzamos. Enélkül pedig az állítás nem is igaz. (2. ábra)



2. ábra

Kihasználtunk viszont egy olyan feltételt, amelyet nem indokoltunk meg: azt, hogy

AOC□
és

BOD□
csúcsszögek. Ez ugyan következik a párhuzamosságból, de ennyi erővel már

AC
és

BD
egyenlőségét is bebizonyíthatjuk.

12. példa. Az előbbi bizonyítást módosítsuk úgy, hogy a

C
és

D
pontot ne

AB
felezőpontjával, hanem egymással kössük össze, és most azt a pontot jelöljük

O
-val, amelyben

CD
az

AB
szakaszt metszi! A bizonyítás egyébként változatlan. Most jogunk van azt mondani, hogy

AOC□
és

BOD□
csúcsszögek, tehát egyenlők. Vajon ez a bizonyítás helyes most már? Nem, most is hibás. Most ugyanis nem tudhatjuk azt, hogy

AB
és

CD
metszéspontja felezi-e az

AB
átmérőt, és így a kör középpontja-e. (Az, hogy

O
-val jelöljük meg, még nem teszi középponttá.)

13. példa. Módosítsuk megint a 11. példában leírt bizonyítást! Következtessünk most az

AOC□
és

BOD□
egybevágóságára abból, hogy

$OA=OB$
,

$OC=OD$
és valamelyikükkel szemközti szögek is egyenlők, mint váltószögek.

Helyes-e a bizonyítás ebben az esetben?

Az a kérdés, mely szögek egyezéséről van szó. Azt felhasználhatjuk, hogy

CAO□$=DBO$□
, mert váltószögek, és nemcsak azt tudjuk, hogy

AC□BD
, hanem azt is, hogy

A
,

O
és

B
egy egyenesen vannak (hiszen

O
-val a 11. példában az

AB
átmérő felezőpontját jelöltük). Ha azonban azt akarnánk kihasználni, hogy

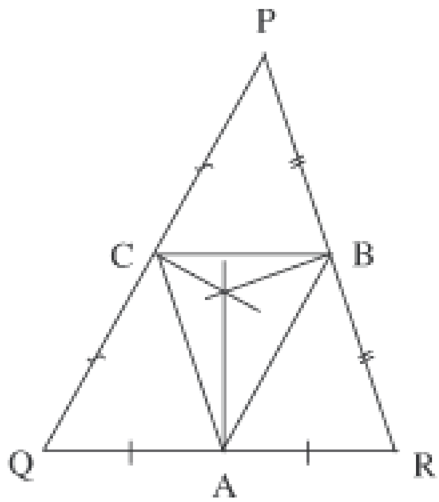
$ACO\square=BDO\square$
, akkor hibásan okoskodnánk, mert nincs jogunk felhasználni, hogy

D
,

O
és

C
is egy egyenesen vannak, amíg azt be nem bizonyítjuk (lásd a 12. példát!).

14. példa. Bebizonyítjuk, hogy bármely háromszög magasságegyenesei egy pontban metszik egymást (3. jegyzet). Induljunk ki abból a tételből, hogy a háromszög oldalfelező merőlegesei egy ponton mennek át (3. ábra); ennek a bizonyítását ismertnek tekintjük.



3. ábra

Az oldalak felezőpontjai háromszöget határoznak meg (ábránkon

$ABC\square$
). Ennek a háromszögnek az oldalai az eredeti háromszögnek középvonalai és így (ismert tétel szerint) párhuzamosak az oldalaival:

$AB\square PQ$
,

$BC\square QR$

,

$CA \parallel RP$
. Ezért a

PQR
háromszög oldalfelező merőlegesei az

ABC
háromszögnek magasságegyenesei (mert átmennek a csúcsain és merőlegesek az oldalaira). De az oldalfelező merőlegesekről már tudtuk, hogy egy pontban metszik egymást. Ebből következik, hogy a magasságegyenesek is egy pontban metszik egymást, hiszen az

$ABC\triangle$
magasságegyenesei ugyanazok az egyenesek, mint a

$PQR\triangle$
oldalfelező merőlegesei (és bármely

$ABC\triangle$
-höz található ilyen tulajdonságú

$PQR\triangle$
).

15. példa. Bebizonyítjuk, hogy bármely négyszög átlói felezik egymást. Legyen

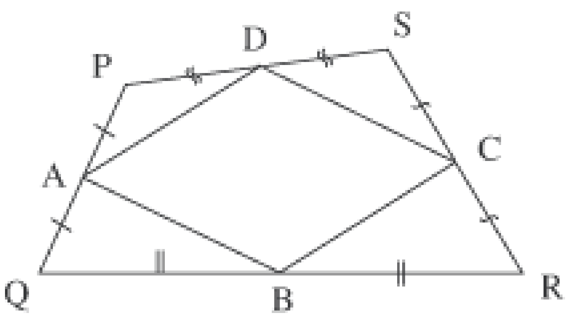
$PQRS$
a szóban forgó négyszög, oldalainak felezőpontjai legyenek

A
,

B
,

C
és

D
(4. ábra).



4. ábra

A háromszögek középvonalaira vonatkozó ismert tétel szerint (lásd a 14. példát!)

$AB \parallel PR$
és

CD□PR
, tehát

AB□CD
. Hasonlóképpen látható be, hogy

BC□AD
. Az

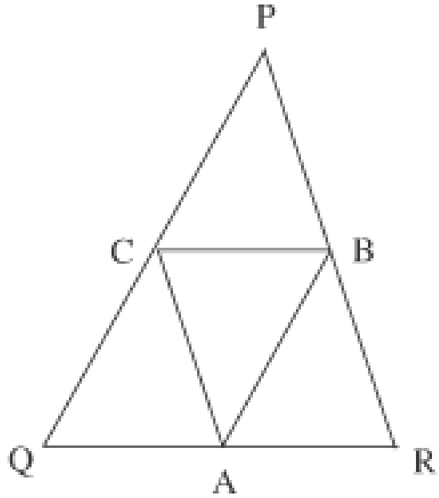
ABCD
négyszög tehát paralelogramma. A paralelogrammákról azonban ismeretes, hogy az átlóik felezik egymást. Tehát bármely négyszög átlói felezik egymást.

Nem helyes! – mondja az olvasó. Abból, hogy a paralelogrammák átlói felezik egymást, miért következnék, hogy bármely négyszög átlói felezik egymást? Ami igaz egy speciális esetben, nem feltétlenül igaz minden esetben.

Joggal tiltakozik az olvasó. De tiltakozott-e a 14. példában szereplő bizonyítás ellen is? Ott ugyanis lényegében ugyanezt a hibát követtük el, ha nem is ilyen átlátszó formában.

Ha bebizonyítunk valamit az olyan négyszögekre, amelyeket úgy kapunk, hogy négyszögek oldalait megfelezzük és a szomszédos oldalak felezőpontjait összekötjük; nem állíthatjuk, hogy azt minden négyszögre bebizonyítottuk. Csak akkor volna helyes ez a bizonyítás, ha azt is megmutatnánk, hogy *bármely* négyszög előállítható valamilyen négyszög szomszédos oldalfelező pontjainak összekötésével. Ez azonban nem is igaz; csak a paralelogrammák állíthatók elő így.

Ha valamit bebizonyítunk az olyan háromszögekre, amelyeket úgy kapunk, hogy háromszögek oldalfelező pontjait kötjük össze; akkor sem állíthatjuk, hogy azt minden háromszögre bebizonyítottuk. Csak akkor volna helyes ez a bizonyítás, ha azt is megmutatnánk, hogy *bármely* háromszög előállítható valamely háromszög oldalfelező pontjainak összekötésével. Ez történetesen igaz ugyan – nem úgy, mint a négyszögekre vonatkozó hasonló állítás –, de nem bizonyítottuk be, sőt nem is utaltunk arra, hogy ez még bizonyításra szorul. Rosszhiszeműen elhallgattunk egy lényeges lépést, s így a háromszögek magasságegyeneseinek metszésére vonatkozó bizonyításunk nem fogadható el.



5. ábra

16. példa. Alakítsuk át a 14. példában szereplő bizonyítást: induljunk ki a belső

ABC

háromszögből, és szerkesszük meg hozzá a külső, PQR háromszöget úgy, hogy az A, B, C csúcsokon át párhuzamosokat húzunk a szemközti oldalakkal.

$AB \parallel PQ$, $BC \parallel QR$, $CA \parallel RP$; három paralelogramma keletkezik: $ABCQ$, $ABPC$, $ARBC$. Tudjuk, hogy a paralelogrammák szemközti oldalai egyenlők, például $AB=QC$ és $AB=CP$. Ezekből következik, hogy $QC=CP$. A C pont tehát felezi a PQ oldalt. Hasonlóan látható, hogy B a PR, A pedig a QR

oldalt felezi.

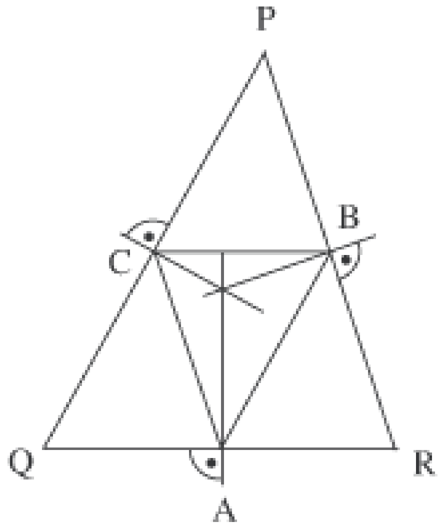
Kiderült, hogy az

ABC
háromszöget a

PQR
háromszög középvonalai határolják. (Ezt most előre nem tudhattuk, más volt a kiindulásunk, mint a 14. példában!)

Be akarjuk bizonyítani, hogy az

ABC
háromszög magasságegyenesei egy pontban metszik egymást. Húzzuk meg ezeket az egyeneseket (6. ábra)!



6. ábra

Az előbbiekből következik (tekintetbe véve a középvonalakra vonatkozó, többször említett tételt is), hogy ezek a magasságegyenesek a

PQR
háromszögnek oldalfelező merőlegesei. De az oldalfelező merőlegesekről tudjuk (ismertnek tettük fel), hogy egy pontban metszik egymást. Ez tehát a magasságegyenesekre is igaz.

Jó ez a bizonyítás? Kifogástalan. Az

ABC
háromszögből indultunk ki, erre semmi megszorítást nem tettünk, tehát amit bizonyítottunk, az *minden háromszögre* igaz. Azt ugyan nem bizonyítottuk be, hogy bármely háromszög előállítható úgy, mint az itt szereplő külső

PQR
háromszög, de erre nem is volt szükség. A bizonyítás akkor is helyes volna, ha a „körülírással" mindig bizonyos speciális háromszögekhez jutnánk; ha minden háromszög oldalfelező merőlegesei egy pontban metszik egymást, akkor ezekre a speciális háromszögekre is ugyanezt mondhatjuk; az a fő, hogy az

ABC
háromszög nem speciális.

17. példa. Egy „bizonyítást" már láttunk arra, hogy

1 mm=1000 km
. Most még kettő következik.

Többet bizonyítunk be, mint amit ígérünk: azt, hogy bármely két egyenesszakasz egyenlő hosszú.

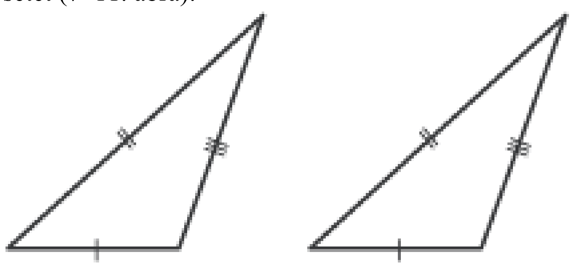
A bizonyításokban ismertnek tekintjük a háromszögekre vonatkozó következő egybevágósági tételeket:

Két háromszög egybevágó, ha megegyeznek

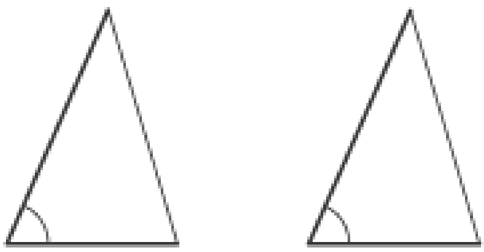
(1) mindhárom oldalukban; (2) két oldalukban és a közbezárt szögükben; (3) két oldalukban és a nagyobbikkal szemben fekvő szögükben; (4) egy oldalukban és két megfelelő helyzetű szögükben.

(A megfelelő helyzetű szögeket így értjük: vagy mindkettő az egyező oldalakon van; vagy az egyik rajtuk, a másik velük szemben és akkor az egyező oldalakon lévő szögek egyenlők, s persze a velük szemben lévők is.)

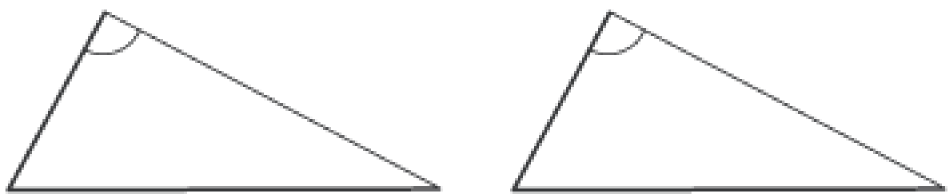
Ábráink mutatják ezt a négy esetet (7–11. ábra):



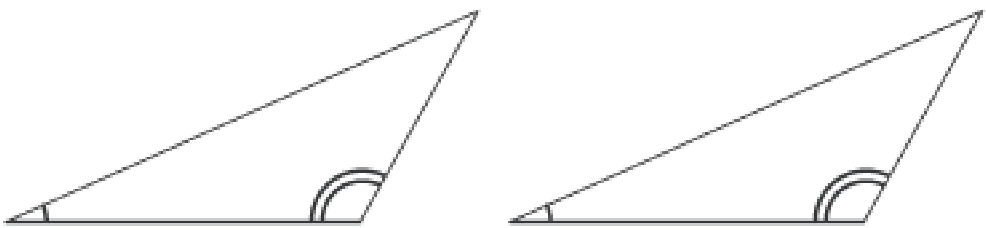
7. ábra



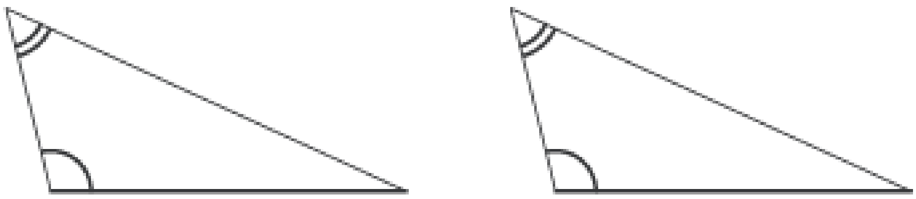
8. ábra



9. ábra



10. ábra

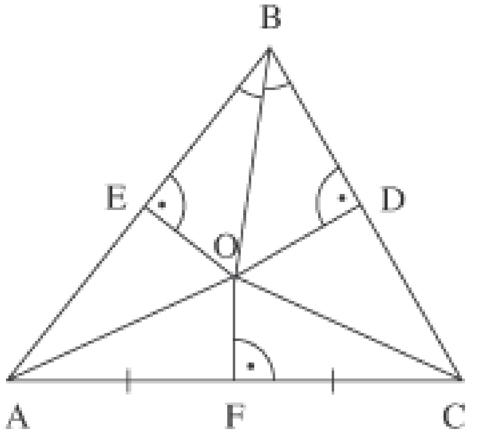


11. ábra

Azt a két tetszőleges szakaszt, amelyeknek az egyenlőségét be akarjuk bizonyítani, helyezzük el úgy, hogy egy háromszög két oldalát alkossák (a 12. ábrán

AB
és

BC
)!



12. ábra

Húzzuk meg a háromszögben az

AC
oldal felezőmerőlegesét és a szemközti szög felezőjét! Ábránkon ezek a háromszögön belül lévő

O
pontban metszik egymást. Kössük össze ezt az

O
pontot

A
-val és

C
-vel, és bocsássunk az

O
-ból merőlegest az

AB
és

BC

oldalra! A jelöléseket az ábra mutatja.

Az *ABC* háromszöget hat kisebb háromszögre bontottuk. Bebizonyítjuk, hogy ezek páronként egybevágóak.

AFO□ □*CFO*□,

mert megegyeznek két oldalukban ($AF=FC$, OF közös) és a közbezárt szögükben (derékszög). Az egybevágóságból következik, hogy $AO=CO$.

BOE□ □*BOD*□,
mert megegyeznek egy oldalukban (BO) és két megfelelő helyzetű szögükben (a B-nél lévő szögek egyenlők, mert szögfelezőt húztunk; E-nél és D-nél derékszögek vannak). Az egybevágóságból következik, hogy $EB=BD$ és $EO=DO$.

AOE□ □*COD*□,
mert megegyeznek két oldalukban ($AO=CO$ és $EO=DO$ az előbbiek szerint) és a nagyobbikkal szemben fekvő szögükben (az E-nél, illetve D-nél lévő derékszögben, 4. jegyzet).

Ennek a két háromszögnek az egybevágóságából következik, hogy

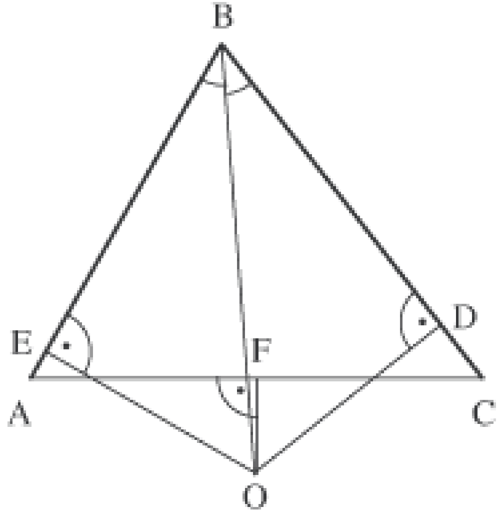
$AE = CD$
. De már az előbb láttuk, hogy

$EB = BD$
, tehát

$AB = CB$
mint két egyenlő szakasz összege.

Ábránkon

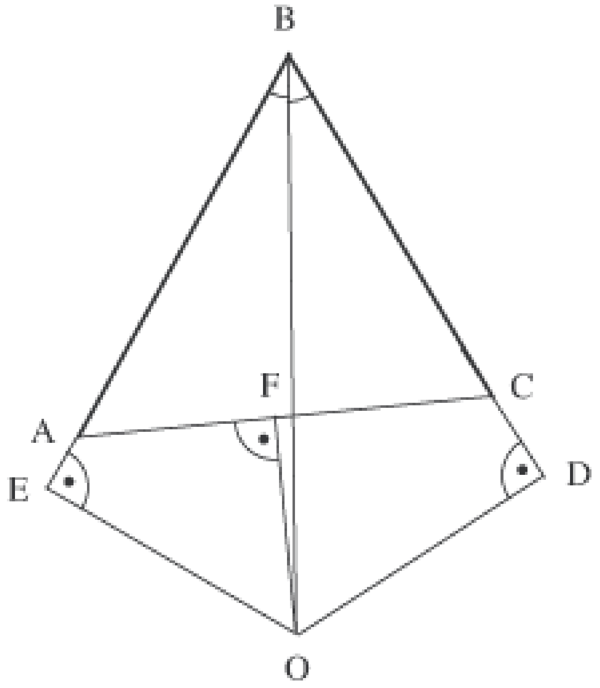
O
a háromszög belsejébe esett. De hátha a háromszögön kívül van (13. ábra)?



13. ábra

A bizonyítás ebben az esetben is szóról szóra ugyanaz.

Az sem változtat a bizonyítás lényegén, ha a metszéspont annyira kinn van, hogy még a merőlegesek talppontjai is az oldalak meghosszabbítására esnek (14. ábra).



14. ábra

A változás csak annyi, hogy akkor végül egyenlő szakaszok kivonásával (nem pedig összeadásával) kapjuk azt, hogy

$AB=BC$
.

Az a határeset is elképzelhető, hogy a metszéspont a háromszög kerületén van. A bizonyítás ekkor még egyszerűbb (az első lépés kimarad, a további gondolatmenet változatlan).

Részletesen végigvizsgáltuk az eseteket, csak egyet felejtettünk ki: azt, amikor a metszéspont kívül van ugyan a háromszögön, *de a belőle húzott merőlegesek nem egyformán viselkednek: az egyiknek a talppontja valamelyik oldalra, a másiké egy oldal meghosszabbítására esik*.

Ez a feledékenység megbosszulta magát: ugyanis éppen ez az az eset, amely *mindig* megvalósul, ha egyáltalán van metszéspontja

AC
felezőmerőlegesének és

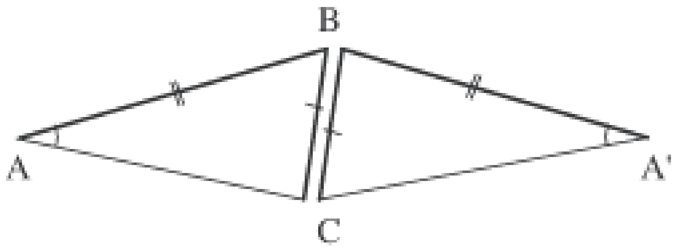
B
szögfelező egyenesének (5. jegyzet).

18. példa. Végül még egy „bizonyítás" következik arra, hogy

1 mm=1000 km
, illetve arra az általánosabb állításra, hogy bármely két szakasz egyenlő.

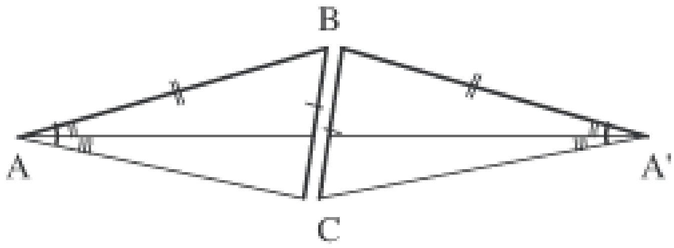
Előkészítésül lássunk egy új tételt a háromszögek egybevágóságára: két háromszög egybevágó, ha megegyeznek két oldalukban és az egyikkel szemközti szögükben.

Ez a tétel a fent (3) alatt ismertetett egybevágósági tétel kiterjesztése. Ott azt a megszorítást tettük, hogy az egyező szögek a nagyobbik oldallal legyenek szemközt; most elejtjük ezt a megszorítást.



15. ábra

Illesszük össze a két háromszöget úgy, ahogy a 15. ábra mutatja! A vastagon kihúzott, egyformán jelölt szakaszok, illetve szögek egyenlők.



16. ábra

Kössük össze az egymásnak megfelelő

A
és

A'
csúcsot (16. ábra)! Az

ABA'
háromszög egyenlő szárú, tehát az ábrán két ívvel jelölt szögek egyenlők. Ha ezeket az egy ívvel jelölt egyenlő szögekből kivonjuk, ismét egyenlő szögeket kapunk (az ábrán ezek háromívesek). Az

ACA'
háromszögben két szög egyenlő, egyenlők tehát a velük szemközti oldalak is:

$AC=A'C$
.

Így hát az

ABC
háromszög és az

$A'BC$
háromszög mindhárom oldalukban megegyeznek, és ezért egybevágók.

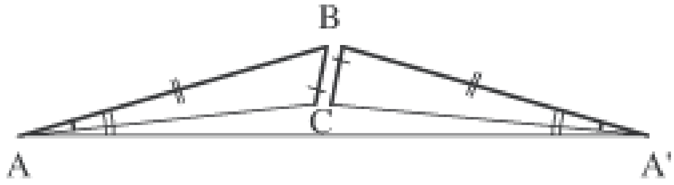
Most az egyíves szögekből ki kellett vonni a kétíves szögeket. Más esetben az

ACA'
háromszög egyenlő szögeit úgy kapjuk, hogy az

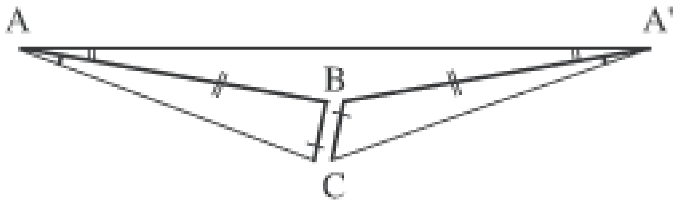
ABA'
háromszög egyenlő szögeiből vonjuk ki (17. ábra), vagy szögeihez adjuk hozzá (18. ábra) az eredeti két egyenlő szöget. Az most semmiképpen sem fordulhat elő, hogy az egyik esetben összeadásra, a másik esetben kivonásra

legyen szükség, mint a 17. példában.

Arra a megállapításra jutottunk tehát, hogy minden megszorítástól mentesen igaz: ha két háromszög megegyezik két oldalában és valamelyikkel szemközti szögében, akkor egybevágóak.



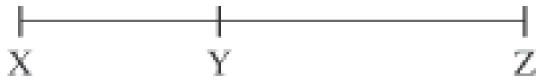
17. ábra



18. ábra

Ennek alapján könnyű bebizonyítani, hogy bármely két szakasz egyenlő. Illesszük a két szakaszt egymás mellé úgy, hogy egy egyenesre essenek és egyetlen közös pontjuk legyen (a 19. ábrán

Y
)!



19. ábra

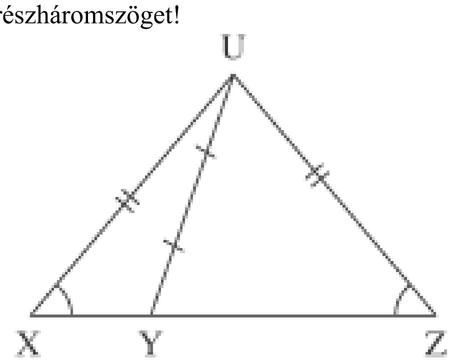
Szerkesszünk egy olyan egyenlő szárú háromszöget, amelynek az alapja a két szakasz egyesítéséből származó

XZ
szakasz! Kössük össze az egyenlő szárú háromszög

U
csúcsát az

Y
ponttal (20. ábra)!

Hasonlítsuk össze a keletkező két részháromszöget!



20. ábra

Megegyeznek két oldalukban ($UX=UZ$, UY közös) és az egyikkel szemközti szögükben (az egyíves szögekben: egyenlő szárú háromszög alapszögei). Tételünk szerint tehát egybevágóak, és így harmadik oldalukban is megegyeznek: $XY=YZ$, bármit mutat is az ábra. Bármely két szakasz egyenlő, $1\ \mathrm{mm}=1000\ \mathrm{km}$.

Ha azonban tételünkből ilyesmi következik, akkor tételünkkel – és így a bizonyításával is – valami baj van. Hol a hiba? Addig minden rendben van, hogy $CAA'\sphericalangle=CA'A\sphericalangle$, mint két egyenlő szög összege vagy különbsége. Ebből azonban nem következik, hogy $AC=A'C$.

Hogyan, hát lehetséges, hogy egy háromszögben két szög egyenlő, és a velük szemközti oldalak mégsem egyenlők?

„Közönséges” háromszögben ez nem lehetséges, de egyenesbe lapuló, „elfajuló” háromszögben igen: a 21. ábra alapján elképzelhetjük, hogy két egyenlő 0°-os szöggel szemközt ebben a határesetben különböző oldalak is lehetnek.

Itt van a rés okoskodásunkban, ezért nem érvényes az az állítólagos „új matematikai tétel”, hogy két oldalukban és valamelyikkel szemközti szögükben megegyező háromszögek mindig egybevágóak. Lehetnek egybevágóak, de lehetnek különbözőek is. Ha megegyeznek ezekben az adataikban, és mégsem egybevágóak, az csak úgy lehetséges, hogy a mondott módon illesztve őket össze, a harmadik oldalpár (amelyekről nem tettük fel, hogy egyenlőek, ábráinkon AC és $A'C$) egy egyenesbe esik. Éppen ilyen eset látható a 20. ábrán; ott XYZ az az elfajuló háromszög, amelynek két szöge egyenlő ugyan (0°-os), a velük szemközti oldalak (XY és YZ) mégsem egyenlőek.

21. ábra

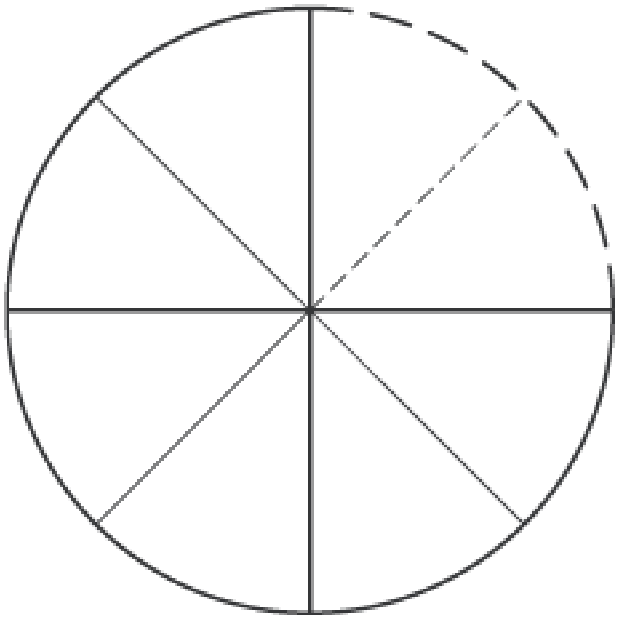
1. JEGYZETEK

1. Miért helyesek az ilyen átalakítások:

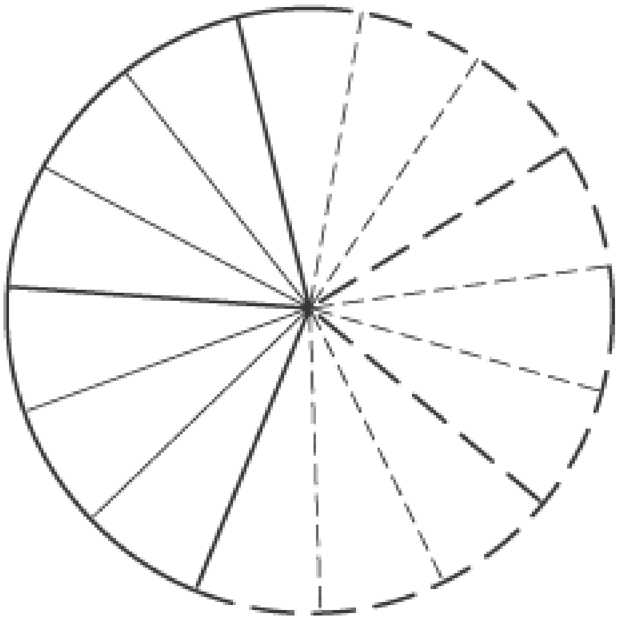
(2□3/2□4)=(3/4) , (2□3/3□5)=(2/5) , (2□6/6□5)=(2/5) , (2□7/7□5)=(2/5) ,
és miért hibásak az ilyenek:

(23/24)=(3/4) , (23/35)=(2/5) , (26/65)=(2/5) , (27/75)=(2/5) ,
még ha az eredményük néha véletlenül jó is? A rajzok mutatják két példán, miért helyesek az előbbiek:

(2□3/2□4)=(6/8)=(3/4) , (2□3/3□5)=(6/15)=(2/5) .



22. ábra



23. ábra

Az első esetben az egészet

2□4
, vagyis 8 egyenlő részre osztottuk és

2□3
, vagyis 6 ilyen részt vettünk. Ez ugyanannyi, mintha 4 egyenlő részre osztjuk az egészet, és 3 ilyen részt veszünk. Általában, ha egy tört számlálóját és nevezőjét ugyanazzal a számmal osztjuk (vagy szorozzuk), akkor a tört értéke ugyanakkora marad.

A következő sorban nem ezt tettük! Ott az egyező számjegyek elhagyása nem osztás, hanem valami más:

(23/24)
-ből úgy lesz

(3/4)
, hogy a számlálójából és a nevezőjéből is kivonunk 20-at; amikor

(23/35)
helyett

(2/5)
-öt írunk, akkor a nevezőjéből kivontunk 30-at, a számlálójából viszont csak 3-at vontunk ki, de azáltal, hogy utolsó előtti számjegy helyett utolsó számjegy lett a 2-ből, még 10-zel is elosztottuk azt, ami a 3 kivonásakor maradt; így a többinél is. Érthető, hogy ezek a bonyolult átalakítások csak kivételesen vezetnek ugyanakkora törtre. Ilyen kivételes törtek még

(16/64)
és

(19/95)
is.

2. A hiba gyökere itt is ugyanaz, mint a 2. és 3. példában: a többjegyű számok másként viselkednek, mint a számjegyeikből alkotott szorzatok vagy összegek. Ott kétjegyű számokat akartunk úgy kezelni, mintha szorzatok lettek volna (egyszerűsíteni egy-egy számjegy kihúzásával), itt az összegekkel akartunk úgy számolni, mintha többjegyű számok lettek volna. Pedig hiába jó például ez a kivonás:

52 14□ 38

Nem lehet alkalmazni a módszerét ebben az esetben:

5+2 1+4□.

Fenn úgy tudjuk elvenni a 4-et, hogy az 5 tízesből egyet felváltunk és a 12-ből veszünk el 4-et. Lenn ilyenről szó sem lehet: nincs 5 tízes, csak 5 egység, nincs mit felváltani.

3. Magasságegyeneseknek nevezzük a háromszög csúcsain áthaladó, a szemközti oldalakra merőleges egyeneseket. Magasságvonalaknak is szokás nevezni ezeket, de a „magasságvonal” elnevezést azokra a szakaszaikra is szokták használni, amelyek a csúcsoktól a szemközti oldalakig vagy meghosszabbításukig terjednek.

4. Honnan tudjuk, hogy

AO
nagyobb, mint

EO
(és így

CO
is nagyobb, mint

DO
)? Onnan, hogy bármely háromszög két különböző szöge közül a nagyobbikkal szemben nagyobb oldal van, mint a kisebbikkel szemben. (Ezt az ismert tételt itt felhasználjuk.) Hát azt honnan tudjuk, hogy az

AOE
háromszögben az

E
-nél lévő szög nagyobb, mint az

A
-nál lévő szög? Onnan, hogy az

E
-nél lévő szög derékszög (mert

O
-ból merőlegest bocsátottunk az

AB
oldalra), s akkor a másik két szög csak kisebb lehet nála. Miért? Mert a háromszög szögeinek összege

180□
, s ha egy szög

90□
, akkor a másik kettőre együttvéve jut csak

90□

.

5. Ha ugyanis a háromszögben

$AB=BC$

, akkor a felezőmerőleges egybeesik a szögfelező egyenessel, nincs metszéspont. Ha viszont

$AB \neq BC$

, akkor abból a feltevésből, hogy mindkét merőleges talppontja az oldalakra, vagy mindkettő az oldalak meghosszabbítására esik, ellentmondásra jutunk (tudniillik arra, hogy

$AB=BC$

, pedig abból indultunk ki, hogy

$AB \neq BC$

). Tehát a fent vázolt bizonyítások hibásak, ha arra akarunk következtetni belőlük, hogy bármely két szakasz egyenlő, de adódik belőlük egy helyes bizonyítás az említett talppontok helyzetére vonatkozólag. Ez a bizonyítás indirekt, vagyis azt bizonyítjuk be, hogy az állítás ellenkezője nem igaz. Direkt bizonyítást is lehet adni arra, hogy az említett talppontok mindig így helyezkednek el.

TUD-E ÖN FEJBEN ÖTÖDIK GYÖKÖT VONNI?

SURÁNYI JÁNOS

Kedves „Ön”, aki ezt a könyvet kézbe vette, nyilván úgy gondolja, a fejezet címéhez hallgatólag hozzáértendő: „…mert ha nem, majd megtanítom”. Elgondolása valóban helyes, de előbb még egyszerűbb számolási ügyességekkel ismerkedünk meg.

1. 1. NÉHÁNY SZÁMOLÁSI KÖNNYÍTÉS

Az alapműveleteket ma már megtanuljuk az általános iskola alsó tagozatában (ha nem is minden erőlködés nélkül), a XV. században azonban az osztást csak az egyetemek felső évfolyamaiban tanították. A lényeges különbséget az okozza, hogy akkor még a római számokat használták, az alapműveletek elvégzése pedig csak a mai, helyiértéken alapuló számírásunk mellett válik egyszerűvé. Könnyen fogalmat alkothatunk arról, hogy mekkora könnyítést jelent ez a számírás, amit ma már elég magától értetődőnek tartunk, ha megpróbáljuk kiszámítani mondjuk a

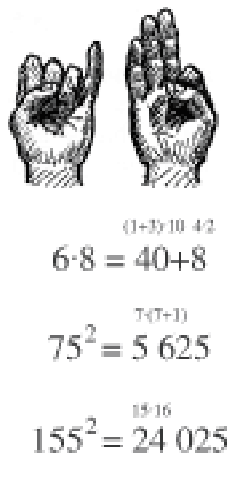
MMDCCXLIII□CCCXCVII

szorzatot (természetesen anélkül, hogy átírnánk arab számjegyekkel). Érthető hát, hogy régebben számos fogás, ügyesség alakult ki a számolás megkönnyítésére. Néhány példa:

a) Ha csak

5□5

-ig tudjuk az egyszeregyet, már össze tudunk szorozni két, 5-nél nagyobb számjegyet a következő módon: Egyik kezünkön kinyújtunk annyi ujjat, amennyivel az egyik tényező több, mint 5, a másikon annyit, amennyivel a másik több 5-nél. Ezután annyi tízeshez, mint a kinyújtott ujjak együtt, hozzáadjuk az egyik és a másik kézen behajtott ujjak számának szorzatát, és megkapjuk a keresett szorzatot.



1. ábra

b) Egy szám négyzetét többféleképpen is kiszámíthatjuk. Ha a szám 5-re végződik, akkor számolhatunk a következő módon is: az 5-ös elhagyásával maradó számot megszorozzuk a következő számmal, és az eredmény után még 25-öt írunk.

c) Egyes népeknél használatos a szorzás ún. orosz módszere, amelynél csak ismételt kétszerezésre és felezésre van szükség (ezek kényelmetlen számírások mellett is, amilyen pl. a római, elég könnyen elvégezhetők). Egymás mellé írjuk a két számot, majd az egyiket (célszerű a nagyobbikat) kétszer vesszük, a másikat felezzük, ha van egy maradék, azt elhagyjuk. Ezt ismételjük addig, míg a felezéssel 1-ig nem jutunk. Ekkor az ismételt duplázással keletkezett oszlopban áthúzzuk azokat a számokat, amelyek mellett a másik oszlopban páros szám áll; a többi összege megadja a keresett szorzatot.

58	~~249~~
29	498
14	~~996~~
7	1992
3	3984
1	$\underline{7968}$

$58 \cdot 249 = 14442$

Számos hasonló fogás és érdekesség ismeretes még, csak illusztrációként említettünk beköszöntőben néhányat. Az olvasó kipróbálhatja az eljárások helyességét példákon. Magyarázatukra még visszatérünk, egyelőre azonban foglalkozzunk egyszerűbb műveletekkel: az összeadással és kivonással!

2. 2. ÖSSZEADUNK ÉS KIVONUNK EGYSZERRE

Több összeadást egyszerre tudunk elvégezni, amint azonban kivonásra is van szükség, minden kivonandót külön kell levonni. Ha ügyesek vagyunk, segíthetünk magunkon annyiban, hogy az összeadandók összegezése után összegezzük a kivonandókat, és összegüket vonjuk le az első összegből.

Egy egyszerű ötlettel azonban még a 3 műveletet is egyetleneggyel helyettesíthetjük. Írjuk az összeadandókat és kivonandókat vegyesen egymás alá, ahogy az összeadásnál szokás, csak – mivel a műveletet számjegyenként végezzük – a kivonandók minden jegyét külön jelöljük meg, pl. föléhúzással. (A 0 jegyeket fölösleges föléhúzni.) Ezután ugyanúgy számolhatunk az „egyesek”-től kezdve oszloponként, mint az összeadásnál, csak a föléhúzott számjegyeket mindig levonjuk.

Előfordulhat, hogy a jegyek összeadása közben a föléhúzott jegy a nagyobb, ha pl. a

$\overline{8}$

, 3 jegyek következnek egymás után. Ekkor 5 kivonandó marad a következő jegyre, a kivonandók elé; megkülönböztetésül az összeadandóktól a „mínusz” szócskát szokás mondani, tehát számolásunk: mínusz 8, mínusz 5. Ezt a következő jegyből levonjuk, ill. ha az is föléhúzott jegy, pl.

$\overline{7}$

, akkor a kivonandó növekszik, tehát mínusz 12 a következő részeredmény. Az is lehet, hogy erre az eredményre egy oszlop végén jutottunk. Ekkor leírjuk az „egyes” helyiértékű jegyet:

$\overline{2}$

-t az eredmény megfelelő helyére, a tízeseket, tehát esetünkben mínusz 1-et átvisszük a következő oszlopba, éppen úgy, mint összeadásnál, de ezúttal természetesen kivonandóként. Ezeket a kivonandó számokat *negatív számnak* szokás nevezni. Néhány példa:

$$\begin{array}{rrr} 769 - 2021 + 953 + 1868 - 1070 = & +\ \ 769 & = \ \ 769 \\ & -2021 & \bar{2}0\bar{2}\bar{1} \\ & +\ \ 953 & 953 \\ & +1868 & 1868 \\ & -1070 = & \bar{1}0\bar{7}0 \\ \hline & & 499. \end{array}$$

$$\begin{array}{r} 342 \\ \bar{7}\bar{7}\bar{8} \\ 231 \\ \bar{5}\bar{4}\bar{9} \\ \underline{1322} \\ 1\bar{4}\bar{3}\bar{2} \end{array}$$

A második feladat egyes oszlopaiban (alulról felfelé haladva) pl. a következő műveletsorokat végezzük: 2, mínusz 7, mínusz 6, mínusz 14, mínusz 12; leírunk $\bar{2}$-t, marad mínusz 1, 1, mínusz 3, 0, mínusz 7, mínusz 3; leírunk $\bar{3}$-t stb.

Hogyan értsük azonban a vegyesen előforduló közönséges (szakkifejezéssel pozitív) és kivonni való, negatív számjegyeket? Úgy, ahogy a szólés érti: „Az egyik tizenkilenc, a másik egy híján húsz”, azaz most bevezetett írásmódunkban $19 = 2\bar{1}$. Ez egyben utal arra is, hogyan lehet átalakítani általában a fellépő negatív számjegyeket.

Kiválasztunk egyet, célszerű balról az elsőt, ez előtt pozitív jegy áll, a megelőző jegyből egyet elveszünk, azt 10 kisebb egységre váltjuk, és ebből vonjuk le a föléhúzott számjegyet. Ezután jobbra haladva, sorra kiküszöbölhetjük a többi föléhúzott számjegyet is:

$$1\bar{4}\bar{3}\bar{2} = 6\bar{3}\bar{2} = 57\bar{2} = 568.$$

Ha a föléhúzott jegy előtt 0 áll, akkor ebből nem tudunk mit fölváltani, ezért balra megyünk az első 0-tól különböző jegyig, azt eggyel csökkentve, egy egységet apróra váltunk, abból 9-et leírunk, egyet ismét felváltunk; ha jobbról újabb 0-jegy következik, akkor ismét leírunk 9-et és 1 egységet tovább váltunk, míg a föléhúzott jegyig nem érünk. Ekkor ezt a jegyet vonjuk le az utolsó felváltásból származó 10 egységből:

$$100\,\bar{7}0\bar{4} = 99\,30\bar{4} = 99\,296.$$

Ha az első számjegy negatív, akkor még ha csupa pozitív kilences szám áll is utána, az nem adhat annyit, mint az első jeggyel írt kivonandó, tehát ez esetben a kivonandók összege nagyobb, mint a kisebbítendőké.

3. 3. SZÁMKITALÁLÁS

Ismeretes rengeteg számkitaláló játék, ezek közül bemutatunk kettőt.

a) Gondolj egy számot, szorozd meg 6-tal, vegyél el belőle 3-at, vedd a dupláját, add össze a számjegyeket, amit kaptál, vedd 3-szor, add össze a számjegyeket, ha az eredmény többjegyű, annak is add össze a jegyeit és ismételd, amíg egyjegyű eredményt nem kapsz, szorozd meg ezt a számot 4-gyel, adj hozzá 13-at! Ha jól számoltál, 49-et kaptál. Növelhetjük a hatást azzal, hogy előre felírjuk egy lapra a végeredményt, és lefordítva letesszük, a végén pedig csak fölemeljük: „Itt az eredmény”.

b) Írj le egy számot, írj ugyanezekkel a jegyekkel, csak más sorrendben egy másikat, és vond ki a nagyobbikból a kisebbet! Csináld ezt meg több számmal is, és mondd meg a végeredményeket:

801, 17 612, 2574, 295 576, 19 998.
Nézd meg a számításodat, a második és negyedik kivonásban biztosan hibáztál!

Vajon honnan lehet ezt tudni? Hiszen nem ismerjük a számokat, amiket a felszólított kivont, ill. amiből az a) esetben a számítások kiindultak. Mindkét esetben a „ 9-es próba” adja a magyarázatot, amellyel az első cikkben ismerkedett meg az olvasó, belátva, hogy egy szám 9-cel osztva annyi maradékot ad, mint amennyit a számjegyeinek összege ad.

Ezt a tényt felhasználják számítások ellenőrzésére is, ugyanis összeg, különbség és szorzat maradéka 9-cel osztva ugyanannyi, mint az egyes számok maradékai összegének, különbségének, illetve szorzatának az osztási maradéka, és ugyanez áll a hatványozásra is, hiszen a hatványozás ismételt szorzás csupa egyenlő tényezővel.

Ennek alapján könnyen kiszámíthatjuk egy művelet vagy műveletsor eredményének osztási maradékát az egyes számok maradékából, és meghatározhatjuk a végeredmény maradékát. Ha a kettő nem egyezik, akkor hiba van a számításban. (De hiba akkor is lehet, ha a 9-cel való osztás maradékával nincs baj.)

A 3.a) pontban leírt játék ezek alapján így érthető: A 6-tal való szorzáskor 3-mal osztható számot kaptunk, ez a 3 elvételével és a kétszerezéssel sem változott meg. (Ezek a lépések egyébként lényegtelenek a feladat szempontjából.) A 3-mal való szorzás után már 9-nek egy többszörösét kapjuk, így a jegyek ismételt összeadásával végül egy egyjegyű és 9-cel osztható számhoz jutunk. Ilyen csak a 9 és a 0, de a számjegyek összege 0 nem lehet, tehát a jegyek összeadásának eredménye csak 9 lehet. A további műveletek eredményét már ismerjük, és így olyan eredményt alakíthatunk ki, amilyent éppen akarunk.

A b) pontban a különbség két tagja ugyanazt a maradékot adja 9-cel osztva, mert ugyanaz a számjegyek összege (ugyanazokból a jegyekből áll mind a két szám), így a különbség osztható 9-cel, tehát kell, hogy a számjegyeinek összege is 9-cel osztható legyen. Ez nem teljesül a második és negyedik eredményre, tehát ezek biztosan hibás számításból eredtek. Megjegyzendő azonban, hogy a másik három számítás is lehet hibás, hiszen sok olyan számolási hiba lehetséges, amely az eredményt 9-cel osztható számmal változtatja meg, és ekkor a számjegyek összege 9-cel osztható marad.

4. 4. A 11-ES PRÓBA

Majdnem ilyen egyszerű eljárás adható a 11-gyel való osztás maradékának megállapítására is: Ha összeadjuk az utolsó („egyes” helyiértékű) jegytől minden második jegyet, és ebből az összegből levonjuk a kimaradó jegyek összegét, a különbség ugyanannyi maradékot ad 11-gyel osztva, mint az eredeti szám, és itt is igaz, hogy összeg, különbség, szorzat, hatvány maradéka ugyanannyi, mint a maradékok összegéé, különbségéé, szorzatáé, ill. a maradék hatványáé. (Ha a kivonandó nagyobb, akkor hozzáadunk a kisebbítendőhöz 11-et vagy annak egy többszörösét, ezzel a 11-gyel való osztás maradéka nem változik.) Ennek alapján „ 11-es próbával” is megvizsgálhatjuk, nem mutatkozik-e valami számolási hiba. Az olvasó megpróbálhatja önállóan megindokolni az eljárást, de megtalálhatja bármelyik, a számelmélet elemeivel foglalkozó könyvben.

A 3.b) eljáráshoz hasonlóan a következőt játszhatjuk: leíratunk egy négy- vagy hatjegyű (általában páros számú jeggyel írt) számot és alá azt a számot, amelyik ebből úgy keletkezik, hogy az első jegyét elhagyjuk és utolsó jegynek írjuk. A két számot összeadatjuk, ez mindig osztható lesz 11-gyel. Mivel a 11-gyel való osztás maradékát a számjegyekből gyorsan ki tudjuk számítani, ez ismét módot ad hibák felfedezésére anélkül, hogy ismernénk az összeadandókat.

5. 5. ÖTÖDIK GYÖKÖT VONUNK FEJBEN

El lehet sajátítani azonban mindezeknél nagyobb számoló művészetet is, például fejben ötödik gyököt vonni bármelyik, legfeljebb 10-jegyű ötödik hatványból. Ez elsősorban azon alapul, hogy minden szám ötödik hatványa ugyanarra a jegyre végződik, mint a hatványozott szám. Ha a szám nem több, mint ötjegyű, akkor ezzel már megvan az ötödik gyöke, mert a legkisebb kétjegyű szám ötödik hatványa:

105=100 000
már hatjegyű, így ennél kevesebb jeggyel írt szám csak egyjegyű számnak lehet az ötödik hatványa.

Ha a szám 5-nél több jegyű, akkor meg kell még határozni a tízes helyiértékű jegyet. Nem lehet ugyanis a szám kettőnél több jegyű számnak az ötödik hatványa, mert a legkisebb háromjegyű szám ötödik hatványa:

1005=10 000 000 000
már 11-jegyű.

A (jobbról) második jegyet megállapíthatjuk, ha tudjuk az egyjegyű számok ötödik hatványáról, hogy hány jegyűek, és milyen jegyekkel kezdődnek. Alább felírtuk az első 9 szám ötödik hatványát, de ezekből elég az első két jegyet és a jegyek számát tudni.

$$
\begin{aligned}
1^5 &= 1\\
2^5 &= 32\\
3^5 &= 243\\
4^5 &= 1\,024\\
5^5 &= 3\,125\\
6^5 &= 7\,776\\
7^5 &= 16\,807\\
8^5 &= 32\,768\\
9^5 &= 59\,049
\end{aligned}
$$

A 10, 20, 30, …, 90 számok ötödik hatványa ezeknél 5-5 jeggyel hosszabb.

Ennek alapján az a szám, amelyiknek az ötödik hatványa

229 345 007
csak 7-re végződhet. Továbbá 40 és 50 közé kell esnie, mivel a felírt szám 9-jegyű és 22-vel kezdődik. Így a felírt szám csak 47-nek lehet az ötödik hatványa, és az valóban ennyi is. Hasonlóan

2 535 525 376
10-jegyű és 25-tel kezdődik, tehát 70 és 80 közé eső és 6-ra végződő számnak, vagyis 76-nak lehet csak az ötödik hatványa, és 76 ötödik hatványa valóban ennyi.

6. 6. AZ ÖTÖDIK HATVÁNY UTOLSÓ SZÁMJEGYE

Vajon miért végződik egy egész szám ötödik hatványa ugyanazzal a számjeggyel, mint maga a szám? A fenti táblázatból látjuk, hogy ez az egyjegyű számokra igaz. Ebből azonban már következik minden egész számra is. Ugyanis a hatványozás ismételt szorzás egyenlő tényezőkkel. Egy szorzat utolsó jegyét a tényezők utolsó jegyének a szorzata adja. Több tényező szorzatánál a végeredmény utolsó jegyét az összes tényező utolsó jegyei szorzatának az utolsó jegye adja. Ha öt egyenlő tényezőt szorzunk össze – vagyis ötödik hatványra emelünk –, akkor a végeredmény utolsó helyén a hatványozott szám utolsó jegye ötödik hatványának utolsó számjegye áll. Azt azonban már tudjuk, hogy ez megegyezik a hatványozott számjeggyel.

7. 7. VÉDEKEZZÜNK A CSAPDÁK ELLEN!

Igen egyszerű és gyors módszert találtunk az ötödik gyökvonásra. Nem is kell hozzá pontosan ismernünk a hatványt, csak az első két számjegyet, az utolsót és a számjegyek számát. Ez azonban nehézséget is rejt magában. Ha valaki próbára akar tenni, vagy egyszerűen hibázott a számolásban, mi az elmondottak alapján akkor is „kitaláljuk”, hogy – ebben az esetben természetesen hibásan – minek az ötödik hatványát diktálták fel. Pl.:

460 065 624.
Módszerünk azt adja, hogy ez csak 54-nek lehet az ötödik hatványa, ha egész szám ötödik hatványa. Azonban

545=459165024
(bár ezt fejben gyorsan aligha tudjuk kiszámítani), s így a diktált szám nem ötödik hatványa egész számnak. A hibalehetőség csökkentésére ismét a 9-es és 11-es próbához fordulhatunk. Az utóbbival különösen egyszerű az ellenőrzés, ugyanis az 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 számok ötödik hatványa 11-gyel osztva rendre 1, 10, 1, 1, 1, 10, 10, 10, 1, 10 maradékot ad, míg a 11-gyel osztható számnak természetesen az ötödik hatványa is osztható 11-gyel. Így mindenekelőtt azt kell megnézni, hogy a diktált szám 11-gyel osztva milyen maradékot ad. Ha ez nem 0, 1 vagy 10, akkor nem ötödik hatvány a feldiktált szám. Ha a három maradék valamelyike adódik, akkor miután alkalmaztuk eljárásunkat, még újra megnézhetjük, hogy a kapott szám 11 szerinti maradékának az ötödik hatványa valóban az előbbi maradékot adja-e.

A 9-es szabály is felhasználható ellenőrzésre, bár ez egy kicsit bonyolultabb, mert meg kell jegyezni az 1–9 számok ötödik hatványának 9-cel való osztási maradékát. Ezek:

a szám	1	2	3	4	5	6	7	8	9
az 5. hatvány maradéka	1	5	0	7	2	0	4	8	0

Itt a 3 és 6 kivételével minden előfordulhat, ezért először azt jó meghatározni, milyen számnak lehet a diktált szám ötödik hatványa, és azután nézni meg, hogy a diktált és a kitalált szám osztási maradéka valóban a táblázat szerint megfelel-e egymásnak (amit könnyű fejben tartani).

8. 8. TÖKÉLETESÍTJÜK TUDOMÁNYUNKAT

Ezt a táblázatot azért is írtuk fel, mert ennek fejben tartásával tovább kápráztathatjuk a hallgatóságot.

– Tessék diktálni egy ötödik hatványt!

–

20…

– Megálljunk! Lehet a további jegyeket felcserélt sorrendben is diktálni.

–

20 111 459.

Most elvesztettük a legnagyobb segítséget, az utolsó számjegy ismeretét. Annyit mindenesetre tudunk, hogy a diktált szám 8-jegyű és 20-szal kezdődik, tehát 20 és 30 közti szám ötödik hatványa. Azt is megállapíthatjuk, hogy a diktált szám 5-öt ad maradékul 9-cel osztva, tehát a keresett szám 2-t, s így csak 20 vagy 29 lehet, de 20 ötödik hatványa öt 0-t tartalmaz (és 7-jegyű), tehát a keresett szám 29. (Valóban ennek ötödik hatványa:

$20\,511\,149$
.)

Ennél az eljárásnál nehézséget okoz, ha a diktált szám 9-cel osztható. Ilyenkor ugyanis az adott 10-es számközben mindegyik 3-mal osztható szám lehet a keresett szám. Egy kis gyakorlattal azonban ilyenkor is megbecsülhetjük az első két számjegy ismeretében, hogy a szóba jövő számok közül melyikről lehet szó. Pl. Az első két jegy 13, a jegyek nagyság szerint: 011223369; a szám 40 és 50 között van, egész közel 40-hez és 3-mal osztható, tehát 42. (Valóban:

$42^5=130\,691\,232$
.) Sok fogással lehetne még finomítani az eljárást, de gondolom, hogy az olvasó sem vágyik elsőosztályú számolóművésszé válni. Ehhez általában külön adottságok és rendkívül sok gyakorlás kell. Mindezek azonban csak kápráztató mutatványokra képesítenek, hasznos tevékenységre alig.

9. 9. SZÁMOLÁSI EGYSZERŰSÍTÉSEK MAGYARÁZATA

Sikerült-e megtalálni az 1. pontban említett számolási eljárások magyarázatát?

α
). Legkönnyebben az utolsóét találhatjuk meg. Ha egy szorzat egyik tényezőjét felezzük, a másikat viszont kétszerezzük, akkor a szorzat értéke nem változik. Az 1. pont számításának első két sorára

$58\cdot 249=29\cdot 498$
. A következő sorban viszont bonyolultabb a helyzet. Tulajdonképpen 28-nak vettük a felét, s így

$28\cdot 498=14\cdot 996,$
tehát a maradék elhagyása miatt a 29 mellett álló 498-cal kisebb lesz a két új szám szorzata, mint az előbbi kettőé:

$29\cdot 498=14\cdot 996+498.$

A következő lépésekben hasonlóan:

$14\cdot 996=7\cdot 1992=3\cdot 3984+1992=1\cdot 7968+3984+1992.$
Így valóban

$58\cdot 249=498+1992+3984+7968.$
Általában is ez a helyzet: valahányszor egy páratlan szám helyett az eggyel kisebb szám felét írjuk le, ezzel a páratlan szám mellett álló számmal kisebb lesz az új két szám szorzata, mint a régi kettőé; így ezeket mind hozzáadva az utoljára duplázott számhoz, kapjuk a kiinduláskor leírt két szám szorzatát.

β
) Az 1.b)-nél általánosabban kiszámíthatjuk pl. a

$62\cdot 68$
szorzatot úgy, hogy a

$6 \cdot 7$
szorzat után leírjuk a

$2 \cdot 8$
szorzatot.

$$62 \cdot 68 = \overbrace{4\,2}^{6 \cdot 7}\overbrace{16}^{2 \cdot 8}$$

Általában alkalmazható az eljárás két olyan szám szorzására, amelyek csak az utolsó jegyben különböznek, és az utolsó jegyek 10-re egészítik ki egymást. Ilyenkor az utolsó jegy elhagyásával maradó számot szorozzuk az 1-gyel nagyobb számmal, és utána írjuk az utolsó jegyek szorzatát, az utóbbit mindig 2-jegyű számnak írva, tehát az

$1 \cdot 9$
-et 09 alakban. Pl.:

$$141 \cdot 149 = \overbrace{210}^{14 \cdot 15}\;\overbrace{09.}^{1 \cdot 9}$$

Az eljárás helyességét a számítás átalakításával láthatjuk be az utolsó előtti példa esetében pl. így:

$62 \cdot 68 = 62 \cdot 60 + 62 \cdot 8 = 62 \cdot 60 + 60 \cdot 8 + 2 \cdot 8.$
Az első két összeadandóban a 60-at kell 62-szer és még 8-szor venni, így

$62 \cdot 68 = 60 \cdot (62+8) + 2 \cdot 8 = 60 \cdot 70 + 2 \cdot 8 = 6 \cdot \cdot 7 \cdot 100 + 2 \cdot \cdot 8 =$

$=4200+16=4216.$
Az átalakításokat éppen a két tényező fönt leírt egyszerű kapcsolata tette lehetővé, így hasonló átalakítások minden, a mondott tulajdonságú szorzatnál elvégezhetők.

Az algebra nyelvén kifejezve, itt az

$(n+k) \cdot (n+l) = n \cdot (n+k+l) + k \cdot l$
azonos átalakítást használjuk fel, ahol

n
a tízeseket jelenti, vagyis

$10t$
alakú és

$k+l=10$
. Így

$n \cdot (n+k+l) = 10t(10t+10) = t \cdot (t+1) \cdot 100$
. Ez az azonosság számos további szorzási egyszerűsítésre ad lehetőséget olyan szorzatoknál, amelyeknél

n
, vagy

$n+k+l$
, vagy – mint esetünkben – a kettő szorzata több 0-ra végződik.

γ) Az 1.a) számítás helyességét a legkevésbé egyszerű látni. Az ott szereplő szorzat a felnyújtott ujjak számának feltüntetésével:

$(5+1)\cdot(5+3)$, a leírt kiszámításmód pedig:

$(1+3)\cdot 10+(5-1)\cdot(5-3)$. Más a szorzatoknál az „1” és „3” helyébe kerül más számú ujj. Általában

k-val és

l-lel jelölve a felnyújtott ujjak számát, azt kell belátni, hogy az

$(5+k)\cdot(5+l)$ szorzás mindig ugyanarra az eredményre vezet, mint a

$(k+l)\cdot 10+(5-k)\cdot(5-l)$ számítás. Egyszerű átalakításokkal, amelyeket konkrét számokkal már az előbbiekben is használtunk, az első szorzat így alakítható át:

$$(5+k)\cdot(5+l)=(5+k)\cdot 5+(5+k)\cdot l=5\cdot 5+k\cdot 5+5\cdot l+k\cdot l,$$

a második kifejezés pedig így:

$$(k+l)\cdot 10+(5-k)\cdot(5-l)=k\cdot 10+l\cdot 10+(5-k)\cdot 5- -(5-k)\cdot l=k\cdot 10+l\cdot 10+5\cdot 5-k\cdot 5-(5\cdot l-k\cdot l)= =k\cdot 10+l\cdot 10+5\cdot 5-k\cdot 5-5\cdot l+k\cdot l= =k\cdot 5+l\cdot 5+5\cdot 5+k\cdot l.$$

A két aláhúzott kifejezés csak a tagok sorrendjében és egyes tagokban a tényezők sorrendjében különbözik. Tudjuk azonban, hogy sem az összeadandók, sem a szorzótényezők felcserélése nem változtat a művelet eredményén, így a kétféle számítás eredménye mindig megegyezik.

10. 10. SZORZÁS RÉSZLETSZORZATOK NÉLKÜL

Említettem, hogy a számolás művészetéhez is bizonyos adottságok kellenek. Ezek különböznek a matematika iránti érzéktől, sőt általában nem is járnak azzal együtt. Nem egyszer különben gyenge szellemi képességekkel párosulnak. Vannak azonban tiszteletreméltó kivételek is. A századfordulón egy *Ferrol* nevű számolóművész[2] pl. azonkívül, hogy újra felfedezte a szorzás – egy igen régen ismert – olyan módját, amely mellett részletszámítások leírása nélkül mindjárt a végeredményt írhatjuk, azt is leírta – és ez előbb nem volt ismeretes –, hogyan lehet ebből egy részletszámítások nélküli osztási eljárást nyerni. A következőkben ezeket az eljárásokat ismertetjük meg.

A $724\cdot 586$ szorzat példáján illusztráljuk, hogyan végezhető el a szorzás úgy, hogy csak a végeredményt írjuk le. Természetesen itt a követhetőség kedvéért a fejben végzett mellékszámítások egy részét is leírjuk.

[2] P. Maennchen: *Geheimnisse der Rechenkünstler*, B. G. Teubner, Leipzig – Berlin, 1913. 33–39. old.

A szorzat utolsó (egyes helyiértékű) jegyét a tényezők egyes helyiértékű jegyeinek szorzatából kapjuk:

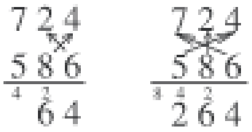
6□4=24

. A 4 a keresett utolsó jegy, és maradt két tízes.

(A nyíl az összeszorzott jegyeket köti össze.) Az átviendő tízeseket az eredmény tízes jegyének helye fölött tüntettük föl, hasonlóan tüntetjük fel a további lépéseket is.

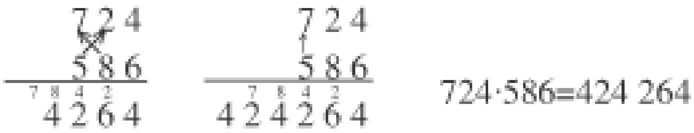
2+8□4+6□2=46

. Így a második jegy 6 és maradt 4 százas. (2. ábra.)



2. ábra

Hasonlóan számíthatjuk a további jegyeket is, ezek számítását külön-külön lépésekben tüntetjük fel. (3. ábra).

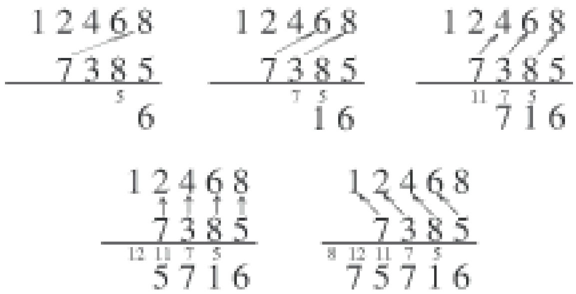


3. ábra

Könnyíthetjük a számítást azzal, ha a szorzó jegyeit fordított sorrendben írjuk a szorzandó alá, így ugyanis nem ellenkező irányban kell haladnunk, amikor az összeszorzandó jegyeket keressük. Lássuk ezt a

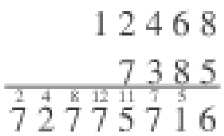
12 468□5837

szorzat példáján:



4. ábra

és hasonlóan haladva tovább:



Ha több számot kell ugyanazzal a számmal megszorozni, érdemes a közös szorzó jegyeit egy külön papírszelet (alsó) szélére írni fordított sorrendben, és azután úgy tolni ezt a papírszeletet a szorzandók felett, egy-egy jeggyel balra, hogy az összeszorzandó jegyek mindig egymás alatt álljanak.

11. 11. OSZTÁS MELLÉKSZÁMÍTÁSOK

NÉLKÜL

Ez az eljárás meg is fordítható, és egy olyan osztási eljáráshoz vezet, amelynél szintén lehet részletszámítások nélkül csak a hányadost írni, még akkor is, ha azt sok tizedesjegyre kell kiszámítanunk.

Fordítsuk meg először első szorzásunkat: osszuk el

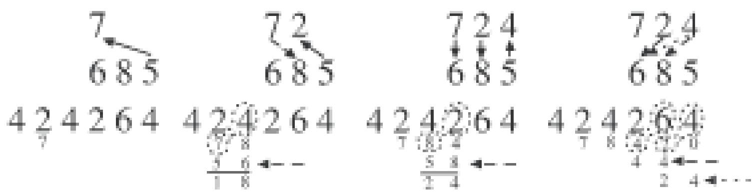
424 264
-et 586-tal. Az osztót az osztandó fölé írjuk, de a könnyebbség kedvéért mindjárt a második szorzás mintájára fordított sorrendben írjuk a jegyeket; a hányadost efölé írjuk. A hányados nyilván 3-jegyű lesz (ti. ha egész szám; különben 3 jegye lesz a tizedesvessző előtt).

A szorzás utolsó lépésében a két tényező legmagasabb helyiértékű jegyét szoroztuk össze – tehát az 5-öt a keresett hányados első jegyével –, és hozzáadtuk az előző lépés maradékát, így kaptuk a szorzat – adott esetben – első két jegyéből álló számot: 42-t; ezt kell tehát osztanunk 5-tel, és megkapjuk a hányados első jegyét. Azonban nem célszerű 8-at venni hányadosul, csak 7-et, mert az előző lépésben két jegypár szorzatához adtunk még maradékot, valószínűtlen tehát, hogy ebből csak 2-t kellett volna átvinni. Ekkor az utolsó előtti lépésből

42–7□5=7
az átvitelből adódott, tehát az utolsó előtti lépésben 74-et kaptunk.

Ez úgy keletkezett, hogy a

7□8
szorzathoz hozzáadtuk a hányados még ismeretlen, második jegyének 5-szörösét és a megelőző lépés maradékát.



5. ábra

Ez a két összadandó tehát együtt

74–7□8=18
. Ebben az 5 megvan 3-szor, de célszerű lesz 2-t venni következő jegynek, mert 3 túl kicsi lenne az előbbi átvitelből származó maradéknak. 2-t választva

18–2□5=8
-at vittünk át az előző részletszámításból, és 2-t írtunk le, tehát 82 volt az előbbi részletszámítás eredménye, ami

7□6
-ból,

2□8
-ból, 5-ször a hányados harmadik jegyéből és az előző lépésből áthozott számból adódott. Így

82–7□6–2□8=24
ötödéből számítandó a hányados következő jegye. Ez már a hányados egyes helyiértékű számjegye, így megpróbálhatunk 4-et választani, különösen, ha várható, hogy a hányados egész szám lesz. Ekkor, mivel az előbbi részletszámítás 46-ot adott, abból

2□6+4□8
-at elvéve kapjuk, hogy az előző részletszámítás eredménye 24, ez megegyezik az osztó és a hányados utolsó jegyeinek szorzatával,

6□4
-gyel, tehát megtaláltuk a pontos hányadost.

12. 12. NEGATÍV SZÁMJEGYEK SEGÍTSÉGÜL VÉTELE

Az eljárás egyik gyöngéje, amint láttuk az, hogy a hányados jegyeinek meghatározása elég bizonytalan. Nem mindig sikerül eltalálni, mennyit célszerű az előző részletszámításból áthozott összeadandóra hagyni. Ezen a bizonytalanságon csökkenthetünk, ha arra gondolunk, hogy az osztó közel van a 600-hoz, ezért első jegyét joggal vehetjük közel 6-nak. A következő jegyből való áthozatot is megbecsülhetjük a megállapítandó új számjegy birtokában, s ha ez túl nagynak ígérkezik, idejében kisebbre vehetjük az új számjegyet.

Aki azonban nem riad vissza attól, hogy negatív jegyekkel is számoljon, amelyekkel már a 2. pontban találkoztunk, annak olyan esetekben sem kell újra számolnia, ha valahol túl nagy jegyet választott. Ami ilyenkor marad, azzal a további számot csökkentenie kell: „negatív áthozat" marad. Ennek a tízszereséhez is hozzáveheti a következő pozitív jegyet, s egy negatív részeredményt kap. Ebben negatív számszor van meg az osztó, tehát a következő számjegy negatív lesz (az ilyeneket most is föléhúzással jelöljük). A következő részeredmény lehet pozitív is, negatív is (célszerű nem nagy negatív részeredményre törekedni); és folytathatjuk az eljárást úgy is, ha vegyesen használunk pozitív és negatív számjegyeket is. A végén a föléhúzott számjegyeket kiküszöbölhetjük a 2. pontban megismert módon.

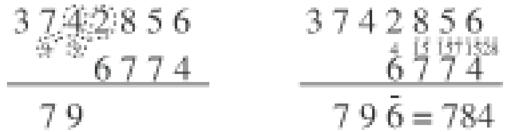
Lássunk erre is egy példát (most már az osztandó alá fogjuk írni az osztót – megfordítva a számjegyek sorrendjét – és az alá a hányadost).

3 742 856
-ot osszuk el 4776-tal! Az osztó első jegyét célszerű felfelé kerekítve 5-nek venni. A hányadosnak 3 egész számjegye lesz. Az 5 a második lépésben

94–7□7=45
-ben 9-szer van meg. A következő lépésben

94–7□7–9□4=9
folytán

92–7□7–9□7=–20
-ból kell az újabb jegyet megállapítani.



6. ábra

Ezt vegyük

–6=6¯
-nak, hogy 4-szeresét levonva elég nagy pozitív maradékot kapjunk a további számjegysorozatok (6. ábra) levonására. (Ha azonban még így is negatív maradékra jutunk, az sem baj, annyi mindenesetre látható lesz már, hogy mennyivel kell a hányadost csökkenteni pozitív és az osztónál kisebb maradék eléréséhez.) Így az előző

áthozatra

$92-7\cdot 7-9\cdot 7-(-6)\cdot 4=4$
, majd

$48-7\cdot 6-9\cdot 7-(-6)\cdot 7=-15$
jut. A következő részeredmény tehát

$-150+5=-145$
, és a további számítás:

$-145-9\cdot 6-(-6)\cdot 7=-157$
,

$-1564-(-6)\cdot 6=-1528$
.

Ezzel azt kaptuk, hogy

$3\,742\,856=4776\cdot 79\bar{6}+\bar{1}\bar{5}\bar{2}\bar{8}=4776\cdot 784-1528= =4776\cdot 783+3248.$

Ebből az is látható, hogy a kapott 784 érték közelebb van a

$(3\,742\,856/4776)$
hányadoshoz, mint 783. – Tizedes jegyeket utólag is számíthatunk az

$\bar{1}\bar{5}\bar{2}\bar{8}$
-at vagy a 3248-at a fenti módon (vagy más úton) tovább osztva.

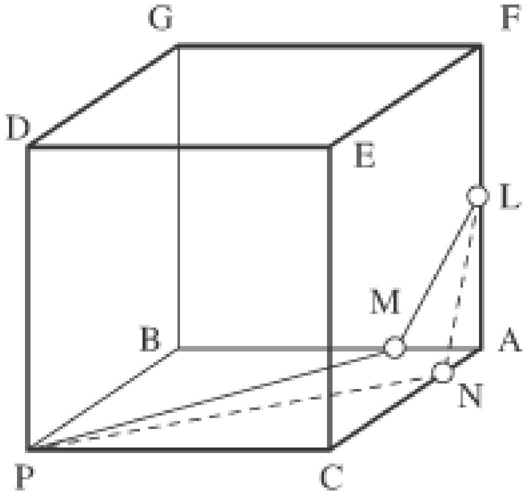
Ami az utolsó három pont eljárásait illeti, a szorzás gyakorlatilag is jól használható, ha kezdetben lassúbbnak is bizonyul, mint a részletszorzatok leírása. Ne feledjük el, hogy az utóbbi eljárást hány éven keresztül gyakoroltuk az iskolában! Az új eljárás mégis már rövid gyakorlás után is eléri a régi gyorsaságát. Kellő gyakorlással az osztás is gyorssá és biztossá tehető, ez azonban már sokkal kevésbé éri meg a hozzá szükséges fáradságot, inkább csak mint ügyesség, „mutatvány” érdekes.

PÓK ÉS LÉGY

(GEODETIKUS VONALAK)

MOLNÁR FERENC

Egy kocka alakú szoba egyik sarkában ül a pók, a tőle legtávolabbi oldalél közepén van a légy (1. ábra). A pók meg akarja fogni a legyet; minthogy azonban nem tud repülni, a falon mászva kénytelen eljutni a légy tartózkodási helyéhez. Erre sokféle lehetőség van. Eljuthat a légyhez pl. úgy, hogy végig a falak határán, azaz a kocka élein mászik, vagy pl. úgy, hogy útja során a szoba több falán (falak közé beleértve a mennyezetet és a padlót is), azaz a kocka több lapján áthalad, és az éleket csak keresztezi. Felvetődik a kérdés, hogyan kell a póknak haladnia, ha a legrövidebb úton akarja elérni a legyet? Ha tudna repülni, akkor nyilván egyenes vonalban repülve jutna el a legrövidebb úton a céljához. Repülni azonban nem tud, így a falon mászva kell megkeresnie a legrövidebb utat. A pók feladata tehát az, hogy a szoba falain, vagyis a kocka felületén haladó és a légyhez vezető utak közül megkeresse a legrövidebbet.



1. ábra

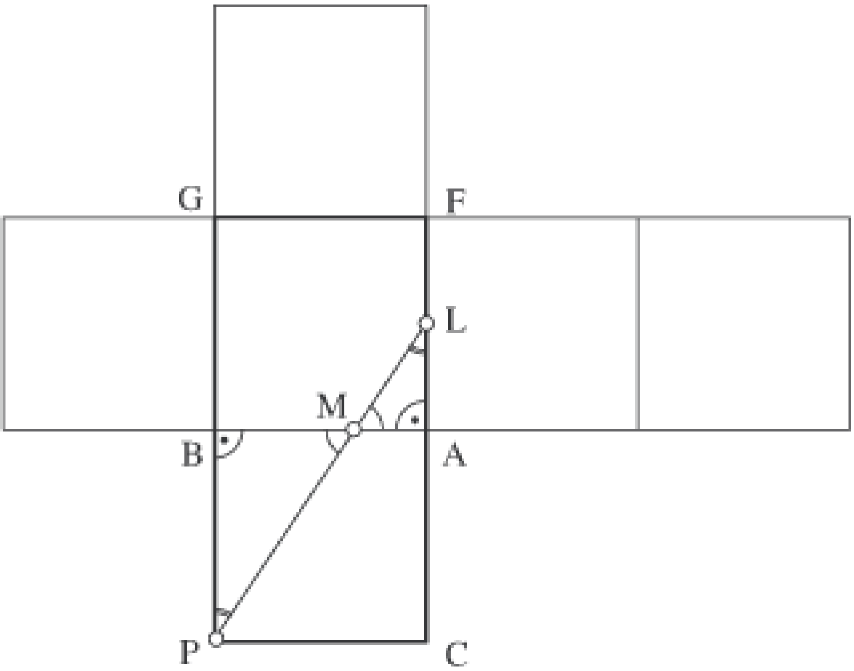
Mielőtt megoldanánk a pók feladatát, nézzünk egy egyszerűbb feladatot! Ha a légy a szobának azon a falán lenne, amelynek egyik sarkában ül a pók, akkor a pók egy síklapon mászva elérheti a legyet. Ekkor nyilván egyenes vonalban kell haladnia, hiszen két pont között a legrövidebb út az egyenes szakasz. A mi esetünkben nem ilyen egyszerű a helyzet, hiszen a pók és a légy nincs a kockának ugyanazon a lapján. Láthatjuk viszont, hogy a kockának két szomszédos (azaz élben találkozó) lapján vannak (pl. az 1. ábrán a

PCAB
és

BAFG
lapon), így a legrövidebb út csak két ilyen szomszédos lapon át vezethet. Készítsük el a kocka felületének olyan, síkba kiterített hálózatát, ahol a

PCAB
és

BAFG
lapok szintén szomszédosak, és jelöljük meg rajta a pók és a légy helyét (2. ábra)!



2. ábra

Így síkbeli ábrához jutottunk, ahol két pontot összekötő legrövidebb vonal az egyenes. A hálózaton tehát a pók és a légy helyét jelölő pontokat (

P
és

L
) összekötő egyenes szakasz mutatja a pók legrövidebb útját. Ha most hálózatunkból ismét kockát készítünk, a hálózaton rajzolt vonalakból a kocka felületén haladó vonalakat kapunk. Nyilvánvaló, hogy a rajzolt vonalak ugyanolyan hosszúak maradnak, hiszen eközben a lapon nem nyúltak és nem zsugorodtak. A pók és a légy helyét a hálózaton összekötő egyenes szakaszból tehát a hálózat összeillesztése után megkapjuk a legrövidebb utat, amelyen a pók eljuthat a légyhez. Ez az út már nem egyenes, hanem, amint azt az 1. ábra mutatja, két egyenes szakaszból álló, úgynevezett törött vonal. A 2. ábráról azt is leolvashatjuk, hogy ez a legrövidebb út a kocka

AB
élét abban az

M
pontban keresztezi, amely ezt az élt

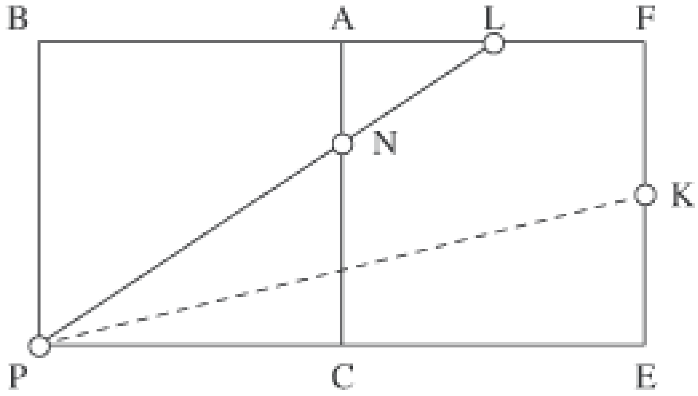
1:2
arányban osztja.3

A pók azonban egy másik utat is választhat, ha a legrövidebben akarja elérni a legyet. Az 1. ábrán berajzolt szaggatott törött vonal mutatja ezt az utat, amely az előbbitől annyiban különbözik, hogy másik két szomszédos lapon át halad, és az

AC
élt keresztezi az

1:2
arányban osztó

N
pontban. Ez az út az előbbivel egyenlő hosszú, hiszen ugyanolyan nagyságú szakaszokból áll (3. ábra). Ezek után felvetődik a kérdés, hogy más szomszédos lapokon át nem juthat-e el a pók ezeknél rövidebb úton is a légyhez?



3. ábra

3Az

LAM
háromszög ugyanis hasonló a

PBM
háromszöghöz (2. ábra), hiszen megfelelő szögeik egyenlők (az

A
és

B
csúcsnál derékszögek vannak, az

M
csúcsnál lévő szögek csúcsszögek, a

P
és

L
csúcsnál lévő szögek váltószögek). Minthogy pedig az

AL
szakasz fele a

BP
szakasznak, a hasonlóság aránya

1:2
, tehát az

AM
szakasz is fele a megfelelő

BM
szakasznak, vagyis az

M
pont az

AB
szakaszt

1:2
arányban osztja.

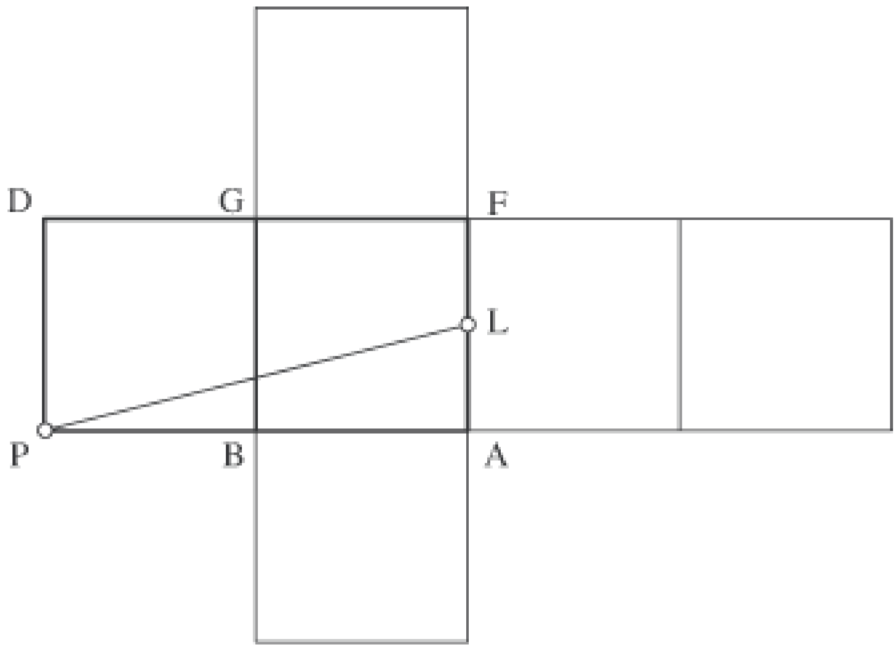
A szomszédos

$PBGD$
és

$BAFG$
lapokon át vezető legrövidebb utat pl. a 4. ábrán látható hálózatba berajzolt egyenes szakasz adja a hálózat összeillesztése után. Könnyű meggyőződni arról, hogy a 4. ábrán szereplő

PL
szakasz hosszabb, mint a 3. ábra

PL
szakasza.



4. ábra

A 4. ábra

PL
szakaszával ugyanis egyenlő hosszú a 3. ábra

PK
szakasza, és ez láthatóan nagyobb az ottani

PL
szakasznál. Aki ismeri Pythagorasz tételét, egyszerű számolással is meggyőződhetik erről. Ha ugyanis a kocka éle

a
hosszúságú, akkor a 3. ábra

PLB
derékszögű háromszögéből Pythagorasz tétele alapján:

$PL=\sqrt{PB^2+BL^2}=\sqrt{a^2+\left(\frac{3a}{2}\right)^2}=\sqrt{a^2+\frac{9a^2}{4}}=\frac{a}{2}\sqrt{13}$, míg a

PKE
derékszögű háromszögből:

$PK = \sqrt{(PE2+EK2)} = \sqrt{((2a)2+((a/2))2)} = = \sqrt{(4a2+(a2/4))} = (a/2)\sqrt{(17)}$.

$\sqrt{(17)}$
viszont nagyobb

$\sqrt{(13)}$
-nál, tehát a 3. ábra

PK
szakasza nagyobb a

PL
szakasznál.

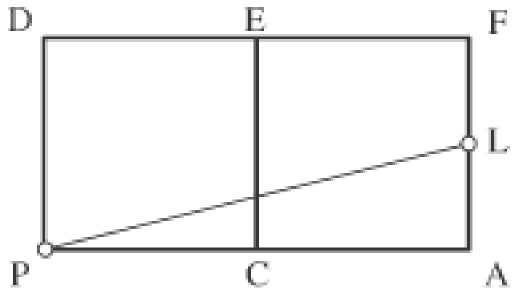
Az 1. ábráról látható, hogy a pók eljuthatna még a légyhez az eddigieken kívül a

BAFG
és

PBGD
lapokkal szemközti szomszédos lapokon keresztül is; az ezeken át vezető legrövidebb út azonban ugyanolyan hosszú, mint amelyik a

BAFG
és

PBGD
lapokon át vezet (5. ábra).



5. ábra

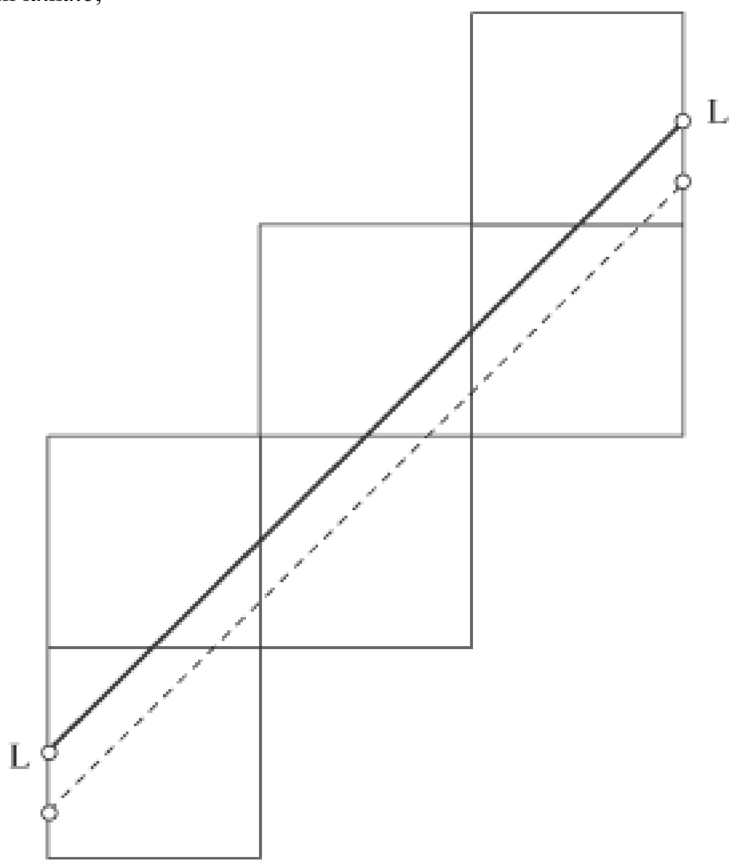
Az 1. ábrán berajzolt törött vonal tehát valóban a pók legrövidebb útja. Figyeljük meg azonban, hogy ez az út csak a kocka felületén haladó utak között a legrövidebb, és nyilván hosszabb a két pontot összekötő egyenes szakasznál. Ha tehát a pók repülhetne, akkor a rövidebb úton is elérhetné a legyet. erre a figyelemre méltó tényre rövidesen visszatérünk még. Miután a pók feladatát megoldottuk, menjünk egy lépéssel tovább! Tegyük fel, hogy a pók, miután megfogta a legyet, a lehető legrövidebb úton végig akarja járni a kockának mind a *hat* lapját úgy, hogy visszatérjen kiindulási helyére, az egyik él felezőpontjába. Kérdés, hogy milyen úton kell haladnia? Az előzőek alapján könnyen megoldhatjuk ezt a feladatot is. Ehhez keresnünk kell a kocka felületének olyan síkba kiterített hálózatát, amelynek határvonalán a pók kiindulási pontja, azaz az 1. ábra

L
pontja, két helyen is előfordul, mégpedig úgy, hogy összeillesztés után ez a két pont a kocka felületén

egybeessék. A hálózat két

L
pontját összekötő egyenes szakasz adja meg összeillesztés után a keresett legrövidebb utat. Azonnal látható, nem mindegy, hogy milyen hálózatot készítünk. Feltételünk ugyanis az volt, hogy a pók a kockának mind a hat lapján át jusson vissza kiindulási helyére. Ezért szükséges, hogy a hálózaton a két

L
pontot összekötő egyenes szakasz mind a hat lapon áthaladjon, továbbá az is, hogy az említett egyenes szakasz végig a hálózaton belül haladjon, egyetlen darabja se legyen a hálózaton kívül. Csak ebben az esetben kapunk ugyanis összeillesztés után a kocka felületén haladó zárt vonalat. Az egyetlen, feltételeinknek megfelelő kockahálózat a 6. ábrán látható;



6. ábra

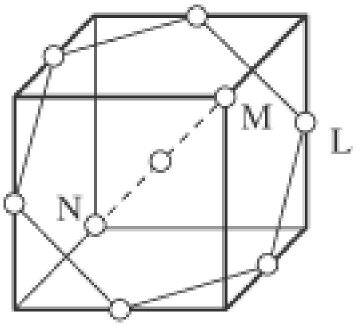
a két

L
pontot összekötő egyenes szakasz mutatja a pók legrövidebb útját. A hálózat összeillesztése után a legrövidebb útvonal hat egyenlő hosszúságú szakaszból álló törött vonal lesz, amely a kocka éleit

45□
-os szögben metszi (7. ábra).4

4Ez a legrövidebb útvonal egy, a kocka felületén elhelyezkedő szabályos hatszög, pontosabban a kockának

MN
testátlójára merőleges síkban lévő szabályos hatszög metszete.



7. ábra

A 6. ábrán az is látható, hogy ha a pók kiindulási helye nem az él középpontja, hanem egy tetszőleges másik pontja, akkor is ugyanolyan hosszú az általa bejárt legrövidebb út. Egy ilyen utat a hálózaton az

LL
szakasszal párhuzamos szakasz ábrázol, amely nyilván egyenlő hosszú az

LL
szakasszal, hiszen annak eltolásával keletkezett. Térjünk vissza most a pók eredeti feladatához, és nézzük meg kissé részletesebben, hogy tulajdonképpen milyen feladatot is oldottunk meg. A feladatot – a póktól és légytől függetlenül – úgy fogalmazhatjuk meg, hogy a kocka felületének két pontját összekötő, a kocka felületén haladó vonalak közül a legrövidebbet keressük. Sőt, továbbmenve, felvethetünk egy sokkal általánosabb problémát is: vegyünk fel egy tetszőleges (nem feltétlenül síklapokból álló) felületen két pontot, és keressük meg azt a vonalat, amely a két pontot a felületen összekötő vonalak között a legrövidebb. Az ilyen vonalat a felület két pontját összekötő *geodetikus vonalnak* nevezzük. Eljutottunk tehát a geometria egyik szép és sok esetben igen nehéz feladatához, felületek geodetikus vonalainak a megkereséséhez. Ha arra gondolunk, hogy a síkban két pontot összekötő legrövidebb vonal az egyenes szakasz, akkor láthatjuk, hogy tetszőleges felületen a geodetikus vonalak ugyanezt a szerepet töltik be.

A geodetikus vonal fogalmának fenti meghatározásában igen lényeges az a kikötés, hogy az a felület két pontját a *felületen összekötő* vonalak között a legrövidebb. Ha ezt a kikötést elhagyjuk, akkor egyszerűen azt mondhatjuk, hogy a két pontot összekötő vonalak közül az egyenes szakasz a legrövidebb. Ez az egyenes szakasz azonban általában nincs a felületen, amint azt már a pók és légy problémájának tárgyalásakor is láttuk. Előfordulhat természetesen, hogy a felület két pontját összekötő egyenes szakasz a felületen van (ilyen pl. a kocka valamelyik lapjának két pontját összekötő szakasz): ebben az esetben természetesen ez az egyenes szakasz a két pontot a felületen összekötő vonalak között is a legrövidebb, azaz geodetikus vonal. A fentiekből mindjárt levonhatjuk azt a következtetést, hogy a felületeken elhelyezkedő egyenesek geodetikus vonalak. A kocka példája mutatja, hogy vannak olyan felületek, amelyeken egyenes vonal húzható. A kockához hasonlóan találhatunk egyenes szakaszokat más síklapokból összetett felületeken is. Ezeket *poliéderfelületeknek* nevezik: ilyen pl. a hasáb és a gúla felülete. Vannak azonban a nem síklapokból álló felületek között is olyanok, amelyeken egyenes található. A legismertebbek ezek közül a hengerpalást és a kúppalást: ezeknek az ún. alkotói egyenesek, vagyis ezeknek a felületeknek geodetikus vonalai. Általában viszont nem mondhatjuk, hogy a felületek geodetikus vonalai egyenesek. Ezért, ha a felület két pontját a felületen összekötő legrövidebb vonal nem egyenes szakasz, akkor az nyilván nem lesz a legrövidebb az összes (persze nemcsak a felületen haladó) vonalak között, amelyek a két pontot összekötik, hiszen ezek között az egyenes szakasz a legrövidebb. Ezt már a pók és a légy példájában is láttuk. Az elmondottakat talán még világosabbá teszi a következő példa: egy terem első sorában ülő legalacsonyabb ember nem feltétlenül a legalacsonyabb a teremben ülő összes ember között, hiszen egy másik sorban ülhet nála alacsonyabb ember is. Hasonló dologról van szó, amikor két pontot egy felületen összekötő vonalak, illetve amikor a két pontot összekötő összes vonal között keressük a legrövidebbet.

Említettük már, hogy a különböző felületek geodetikus vonalainak megkeresése általában nem könnyű feladat. Mégis bizonyos felületeken egyszerűen meghatározhatjuk a geodetikus vonalakat, ha még egyszer figyelmesen megnézzük a pók feladatának megoldását. Ezt a feladatot azért tudtuk aránylag könnyen megoldani, mert a kocka felületét síkba terítettük ki, és felhasználhattuk, hogy síkban két pont között a legrövidebb út az egyenes. Ebből már látható, hogy ha egy felület síkba kiteríthető (szokás azt is mondani, hogy síkba fejthető), akkor a felület két pontját összekötő legrövidebb felületi vonalat a következőképpen kaphatjuk meg: a síkba kiterített felületen berajzoljuk a két pontot összekötő egyenes szakaszt, és ez a felület visszahajlítása során a keresett legrövidebb vonalba megy át. A síkba kiterített felületen berajzolt egyenesek lesznek tehát visszahajlítás után az eredeti felület geodetikus vonalai.

Az említett eljárást alkalmazhatjuk minden, síklapokból öszszetett felületre, azaz a poliéderekre, hiszen ezek síkba kiteríthetők. Azonban más felületeket is tudunk síkba fejteni. Ilyenek pl. a hengerpalást és a kúppalást, ezeknek a geodetikus vonalait is könnyen meg tudjuk tehát határozni. Ha az egyenes körhenger[5] palástján felvett két pontot (

A
és

B
) összekötő legrövidebb vonalat akarjuk meghatározni, akkor eljárhatunk úgy, hogy felvágjuk a hengerpalástot az egyik ponton átmenő alkotó mentén. A kiterített hengerpalást egy téglalap lesz (amelynek alapja a henger alapkörének kerülete, magassága a henger magassága): itt megrajzoljuk az

A
és

B
pontokat összekötő egyenes szakaszt, ebből kapjuk meg visszahajlítás után a két pontot a hengerpaláston összekötő legrövidebb vonalat. A kapott vonalat a

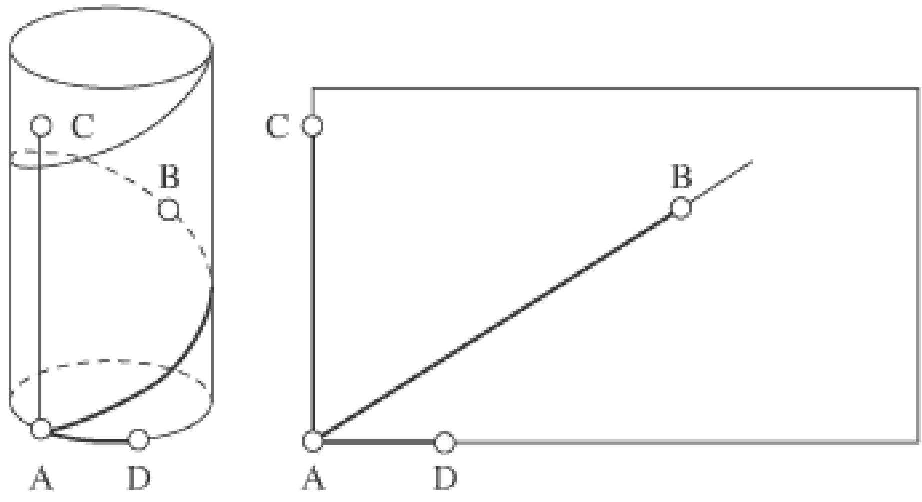
B
ponton túl is megrajzolhatjuk, ha előzőleg a síkbeli ábrán meghosszabbítjuk az

AB
szakaszt. Attól függően, hogy az

A
és

B
pontok hol helyezkednek el a hengerpaláston, különböző geodetikus vonalakhoz juthatunk.

[5] Egyenes körhengernek nevezzük az olyan hengert, amelynek alapja (vezérvonala) kör, alkotói a kör síkjára merőlegesek.



8. ábra

Ha a két pont nincs ugyanazon az alkotón, és nincs rajta a henger alapkörén vagy egy azzal párhuzamosan elhelyezkedő körön, akkor az

AB
egyenes nem lesz párhuzamos a téglalap egyik oldalával sem. Egy ilyen egyenesből visszahajlítás során a hengerpaláston kapott vonalat *csavarvonalnak* nevezzük (8. ábra). A csavarvonal tehát a hengernek geodetikus vonala. Az ábráról azonnal leolvasható a csavarvonalnak egy érdekes tulajdonsága. Ha berajzolnánk a 8. ábra téglalapján a függőleges egyeneseket, akkor láthatnánk, hogy az

AB
egyenes a függőlegesek mindegyikével ugyanazt a szöget zárja be. Mivel visszahajlítás során ezek a függőleges egyenesek a henger alkotóiba mennek át, elmondhatjuk, hogy a csavarvonal a henger mindegyik alkotóját ugyanakkora szögben metszi.

Nézzük most, hogy milyen más geodetikus vonalakhoz juthatunk a két pont más elhelyezkedése esetén. Ha a két pont ugyanazon az alkotón van (pl. a 8. ábrán az

A
és

C
pont), akkor az alkotó mentén vezet a legrövidebb út az egyik pontból a másikba, azaz az alkotók is geodetikus vonalak. Ezt már eddig is tudtuk, hiszen az alkotók a hengerpaláston elhelyezkedő egyenesek. Két ilyen pontot a téglalapon függőleges egyenes szakasz köt össze. Ha viszont a hengerpalástot határoló körvonalon veszünk fel két pontot (pl. a 8. ábrán

A
és

D
), akkor a paláston ezen a körvonalon vezet a legrövidebb út, hiszen síkba fejtéskor ez a kör a téglalap alapjába, vagyis egyenesbe megy át. Ugyancsak látható, hogy az alapkörrel együtt a vele párhuzamos, azaz a henger tengelyére merőleges síkmetszetek (amelyek az alapkörrel egybevágó körök) szintén geodetikus vonalak, hiszen síkba fejtéskor a téglalap alapjával párhuzamos egyenesekbe mennek át. Íly módon eljutottunk az egyenes körhenger összes geodetikus vonalához. Ezek tehát az alkotók, a tengelyre merőleges síkmetszetként adódó körök és a csavarvonalak. A hengerpaláston két pontot összekötő legrövidebb út tehát mindig egy ilyen vonalon vezet.

Ha viszont a hengerpalást egy

E

pontjából kiinduló, és a palást körüljárása után a kiindulási pontba visszavezető utak közül keressük a legrövidebbet, akkor a következőképpen járhatunk el: a hengerpalástot az

E

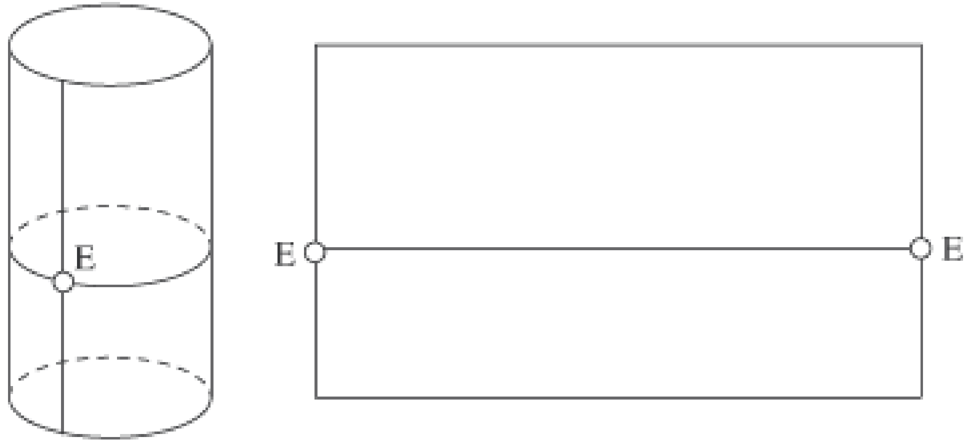
ponton átmenő alkotó mentén felvágva, síkba kiterítjük. A kapott téglalap mindkét függőleges oldalán megtalálható az

E

pont, az alaptól azonos távolságra. Ezt a két

E

pontot összekötő egyenes szakasz adja visszahajlítás után a keresett legrövidebb utat. Mivel az említett egyenes szakasz a téglalap alapjával párhuzamos, visszahajlítás után az alapkörrel párhuzamos (azaz a tengelyre merőleges síkban lévő) körbe megy át (9. ábra).



9. ábra

Ez a kör lesz tehát a palástot körüljáró legrövidebb út.

Nézzük most ez utóbbi problémát az egyenes körkúp6 palástján felvett

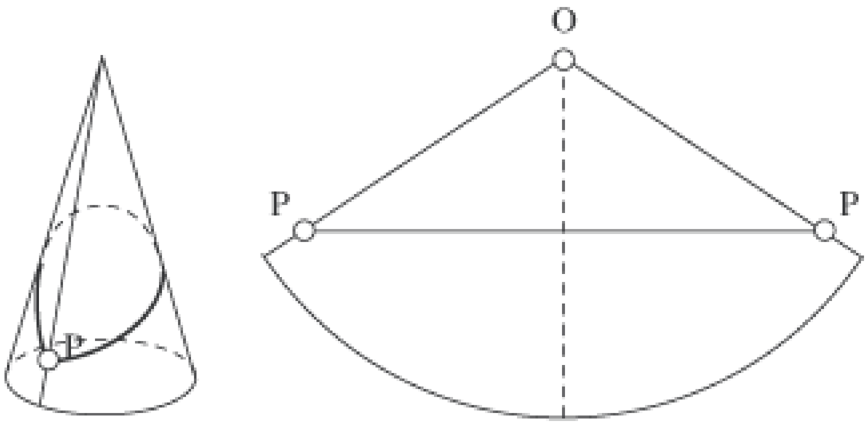
P

pont esetén! A probléma megoldásához vágjuk fel a kúppalástot a

P

ponton átmenő alkotó mentén, majd terítsük ki síkba: ily módon egy körcikket kapunk (10. ábra).

6*Egyenes körkúpnak* nevezzük az olyan kúpot, amelynek alapja (vezérvonala) kör, csúcsa az alapkör középpontjában az alapkör síkjára állított merőlegesen van.



10. ábra

A

P
pont a körcikknek mindkét sugarán megtalálható, mégpedig az

O
csúcstól egyenlő távolságban. Nyilván a két

P
pontot összekötő egyenes szakasz adja visszahajlítás után a palástot körüljáró és a kiindulási pontba visszavezető legrövidebb utat.

Ez az út egy hurokszerű vonal. Az eredmény kissé meglepő, hiszen a szemlélet és a henger példája alapján azt várhattuk volna, hogy a kúp tengelyére merőleges síkban lévő kör a legrövidebb út. (Ezek a körök azonban nem geodetikus vonalak, hiszen a kiterített kúppaláston a körcikk alapjával koncentrikus körívek lesznek.) Az ábráról leolvasható, hogy a kapott hurokszerű vonal visszaérkezéskor ugyanolyan szögben metszi a

P
ponton átmenő alkotót, mint kiinduláskor, az átellenes alkotóhoz pedig derékszögben hajlik.

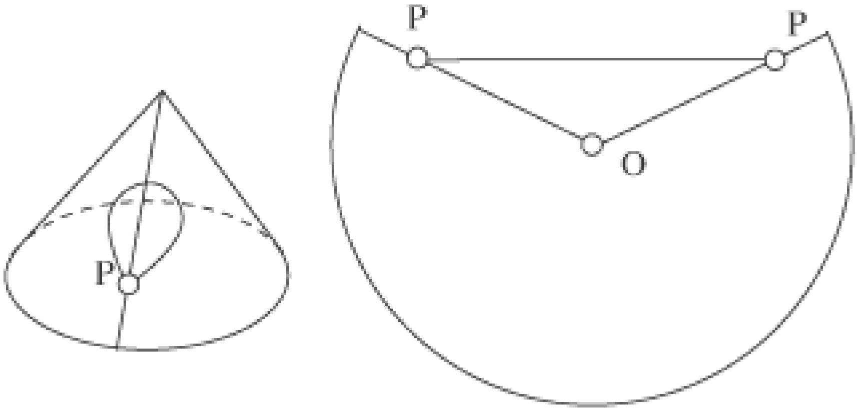
Ha figyelmesebben megnézzük a 10. ábrát, láthatjuk, hogy okoskodásunk csak abban az esetben helytálló, ha a körcikk csúcsánál lévő szög az egyenes szögnél, azaz

180°
-nál kisebb. Ha ez a szög homorú, vagyis

180°
-nál nagyobb (esetleg pontosan

180°
: ebben az esetben a kúppalástból síkbafejtés után egy félkört kapunk), akkor a két

P
pontot összekötő egyenes szakasz a körcikken kívül halad, és így visszahajlítás után nem lesz a paláston, vagyis nem kapjuk meg a fenti hurokszerű vonalat (11. ábra).



11. ábra

Ebben az esetben a keresett legrövidebb út az, amely a

P

pontból az alkotó mentén a kúp csúcsához vezet, majd onnan ugyanezen az alkotó mentén vissza a

P

pontba. Nem mondhatjuk persze, hogy ez az út körüljárja a kúppalástot, viszont az ábráról látható, hogy minden olyan út, amely a

P

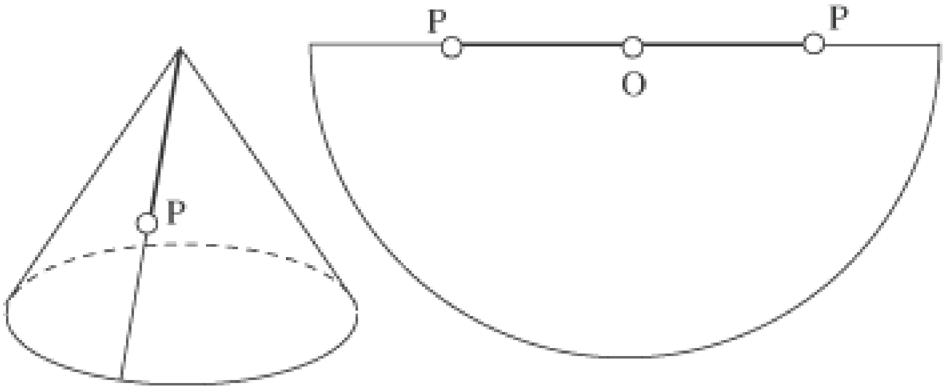
pontból kiindulva körbejárja a palástot, nagyobb a fenti útnál. Azt a meglepő tényt kaptuk tehát, hogy ilyenkor a palástot ténylegesen körbejáró utak között tulajdonképpen nincs is legrövidebb. Ha a körcikk csúcsánál lévő szög

180°

, akkor a két

P

pontot összekötő egyenes szakasz a félkör átmérőjén halad. Visszahajlítás után most sem kapjuk meg a fenti hurokszerű vonalat (12. ábra).

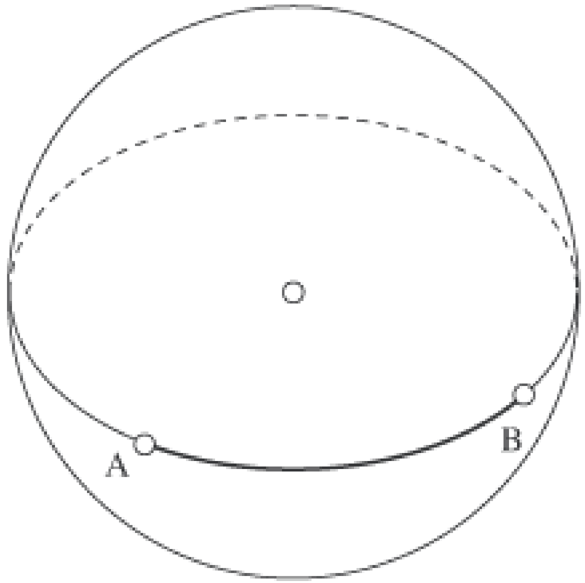


12. ábra

Ebben az esetben is ugyanúgy juthatunk el a keresett legrövidebb úthoz, mint az előbbi esetben.

Az egyenes körhenger és körkúp csak igen speciális esetei a henger- és kúpfelületeknek, amelyek mind síkba fejthetők. Általános *hengerfelületnek* nevezzük az olyan felületet, amely úgy keletkezik, hogy valamely síkbeli vonal mentén egy olyan egyenest mozgatunk önmagával párhuzamosan, amely nincs benne a vonal síkjában. Ha ez a síkbeli vonal kör, és az egyenes a kör síkjára merőlegesen mozog, akkor egyenes körhenger keletkezik. Általános *kúpfelülethez* viszont úgy jutunk, hogy egy síkbeli vonal mentén egy egyenest mozgatunk úgy, hogy állandóan áthaladjon egy rögzített ponton is (ezt a pontot a vonal síkján kívül vettük fel). Ha a síkbeli vonal kör, és a rögzített pont a kör középpontjában, annak síkjára állított merőlegesen van, akkor egyenes körkúp keletkezik. Ezeket a felületeket egyik egyenesük (alkotójuk) mentén történő felvágás után síkba teríthetjük, és így geodetikus vonalaikat az előbbiek szerint könnyen meghatározhatjuk. Sőt a henger- és kúpfelületeken kívül is vannak még további felületek, amelyek síkba teríthetők; ezeknek vizsgálata azonban már a geometria magasabb fejezeteibe tartozik.

Vajon mik a geodetikus vonalai a harmadik legismertebb görbe felületnek, a gömbnek? A gömbfelületen nem ilyen egyszerű a helyzet, ugyanis ezt nem lehet kiteríteni a síkba. Ez szemléletesen azt jelenti, hogy pl. papírból nem tudjuk elkészíteni hajlítással a gömbfelület modelljét. A gömbfelület geodetikus vonalait más úton kell tehát meghatározni. Ennek részletezésébe nem bocsátkozhatunk, csak megemlítjük, hogy a gömbfelületen az ún. főkörök a geodetikus vonalak. A *főkörök* a gömb középpontján átmenő síkoknak a gömbfelülettel alkotott metszésvonalai: a gömb sugarával egyenlő sugarú körök. A földgömbön pl. a hosszúsági körök (délkörök) mind főkörök. A gömbfelület két pontját összekötő legrövidebb vonalat tehát a következőképpen kaphatjuk meg: a két pont és a gömb középpontja által meghatározott síkkal elmetsszük a gömbfelületet. A kapott főkörnek a két pontot összekötő nem hosszabbik íve lesz a keresett legrövidebb vonal (13. ábra).



13. ábra

Hangsúlyozzuk, hogy a két ponton átmenő főkörnek a két pont által határolt darabjai közül a *kisebbet* kell vennünk. Itt említjük meg, hogy a geodetikus vonal elnevezést nemcsak szűken két pontot összekötő legrövidebb vonalra szokás érteni, hanem általánosabban azokra a vonalakra, amelyek valamely felületen az egyenes szerepét játsszák. A gömb főkörei ilyen értelemben geodetikus vonalak, és ilyen értelemben lesznek a síkba kiterített felület egyeneseiből is visszahajlítás során az eredeti, síkba fejthető felület geodetikus vonalai. A fenti példa azt a figyelemre méltó tényt mutatja, hogy a geodetikus vonal elnevezést ily módon használva, a felület két pontját összekötő legrövidebb vonal mindig geodetikus vonal, de egy geodetikus vonaldarab nem feltétlenül a legrövidebb a végpontjait összekötő és a felületen haladó vonalak között. Erre más példát is mutathatunk. Az egyenes körhengernek két, ugyanazon az alkotón lévő pontját összeköthetjük pl. az alkotón kívül csavarvonallal is, amely geodetikus vonal, és mégis hosszabb, mint a két pontot összekötő alkotó-szakasz. Ha tehát két pontot több geodetikus vonaldarab köt össze, akkor ezek közül a legrövidebbet kell kiválasztanunk. Ezzel a legrövidebb geodetikus vonaldarabbal mérhetjük a felületen a két pont távolságát, annak megfelelően, hogy síkban két pont távolsága az összekötő egyenes szakasz hossza.

Földünkön két tetszőleges helyet összekötő legrövidebb út mindig egy főkörív. Két földrajzi hely egymástól való távolságát ezzel a legrövidebb főkörívvel mérik, és a két hely által meghatározott *ortodrómának* nevezik. Ha két földrajzi hely ugyanazon a hosszúsági körön van, akkor az ortodróma a hosszúsági körnek egy íve. Ha viszont két, ugyanazon a szélességi körön lévő földrajzi helyet tekintünk, az ezeket összekötő legrövidebb út nem a szélességi kör mentén vezet – amint azt első pillanatra a szemlélet alapján várni lehetne –, hiszen a szélességi körök általában nem főkörök. Az egyetlen szélességi kör, amely egyúttal főkör is, az egyenlítő. Az egyenlítő egyik pontjából a másikba tehát az egyenlítő mentén vezet a legrövidebb út. (Ezzel kapcsolatban érdemes megemlíteni, hogy mialatt egy szélességi kör mentén körbejárunk, az Északi, illetőleg a Déli saroktól ugyanolyan távolságban maradunk, hiszen a sarkoktól egy szélességi kör pontjaihoz vezető főkörívek – amelyek egyúttal hosszúsági körök ívei is – egyenlő hosszúak.)

Az előbbiek szerint, ha valaki a legrövidebb úton akar eljutni az egyik földrajzi helyről egy másikra, akkor a két hely által meghatározott ortodróma mentén kell haladnia. A repülőgépek és tengerjáró hajók azonban biztonsági okokból általában mégsem az ortodróma mentén, hanem egy másik vonal, a két helyet összekötő *loxodróma* mentén haladnak. A loxodróma, vagy más szóval az *állandó irány vonala* olyan vonal, amely úgy köt össze a földgömbön két földrajzi helyet, hogy közben minden délkörrel ugyanakkora szöget zár be. A repülőgépek és hajók ezt a vonalat egyszerűbben tudják követni, mert így csak egyszer kell beállniuk az iránytű segítségével a megfelelő irányba, és azután már csak arra kell ügyelniük, hogy ne térjenek el ettől az iránytól. Így hosszabb úton haladnak ugyan, de kisebb az eltévedés veszélye.

Még egy érdekes tényre mutatunk rá a gömbfelülettel kapcsolatban. A gömbfelületeken a főkörök, mint geodetikus vonalak játsszák azt a szerepet, amit a síkban az egyenesek. Ez többek között azt jelenti, hogy ha Földünkön valaki tudomása szerint egyenes vonalban halad, akkor mozgását egy másik bolygó élőlénye úgy látja, mintha egy főkör mentén haladna, azaz valójában nem egyenes vonalban halad. Ennek alapján a síkbeli háromszögek mintájára alkothatunk a gömbfelületen olyan háromszögeket, amelyeknek oldalai főkörívek. Ezeknek a *gömbháromszögeknek* nevezett gömbi alakzatok több olyan tulajdonsággal rendelkeznek, mint a síkbeli háromszögek. Ezekből az észrevételekből kiindulva, a síkgeometriához hasonlóan fel lehet építeni a gömbfelületen is egy geometriát. Ebben az ún. *gömbi geometriában* pl. a főkörök veszik át az egyenesek, a gömbháromszögek a síkbeli háromszögek szerepét. Ez a geometria azonban több lényeges tekintetben eltér a síkgeometriától. Be lehet bizonyítani pl., hogy a gömbháromszög szögeinek összege a síkbeli háromszögtől eltérően nem

180□

, hanem mindig nagyobb ennél. Látható továbbá, hogy a gömbi geometriában nincsenek párhuzamos egyenesek, hiszen bármely két főkör a gömb két átellenes pontjában metszi egymást. Éppen az utóbbi tény miatt mondjuk azt, hogy a gömbi geometria a síkbeli, euklidészi geometriától eltérően az ún. *nem euklidészi geometriák* egyik fajtája. Ezek után gondolhatná valaki, hogy mi tulajdonképpen a Földön élünk, amely igen jó közelítéssel gömb alakú, tehát a gömbi geometriát kellene használnunk. Ennek ellenére mi mégis a síkbeli, euklidészi geometriát használjuk! Ez azonban mégsem okoz bajt, mert egy nagy sugarú gömb felületének kis darabja már igen jó közelítéssel síknak tekinthető. Földünk is ilyen nagy sugarú gömb, és így gyakorlatilag teljesen kielégítő, ha az euklidészi geometriát használjuk. Nem kell tehát például a síkbeli háromszögek helyett gömbháromszögekkel dolgoznunk, hiszen a gyakorlatban előforduló háromszögek oldalai nagy sugarú főkörök kis ívei, amelyek már igen jó közelítéssel egyenes szakaszok. Hasonlóan, ha Földünkön tudomásunk szerint egyenes vonalban haladva, nem túlságosan nagy távolságot teszünk meg, akkor szintén egy nagy sugarú főkörnek igen kis ívén, tehát gyakorlatilag egyenes szakaszon haladunk végig. Megnyugtathatjuk tehát magunkat abban a tekintetben, hogy bár Földünkön nagy méretekben a gömbi geometria érvényes, a gyakorlatban előforduló méretek mellett nem követünk el számottevő hibát, ha az euklidészi geometriát használjuk.

A HÉRÓN-FÉLE HÁROMSZÖKEGRŐL

VADKERTY TIBOR
HÓDI ENDRE

1. Az alábbiakban bizonyos háromszögek oldalainak, szögeinek, illetve területének mérőszámára vonatkozólag teszünk kijelentéseket. Egyszerűség – és rövidség – kedvéért egyrészt oldalak mérőszáma helyett csak oldalakat fogunk írni stb., másrészt pl. az oldalak mérőszámát ugyanazokkal a szimbólumokkal (betűkkel) jelöljük, mint magukat az oldalakat; hasonlóan járunk el a szögek és a terület esetében is. Az effajta pongyolaság egyébként szinte általánosan elterjedt a gyakorlatban.

2. Pitagorasz tétele (bármely derékszögű háromszögben a két befogó négyzetének összege egyenlő az átfogó négyzetével) legalább harmadfél ezer éve ismert, sőt [1] szerint a babiloniak több, mint 1000 évvel Pitagorasz előtt több numerikus példát tudtak ugyanerre az összefüggésre. Amit ezzel kapcsolatban mégis szeretnénk kiemelni, az a következő: geometriai úton könnyen bizonyítható, hogy Pitagorasz tétele *minden derékszögű háromszögre* érvényes, olyanra is, amelynek egyik oldala sem racionális; ilyen pl. a $\sqrt{2}$, $\sqrt{3}$ és $\sqrt{5}$ oldalakkal rendelkező.

3. Legutóbbi mondatunkkal szembeállíthatjuk azt a feladatot, amelyet [2] az egyik legrégibb számelméleti problémának nevez:

Határozzuk meg az összes olyan derékszögű háromszöget, melyeknek oldalai egész számok, vagyis adjuk meg az

$$x^2+y^2=z^2 \qquad (1)$$

egyenlet pozitív egész megoldásait!

Hogy (1)-nek van megoldása a pozitív egész számok körében, arra legegyszerűbb példa az $x=4$, $y=3$, $z=5$, illetve rövidebb jelöléssel a (4, 3, 5) számhármas. Jogos tehát, hogy az (1) egyenletet kielégítő, pozitív egész számokból álló (x, y

, z) számhármasokat *pitagoraszi számhármasoknak*, a nekik megfelelő derékszögű háromszögeket pedig még rövidebben *P-háromszögeknek* nevezzük.

Célszerű bevezetni még a *primitív* pitagoraszi számhármas fogalmát. Egy pitagoraszi számhármast akkor mondunk primitívnek, ha elemeinek nincs 1-nél nagyobb közös osztója. Könnyen belátható, hogy egy primitív pitagoraszi számhármas elemei páronként is viszonylagos törzsszámok. Világos továbbá, hogy ha (x, y, z) primitív pitagoraszi számhármas, akkor (kx, ky, kz) pitagoraszi számhármas, és megfordítva: ha (kx, ky, kz) pitagoraszi számhármas, ahol k egynél nagyobb egész szám, egyszersmind a kx, ky, kz számok legnagyobb közös osztója, akkor (x, y, z) primitív pitagoraszi számhármas.

[2]-ben azután megtalálható, hogy az (1) egyenletet azok és csak azok a primitív (x, y, z) számhármasok elégítik ki, melyekre

(2) $$x=2uv,\ y=u^2-v^2,\ z=u^2+v^2;$$

ahol az u, v pozitív egész számoknak a következő feltételeket kell teljesíteniük:

(3) $$u>v,\ (u,v)=1,\ u\not\equiv v \pmod 2.$$

A fentebb mondottak értelmében így az (1) egyenletnek eleget tevő összes, pozitív egész számokból álló számhármast, vagyis az összes pitagoraszi számhármast – elvileg – elő tudjuk állítani, át tudjuk tekinteni őket, akkor is, ha már a primitív pitagoraszi számhármasokból is nyilvánvalóan (megszámlálhatóan) végtelen sok van.

A *primitív P-háromszög* fogalma önként adódik az előbbiek alapján. Könnyen belátható, hogy egyrészt nincs egyenlő szárú primitív P-háromszög, másrészt minden primitív P-háromszög területe páros szám.

4. Ugyancsak [2]-nek 73. oldaláról idézzük a következőket:

„Megoldottuk a pitagoraszi háromszögek (P-háromszögek) meghatározásának problémáját. Ezután felvetődik az idevágó kérdések tanulmányozása. A problémakör egyik természetesen adódó kiterjesztése a Hérón alexandriai görög matematikusról elnevezett háromszögekhez (H*-háromszögek*) kapcsolódik. Most is – mint az előzőkben – az oldalak egész számok, de az egyik szög 90°-os volta helyett azt kötjük ki, hogy a szóban forgó háromszög területe is egész szám legyen. Világos, hogy a P

-háromszögek egyszersmind

H
-háromszögek is.

Egy adott háromszögről legegyszerűbb a

$t=\sqrt{s(s-a)(s-b)(s-c)}$
Hérón-féle területképlet – ahol

a
,

b
és

c
a háromszög oldalait,

$2s$
pedig a kerületét jelöli – alkalmazásával dönteni el, hogy vajon

H
-háromszög-e vagy sem. Bár elég sok Hérón-féle háromszögről tudunk, *nincsen olyan általános képletünk, amely meghatározná az összes ilyen típusút*. (Kiemelés e cikk szerzőitől.) Az első néhány (nem derékszögű) példa:

$a=20$
,

$b=15$
,

$c=7$
;

$t=42$
, rövidebben: (20, 15, 7; 42), azután: (17, 10, 9; 36), (15, 14, 13; 84); (50, 41, 39; 780).”

5. Ennek a cikknek az a célja, hogy megadjon egy képletcsoporton alapuló eljárást, amellyel *minden primitív H-háromszög előállítható*, mégpedig pontosan egyszer. (A primitív

H
-háromszög fogalma a primitív

P
-háromszögéhez hasonlóan értelmezhető. Hogy van primitív

H
-háromszög, azt jól mutatják az iménti nem derékszögű példák is, hiszen valamennyi ilyen. Primitív

H
-háromszögekben – természetesen csak a nem derékszögűekben – azonban két oldalnak már lehet 1-nél nagyobb közös osztója: ezt példázza (20, 15, 7; 42). Az eddig megismert – derékszögű és nem derékszögű – példákon azt is megfigyelhetjük, hogy a primitív

H

-háromszögekben két oldal páratlan, míg egy páros, éppen úgy, mint a primitív

P
-háromszögek esetén. Nem derékszögű primitív

H
-háromszögekre nem magától értetődő e tulajdonság érvényben maradása, ezért a későbbiekben bebizonyítjuk majd, hogy előbbi megállapításunk csakugyan igaz minden primitív

H
-háromszögre vonatkozólag.)

6. Nincs egyenlő oldalú

H
-háromszög. Ha ugyanis lenne ilyen, akkor kellene, hogy

a
oldala egész legyen. Ha viszont

a
egész szám, abban az esetben a háromszög területe:

$t=(\sqrt{3}/4)\square a2$
irracionális szám. Ez pontosan ugyanolyan gondolatmenettel igazolható, amilyennel

$\sqrt{2}$
irracionalitását szokták bizonyítani.

7. Egyenlő szárú

H
-háromszög viszont biztosan létezik, amint azt a (6, 5, 5; 12) példa is mutatja. Könnyen belátható, hogy ez a

H
-háromszög hegyesszögű. Nem volna nehéz példát adni tompaszögű egyenlő szárú

H
-háromszögre sem, azt viszont már a **3.** pontban megállapítottuk, hogy nincs egyenlő szárú primitív

P
-háromszög, tehát semmilyen egyenlő szárú

P
-háromszög sincs, ennélfogva végső soron derékszögű egyenlő szárú

H
-háromszög sem létezik.

Teljesen világos továbbá, hogy ha bármely primitív

P
-háromszöget tükrözünk valamelyik befogójára és a tükörképet hozzáillesztjük az eredetihez, akkor – hegyes- vagy tompaszögű – egyenlő szárú primitív

H
-háromszöget kapunk. (A primitivitáshoz csupán azt kell meggondolnunk, hogy a „volt” páratlan átfogónak és

bármelyik „volt” befogó kétszeresének nem lehet 1-nél nagyobb közös osztója, ha eredetileg nem volt.) Az is nyilvánvaló, hogy a kapott egyenlő szárú primitív

H
-háromszög területe páros, sőt 4-gyel is osztható szám.

A most mondottakból következik, hogy van tehát egyenlő szárú primitív

H
-háromszög. Az általunk előállított háromszögekben az alap mindig páros, a szárak pedig páratlanok. Azt állítjuk, hogy bármely egyenlő szárú primitív

H
-háromszögben így van ez. Ha ugyanis mindhárom oldal páratlan lenne, akkor Hérón képletéből nem adódna egész szám a háromszög területére. Hasonló indokolással zárhatjuk ki azt az esetet is, amikor a szárak párosak és az alap páratlan. Hátravan még annak az esetnek a taglalása, amikor mindhárom oldal páros.

Azt mondhatja valaki, hogy az ilyen

H
-háromszög nem primitív. Nos, rendben van. Osszuk el tehát az oldalakat legnagyobb közös osztójukkal! A páratlan közös osztókkal való osztás nyilván változatlanul hagyja az oldalak paritását, azonkívül azt is könnyű belátni, hogy ha a kiindulásul szolgáló háromszög területe egész szám volt, akkor az oldalakat legnagyobb páratlan közös osztójukkal osztva, az így nyert háromszög is

H
-háromszög marad. Osszuk most el az oldalakat legnagyobb páros közös osztójukkal! Lényegében háromféle eredményhez juthatunk: vagy egy oldal – az alap – marad páros, a szárak páratlanok; ez az az eset, amelyet megengedhetünk. Érdekesebbek azonban azok, amikor az eredeti oldalak legnagyobb páros közös osztójával osztva két páros és egy páratlan oldalhoz, illetve három páratlan oldalhoz jutunk. Azt már láttuk, hogy nincs ilyen tulajdonságú egyenlő szárú primitív

H
-háromszög, de az még elképzelhető lenne, hogy a kétszer akkora oldalakkal rendelkező háromszög már egyenlő szárú

H
-háromszög (ha nem is primitív). Megmutatjuk, hogy ez sem lehetséges.

A két eset együtt tárgyalható, mert lényegük az, hogy az egyenlő szárú háromszög alapja egy páratlan szám kétszerese, szárai pedig párosak. Az alaphoz tartozó magasság két egybevágó derékszögű háromszögre osztja az eredetit. Így mindegyiknek páros az átfogója és páratlan az egyik befogója. Ami a másik befogót, vagyis az egyenlő szárú háromszög alapjához tartozó magasságot illeti, az az

$ma=(2t/a)$
képlet miatt mindenesetre racionálisc, a derékszögű részháromszögekre érvényes Pitagorasz-tétel következtében pedig egész szám is. Ha egész szám, akkor vagy páros, vagy páratlan. Egyik sem lehet azonban, mivel sem egy páros szám négyzetének és egy páratlan négyzetének összege, sem két páratlan szám négyzetének összege nem osztható 4-gyel. ezért nem lehet egyenlő egy páros szám négyzetével. Ezzel beláttuk, hogy csakugyan minden egyenlő szárú primitív

H
-háromszögben az alap páros szám, a szárak pedig páratlanok.

Ha tehát megfordítva, valamely tetszés szerinti egyenlő szárú primitív

H
-háromszögben meghúzzuk az alaphoz tartozó magasságot, akkor az előbbi gondolatmenettel könnyen bizonyítható, hogy ez a magasság két egybevágó primitív

P
-háromszögre osztja azt. Ebből következik, hogy az összes egyenlő szárú primitív

H
-háromszöget megkapjuk, ha a primitív

P
-háromszögeket rendre tükrözzük egyik, majd másik befogójukra, és a tükörképet mindkét esetben hozzáillesztjük az eredeti háromszöghöz. Ehhez csatlakozva azt sem nehéz igazolni, hogy egy bizonyos primitív

P
-háromszögből kiindulva mindig két egyenlő szárú primitív

H
-háromszöghöz jutunk, mégpedig egy hegyesszögűhöz a hosszabbik, és egy tompaszögűhöz a rövidebbik befogóra való tükrözéskor. Célszerű ismételten megemlíteni azt a tényt, hogy minden egyenlő szárú primitív

H
-háromszögben az alap páros szám, míg a szárak páratlanok, továbbá, hogy az egyenlő szárú primitív

H
-háromszög területe minden esetben páros, sőt 4-gyel is osztható szám. Végül legyen szabad kiemelnünk azt a megállapítást, hogy az egyenlő szárú primitív

H
-háromszögekben az alaphoz tartozó magasság is egész szám.

8. Ezek után térjünk rá a nem egyenlő szárú

H
-háromszögek vizsgálatára! Mindenekelőtt azt figyelhetjük meg, hogy az ilyen háromszögek között van derékszögű is, hegyesszögű is és tompaszögű is. Már említettük a **4.** pontban, hogy a

P
-háromszögek – közöttük nem lehet egyenlő szárú – egyúttal

H
-háromszögek is. Az ugyanitt felsorolt konkrét

H
-háromszögekről könnyű eldönteni, hogy az utolsó kettő hegyesszögű, az első kettő pedig tompaszögű – és nyilvánvalóan egyik sem egyenlő szárú.

További vizsgálatainkból a derékszögű

H
-háromszögeket nyugodtan kihagyhatnánk, hiszen a mondottak értelmében előállításuk kérdése tisztázódott a

P
-háromszögek meghatározása problémájának megoldásával. Hogy mégsem így járunk el, annak megvan az oka. Erre kívánunk rávilágítani a következőkben.

Az egyenlő szárú primitív

H
-háromszögek tárgyalását az könnyítette meg nagymértékben, hogy az alaphoz tartozó magasságuk egész szám volt, következésképpen ennek a magasságnak a meghúzásával két (egybevágó) primitív

P
-háromszögre lehetett osztani az eredetit. Hátha ferdeszögű (nem egyenlő szárú)

H
-háromszögekben, esetleg már az ilyen primitívekben is, legalább az egyik magasság mindig egész, ezáltal az ilyen esetekben is lehetővé válna az eredeti háromszög előállítása két – lehetőleg primitív –

P
-háromszög „összegeként” (amikor az eredeti

H
- -háromszög hegyesszögű volt), illetve esetleg „különbségeként” (amikor az tompaszögű volt)?

Egyszerűen ellenőrizhetjük, hogy a **4.** pontban közölt példákban mindenkor pontosan egy magasság, mégpedig rendre a 7, 9, 14, illetve 39 oldalhoz tartozó egyenlő egész számmal. Ez elegendő is lenne, csakhogy …Csakhogy nem mindegyik (primitív)

H
-háromszög teszi meg nekünk ezt a szívességet. Mutatóba ideírtunk néhány „renitens” primitív

H
-háromszöget: (39, 35, 10; 168), (91, 75, 26; 840), (45, 40, 13; 252), (35, 34, 15; 252), (60, 55, 17; 462). Annak ellenőrzését, hogy ezeknek a háromszögeknek egyik magassága sem egész, az igen tisztelt Olvasóra bízzuk. Megjegyezzük továbbá, hogy imént megadott példáink között egyaránt található hegyesszögű háromszög és tompaszögű is. Végül megemlítjük azt a sejtésünket, hogy végtelen sok olyan primitív

H
-háromszög létezik, amelyekben egyik magasság sem egész szám.

9. Hogyan segítsünk hát magunkon, hogyan tegyük áttekinthetővé a ferdeszögű, nem egyenlő szárú (nem feltétlenül primitív)

H
-háromszögek tárgyalását?

Hátha nincs is mindig egész számmal egyenlő magassága egy ferdeszögű, nem egyenlő szárú (nem feltétlenül primitív)

H
-háromszögnek, de *legnagyobb oldala* biztosan van. Sőt ez a tulajdonsága megvan *bármely*

H
-háromszögnek, úgyhogy most következő meggondolásainkat egy tetszés szerinti

H
-háromszögre fogjuk végezni. Az nem okoz gondot, hogy a **7.** pontban már teljesen elintéztük a ferdeszögű, egyenlő szárú primitív

H
-háromszögek, ezzel egyszersmind valamennyi ilyen

H
-háromszög vizsgálatát, mert azok a megállapítások, amelyeket egy tetszés szerinti

H
-háromszögre vonatkozólag teszünk majd, természetesen az előbbiekre is, valamint a derékszögű

H
-háromszögekre (a

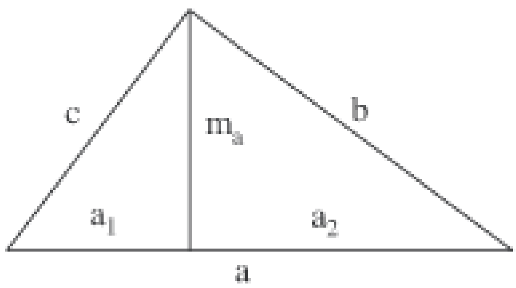
P
-háromszögekre) is érvényesek. Egyenlő szárú

H
-háromszögek esetén lehetséges, hogy a szárak nagyobbak az alapnál, ekkor tehát két legnagyobb oldal van, közülük bármelyiket vehetjük legnagyobbnak.

A legnagyobb oldalhoz tartozó magasság most is két derékszögű háromszögre osztja az eredeti

H
-háromszöget, ezek azonban általában nem egybevágók. Az

m_a
magasságról ugyan nem tételezhetjük fel, hogy egész, az azonban bizonyos, hogy racionális, miként ez látszik a **7.** pontban felírt képletből.



1. ábra

Azt állítjuk, hogy az

m_a
magasság racionális szakaszokra osztja az

a
oldalt. Tegyük fel ugyanis, hogy ez nem így lenne, hanem pl.

a_1
irracionális volna! Ez a bal oldali derékszögű részháromszögre érvényes Pitagorasz-tétel miatt azt jelentené, hogy

$a_1=\sqrt{r}$
, ahol

r
olyan racionális számot jelöl, amely nem négyzete egyetlen racionális számnak sem. Ekkor tehát

$a_2=a-\sqrt{r}$
. Alkalmazzuk most Pitagorasz tételét a jobb oldali derékszögű részháromszögre:

(4)

$$m_a^2+(a-\sqrt{r})^2=b^2,$$

ahonnan

$\sqrt{r}=(1/2a)(a^2-b^2+m_a^2+r)$. Ez azonban nyilvánvaló ellentmondás, hiszen feltevésünk szerint $\sqrt{r}$ irracionális, tehát nem lehet egyenlő egy racionális számmal.

Arra az eredményre jutottunk, hogy m_a-n kívül a_1 és a_2 is racionálisak. Legyen

$m_a=(p/q)$, $a_1=(p_1/q_1)$, $a_2=(p_2/q_2)$, ahol $(p,q)=(p_1,q_1)=(p_2,q_2)=1.$ Ekkor egyrészt

$((p/q))^2+((p_1/q_1))^2=c^2,$ azaz

(5)

$$(pq_1)^2+(p_1q)^2=(cqq_1)^2;$$

másrészt

$((p/q))^2+((p_2/q_2))^2=b^2,$ azaz

(6)

$$(pq_2)^2+(p_2q)^2=(bqq_2)^2.$$

(5) és (6) szerint (pq_1, p_1q, cqq_1), valamint (

$pq2$, $p2q$, $bqq2$) pitagoraszi számhármasok (nem feltétlenül primitívek), azért a derékszögű részháromszögek oldalaiból képzett (m_a, a_1, c) és (m_a, a_2, b) számhármasok egy-egy pitagoraszi számhármasból származtathatók, mégpedig oly módon, hogy a bennük szereplő számokat megszorozzuk ugyanazzal a racionális számmal:

$(1/qq1)$-gyel, illetve

$(1/qq2)$-vel.

Világos továbbá, hogy ha van olyan pitagoraszi számhármas, amelynek elemeit rendre ugyanazzal a racionális számmal szorozva megkapjuk, pl. a bal oldali derékszögű részháromszög oldalait, akkor van ilyen tulajdonságú primitív pitagoraszi számhármas is, hiszen minden pitagoraszi számhármas a primitívekből állítható elő az elemeknek valamely pozitív egész számmal történő szorzása útján. Végül az is nyilvánvaló, hogy ha bármely

P-háromszög oldalait megszorozzuk ugyanazzal a racionális számmal, akkor egyrészt a kapott háromszög *hasonló* lesz az eredetihez, másrészt annak oldalai mindenesetre racionális számok lesznek.

Ezek után jelen pontbeli fejtegetéseink eredményét így foglalhatjuk össze:

Ha bármely

H-háromszöget legnagyobb, illetve egyik legnagyobb oldalához tartozó magassággal két derékszögű háromszögre osztunk, akkor e részháromszögek mindegyike egy-egy primitív

P-háromszögből származtatható, mégpedig úgy, hogy annak oldalait rendre megszorozzuk ugyanazzal a racionális számmal. Lényeges megjegyezni, hogy a két primitív

P
-háromszög általában különböző; hogy a részháromszögek oldalainak előállításához felhasznált két állandó racionális szorzónak sem kell megegyeznie egymással; végül emlékeztetünk arra, hogy ha egy háromszög mindegyik oldalát ugyanazzal a számmal szorozzuk, akkor az így kapott háromszög hasonló az eredetihez.

Konkrét példaként tekintsük a (15, 14, 13; 84) primitív

H
-háromszöget! Ebben:

$a=15$

,

$b=14$

,

$c=13$

,

$m_a=11{,}2$

;

$a_1=6{,}6$

;

$a_2=8{,}4$
. A bal oldali derékszögű részháromszög az (56, 33, 65) primitív

P
-háromszögből állítható elő, ha annak oldalait rendre megszorozzuk 0,2-del. Ezzel szemben a jobb oldali derékszögű részháromszög a (4, 3, 5) primitív

P
-háromszöghöz hasonló; oldalai 2,8-szer akkorák, mint annak oldalai.

10. Most már be tudjuk váltani az **5.** pontban tett ígéretünket: megadunk egy képletcsoporton alapuló eljárást, amellyel minden primitív

H
-háromszög előállítható, mégpedig pontosan egyszer.

Figyelembe véve az előbbi pontban kapott eredményünket, vegyünk először egy „bal oldali” (

x_1

,

y_1

,

z_1
) primitív

P
-háromszöget, majd egy tőle különböző „jobb oldali” (

x_2

, y_2, z_2), ugyancsak primitív P-háromszöget, ahol az x_i, y_i, z_i ($i=1,2$) számokat a (2) alatti formulákkal képezzük, tehát x_i és y_i közül mindig x_i páros, y_i páratlan, és természetesen z_i is páratlan!

Nagyítsuk ezek után a bal oldali háromszöget (lineárisan) x_2-szörösére, a jobb oldalit pedig x_1-szeresére; így jutunk rendre az (x_2x_1, x_2y_1, x_2z_1) és (x_1x_2,

$x_1 y_2$, $x_1 z_2$) P-háromszögekhez, amelyek persze már nem primitívek. Illesszük össze őket egyenlő befogójuk mentén egyetlen H-háromszöggé! Ennek oldalai a következők lesznek: $x_2 y_1 + x_1 y_2$, $x_1 z_2$ és $x_2 z_1$. Hasonlítsuk össze az oldalakat! Ha közülük $x_2 y_1 + x_1 y_2$ a legnagyobb vagy az egyik legnagyobb, akkor tovább folytatjuk meggondolásainkat, ellenkező esetben nem, mert akkor ebből a konstrukcióból nem kapunk olyan H-háromszöget, amelyben a részháromszögek nem összeillesztett – általában egymástól különböző – befogóinak összege adja a leghosszabb, vagy legalább az egyik leghosszabb oldalt.

Tegyük fel azonban, hogy az $x_2 y_1 + x_1 y_2$, $x_1 z_2$ és $x_2 z_1$ oldalakkal rendelkező H-háromszögben a sorrendben elsőnek felírt oldal leghosszabb (ebbe beleértjük mindkét lehetőséget, amelyeket eddig külön szoktunk választani)! Nyilvánvaló, hogy ha ezt a háromszöget $(1/d_1)$-szeresére kicsinyítjük, ahol $d_1 = (x_1, x_2)$, akkor még mindig H-háromszöghöz jutunk, hiszen x_1 is,

x_2 is osztható d_1-gyel (emiatt egész számok a kicsinyítéssel nyert, röviden: kicsinyített háromszög oldalai), továbbá egész a legnagyobb oldalhoz tartozó magassága is, ami biztosítja a terület egész voltát.

A kicsinyített háromszög oldalai közül mindig egy páros, kettő pedig páratlan. Ha ui. mind (x_1/d_1), mind (x_2/d_1) páratlan, akkor nyilván a leghosszabb oldal páros, a többiek páratlanok, ha viszont pl. (x_1/d_1) páros (természetesen ekkor (x_2/d_1)-nek páratlannak kell lennie, különben d_1 nem lenne x_1 és x_2 legnagyobb közös osztója), akkor az $(x_1 z_2/d_1)$ oldal páros, a többi páratlan és hasonlóan intézhető el az az eset is, amikor (x_2/d_1) páros.

Az $x_2 y_1 + x_1 y_2$, $x_1 z_2$ és $x_2 z_1$ számok D_1 legnagyobb közös osztója lehet ugyan nagyobb is d_1-nél, de az előbbi bekezdésben mondottak miatt akkor is csupán d_1-nek valamely páratlan többszörösével egyenlő. (Páratlan számnak nem lehet páros osztója, és az eredeti H

-háromszög oldalait

d_1
-gyel osztva két páratlan számhoz is jutottunk.) A **7.** pontban említettük már, hogy egyrészt a páratlan közös osztókkal való osztás változatlanul hagyja az oldalak paritását, másrészt azt is könnyű belátni, hogy ha a kiindulásul szolgáló háromszög

H
-háromszög volt, akkor az oldalakat legnagyobb páratlan közös osztójukkal osztva, a kapott háromszög is

H
-háromszög marad.

A mondottakból következik, hogy ha az

$x_2y_1+x_1y_2$
,

x_1z_2
és

x_2z_1
oldalakkal rendelkező

H
-háromszögben a sorrendben elsőnek felírt oldal leghosszabb, akkor az a háromszög, amelynek oldalai rendre

$(1/D_1)(x_2y_1+x_1y_2)$, $(1/D_1)\cdot x_1z_2$ és $(1/D_1)\cdot x_2z_1$,

ahol $D_1=(x_2y_1+x_1y_2, x_1z_2, x_2z_1)$,
primitív

H
-háromszög. Ennek a primitív

H
-háromszögnek bal oldali derékszögű részháromszöge az (

x_1
,

y_1
,

z_1
) primitív

P
-háromszögből keletkezik, ha annak oldalait rendre megszorozzuk

(x_2/D_1)
-gyel. Ezzel szemben a jobb oldali derékszögű részháromszög az (

x_2
,

y_2

,

$z2$
) primitív

P
-háromszögből áll elő, ha annak oldalait sorjában

$(x1/D1)$
-gyel szorozzuk.

A bal oldali (

$x1$
,

$y1$
,

$z1$
) primitív

P
-háromszögből és a tőle különböző jobb oldali (

$x2$
,

$y2$
,

$z2$
), ugyancsak primitív

P
-háromszögből még három, a fentitől – és egymástól is – különböző primitív

H
-háromszöget tudunk létrehozni az imént ismertetetthez hasonló eljárással. A részleteket mellőzve csupán a keletkező primitív

H
-háromszögek oldalait írjuk le:

$(1/D2)(y1y2+x1x2)$, $(1/D2)$□$x1z2$ és $(1/D2)$□$y2z1$,

ahol $D2=(y1y2+x1x2,x1z2,y2z1)$;

$(1/D3)(x1x2+y1y2)$, $(1/D3)$□$y1z2$ és $(1/D3)$□$x2z1$,

ahol $D3=(x1x2+y1y2,y1z2,x2z1)$;

$(1/D4)(x1y2+x2y1)$, $(1/D4)$□$y1z2$ és $(1/D4)$□$y2z1$,

ahol $D4=(x1y2+x2y1,y1z2,y2z1)$.

Ezek közül a primitív

H-háromszögek közül is csak azokat tartjuk meg, amelyekben az első helyen álló oldal a legnagyobb.

A kiindulásul szolgáló bal és jobb oldali primitív P-háromszögek felcserélése nyilván nem vezet újabb primitív H-háromszögekhez, csupán olyanokhoz, amelyek úgy származtathatók, hogy egy már előzőleg létrehozott primitív H-háromszöget tükrözünk legnagyobb oldalának felezőmerőlegesére, ezért természetesen a tükrözöttel egybevágó háromszöget kapunk.

A különböző (x_1, y_1, z_1) és (x_2, y_2, z_2) primitív P-háromszögek hegyesszögeinek nagysági viszonyaitól függően lehetséges, hogy mind a négy esetben olyan primitív H-háromszög keletkezik a megfelelően felnagyított részháromszögek egyenlő befogójuk mentén történő összeillesztésével, majd ennek a háromszögnek esetleges kicsinyítésével, amelyben a részháromszögek nem összeillesztett befogóinak összege, illetőleg ennek alkalmas kicsinyítése adja a leghosszabb oldalt. Annyi azonban mindenképpen bizonyos, hogy a négy eset közül legalább kettőben a mondott tulajdonságú primitív H-háromszöghöz jutunk.

Ugyancsak az (x_1, y_1,

z_1
) és (

x_2
,

y_2
,

z_2
) primitív

P
-háromszögek hegyesszögeinek nagysági viszonyait figyelembe véve láthatjuk be, hogy a fenti négy esetben valóban csupa különböző primitív

H
-háromszöget kapunk, legfeljebb nem az az oldal lesz feltétlenül a legnagyobb, amely a derékszögű részháromszögek nem egymás mellé illesztett befogóiból tevődött össze. Tekintettel arra, hogy ezek a meggondolások elég egyszerűek, elvégzésüket az igen tisztelt Olvasóra bízzuk.

Világos továbbá, hogy ha másik pár primitív

P
-háromszögből indulunk ki: (

x_1'
,

y_1'
,

z_1'
)-ből és (

x_2'
,

y_2'
,

z_2'
)-ből, ahol legfeljebb az egyik vesszőtlen háromszög lehet azonos valamely vesszőssel, azonkívül a két vesszős háromszög ismét különböző, akkor felhasználásukkal a fentebb ismertetett eljárás újabb négy primitív

H
-háromszöghöz vezet, amelyek egymástól is, a korábbiakban kapottak mindegyikétől is különböznek. Legfeljebb nem az az oldaluk lesz feltétlenül a legnagyobb, amely a derékszögű részháromszögek nem egymás mellé illesztett befogóiból tevődött össze. A **9.** pontban ugyanis láttuk, hogy ha bármely

H
-háromszöget (így természetesen bármely primitív

H
-háromszöget is) a legnagyobb oldalához tartozó magassággal két derékszögű háromszögre osztunk, akkor ezek a részháromszögek – a sorrendtől eltekintve – pontosan egy pár primitív

P
-háromszögből származtathatók – azok oldalait rendre megszorozva ugyanazzal a racionális számmal, amely azonban nem okvetlenül egyenlő a két részháromszög esetében. Bár ebben a pontban azt tapasztaltuk, hogy ugyanabból a primitív

P
-háromszögpárból kedvező esetben négy különböző primitív

H
-háromszög is állítható elő, ahol a legnagyobb oldalhoz tartozó magasság az adott primitív

P
-háromszögpár elemeihez hasonló részháromszögeket hoz létre, a négy esetben azonban jól megkülönböztethető egymástól a részháromszögek összeillesztésének módja, akár a részháromszögek hegyesszögeit, akár a primitív

P
-háromszögek oldalainak megszorzására használt, egyértelműen meghatározott állandó racionális számokat tekintjük is.

Az elmondottakból mindenesetre kiderül, hogy ha két különböző primitív

P
-háromszögből indulunk ki, azokból az ismertetett módon négy különböző primitív

H
-háromszög állítható elő. Ezek közül legalább kettő olyan – de lehetséges, hogy mind a négy –, ahol a legnagyobb oldalhoz tartozó magasság bontja a

H
-háromszöget a kiindulásul szolgáló háromszögekhez hasonló részháromszögekre. Különböző primitív

P
-háromszögpárokból csupa különböző primitív

H
-háromszögnégyest származtathatunk. Különbözőnek tekintünk két primitív

P
-háromszögpárt, ha legfeljebb egy elemük közös, de ugyanannak a párnak az elemei feltétlenül különböznek egymástól. Két primitív

H
-háromszögnégyest pedig akkor mondunk különbözőnek, ha a két háromszögnégyesnek összesen nyolc eleme páronként is különböző.

11. Engedjük meg végül, hogy a kiindulásul vett bal és jobb oldali primitív

P
-háromszögek azonosak legyenek; jelöljük őket egyszerűen (

x
,

y
,

z)-vel! Az összetételükből létrehozható primitív H-háromszögek közül kétfélével foglalkoztunk már a **7.** pontban. Megállapítottuk, hogy egyrészt az összes egyenlő szárú primitív H-háromszöget megkapjuk, ha a primitív P-háromszögeket rendre tükrözzük egyik, majd másik befogójukra, és a tükörképet mindkét esetben hozzáillesztjük az eredeti háromszöghöz, másrészt egy bizonyos primitív P-háromszögből kiindulva mindig két egyenlő szárú primitív H-háromszöghöz jutunk, mégpedig egy hegyesszögűhöz a hosszabb, és egy tompaszögűhöz a rövidebb befogóra való tükrözéskor.

Nem foglalkoztunk még azzal az esettel, amikor pl. a bal oldali (x, y, z) háromszöget lineárisan y-szorosára, a jobb oldalit pedig x-szeresére nagyítjuk, majd az így nyert (xy, y^2, yz) és (x^2, xy, xz)

P-háromszögeket összeillesztjük xy egyenlő befogójuk mentén egyetlen H-háromszöggé, amelynek oldalai rendre: y^2+x^2, xz és yz. Mivel (x, y, z) mindenesetre P-háromszög, ezért $y^2+x^2=z^2$, a kapott H-háromszög oldalai tehát y^2, xz és yz, amelyeket legnagyobb közös osztójukkal, z-vel osztva mintegy „visszakapjuk" primitív H-háromszögként a kiindulásul vett (x, y,

z
) primitív

P
-háromszöget. Érdekes, hogy ebben az esetben a bonyolultabb összetétellel állítjuk elő az alapelemül alkalmazott primitív

P
-háromszöget, amiben persze nincs semmi csodálatos, hiszen derékszögű háromszögben a legnagyobb oldalhoz, az átfogóhoz tartozó magasság az eredetihez hasonló részháromszögekre osztja azt, és az is világos, hogy másféle háromszög nem rendelkezhet ezzel a tulajdonsággal. Végül az is nyilvánvaló, hogy ha a bal oldali (

x

,

y

,

z
) részháromszöget nagyítanánk lineárisan

x
-szeresére, a jobb oldalit pedig

y
-szorosára, akkor az imént követett eljárással újra az eredeti (

x

,

y

,

z
) primitív

P
-háromszöget nyerjük eredményül, csak az előbbihez képest átfogójának felezőmerőlegesére tükrözött állásban. Azt szinte felesleges is említenünk, hogy másik két azonos primitív

P
-háromszögből kiindulva, nemcsak különböző hegyes- és tompaszögű egyenlő szárú, hanem különböző derékszögű primitív

H
-háromszöghöz is jutunk.

12. Úgy véljük, az elmondottakból eléggé kitűnik, hogy a **10.** és **11.** pont eljárásai különböző kiinduló primitív

P
-háromszögpárokból csupa különböző primitív

H
-háromszögnégyest állítanak elő, ha a primitív

P

-háromszögpárok elemei nem azonosak, és ugyancsak csupa különböző primitív

H
-háromszöghármast, ha a primitív

P
-háromszögpárok elemei azonosak. Azt is tisztáztuk, hogy mikor tekintünk különbözőnek két primitív

P
-háromszögpárt, továbbá két primitív

H
-háromszögnégyest, hasonlóan két primitív

H
-háromszöghármast. Említettük, hogy egy-egy primitív

H
-háromszögnégyesből, illetve

H
-háromszöghármasból csak azokat a háromszögeket – esetenként legalább kettőt – használjuk fel az összes primitív

H
-háromszög meghatározására, amelyekben a legnagyobb oldalhoz tartozó magasság bontja a

H
-háromszöget a kiindulásul szolgáló háromszögekhez hasonló részháromszögekre.

Azt állítjuk, hogy „végighaladva" az azonos és nem azonos elemekből álló különböző primitív

P
-háromszögpárok (végtelen) sorozatán, a **10.** és **11.** pont eljárásainak – amelyek lényegében azonos szerkezetűek – utasításait követve csakugyan megkapunk minden primitív

H
-háromszöget, mégpedig pontosan egyszer. Az ismétlődést már kizártuk az eddigiek során. Tegyük fel azonban, hogy találtunk olyan primitív

H
-háromszöget, amely nem szerepel az általunk előállítottak között!

Ez nem lehetséges, hiszen a **9.** pontban láttuk, hogy ha bármely

H
-háromszöget (így természetesen bármely primitív

H
-háromszöget is) a legnagyobb oldalához tartozó magassággal két derékszögű háromszögre osztunk, akkor ezek a részháromszögek – a sorrendtől eltekintve – pontosan egy pár primitív

P
-háromszögből származtathatók – azok oldalait rendre megszorozva ugyanazzal a racionális számmal, amely azonban nem okvetlenül egyenlő a két részháromszög esetében. De az összes primitív

P

-háromszögpár között ennek a bizonyosnak is elő kellett fordulnia, és eljárásainkban mindegyik primitív

P
- -háromszögpárból létrehoztunk valamennyi lehetséges különböző primitív

H
-háromszöget: négyet, ha a primitív

P
-háromszögpár elemei különbözők voltak, és hármat, ha elemei megegyeztek. Ennélfogva annak a primitív

H
-háromszögnek, amelyet hiányoltunk az általunk előállítottak sorozatából, szükségképpen mégiscsak szerepelnie kell abban. Ez ellentmondás, amit csupán feltevésünk elvetésével lehet feloldani. Tehát a **10.** és **11.** pontban leírt eljárásaink csakugyan megadnak minden primitív

H
-háromszöget, mégpedig pontosan egyszer. Ezt kívántuk bizonyítani.

Az összes

H
-háromszög mármost ugyanúgy kapható meg az összes primitív

H
-háromszögből, ahogyan az összes

P
-háromszög az összes primitív

P
-háromszögből.

A

H
-háromszögeknek, speciálisan a primitíveknek számos érdekes tulajdonsága van; egy részük levezethető, illetőleg bizonyítható a fentiekben tárgyaltak alapján, más részük további vizsgálatokat igényel.

IRODALOM

[1] [1] Szabó Árpád: *A görög matematika kibontakozása*, Magvető Kiadó, Budapest, 1978.

[2] [2] Oystein Ore: *Bevezetés a számelmélet világába*, Gondolat Kiadó, Budapest, 1977.

[3] [3] Kelemen József: A Hérón-féle háromszögekről, *A Matematika Tanítása* **XXVI** (1979/6.)

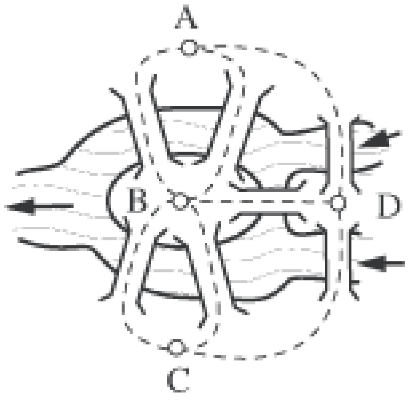
[4] [4] Surányi János: Hérón-féle háromszögekről, *A Matematika Tanítása* **XXIX** (1982/5.)

[5] [5] *Mathematics and Informatics quarterly*, 3/96, **Volume 6**, Singapore.

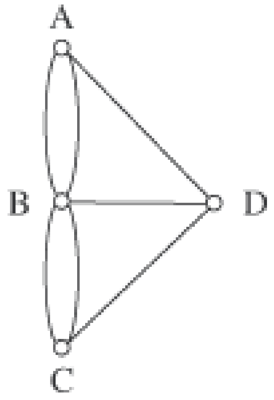
A KÖNIGSBERGI HIDAK, A KILENC ÖSVÉNY ÉS MÁS GRÁFELMÉLETI PROBLÉMÁK

GALLAI TIBOR

Königsbergi hidak. Euler-vonal. A XVIII. század 30-as éveiben Königsberg város polgárai a következő kérdést vetették fel: Königsbergnél a régi és új Pregel összefolyásánál a Pregel folyó egy szigetet alkot, s itt annak idején 7 híd állt. Milyen sorrendben kell az egyes hidakon áthaladni, hogy minden hídon egyszer és csakis egyszer menjünk keresztül? (1. ábra.)



1. ábra



2. ábra

A kérdésre csak tagadó válasz adható: a hidakat nem lehet a kívánt módon bejárni. Ezt szeretnénk most belátni.

Először a térképvázlatot egy egyszerűbb képpel helyettesítjük. Minden partot egy-egy ponttal (

A

,

B

,

C

,

D

) jelképezünk, a hidakat pedig egy-egy, a pontokat összekötő vonallal (2. ábra). A kapott ábrát *gráfnak* nevezik. Ebben lényegtelen a pontok helyzete és a vonalak alakja, csak az a fontos, hogy melyik két pontot hány vonal köt össze.

A

,

B

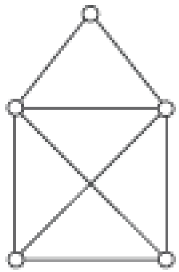
,

C

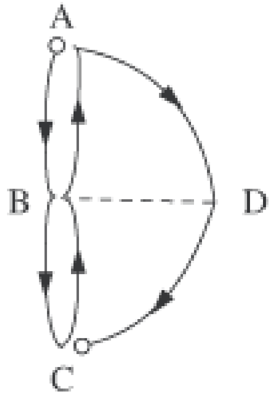
,

D
a gráf *csomópontjai* vagy szögpontjai (a továbbiakban ezeket többnyire csak pontoknak nevezzük), a csomópontokat összekötő vonalak a gráf *élei*. (Egy élnek csak a végpontjai csomópontok!)

Problémánkat most így fogalmazhatjuk meg: hogyan lehet a ceruzát felemelés nélkül úgy végigvezetni a gráf élein, hogy mindegyiken egyszer és csakis egyszer haladjunk keresztül, vagy még rövidebben, hogyan lehet az ábrát egyetlen összefüggő *vonallal* megrajzolni? Ehhez hasonló az a közismert feladat, amely a 3. ábrán



3. ábra



4. ábra

látható „ház" egyetlen vonallal való megrajzolását követeli meg. Itt a gráf 5 csomópontból és 8 élből áll. (A négyzet átlóinak metszéspontját tehát nem tekintjük csomópontnak, ott áthaladva nem szabad irányt változtatni, a teljes átlók alkotnak csak egy-egy élt.)

A 2. ábra megrajzolásával próbálkozva egy híján sikerült minden élt egyetlen vonalba belefoglalni (4. ábra). Figyeljük meg itt, hogy egy olyan pontnál, amelyik sem nem kezdőpont, sem nem végpont (

B
és

D
), ugyanannyiszor lépünk be, mint ahányszor kifutunk, a kezdőpontnál (

A
) eggyel többször lépünk ki, a végpontnál (

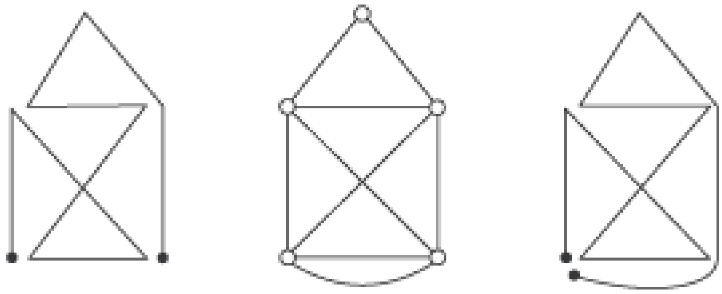
C
) pedig eggyel többször lépünk be (mint ahányszor be-, ill. kifutottunk). Nem nehéz utánagondolni, hogy megállapításaink nemcsak a 4. ábrára, hanem minden olyan, egyetlen vonallal megrajzolt gráfra is érvényesek, amelynél a vonal kezdőpontja nem esik egybe a végpontjával. A rövidség kedvéért nevezzünk egy vonalat *nyitottnak*, ha kezdő- és végpontja különböző, *zártnak* pedig, ha e két pont egybeesik. Ekkor megállapításainkból következik, hogy bármely nyitott vonallal megrajzolt gráfban a kezdő- és végponttól különböző pontokhoz páros számú, a kezdő- és végponthoz pedig páratlan számú él illeszkedik. Egy ponthoz illeszkedő élek számát a pont *fokának* nevezik. Egy pontot, amelynek foka páros, ill. páratlan, röviden *páros*, ill. *páratlan* pontnak fogunk mondani. Az előzőek most így fogalmazhatók:

egyetlen nyitott vonallal megrajzolt gráfban a kezdő- és a végpont páratlan, a többi pont pedig páros.

Ugyanilyen módon látható be a következő állítás is:

egyetlen zárt vonallal megrajzolt gráfban valamennyi pont páros.

Állításaink alapján megállapíthatjuk, hogy a königsbergi hidak gráfja (2. ábra) nem rajzolható meg sem zárt, sem pedig nyitott vonallal. Ebben a gráfban ugyanis mind a négy pont páratlan. A 3. ábra gráfja nem rajzolható meg zárt vonallal, mert a két alsó pont páratlan. Ha nyitott vonallal megrajzolható, a kezdő- és végpontnak az alsó két pontnak kell lennie. Néhány próbálkozás után valóban sikerült megfelelő nyitott vonalat találni (5. ábra). Ha az alsó két pontot még egy újabb éllel is összekötjük, akkor az új gráfban (6. ábra) már minden pont páros lesz, és az 5. ábra alapján azonnal látható, hogy az új gráf megrajzolható egyetlen zárt vonallal (7. ábra).

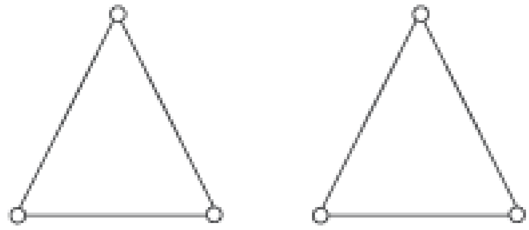


5. ábra   6. ábra   7. ábra

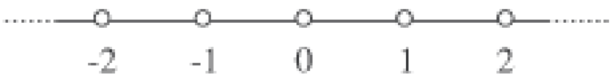
Az elmondottakból sejthető a következő, *Eulertől* származó tétel (a tétel bizonyítását nem közöljük):

Minden véges, összefüggő és csupa páros pontot tartalmazó gráf megrajzolható egyetlen zárt vonallal.

A gráf összefüggő volta nyilvánvalóan szükséges az egyetlen vonallal való megrajzolhatósághoz. (A 8. ábrán látható, két különálló háromszögből álló gráfot nem lehet egyetlen vonallal megrajzolni.)

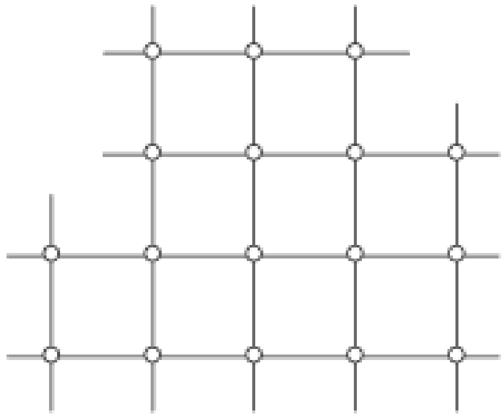


8. ábra



9. ábra

Fel kell tenni a gráf végességét is (véges a gráf, ha véges sok éle és véges sok csomópontja van), mert hiszen minden zárt és nyitott vonal véges sok élből áll, egy végtelen gráf pedig általában végtelen sok élt tartalmaz. A 9. ábrán a végtelen gráf csomópontjai a számegyenesnek az egész számokhoz tartozó pontjai, élei a szomszédos egészeket összekötő szakaszok. A 10. ábrán a „végtelen kockás papiros" gráfját, ill. annak egy részletét tüntettük fel.



10. ábra

Megemlítjük, hogy az olyan zárt vonalakat, amelyek egy gráf minden élét tartalmazzák, a gráf *Euler-vonalainak* nevezzük.

Euler tételéből következtetést vonhatunk le olyan összefüggő és véges gráfok megrajzolhatóságára, amelyekben két páratlan pont van, és amelyekben valamennyi többi pont páros. Ha egy ilyen gráfban a két páratlan pontot egy új éllel összekötjük, akkor olyan összefüggő és véges gráfhoz jutunk, amelyben minden pont páros. (Vö. a 3. és 6. ábrával!) Euler tétele szerint ez megrajzolható egyetlen zárt vonallal (7. ábra). Ha ebből a zárt vonalból elhagyjuk az új élt, olyan nyitott vonal marad vissza, amely az eredeti gráf minden élét tartalmazza, és amelynek kezdő- és végpontja a két páratlan pont (5. ábra). Ezzel beláttuk a következő állítás helyességét:

Minden olyan véges és összefüggő gráf, amelyben két páratlan pont van, és amelyben a többi pont páros, megrajzolható egyetlen nyitott vonallal. A nyitott vonal kezdő- és végpontja a két páratlan pont.

Hasonló meggondolásokkal látható be Euler tételéből a következő, általánosabb tétel is: ha egy véges és összefüggő gráfban a páratlan pontok száma $2k$ (k tetszőleges pozitív egész szám), akkor a gráf k számú nyitott vonallal megrajzolható. E nyitott vonalak mindegyikének kezdő- és végpontja egy-egy páratlan pont. Nem nehéz azt sem belátni, hogy az említett tulajdonságokkal rendelkező gráf nem rajzolható meg k-nál kevesebb számú vonallal. (Pl. a 2. ábra gráfja két vonallal megrajzolható, eggyel nem.)

Az említett tételekkel a véges és összefüggő gráfok mindegyikét érintettük. Egyszerűen megmutatható ugyanis, hogy véges gráf nem tartalmazhat 1 vagy 3 vagy 5… páratlan pontot.

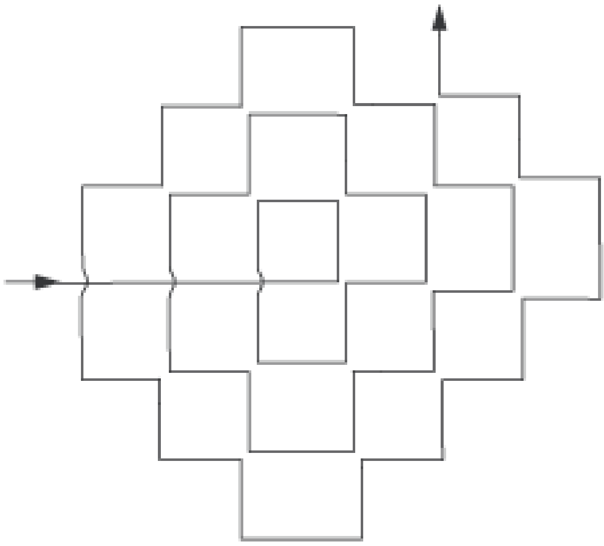
Másképpen kifejezve:

Minden véges gráfban a páratlan pontok száma páros.

A végesség kikötése lényeges, mert pl. az a gráf, amelynek csomópontjai a számegyenes pozitív egész számokkal megjelölt pontjai, élei pedig a számegyenes e pontokat összekötő szakaszai, egyetlen páratlan pontot tartalmaz (11. ábra).

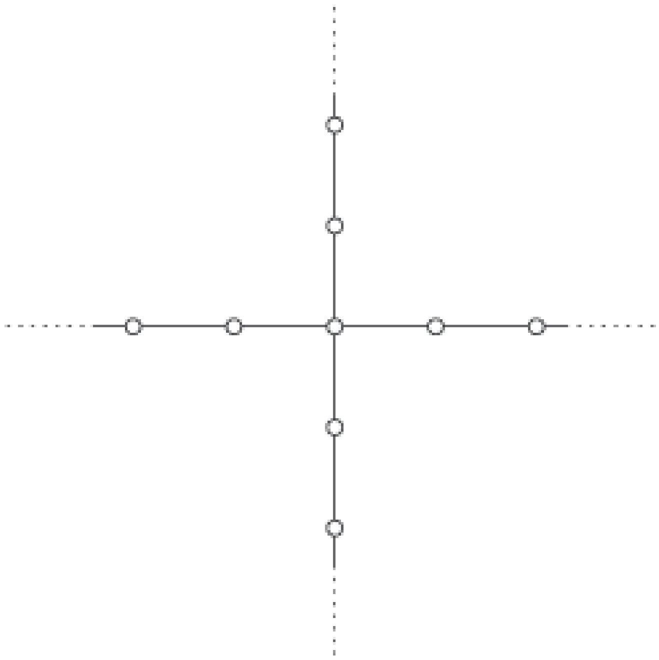
11. ábra

A végtelen gráfokkal kapcsolatban is felvethető az a kérdés, hogy melyek „rajzolhatók meg” egyetlen, mindkét irányban végtelen vonallal. Nyilvánvalóan megrajzolható ilyen módon a 9. ábra gráfja. A 12. ábra mutatja, hogy a 10. ábra gráfjával ugyanez a helyzet.



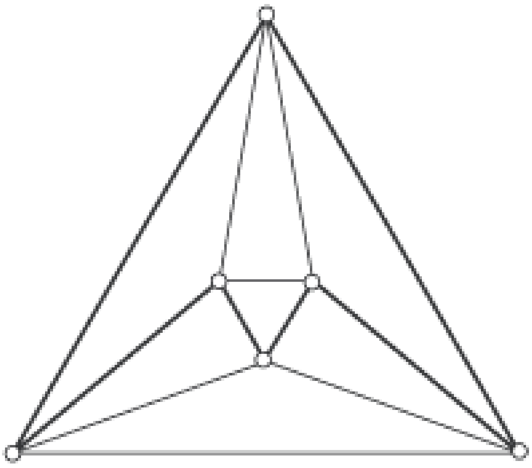
12. ábra

A megrajzolhatósághoz itt is szükséges a gráf összefüggő volta, valamint az, hogy minden pont páros legyen (feltételezzük, hogy minden ponthoz csak véges sok él illeszkedik). A 13. ábra gráfja (ennek csomópontjai egy derékszögű koordináta-rendszer két tengelyén fekvő, egész koordinátájú pontok, élei pedig a tengelyeknek e pontokat összekötő szakaszai) azonban azt mutatja, hogy ez a két feltétel nem elégséges ahhoz, hogy egy végtelen gráf a kívánt módon megrajzolható legyen. Megadhatók a megrajzolásnak szükséges és elégséges feltételei is; ezek azonban összetettebbek.



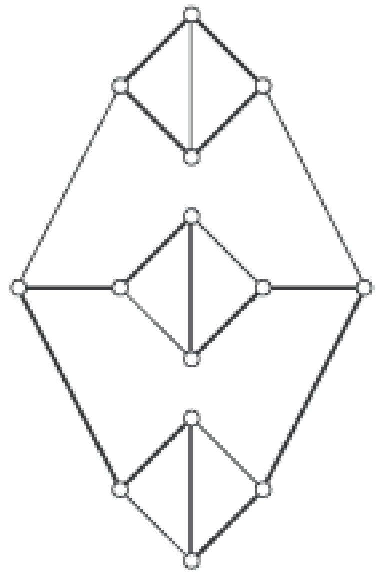
13. ábra

Hamilton-vonal. Másodfokú faktorok. Az eddig tárgyalt problémákban szereplő vonalak „keresztezhették” önmagukat (l. pl. a 7. ábrát!). A gráfelméletben fontos szerepet játszanak az önmagukat nem keresztező vonalak is. Ezek közül a zártakat *körutaknak*, röviden *köröknek* nevezzük. Ha egy kör 3, 4 stb. élt tartalmaz, akkor háromszögnek, négyszögnek stb. is mondjuk.



14. ábra

A 14. ábrán a vastagon rajzolt élek olyan kört (hatszöget) alkotnak, amely a gráf minden csomópontján átmegy. Egy ilyen kört a gráf *Hamilton-vonalának* nevezünk. Nem minden gráfnak van Hamilton-vonala. Pl. a 15. ábrán látható gráfnak nincs.



15. ábra

A Hamilton-vonalak létezésére nem ismerünk olyan feltételeket, melyek szükségesek és elégségesek is. A következőkben ismertetünk két egyszerűen megfogalmazható elégséges feltételt (ezeknek bizonyítása lényegesen nehezebb, mint az előzőekben említett *Euler*-tételé).

Ha egy gráfban minden pont foka

$\geq k$
(

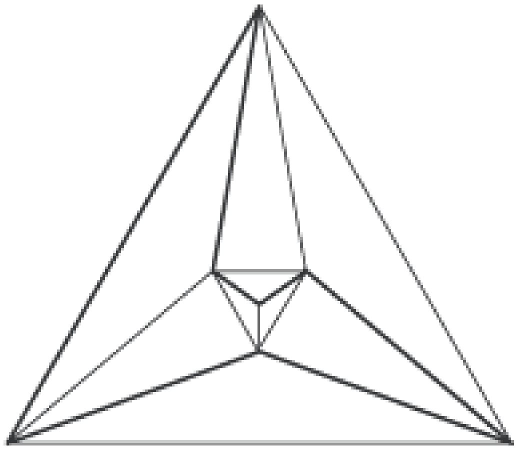
k
tetszőleges, 1-nél nagyobb egész szám) és a gráf nem tartalmaz

$2k$
-nál több pontot, akkor van Hamilton-vonala. (*Dirac* tétele.)

(E tétel alkalmazható pl. a 14. ábra gráfjára.)

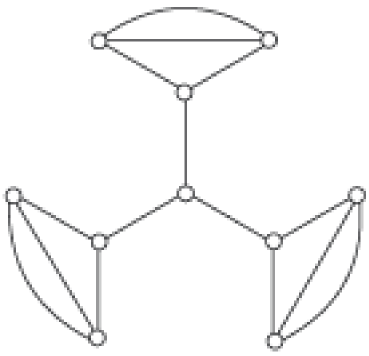
Ha egy háromszöget úgy bontunk szét háromszögekre, hogy bármely két háromszögnek (az eredetit is beleértve) *vagy nincsen közös pontja, vagy egyetlen csúcsuk, vagy egy egész oldaluk közös* (l. 14. ábra), *továbbá ha az eredeti háromszöget nem tekintve, az ábra egyetlen háromszögén belül sincsenek további háromszögek* (a 16. ábra felosztása nem ilyen), *akkor annak a gráfnak, amelynek csomópontjai a háromszögek csúcsai, élei pedig a háromszögek oldalai, van Hamilton-vonala.* (*Whitney* tétele.)

Ez a tétel is alkalmazható a 14. ábra gráfjára. Mindkét tételben a feltételek nem „szükségesek”, tehát léteznek olyan gráfok, amelyek e feltételeknek nem tesznek eleget, és mégis van Hamilton-vonaluk. (A 16. ábra gráfja is ilyen.)



16. ábra

Említettük, hogy a 15. ábra gráfjának nincsen Hamilton-vonala, tehát nem lehet a gráf valamennyi csomópontját belefoglalni egyetlen körbe. Lehetséges azonban a csomópontokat két olyan körbe foglalni, amelyeknek nincs közös pontjuk. (A megvastagított élek két ilyen kört alkotnak.) Egy tetszőleges gráfnál a köröknek olyan rendszerét, amelynek körei együtt a gráfnak minden csomópontját tartalmazzák, de amelyek közül bármelyik kettőnek nincs közös pontja, a gráf *másodfokú faktorának* nevezzük. (A Hamilton-vonal is másodfokú faktor!) Másodfokú faktora sincs minden gráfnak, még olyanoknak sem mindig, amelyekben – ilyen volt a 15. ábra gráfja is – minden pont foka ugyanakkora. A 17. ábrán levő gráfban minden pont harmadfokú, de a gráfnak nincsen másodfokú faktora. Mint az alábbi, bizonyítás nélkül közölt (Petersentől származó) tételek mutatják, egy gráfnak az a „szabályossága”, hogy minden pontjának ugyanakkora a foka, esetleges egyszerűbb kiegészítő feltételekkel együtt biztosítja másodfokú faktor létezését.



17. ábra

Ha egy véges gráf minden pontjának ugyanakkora a foka, és ez a fokszám pozitív páros szám, akkor a gráfnak van másodfokú faktora.

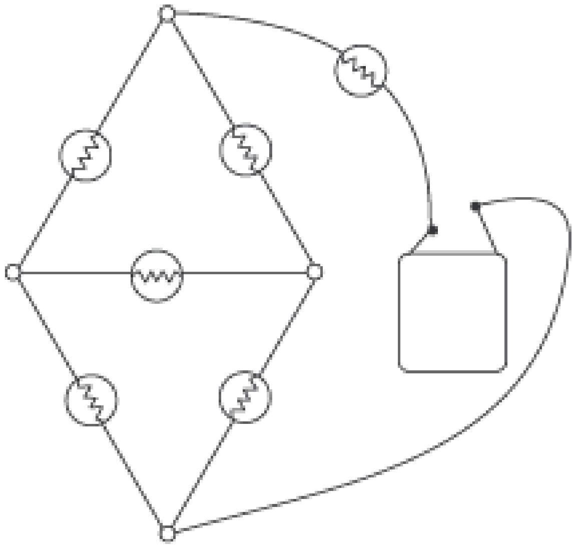
(Ez a tétel alkalmazható pl. a 14. ábra gráfjára.)

Ha egy véges és összefüggő gráf minden pontjának ugyanakkora a foka, és ez a fokszám 1-nél nagyobb páratlan szám, továbbá, ha a gráfnak nincsen hídja, akkor a gráfnak van másodfokú faktora. (Hídnak nevezzük egy összefüggő gráfnak olyan élét, amelynek elhagyása két össze nem függő részre bontja a gráfot.)

A 17. ábra gráfjának három hídja van. Az utóbbi tétel alkalmazható a 15. ábra gráfjára.

Megjegyezzük, hogy az utolsó két tétel feltételeinek teljesülése sem szükséges a másodfokú faktor létezéséhez. Pl. a 3. ábra gráfjának van másodfokú faktora, pedig nem teljesíti e tételek feltételeit.

Elektromos hálózatok. A gráfok köreinek vizsgálata igen fontos a gráfelmélet elektrotechnikai alkalmazásaiban. A múlt század 40-es éveiben *Kirchhoff* dolgozta ki az elektromos hálózatok számítási eljárásának alapjait. Ezekben felhasználta a gráfok köreinek néhány tulajdonságát. Egy elektromos hálózat ágai a gráf éleinek, összekapcsolási pontjai a gráf csomópontjainak tekinthetők. A továbbiakban csak olyan hálózatokról lesz szó, amelyekben csupán ohmikus ellenállások (pl. izzólámpák) és telepek (zseblámpaelemek) fordulnak elő. A 18. ábrán egy 6 ágból és 4 csomópontból álló hálózat látható.



18. ábra

Telep csak az egyik ágban van; ennek elektromotoros erejét (feszültségét)

E

-vel jelöljük (egy zseblámpaelemnél ez 4,5 volt). A telepet tartalmazó ág ellenállása, a telep belső ellenállásával együtt legyen

R_0

(

R_0

az ohmok számát jelöli), a többi ágban elhelyezett ellenállások nagysága legyen rendre

R_1

,

R_2

,

R_3

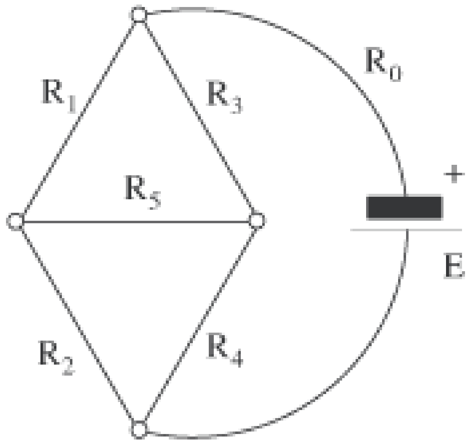
,

R_4

és

R_5

.



19. ábra

A 19. ábrán az izzókat jelző köröket elhagytuk, csak az ellenállások értékét írtuk a megfelelő gráfélek mellé, az elem rajzát pedig a szokásos jelzéssel helyettesítettük.

A legegyszerűbb számítási feladatokban megadják az

E
feszültség és az

$R0$
,

$R1$
, …,

$R5$
ellenállások nagyságát, és ezekből kell meghatározni az egyes ágakban folyó áramok erősségét (intenzitását), valamint az áramlási irányokat. A továbbiakban azt akarjuk megmutatni, hogyan segíti elő a hálózati gráf tulajdonságainak vizsgálata az intenzitások meghatározását.

Az áramlási irányok meghatározása nem okoz külön gondot. Vegyünk fel ugyanis minden egyes ágban tetszőlegesen egy-egy áramlási irányt, és ezzel együtt tekintsük a meghatározandó

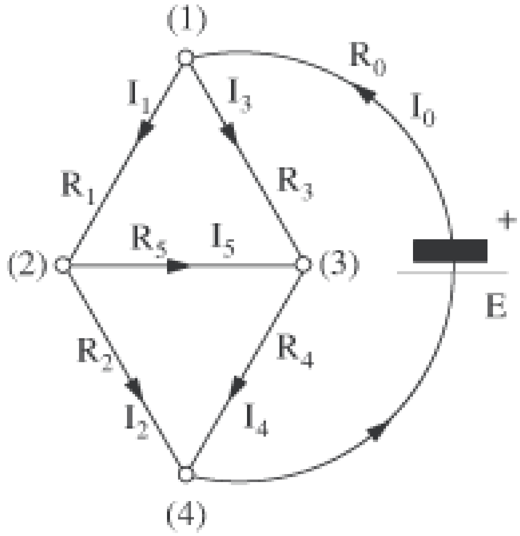
$I0$
,

$I1$
, …,

$I5$
intenzitásokat előjeles (pozitív vagy negatív) mennyiségeknek, olyan értelemben, hogy ha egy ágban, pl.

$R1$
-ben, a tényleges áramlás iránya megegyezik a felvett iránnyal, akkor

$I1$
-et pozitívnak, ha ellentétes vele, akkor negatívnak tekintjük.



20. ábra

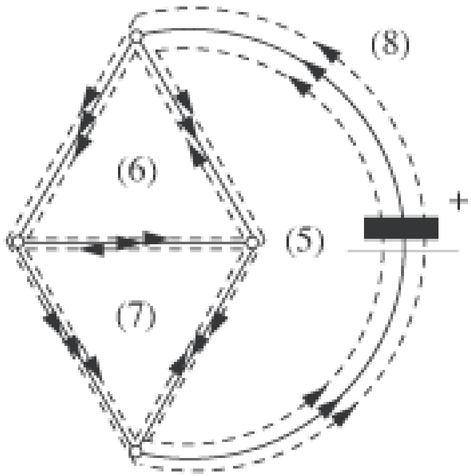
(20. ábra. Pl. ha a számítások az $I_1=-0,5$ amper értéket szolgáltatják, akkor az R_1 ágban a tényleges áramlás iránya a (2)-es csomóponttól az (1)-es felé mutat.)

6 ismeretlent (I_0, I_1, …, I_5) kell meghatároznunk. Ehhez 6 egyenletre van szükségünk, mégpedig olyan 6 egyenletre, amelyek *függetlenek* egymástól, vagyis ahol egyik sem „következménye” a másik ötnek. (Pl. az $x+y=5$ és a $2x+2y=10$ egyenletekből nem lehet egyértelműen kiszámítani az x és y ismeretleneket. Ezek az egyenletek nem függetlenek egymástól, a második „semmi újat” nem mond az elsőhöz képest; az elsőt kielégítő, bármelyik x, y számpár a másodikat is kielégíti.) Az egyenleteket az ún. *Kirchhoff*-féle törvények alapján lehet felírni. Az első

törvény azt mondja ki, hogy bármely csomópontnál a befutó áramok intenzitásainak összege egyenlő a kifelé haladók intenzitásainak összegével. Ennek alapján minden csomóponthoz felírhatunk egy-egy egyenletet:

(1)	$I_0=I_1+I_3$
(2)	$I_1=I_2+I_5$
(3)	$I_3+I_5=I_4$
(4)	$I_2+I_4=I_0$

A második *Kirchhoff*-törvény a hálózat köreire vonatkozik (és lényegében az Ohm-törvénynek egyszerű következménye). Válasszunk ki a gráfban egy kört, és lássuk el tetszőleges befutási iránnyal (21. ábra szaggatott vonalai)!



21. ábra

A kör által tartalmazott minden egyes ág (él) ellenállását szorozzuk meg az ágban folyó áram intenzitásával, és lássuk el ezeket a szorzatokat

+

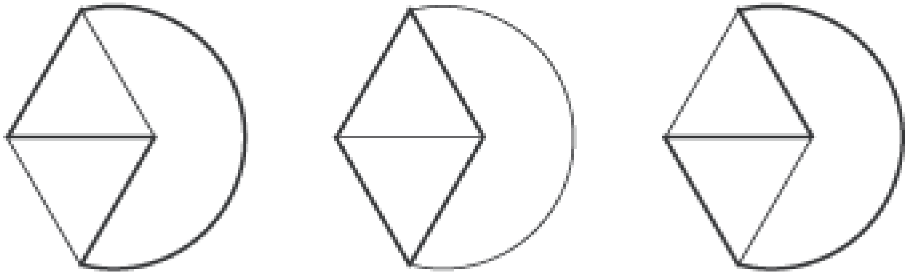
vagy

–

előjellel aszerint, hogy az ág irányítása megegyező vagy ellenkező a kör befutási irányával. A Kirchhoff-törvény szerint ekkor a kapott értékek összege megegyezik a kör ágaiban elhelyezkedő telepek (megfelelő előjellel vett) feszültségeinek összegével. A 21. ábrán négy irányított kört tüntettünk fel. Az ezekhez tartozó egyenletek:

(5)	$R_0I_0+R_3I_3+R_4I_4=E$
(6)	$R_1I_1+R_5I_5-R_3I_3=0$
(7)	$R_2I_2-R_4I_4-R_5I_5=0$
(8)	$R_0I_0+R_1I_1+R_2I_2=E$

(Egy kör befutási irányának megváltoztatása a körhöz tartozó egyenletben a tagok előjelének megváltozását eredményezi.) Ha a befutási irányoktól eltekintünk, akkor gráfunknak a feltüntetett 4 körön kívül (ezek mind háromszögek) még három köre van (ezek négyszögek, l. 22. ábra).



22. ábra

Irányításuk után ezekhez is felírható egy-egy egyenlet. Így 7 kör-egyenlet és 4 csomóponti egyenlet, tehát összesen 11 egyenletünk lesz. Ezek közül kell 6 függetlent kiválasztani. A választásban nem szerepelhet mind a négy csomóponti egyenlet. Könnyen ellenőrizhető ugyanis, hogy közülük az első hármat összeadva éppen a negyediket kapjuk. De nem szerepelhet mind a négy felírt kör-egyenlet sem, mert itt is igaz, hogy közülük az első három összege éppen a negyediket (a (8)-as egyenletet) adja.

A csomóponti egyenletekre belátható, hogy közülük bármely három független egymástól; sőt általánosan is igaz, hogy bármilyen összetett hálózat esetén, ha a hálózat

p

számú csomópontot tartalmaz, akkor bárhogyan is választunk ki közülük

p–1

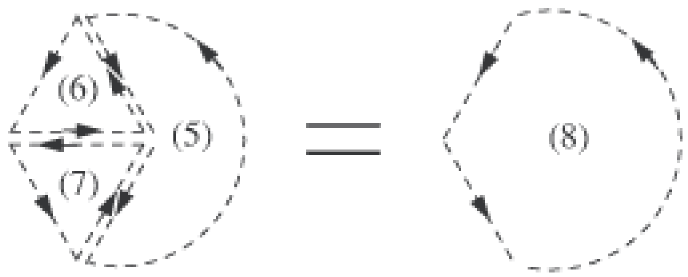
csomópontot, az ezekhez tartozó egyenletek függetlenek egymástól (a

p

-edik csomóponthoz tartozó viszont következménye a többinek).

Nehezebb feladat a 7 kör-egyenlet közül kiválasztani 3 függetlent (a 3 csomóponti egyenlettel együtt így juthatnánk 6 független egyenlethez). Fő célunk éppen az, hogy példánkon keresztül megmutassuk, miként lehet kiválasztani a hálózati gráf vizsgálata révén (tetszőleges hálózat esetén is) a szükséges számú megfelelő kört.

A kör-egyenletek között fennálló kapcsolatok gyökere az irányított gráfkörök közti kapcsolatban keresendő. Hogy az (5)-, (6)- és (7)-tel jelzett egyenletek összege éppen a (8)-as egyenletet fogja adni, az kitűnik a 23. ábrából,

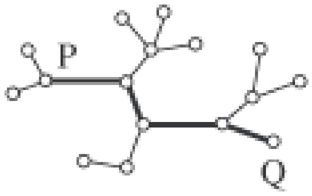


23. ábra

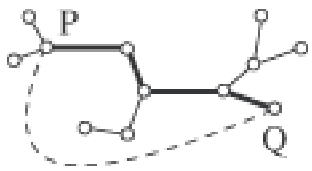
ha annak bal oldali részén az ellentétesen befutott éleket „megsemmisítetteknek” tekintjük.

Emiatt elegendő az egyenletek közti kapcsolatok helyett a gráfkörök közti kapcsolatokat vizsgálni. E vizsgálathoz először meg kell ismernünk az olyan gráfoknak néhány tulajdonságát, amelyekben egyáltalán nincsen kör.

A kör nélküli összefüggő gráfok rajza fához hasonlít, amiért is az ilyen gráfokat *fáknak* nevezzük (24. ábra).



24. ábra



25. ábra

A szemlélet alapján világosak a fák következő tulajdonságai:

1. Egy fa két pontját egyetlen olyan úttal lehet csak összekötni, amelyik a fa éleiből van összerakva. (A 24. ábrán a megvastagított élek mutatják a

   P
   és

   Q
   pontokat összekötő utat.)

2. Ha egy fa két pontját egy új éllel kötjük össze, akkor olyan gráfot kapunk, amelyben pontosan egy kör van (25. ábra).

3. Minden fa tartalmaz *végéleket* (a végél olyan él, amelynek egyik végpontjához csak ez az él illeszkedik).

   A 3. tulajdonságból egyszerűen belátható (a végélek egymás utáni elhagyásával végül is egy egyélű fához jutunk), hogy

4. *minden fában a pontok száma 1-gyel több az élek számánál:*

   $p=é+1$
   . (

   p=
   pontok száma,

   é=
   élek száma)

Ezek után tekintsünk egy tetszőleges, köröket is tartalmazó, összefüggő gráfot (pl. a 19. ábra gráfját). Ha a gráf egyik körének valamelyik élét elhagyjuk, összefüggő gráf marad vissza, és ez a gráf tartalmazza az eredetinek minden csomópontját. Ha az ilyen élelhagyásokat mindaddig ismételjük, ameddig csak kör van a gráfban, akkor olyan kör nélküli gráfhoz jutunk, amelyik összefüggő, és az eredetinek minden csomópontját tartalmazza (26. ábra).

26. ábra

A visszamaradó gráf tehát olyan fa, amely az eredeti gráfnak része, és annak minden csomópontját tartalmazza.

Az ilyen fákat az eredeti gráf fa alakú vázainak, röviden *favázainak* nevezzük. Egy gráfnak általában sok faváza van. (Sokféleképpen választhatjuk ki a köröket, és azokban az elhagyásra szánt éleket.) Minden faváznak azonban ugyanannyi az éle, ti. pontosan eggyel kevesebb, mint ahány csomópontja van az eredeti gráfnak (lásd 4.), s ezért az eredetiből mindig ugyanannyi

$é-(p-1)$
(

$é=$
az eredeti gráf éleinek száma,

$p=$
az eredeti gráf csomópontjainak száma,

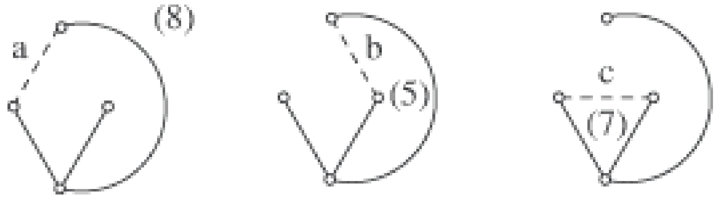
$p-1=$
a faváz éleinek száma) számú élt kell elhagyni ahhoz, hogy egy faváz maradjon vissza.

(A 26. ábrán

$é=6$
,

$p=4$
,

$é-(p-1)=3$.
) A fák 2.) alatt említett tulajdonsága miatt, ha egy favázhoz visszatesszük bármelyik elhagyott élt, olyan gráfot kapunk, amelynek pontosan egy köre van. Tehát minden elhagyott él a favázzal együtt az eredeti gráfnak pontosan egy körét határozza meg (27. ábra).



27. ábra

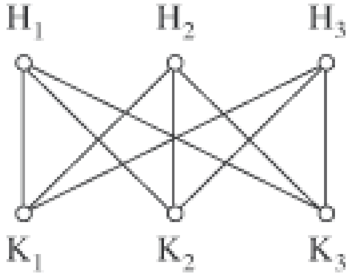
Bebizonyítható, hogy egy faváz és a hozzá tartozó elhagyott élek éppen olyan köröket határoznak meg, amelyekhez tartozó egyenletek függetlenek. Ilyen módon bármely faváz segítségével kiválasztható

$é-(p-1)$
számú független kör-egyenlet. Igazolható, hogy ezek és

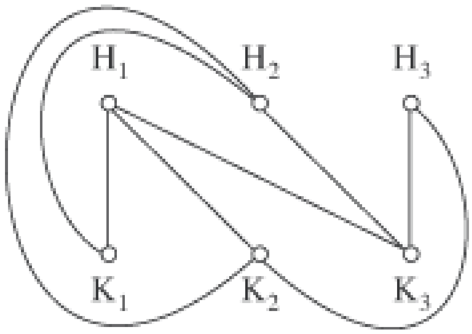
$p-1$
számú csomóponti egyenlet együttvéve is olyan egyenletrendszert alkotnak, amelyben egyik egyenlet sem következménye a többinek. Így tehát éppen

$é-(p-1)+(p-1)=é$
számú független egyenlethez jutunk; annyi egyenlethez, ahány éle van a gráfnak, azaz ahány ismeretlen intenzitást kell meghatároznunk. Ezekből az egyenletekből kiszámíthatók az áramintenzitások. (Példánkban a (8), (5), (7), (1), (2) és (3) egyenletekből.)

A „9 ösvény” vagy a „3 ház – 3 kút” problémája. Három házból vezessünk három kúthoz kilenc ösvényt úgy, hogy minden házból minden kúthoz egy-egy ösvény fusson, és az ösvények ne keresztezzék egymást. Ha a házakat és kutakat egy-egy ponttal, az ösvényeket egy-egy vonallal szemléltetjük, akkor az a feladatunk, hogy a 28. ábrán látható „3 ház – 3 kút-gráfot” oly módon rajzoljuk le, hogy az élek közül semelyik kettő se keresztezze egymást. (A gráf 6 csomópontból és 9 élből áll.)



28. ábra

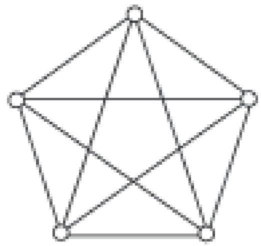


29. ábra

A 29. ábra mutatja, hogy nyolc „ösvényt” még el lehet helyezni kereszteződés nélkül. Ezen az ábrán a hiányzó kilencedik élt, a

*H*3*K*1

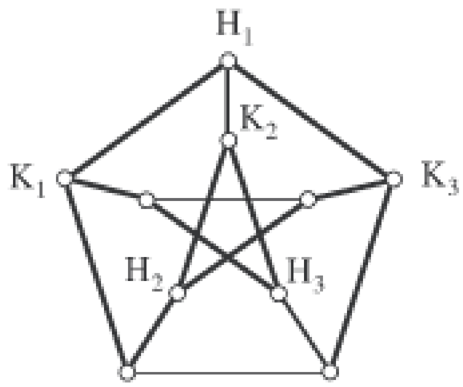
élt csak úgy lehet megrajzolni, hogy az valamelyik, már megrajzolt élt átmetszi. Bebizonyítható, hogy másféleképpen sem lehet mind a 9 élt úgy megrajzolni síklapon, hogy azok közül legalább kettő ne messe egymást. Ezt röviden úgy fejezzük ki, hogy a 3 ház – 3 kút-gráf *nem síkra rajzolható*. Ugyancsak igaz, hogy gömbfelületen sem lehet átmetszés nélkül a szóban forgó gráfot megrajzolni. Nem nehéz azonban belátni, hogy egy gyűrűfelületen (mentőöv), valamint egy Möbius-szalagon (egyszer megcsavart szalag, lásd a térképszínezésről szóló és a Csalafinta felületek című cikket!) a 3 ház – 3 kút-gráf megrajzolható úgy, hogy a 9 él ne keresztezze egymást. A térképszínezésről szóló cikkben szerepel a „teljes ötszög” gráfja (30. ábra).



30. ábra

Ebben öt csomópont van, és mindegyik csomópontot egy-egy él köti össze a többi néggyel. Ez a gráf sem rajzolható sem a síkra, sem a gömbfelületre, vagyis nem rajzolható meg úgy ezeken a felületeken, hogy élei ne messék egymást.

A 3 ház – 3 kút-gráf és a teljes ötszög-gráf a „legegyszerűbb”, síkra nem rajzolható gráfok. Igazolható, hogy minden, síkra nem rajzolható gráf tartalmaz vagy a 3 ház – 3 kút-gráfhoz, vagy a teljes ötszög-gráfhoz „hasonló” szerkezetű gráfot. (*Kuratowski* tétele.) A 31. ábrán



31. ábra

egy 10 csomópontból és 15 élből álló, síkra nem rajzolható gráfot tüntettünk fel. A vastag élek egy, a 3 ház – 3 kút-gráfhoz hasonló szerkezetű gráfot alkotnak benne.

A GALTON-DESZKA

BOGNÁR JÁNOSNÉ

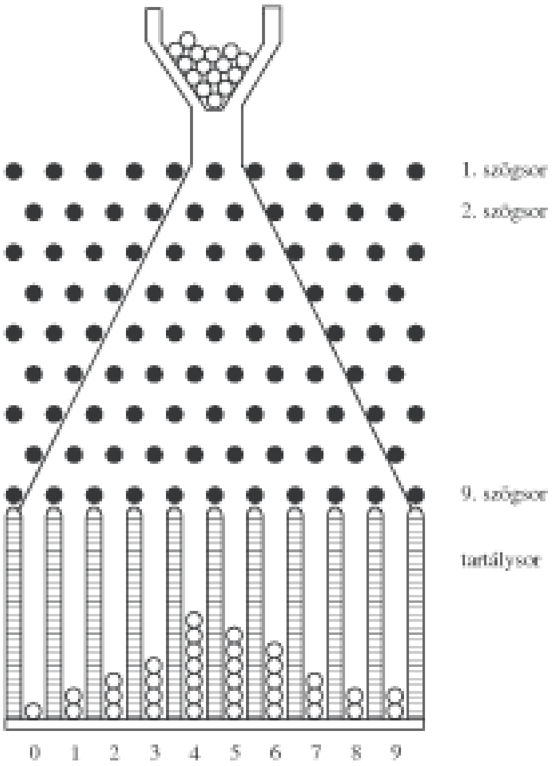
1. BEVEZETÉS

E sorokban szeretnénk röviden összefoglalni (a teljességre való törekvés nélkül), hogy az itt ismertetésre kerülő ún. Galton-deszka segítségével, illetve annak módosított változataival a valószínűségszámítás milyen különböző törvényeit szemléltethetjük, illetve milyen, a szemléltetésen túlmenő felhasználása lehetséges a Galton-deszka egyes módosított alakjainak.

A valószínűségszámítás alapfogalmainak részletes ismertetésére itt nincs lehetőségünk; úgy hisszük azonban, hogy erre nincs is szükség, hiszen több magyar nyelvű munka áll az érdeklődők rendelkezésére. (Elsősorban az [1] könyvet említjük meg, amely kevés matematikai előismerettel rendelkezők számára igen jó bevezetést nyújt a valószínűségszámítás alapfogalmaiba.)

2. A GALTON-DESZKA LEÍRÁSA

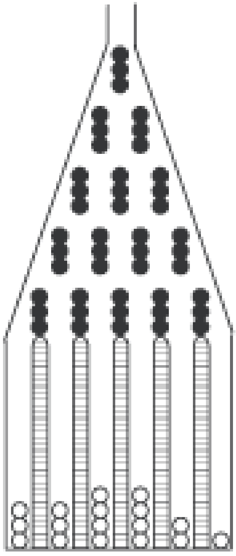
A Galton-deszka eredeti alakjában olyan deszka, amelyre egymással párhuzamos sorokba rendezett szögek vannak elhelyezve (szögsorok), mégpedig úgy, hogy egy adott szögsor szögei mindig a megelőző sor szögei közti intervallumok középpontjai alá esnek egymástól egyenlő távolságban. Az általában függőlegesen vagy lejtősen felállított deszkára egy, az első szögsor középső szöge felé, a szögsorokra merőlegesen irányított tölcséren keresztül apró golyókat lehet bocsátani, amelyek átmérője egyforma és csak kevéssel kisebb, mint a szögek közti távolság. A leguruló golyók nekiütközve az első szögsor szögének, ott véletlenszerűen jobbra vagy balra térnek el. Akármelyik irányba is tért el egy leguruló golyó, a szögek közti „csatornákon” továbbjutva ismét beleütközik a következő szögsor valamelyik szögébe, ahol ismét véletlenszerűen jobbra vagy balra tér el s így tovább, míg végül a deszka utolsó szögsorán való ütközés után a golyó a deszka alján levő tartálysor valamelyik tartályába kerül. (1. ábra)



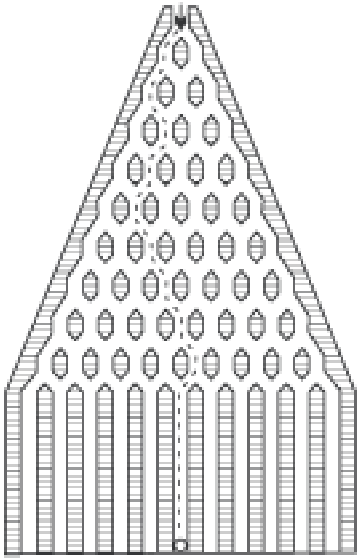
1. ábra

Ha kizárjuk azokat a fizikailag lehetséges eseteket, amikor egy golyó valamelyik szögnek ütközve, attól olyan nagy impulzussal pattan jobbra vagy balra, hogy nem az illető szög mellett közvetlenül elhelyezkedő „csatornák” egyikén folytatja útját a tartálysor felé, hanem egy távolabbi csatornán, akkor nyilvánvaló, hogy a leguruló golyók az egész deszkának csak egy szabályos háromszög alakú részében tartózkodhatnak egyáltalán, úgyhogy ha Galton-deszkáról beszélünk, mindig ilyen szabályos háromszög alakú szerkezetre gondolunk. A Galton-deszka szögsorai tehát rendre 1, 2, 3, 4…szöget tartalmaznak.

A továbbiakban feltesszük, hogy egy golyónak az egyes sorokon való jobbra, balra téréseit nem befolyásolja az a körülmény, hogy a megelőző soron jobbra vagy balra tért-e el, más szóval az az esemény, hogy mondjuk az i-edik sorban jobbra vagy balra tért-e el, független attól, hogy az $(i–1)$-edik sorban merre tért el. Mivel az ilyen szögekből álló Galton-deszka esetében ez a feltétel nem teljesül minden további nélkül, hanem a szögeken való ütközés olyan, hogy ha egy golyó valahol jobbra tért el, akkor a következő soron is „inkább” tér el jobbra, mint balra; ezért ennek kiküszöbölésére az eredeti szöges Galton-deszkán a következő módosítást végezték: egy-egy szög helyére három, közvetlenül egymás alatt elhelyezett szög került (2. ábra);



2. ábra



3. ábra

ezáltal a golyók ütközés után egy hosszabb csatornába kerülnek, ahol „lefékeződnek”, a következő ütközésig elvész az előző ütközés hatása, s az egyes ütközéseknél létrejövő eltérítések függetlensége már feltételezhető. Még nagyobb pontossággal érhető el e feltétel teljesülése, ha a szögek (ill. szöghármasok) helyén hatszög alakú ékek vannak (3. ábra). Ha a Galton-deszka szabályos (az ugyanabban a sorban levő ékek közti csatornák párhuzamosak a háromszög szimmetriatengelyével, minden ék a felette, ill. alatta levő csatorna középvonalába esik), akkor a golyók minden ütközésnél egyforma, vagyis

(1/2)
–

(1/2)
valószínűséggel térnek el jobbra és balra. Egy golyó valamilyen véletlenszerű lefutását *útnak* fogjuk nevezni, az egyik sorból a másik sorba való jutást pedig *lépésnek*. A 3. ábra 10 éksorból álló Galton-deszkáján egy lehetséges véletlen utat ábrázoltunk.

3. EGY JÁTÉK A GALTON-DESZKÁVAL. A GALTON-DESZKA ÉS A BINOMIÁLIS ELOSZLÁS

Álljon most a Galton-deszka

N
éksorból, és az utolsó éksor ékei közti csatornák alatt, az

$N+1$
-edik éksor helyén álljanak tartályok. Mivel az első sorban 1 ék, a második sorban 2 ék, végül az

N
-edik sorban

N
ék (és

$N+1$
csatorna) van, az

$(N+1)$
-edik sor ékei helyét betöltő tartálysor

$N+1$
tartályból áll. Számozzuk meg ezeket balról jobbra, mégpedig a későbbiek kedvéért 0-val kezdjük a számozást. Tehát a tartályok sorszámai rendre 0, 1, 2, …,

N
.

Képzeljük el, hogy két játékos,

A
és

B
, a következő játékot játssza:

A
fogad, hogy egy legurított golyó a

k
-adik tartályba fog esni

$(k=0,1,2,\ldots,N)$
. Ha eltalálta, kap

B
-től

xk
fillért; ha nem találta el, vagyis ha a golyó az

l
-edik

$(l\neq k)$
tartályba esik, akkor

A
fizet

B
-nek

$yl(xk)$
fillért. (A

B
-nek fizetendő összeg általában függjön az

xk
-tól!) A következő játszmában

A
és

B
szerepet cserélnek. Kérdés, hogyan kell az

$yl(xk)$
értékeket megválasztanunk, hogy a játék méltányos legyen, azaz hogy az

A

(B)
játékos átlagos nyereménye (pontosabban:

A
nyereményének várható értéke minden tipp esetén) 0 fillér legyen.

Ahhoz, hogy ezt meg tudjuk határozni, elsősorban az érdekel bennünket, hogy egy legurított golyó melyik tartályba jut, és milyen „biztonsággal” lehet ezt megjósolni; vagyis ismernünk kell, hogy az eseteknek kb. hányad részében, pontosabban *mekkora valószínűséggel* esik egy golyó egy megadott, mondjuk

k
-adik tartályba. Ez a valószínűség a tartály sorszámán kívül nyilván függ a Galton-deszka éksorainak számától is.

N
éksor esetén jelöljük ezt a valószínűséget

$PN(k)$
-val!

$PN(k)$
tehát annak a valószínűségét jelenti, hogy egy

N
éksorból álló Galton-deszkán egy legurított golyó a

k
-adik tartályba kerül

$(k=0,1,2,\ldots,N)$
.

Egy adott golyónak minden sorban két lehetősége van a továbbjutásra: jobbra vagy balra tér el, így a deszkán való lehetséges lefutásainak a száma (a lehetséges utak száma)

$2\square 1\square 2\square 2\square \ldots \square 2\square N=2N.$
Mivel feltettük, hogy a Galton-deszka szabályos, és az egyes éksorokon való eltérítések függetlenek egymástól, így minden egyes útnak

$(1/2N)$
a valószínűsége. Ebből azonban még korántsem következik, hogy minden tartályba ugyanolyan valószínűséggel eshet egy golyó, hiszen több különböző (kedvező) úton is eljuthat ugyanabba a tartályba. Ha pl. mind az

N
sorban balra tér el a golyó, akkor nyilván a 0-adik tartályba jut; ha közben valahol egyszer jobbra tért el, akkor

az első tartályba, ha közben (bármelyik) két sornál jobbra tért el, akkor a második tartályba és így tovább, általában a balról számított

k
-adik tartályba azokon az utakon juthat a golyó, amelyeken összesen

k
-szor tért el jobbra és

$(N-k)$
-szor balra. Az

N
éksor közül az a

k
sor, ahol jobbra tért el,

$(N(N-1)(N-2)\cdot\ldots\cdot(N-k+1)/1\cdot 2\cdot 3\ldots k)=(Nk)$
-féleképpen választható ki, így a

k
-adik tartályba vezető utak száma

(Nk)
, amiből következik, hogy

(1) $$PN(k)=((Nk)/2N),\ (k=0,1,2,\ldots,N).$$

Ha pl.

$N=8$
, akkor az egyes tartályokba való jutás valószínűségei rendre a következők:

$(1/256)$
,

$(8/256)$
,

$(28/256)$
,

$(56/256)$
,

$(70/256)$
,

$(56/256)$
,

$(28/256)$
,

$(8/256)$

,

(1/256)

.

Megjegyezzük, hogy az (1) formulát más úton is beláthatjuk, mégpedig az éksorok számára,

N

-re vonatkozó teljes indukcióval.

$N=1$

-re ugyanis nyilván igaz (1), hiszen

$P1(0)=(1/2)=((10)/21)$

és

$P1(1)=(1/2)=((11)/21)$

. Ha már ismerjük egy

N

éksorú Galton-deszkára, hogy egy legurított golyó mekkora valószínűséggel esik az egyes tartályokba, akkor ebből könnyen megkaphatjuk egy

$N+1$

éksorból álló Galton-deszka esetére is a megfelelő valószínűségeket. Az

$N+1$

éksorú Galton-deszka

k

-adik

$(k=1,2,\ldots,N)$

tartályába ugyanis kétféleképpen juthat egy golyó: vagy az

N

-edik sor

$(k–1)$

-edik csatornáján halad keresztül és azután jobbra tér el, vagy a

k

-adik csatornán és azután balra tér el. Mivel az

$(N+1)$

-edik sor 0-adik, ill.

$(N+1)$

-edik tartályába csak egyféleképpen juthat egy golyó az

N

-edik sorból, ez összhangban lesz a

$PN(–1)=PN(N+1)=0$

értelmezéssel. Így

$PN+1(k)=PN(k–1)(1/2)+PN(k)(1/2)=((Nk–1)+(Nk)/2N+1)=((N+1k)/2N+1).$

Ezzel egyben egyszerű számolási eljárást kaptunk a

$P_{N+1}(k)$
valószínűségek kiszámítására a

$P_N(k)$
valószínűségek ismeretében.

Vizsgáljuk meg kissé közelebbről ezeket a valószínűségeket! Az, hogy egy golyó a

k
-adik tartályba

$P_N(k)$
valószínűséggel juthat, azt jelenti, hogy a leengedett golyóknak kb.

$P_N(k)$
-ad része

$(100 \cdot P_N(k)\%$-a)
kerül a

k
-adik tartályba. Ha nagyszámú, mondjuk

R
darab golyót gurítunk le a Galton-deszkán és

r_k
-val jelöljük ezek közül a

k
-adik tartályba eső golyók számát (amely szám a golyók átmérőjével mint egységgel mérve, egyben azt a magasságot is megadja, ameddig a

k
-adik tartály meg lesz töltve), akkor az

(r_k/R)
hányados, vagyis annak az eseménynek a relatív gyakorisága, hogy egy legurított golyó a

k
-adik tartályba jut, az esetek legnagyobb részében (nagyon sokszor legurítva

R
számú golyót, ezek közül „legtöbbször”) csak kevéssel fog eltérni a szóban forgó esemény valószínűségétől, a

$P_N(k)$
-tól, így

r_k
csak kevéssel fog eltérni az

$R \cdot P_N(k)$
-tól. Tehát a golyók a tartályokat az odaesés valószínűségével arányos magasságig fogják megtölteni. Mivel a

$P_N(k)$

valószínűségek ún. binomiális eloszlást alkotnak7

$(p=(1/2))$
, jól tanulmányozhatjuk a Galton-deszka segítségével a binomiális eloszlás képét. Megfigyelhetjük, hogy a középső tartályokba jut a legtöbb golyó, a deszka szélei felé haladva pedig egyre kevesebb; s hogy a Galton-deszka szimmetriatengelyére szimmetrikus tartályokba (

k
-adik és

$(N–k)$
-adikba) kb. ugyanannyi golyó kerül

(hiszen $(Nk)=(NN–k)$)
. Már aránylag kevés golyó legurítása esetén is észrevehető az a tény, hogy páros

N
esetén éppen az

$(N/2)$
-edik sorszámú középső tartályba, páratlan

N
esetén pedig az

$(N–1/2)$
-edik és

7A

$P0$
,

$P1$
, …,

PN
számokról azt mondjuk, hogy

N
-ed rendű,

p
paraméterű binomiális eloszlást alkotnak, ha

$Pk=(Nk)pk(1–p)N–k$ $(k=0,1,2,…,N)$,
vagy másképpen: egy

ξ
valószínűségi változóról azt mondjuk, hogy (

N
-ed rendű,

p
paraméterű) binomiális eloszlású, ha lehetséges értékei 0, 1, 2, …,

N
és

k
értékét

Pk
valószínűséggel veszi fel.

$(N+1/2)$
-edik sorszámú két középső tartályba esik a legtöbb golyó (legvalószínűbb tartályok).

Hogy az olvasó a későbbi konkrét számpéldák eredményeit könnyebben követhesse, illetve ellenőrizhesse, megemlítjük azt a fontos tételt, hogy a binomiális eloszlás tagjai bizonyos feltételek mellett a

$\square(x)=(1/\sqrt{(2\pi)})e{-}(x2/2)$
ún. Gauss-féle függvény (haranggörbe) segítségével közelíthetők meg (utóbbi értékei táblázatokból kikereshetők; ilyen táblázat található pl. a [2], [3] könyvekben), mégpedig a mi esetünkben és jelöléseinkkel, ha

N
elég nagy és

k
közel van az

$(N/2)$
-höz, valamint

$(\square 2k-N\square/\sqrt{N})=x$
egy

N
-től nem függő korlát alatt marad, akkor

$(Nk)(1/2N)$
közelítőleg

$(2\square(x)/\sqrt{N})$
-nel lesz egyenlő.

Ezek után könnyen megvizsgálhatjuk azt a kérdést, hogy a tartályok befogadóképességének és az egyszerre legurítható golyók számának milyen kapcsolatban kell állniuk? (Ezzel arra a kérdésre kapunk választ, hogy meddig játszhat egymással

A
és

B
a tartályok kiürítése nélkül.)

Ha egy-egy tartályban mondjuk

S
golyó fér el, akkor nyilván legfeljebb annyi golyót szabad leengednünk (

$R\mathrm{max}$
), amennyiből a középső, legvalószínűbb tartályba (tartályokba) legfeljebb

S
golyó jut. Ha tehát a Galton-deszka

$N=2M$
éksorból áll, akkor

$rM\approx R\mathrm{max}P2M(M)\leq S$

, azaz

$R\max \leq (S/P2M(M))$
-nek kell teljesülni. Ha pl. 100 golyó fér egy tartályba és a Galton-deszka 8 éksorból áll, akkor legfeljebb 366 golyót lehet egyszerre legurítani, 16 éksor esetén 509 golyót, 36 éksor esetén 757 golyót.

Megvizsgálhatjuk azt a kérdést is, hogy hány golyót kell legurítanunk ahhoz, hogy nagy valószínűséggel ne maradjanak üres tartályok (a széleken). Ha ui. az éksorok száma elég nagy, és összesen nem túl sok golyót gurítunk le, akkor néhány szélső tartályba nagy valószínűséggel egyáltalán nem gurul golyó. Ha pl.

$N=25$
, akkor annak valószínűsége, hogy egy golyó a szélső 4-4 tartály valamelyikébe jusson,

$p=2((250)+(251)+(252)+(253)/225)=(1+25+300+2300/224)=(1313/223)$.
Akkor annak a valószínűsége, hogy 100 legurított golyó közül egy se kerüljön az említett tartályokba,

$(1000)p0(1-p)100=(1-(1313/223))100>0{,}98$,
vagyis 98 % valószínűséggel csak a középső 18 rekeszbe jut egyáltalán golyó. Vagy pl.

$N=16$
esetén, ha legfeljebb kb. 44 golyót gurítunk le, 99% valószínűséggel üres lesz legalább a két szélső tartály.

Térjünk most vissza a már említett játékhoz! Mivel azt akarjuk, hogy

A
átlagos nyereménye (bármelyik tartályra is fogad) 0 legyen, így az

$yl(xk)$
számokra az

(2)

$$xkPN(k)-\sum l\neq kyl(xk)PN(l)=0 \quad (k=0,1,2,\dots,N)$$

egyenlőségnek kell fennállni. Ha semmi további kikötést nem teszünk, akkor az

$yl(xk)$
számokat még nagyon sokféle módon választhatjuk (pontosabban: egy kivételével a többit tetszőlegesen). Ha azonban még azt is megköveteljük, hogy – szemléletesen szólva – akárhova is esik a golyó (a

k
-adik tartályon kívül), átlagosan mindig ugyanannyit kapjon

B
, vagyis az

$yl(xk)PN(l)$
(amit jelöljünk

$cN(xk)$
-val) ne függjön

l
-től, csak

xk
-tól (és természetesen

N
-től), akkor ebből (2) alapján következik, hogy

(3)

$$x_k P_N(k) = N c_N(x_k), \text{ azaz } y_l(x_k) = (x_k P_N(k)/N P_N(l)) \quad (k=0,1,2,\ldots,N, l \neq k).$$

Az

x_k
egész számok választására is igen sokféle lehetőségünk van (választhatjuk akár valamennyit egyenlőnek is, bár ez nem célszerű), mindenesetre úgy érdemes megadnunk őket, hogy az

$y_l(x_k)$
számok is egészek legyenek. Egy lehetséges választást pl. az alább táblázatban találunk

$(N=8)$
:

	y_0	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
$x_0 = 2240$	.	35	10	5	4	5	10	35	280
$x_1 = 1120$	1120	.	40	20	16	20	40	140	1120
$x_2 = 1120$	3920	490	.	70	56	70	140	490	3920
$x_3 = 1120$	7840	980	280	.	112	140	280	980	7840
$x_4 = 224$	1960	245	70	35	.	35	70	245	1960
$x_5 = 1120$	7840	980	280	140	112	.	280	980	7840
$x_6 = 1120$	3920	490	140	70	56	70	.	490	3920
$x_7 = 1120$	1120	140	40	20	16	20	40	.	1120
$x_8 = 2240$	280	35	10	5	4	5	10	35	.

A számpélda és játék további elemzését, valamint lehetséges módosításait az olvasóra bízzuk.

4. A GALTON-DESZKA ÉS A POISSON-ELOSZLÁS

A Galton-deszka alkalmas az ún. Poisson-féle eloszlás szemléltetésére is.[8] Vizsgáljuk meg először, hogy

R
számú golyót legurítva az

N
éksorból álló Galton-deszkán, ebből átlagban hány darab esik a

k
-adik tartályba! Jelöljük az odaeső golyók számát mint valószínűségi változót

ξ_k
-val. Láttuk már, hogy annak a valószínűsége, hogy egy legurított golyó a

k
-adik tartályba jusson,

$P_N(k)=\binom{N}{k}(1/2^N)$
-nel egyenlő; akkor a binomiális eloszlás már megismert képletével felírhatjuk annak a valószínűségét, hogy

R
golyó közül

r
darab jusson a

k
-adik tartályba (

$P_N(k)=P(k)=p_k$
jelöléssel)

(4)

$$P(\xi_k=r)=\binom{R}{r}p_k^r(1-p_k)^{R-r}\ (k=0,1,2,\ldots,N),\ (r=0,1,2,\ldots,R).$$

Felhasználjuk azt a tételt, hogy a binomiális eloszlás nagy

R
és kis

p_k
esetén jól közelíthető a

$(\lambda_k^r/r!)e^{-\lambda_k}(r=0,1,2,\ldots)$
Poisson-eloszlással, ahol most

$\lambda_k=Rp_k$
. Ha tehát a legurított golyók száma elég nagy, és a középtől elég távol eső tartályokban vizsgáljuk az odaeső golyók számának véletlen ingadozását (amikor tehát teljesül, hogy

p_k
kicsiny), akkor ezzel a Poisson-eloszlásról nyerhetünk képet; ekkor ugyanis alkalmazható az említett közelítés:

8Egy

ξ
valószínűségi változót Poisson-eloszlásúnak mondunk, ha lehetséges értékei

$0,1,2,\ldots$
és a

k
értéket

$(\lambda^k/k!)e^{-\lambda}$

$(k=0,1,2,\ldots)$
valószínűséggel veszi fel, ahol

λ
pozitív állandó (az eloszlás várható értékét jelenti).

$$P(\xi_k=r)\approx\frac{(Rp_k)^r}{r!}e^{-Rp_k}.$$

Például egy 8 éksorból álló Galton-deszkán végezve a kísérletet, és a 0-adik tartályt vizsgálva $\left(p_0=\frac{1}{256}\right)$, ha az egyszerre legurított golyók száma 512 (tehát $R=512$), akkor a mondott Poisson-féle közelítés alapján annak a valószínűsége, hogy a vizsgált tartályba ne jusson egyetlen golyó sem, 0,135; hogy pontosan 1 golyó jusson oda, 0,271; hogy 2 golyó, 0,271; hogy 3 golyó, 0,180; hogy 4 vagy annál több golyó, annak 0,143 a valószínűsége. Ez tehát azt jelenti, hogy elég sokszor megismételve a kísérletet, mondjuk 1000 alkalommal legurítva mindig 512 golyót, az esetek közül kb. 135-ször üresnek találjuk a 0-adik tartályt stb.

Módosítsuk most kísérletünket úgy, hogy az egy-egy alkalommal legurított golyók száma maga is valószínűségi változó legyen (jelöljük ν-vel), mégpedig ez is Poisson eloszlású, azaz

$$P(\nu=R)=\frac{\mu^R}{R!}e^{-\mu}\quad(R=0,1,2,\dots)$$

(ahol μ az eloszlás várható értéke, azaz az egyszerre legurított golyók átlagos száma). Ekkor könnyen kimutatható, hogy az egyes tartályokba jutó golyók száma is Poisson-eloszlású (nemcsak közelítőleg, hanem pontosan), mégpedig annak a valószínűsége, hogy a k-adik tartályba r golyó jut

$$P(\xi_k=r)=\frac{(\mu p_k)^r}{r!}e^{-\mu p_k}\quad(r=0,1,2,\dots).\tag{5}$$

[9]

Az is[9] igaz, hogy a tartályokban összegyűlő golyók száma csakis ebben az esetben lesz Poisson-eloszlású, tehát ha nem Poisson-eloszlás szerint vesszük a legurítandó golyókat, akkor a tartályokban levő golyók száma sem lesz Poisson-eloszlású [5].

[9] Ez a teljes valószínűség tételéből azonnal következik, ugyanis

$$P(\xi_k=r)=\sum_{R=r}^{\infty}P(\xi_k=r\mid\nu=R)P(\nu=R)=$$

$$=\sum_{R=r}^{\infty}\frac{R!}{r!(R-r)!}p_k^r(1-p_k)^{R-r}\frac{\mu^r\mu^{R-r}}{R!}e^{-\mu}=$$

$$=\frac{(\mu p_k)^r}{r!}e^{-\mu}\sum_{R=r}^{\infty}\frac{[(1-p_k)\mu]^{R-r}}{(R-r)!},$$

legyen $R-r=s$ és felhasználva, hogy $\sum_{s=0}^{\infty}\frac{x^s}{s!}=e^x$, akkor

$$\sum_{R=r}^{\infty}\frac{[(1-p_k)\mu]^{R-r}}{(R-r)!}=e^{\mu}\cdot e^{-p_k\mu},$$

amiből következik (5).

Képzeljük most el, hogy egy olyan Galton-deszka áll rendelkezésünkre, amelynek tartályai eltolhatók, és módosítsuk az előbbi kísérletünket a következőképpen: toljunk végig egy tartályt

(T)
az

N
éksorú Galton-deszka tartálysorán a 0-iktól az

N
-edikig, és minden egyes helyzetben gurítsunk le bizonyos számú golyót, mégpedig a

k
-adik tartály helyére érve

R_k
számú golyót (

R_k
lehet 0 is), és vizsgáljuk ebben a végighaladó

T
tartályban összesen felgyűlő golyók számát! Bebizonyítható [5], hogy ha az éksorok száma elég nagy és az egyes lépésekben legurított golyók számára bizonyos megszorítást teszünk, akkor a

T
-ben felgyűlő golyók száma (jelöljük

ξ_N
-nel) közelítőleg Poisson-eloszlású lesz. Ha speciálisan minden lépésnél ugyanannyi, mondjuk

λ
számú golyót gurítunk le, akkor éppen ez a

λ
lesz az eloszlás várható értéke; vagyis ebben az esetben annak a valószínűsége, hogy

T
-ben éppen

r
számú golyó legyen, határértékben (

N
növekedésével)

$(\lambda^r/r!)e^{-\lambda}$
-val lesz egyenlő

$(r=0,1,2,\dots)$.

Ha most úgy módosítjuk a kísérletet, hogy a

T
tartály különböző helyzeteiben leengedett golyók száma is valószínűségi változó legyen, mégpedig ismét Poisson-eloszlású (általában lépésenként különböző

λk
-kal), vagyis annak a valószínűsége, hogy a

k
-adik eltolásnál

x
golyót gurítunk le,

$(\lambda kx/x!)e–\lambda k$
-val egyenlő

$(x=1,2,\ldots)$
, akkor

ξN
már véges

N
esetén is Poisson-eloszlású lesz (nemcsak határértékben), sőt

ξN
csakis ebben az esetben lesz Poisson-eloszlású.

Az eddigiekben mindig *egy* adott tartályban felgyűlő golyók számát, illetve annak eloszlását vizsgáltuk. Válasszuk most ki az

N
éksorú Galton-deszka *két* tetszőleges tartályát, mondjuk az

i
-ediket és

j
-ediket és figyeljük meg, hogy

R
darab legurított golyó közül hány golyó kerül az

i
-edik, ill.

j
-edik tartályba.

Jelöljük ezek számát mint valószínűségi változókat

ξi
-, ill.

ξj
-vel. A

ξi
és

ξj

valószínűségi változók nyilván nem függetlenek, hiszen az összesen legurított golyók számának rögzítése esetén, ha az egyik tartályba nagyon sok golyó jutott, akkor a másikba már csak kevesebb juthatott. A

ξ_i
és

ξ_j
együttes eloszlása (vagyis azoknak a valószínűségeknek az összessége, hogy

$\xi_i r$
-rel és

$\xi_j s$
-sel egyenlő) ún. trinomiális eloszlás, tehát

(6) $$P(\xi_i=r,\xi_j=s)=(R!/r!\,s!\,(R-r-s)!)p_i^r p_j^s(1-p_i-p_j)^{R-r-s}.$$

Ha azonban a legurítandó golyók számát nem rögzítjük, hanem az is valószínűségi változó

(ν)
, akkor igaz az az érdekes tétel [5], hogy

ξ_i
és

ξ_j
akkor és csak akkor függetlenek, ha a legurított golyók száma Poisson-eloszlású, mégpedig, ha

$P(\nu=R)=(\lambda^R/R!)e^{-\lambda}$ ($R=0,1,2,\dots$), akkor

$P(\xi_i=r,\xi_j=s)=((\lambda p_i)^r/r!)e^{-\lambda p_i}\cdot((\lambda p_j)^s/s!)e^{-\lambda p_j}$
lesz (utóbbi (5)-höz hasonlóan következik a teljes valószínűség tételéből), vagyis

ξ_i
és

ξ_j
külön-külön is Poisson-eloszlásúak lesznek:

(7) $$P(\xi_i=r)=((\lambda p_i)^r/r!)e^{-\lambda p_i} \text{ és } P(\xi_j=s)=((\lambda p_j)^s/s!)e^{-\lambda p_j}.$$

Ha nemcsak két tartályt vizsgálunk egyidejűleg, hanem mindegyiket, és rendre

ξ_0
,

ξ_1
, …,

ξ_N
-nel jelöljük

R
legurított golyó közül a 0-adik, első, második, …,

N
-edik tartályba jutó golyók számát, akkor

ξ_0
,

ξ_1
, …,

ξ_N
együttes eloszlása általában polinomiális eloszlás, azaz

$P(\xi_0=r_0,\xi_1=r_1,\ldots,\xi_N=r_N)=(R!/r_0!\ r_1!\ \ldots r_N!)p_0^{r_0}p_1^{r_1}\ldots p_N^{r_N}$

$(r_0+r_1+\ldots+r_N=R)$
és a

ξ_0
,

ξ_1
, …,

ξ_N
akkor és csak akkor lesznek függetlenek, ha a leengedett golyók száma Poisson-eloszlású.

Módosítsuk most a tartályok eltolásában álló, már említett kísérletet a következőképpen: jelöljük

T_0
-val a 0-adik tartályt, és

T_k
-val a

k
-adik tartályt, és toljuk „vissza” a tartálysort úgy, hogy

T_k
kerüljön a

T_0
-helyére! Gurítsunk le ebben a helyzetben bizonyos számú golyót, majd jobbra tolva a tartálysort egy egységgel (egy tartállyal), ismét gurítsunk le valamennyi golyót, és így tovább, összesen

$N+k$
eltolást végezve, a

T_0
is végighalad mind az

$N+1$
csatorna alatt. Most is igaz lesz az a tétel, hogy a

T_0

-ban és

Tk
-ban felgyűlő golyók száma akkor és csak akkor lesz egymástól független, ha az egyes helyzetekben leengedett golyók száma Poisson-eloszlású valószínűségi változó.

Egy további érdekessége az eltolható tartálysorral rendelkező Galton-deszkának, hogy segítségével a már említett Gauss-féle görbék szuperpozícióját tudjuk előállítani, mégpedig oly módon, hogy a tartálysor két különböző helyzetében gurítunk le golyókat (minden helyzetben elég nagy számú golyót). Megfigyelhetjük, hogy ha elég sok tartállyal toljuk arrébb a tartálysort, akkor a keletkező görbének már két „csúcsa” (két maximuma) lesz.

5. VÉLETLEN SZÁMOK ELŐÁLLÍTÁSA A GALTON-DESZKA SEGÍTSÉGÉVEL

Ha 0-tól 2-nek valamely pozitív egész kitevőjű hatványáig terjedő számokból akarunk véletlenszerűen választani egy számsorozatot (véletlen számok), akkor ilyen számsorozatot a Galton-deszka segítségével a következőképpen állíthatunk elő: tegyük fel, hogy a Galton-deszka szabályos és

N
éksorból áll. Gurítsunk le egy golyót és rögzítsük le valahogyan az útját! Egy ilyen „véletlen” út jobbra-balra térések véletlen sorozata. Rendeljünk 0-t minden balra téréshez és 1-et minden jobbra téréshez, és az így kapott jegyeket írjuk egymás után olyan sorrendben, ahogyan a lépések egymást követték. Ekkor egy 0-ákból és 1-esekből álló véletlen sorozatot kapunk, amely megfelel egy 0 és

$2N–1$
közé eső egész szám (a határokat is beleértve) 2-es számrendszerbeli előállításának, vagyis az így nyert 0–1-sorozat egy 0 és

$2N–1$
közé eső véletlen számot állít elő. A golyó legurítását tetszőleges sokszor megismételve, tetszőlegesen sok számot tartalmazó véletlen számsorozatot tudunk előállítani.

6. A GALTON-DESZKA ÉS A MARKOV-LÁNCOK

A Galton-deszkával a Markov-láncok fogalmát is szemléltethetjük. Ha ui. az eredeti, szöges Galton-deszkát használjuk (vagy az ékekből álló Galton-deszkán megkopnak az ékek), akkor el kell ejtenünk azt a feltevésünket, hogy az egyes sorokon való ütközéseknél létrejövő eltérítések függetlenek egymástól. Ebben az esetben ezek ún. *Markov-láncot* alkotnak.

Az olyan kísérletsorozatokat, ahol az

$(n+1)$
-edik kísérlet eredménye csak az

n
-edik kísérlet eredményétől függ, azonban az első, második, …,

$(n–1)$
-edik kísérlet eredményétől közvetlenül nem függ, csak közvetve azáltal, hogy ezek befolyásolták az

n-edik kísérlet eredményét, Markov-láncnak nevezzük. Pontosabban, ha egy kísérlet lehetséges eredményei az A_0, A_1, A_2, …események, a kísérletsorozat akkor alkot Markov-féle láncot, ha a mellett a feltevés mellett, hogy bizonyos előző kísérletek megadott eredményre vezettek, annak a feltételes valószínűsége, hogy az $n+1$-edik kísérlet eredménye mondjuk az A_k esemény, csak a legutolsó adott eredményű kísérlet eredményétől függ. Valószínűségi változók sorozatára is értelmezhetjük a fogalmat. Jellemezze az n-edik kísérlet eredményét a ξ_n valószínűségi változó, és legyen a $\xi_n=k$, ha az n-edik kísérlet eredménye az A_k esemény $(n=1,2,\dots;k=0,1,2,\dots)$. Feltételünket képlettel így adhatjuk meg:

$$P(\xi_{n+1}=k \mid \xi_1=j_1,\xi_2=j_2,\dots,\xi_n=j_n)=$$

$$=P(\xi_{n+1}=k \mid \xi_n=j_n).$$

A $P(\xi_{n+1}=k \mid \xi_n=j_n)$ feltételes valószínűséget egylépéses átmenet-valószínűségnek nevezzük, ugyanis az A_j-ből az A_k-ba való „átmenet” egy lépésben való létrejöttének a valószínűségét adja meg. A rövidség kedvéért vezessük be erre a P_{jk}^{n} jelölést. Az olyan Markov-láncokat, amelyeknél az átmenet-valószínűség nem függ az

n-től (ekkor tehát jelölhetjük P_{jk}-val), *homogén* Markov-láncnak nevezzük. Ha a kísérleteket időben egymás után végbemenőknek képzeljük, az $n=1,2,\ldots$ index jelenti az időpontokat, akkor tehát ξ_n az n időpontban végrehajtott kísérlet eredményét jelenti. Az A_0, A_1, A_2, …eseményeket pedig *állapotoknak* szokás nevezni.

A Galton-deszka esetében az időpontoknak a sorok felelnek meg, az állapotoknak pedig a balra, ill. jobbra térés (A_0,A_1). Legyen most $\xi_n=0$ vagy 1 aszerint, hogy az n-edik éksoron ütközve balra vagy jobbra tér-e el egy legurított golyó, akkor azt mondhatjuk, hogy a ξ_n valószínűségi változók sorozata Markov-láncot (mégpedig homogén Markov-láncot) alkot. Itt az átmenet-valószínűségeket tehát jelölhetjük P_{00}, P_{01}, P_{10}, P_{11}-el. Tehát pl. P_{10} a balra térés valószínűsége, feltéve hogy az előző sorban jobbra térés történt. Feltehető, hogy $P_{01}=P_{10}$ és

$P00=P11$

.

A

$\xi 1+\xi 2+\ldots+\xi N$
megadja, hogy az

N
éksorú Galton-deszkán legurítva egy golyót, az hányadik tartályba jut. Érdekes megfigyelni, hogy a golyók a tartályokban ebben az esetben is a Gauss-féle görbének megfelelően oszlanak el, ami a Markov-láncok elméletéből jól ismert tétel kísérleti alátámasztása. Ekkor azonban, attól függően, hogy

$P01<(1/2)$
vagy

$P01>(1/2)$
, a görbe lapultabb vagy „csúcsosabb” lesz, mint a

$P01=(1/2)$
esetben.

7. EGYENLETES ELOSZLÁS ELŐÁLLÍTÁSA GALTON-DESZKÁVAL

Legyen most a Galton-deszka ismét szabályos, és módosítsuk a következőképpen: az

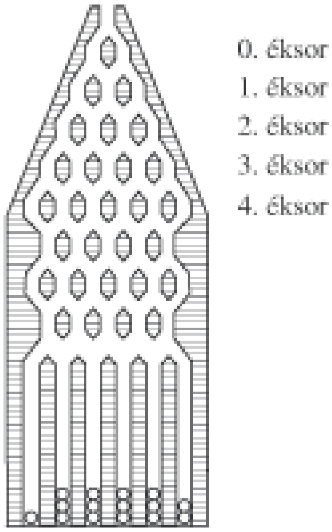
N
-edik éksortól kezdve (az éksorok számozását kezdjük most 0-val) vágjuk le a Galton-deszkát a két „szélén” úgy, hogy ettől kezdve az összes további sorban az ékek száma felváltva

$N+1$
, ill.

N
legyen (mivel az éksorok számozását 0-val kezdtük, az

N
-edik éksorban most nyilván

$N+1$
ék van) és tegyük fel, hogy még igen sok további éksort tartalmaz Galton-deszkánk (l. 4. ábra).



4. ábra

Határozzuk meg ennél a módosított Galton-deszkánál, hogyan oszlanak meg a golyók a tartályokban. Jelentse

ξn
a golyó helyzetét, vagyis a Galton-deszka tengelyétől való vízszintes távolságát az

n
-edik szögsor elérésének pillanatában. (Egységnek itt most két, egy sorban levő szomszédos ék távolságának felét vesszük.)

$(\xi 0=0)$
. A

ξn
változók Markov-láncot alkotnak. Itt most a lehetséges állapotok egy leengedett golyó lehetséges helyzetei, vagyis az

n
-edik szögsor elérésének pillanatában a lehetséges

Aj
állapotok azt jelentik, hogy a golyó a Galton-deszka tengelyétől az

n
-edik éksorra érve

j
távolságra van

$(-N\leq j\leq N)$
. Könnyű látni, hogy ha

n
páros, akkor

ξn
is páros; ha

n

páratlan, akkor

ξn
is páratlan. Az átmenet-valószínűségek

n
-től nem függnek, jelen esetben is homogén Markov-lánccal van dolgunk.

$Pj,j+1=P(\xi n+1=j+1 \mid \xi n=j)$ =(1/2), ha $-N<j<N$, $Pj,j-1$ =(1/2), ha $-N<j<N$; és $P-N,-N+1$ =1 és $PN,N-1=1$; minden más

j
,

k
-ra

$Pjk=0$
.

Meg lehet mutatni a Markov-láncok elméletének néhány egyszerű tételét felhasználva, hogy akármilyen

N
esetén, ha az összes sor száma páros és elég nagy, akkor minden tartályban kb. ugyanannyi golyó lesz. Ha a sorok száma viszont páratlan és elég nagy szám, akkor a két szélső tartályt kivéve a többi tartályban kb. ugyanannyi golyó lesz, a két szélsőben pedig feleannyi, mint a középső tartályokban. A kísérleti eredmények igen jó egyezést mutatnak az elméletivel.

Ezzel a módosítással tehát a Galton-deszkával sikerült ún. egyenletes eloszlást létrehozni.

További alkalmazásokra és a Galton-deszka különféle módosításaira nézve utalunk a [4], [6], [8], [9] munkákra.

IRODALOMJEGYZÉK

[1] [1] Gnyegyenko, B. V. – Hincsin, A. Ja.: *Bevezetés a valószínűségszámításba.* Művelt Nép Kiadó, 1954.

[2] [2] Prékopa András: *Valószínűségelmélet.* Műszaki Könyvkiadó, 1962.

[3] [3] Rényi Alfréd: *Valószínűségszámítás.* (Egyetemi tankönyv) Tankönyvkiadó, 1966.

[4] [4] Bernstein, F.: Verallgemeinertes Galtonbrett zur Durchführung von Funktionaltransformationen. *Zeitschr. Phys.* **77** (1932), 104–113. old.

[5] [5] Medgyessy Pál: A Galton-féle deszkával kapcsolatos néhány problémáról. *MTA Alkalmazott Matematikai Intézetének Közleményei* **1** (1952), 165–174. old.

[6] [6] Medgyessy Pál: Kiegészítés „A Galton-féle deszkával kapcsolatos néhány problémáról” című dolgozathoz. *MTA Alkalmazott Matematikai Intézetének Közleményei* **2** (1953), 233–237. old.

[7] [7] Rényi Alfréd: *Szakkörökben elvégezhető valószínűségszámítási kísérletekről. Előadások az iskolai matematika köréből.*

[8] [8] Schulz, G.: Zur Theorie des Galtonschen Brettes. *Zeitschr. Phys.* **92** (1934), 747. old.

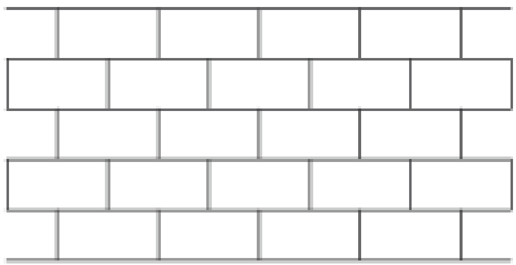
[9] [9] Seitz, W. – Hamacher, K. – Odenhausen: Untersuchungen über das Galtonbrett. *Naturwiss.* **22** (1934), 494. old.

PARKETTÁK A GEOMETRIA SZEMSZÖGÉBŐL

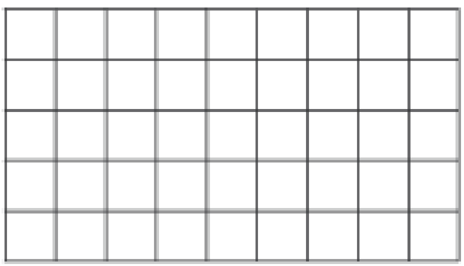
REIMAN ISTVÁN

A parkettázás története szinte egyidős az építkezések történetével. Az épületek padozatát régen kődarabokkal rakták ki; ez eleinte találomra egymás mellé rakott laposabb kődarabokból állott; később már faragtak ezeken a köveken, hogy az összeillesztésnél a padozat minél nagyobb része legyen fedett és minél kevesebb legyen a hézag. A fejlődés további folyamán a hézagokat is igyekeztek kiküszöbölni a padozatból a felhasznált kődarabok faragásával. (Egyébként a falak építésére felhasznált építőelemeknél, köveknél is megtalálható ez a fejlődési folyamat.)

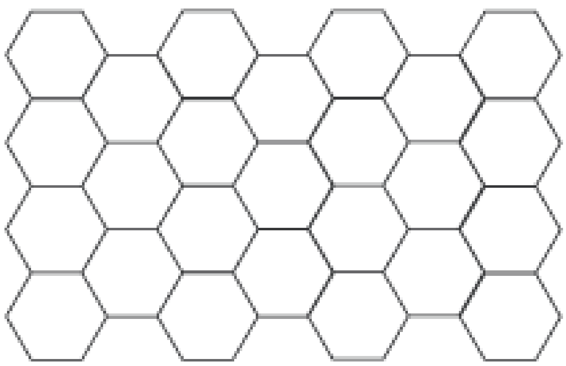
Gyakorlati tapasztalatok alapján rájöttek arra, hogy egyszerűbb a padozat beborítása, ha egyforma, azonos alakú és nagyságú fedőköveket használnak; különösen akkor, ha ezeket pl. agyagból égetik. Néhány ilyen parkettázást mutatnak az ókori lakóépületekből az 1–5. ábrák.



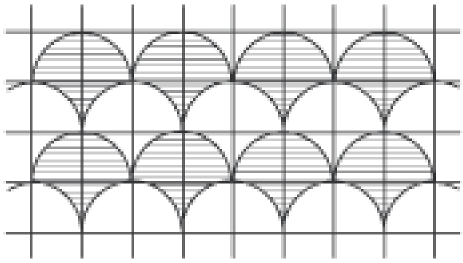
1. ábra



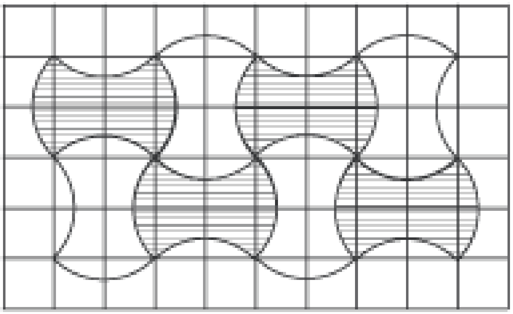
2. ábra



3. ábra



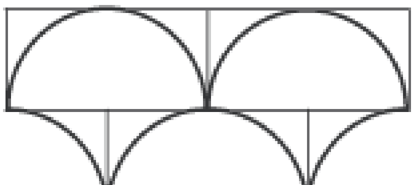
4. ábra



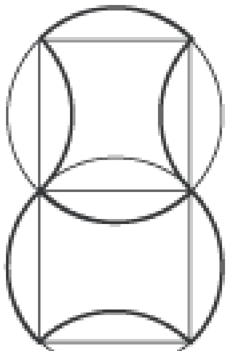
5. ábra

Már ebből az öt, évezredekkel ezelőtt ismert parkettázási módból is a fellépő formák nagy gazdagságára következtethetünk. És itt lép be gondolatsorunkba a matematika: felvetődik ti. a kérdés, hogy egyáltalán milyen parkettázások készíthetők. Bevezetőül tisztázzuk az alábbiakban felhasználandó fogalmakat. A matematikai tárgyalás egyszerűsítése céljából parkettázáson a síknak síkidomokkal való egyrétű és hézagtalan lefedését fogjuk érteni. *Egyrétűnek* nevezzük a lefedést, ha a sík minden pontját legfeljebb egy lefedő idom belső pontja fedi le. (Egy pontot több fedőidom határpontja is fedhet.) *Hézagtalan* a lefedés, ha a sík minden pontját lefedi legalább egy fedőidom belső- vagy határpontja. Parkettázásnál a teljes sík lefedésére gondolunk.

Szorítkozzunk először a már említett, a gyakorlat által felvetett kérdésre: milyen parkettázás készíthető egybevágó idomokból, azaz: *milyen egybevágó idomokból készíthetők parketták*. Nyilvánvaló például, hogy körből nem készíthető parkettázás, de bizonyos körívekkel határolt idomokból már igen, mint azt a 4–5. ábrák mutatják. A probléma némi egyszerűsítéséhez jutunk, ha észrevesszük, hogy pl. a 4. és 5. ábrák az 1 és 2. ábrákból származtathatók úgy, hogy azokba bizonyos köríveket rajzolunk be (6. és 7. ábra).



6. ábra



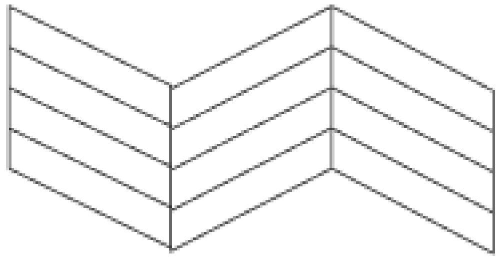
7. ábra

Kézenfekvőnek látszik tehát az a gondolat, hogy minden parkettázást sokszögekkel való parkettázásra vezessünk vissza.

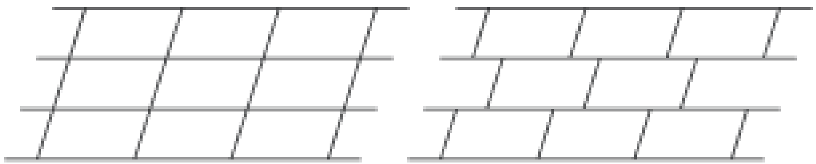
Vizsgáljuk először az *egybevágó konvex sokszögekkel* létesíthető parkettákat (konvex egy sokszög, ha minden szöge kisebb

180□

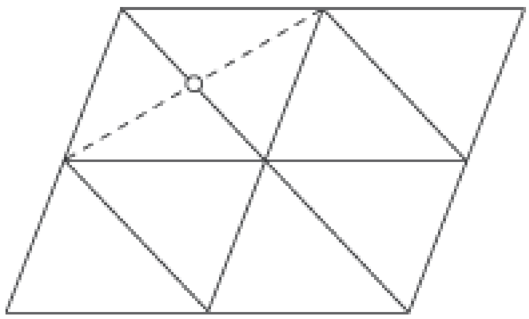
-nál). Ilyet mutatott az eddigiek folyamán az 1., 2., 3. ábra. Nagyon könnyű belátni, hogy tetszőleges paralelogrammából többféleképpen is készíthető parketta! (8. ábra).



8. ábra



Ebből közvetlenül következik, hogy parkettázás készíthető minden háromszögből is. Tükrözzünk ui. egy háromszöget egyik oldalának felezőpontjára, így paralelogrammát kapunk. E paralelogrammával – tehát akkor a háromszöggel is – lefedhető a sík (9. ábra).



9. ábra

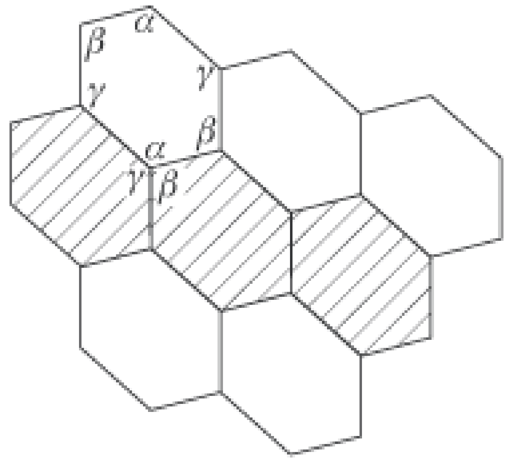
A 3. ábrán egy szabályos hatszögből készített parkettát láthatunk. Megfigyelhetjük azonban, hogy a szabályos hatszögnek a parkettázás szempontjából leglényegesebb tulajdonsága középpontosan szimmetrikus volta; ti. minden középpontosan szimmetrikus hatszögből készíthető lefedés. Ennek igazolása leolvasható a 10. ábráról: a hatszögekből szalagokat készíthetünk; ezek hézagtalanul egymás mellé illeszthetők, mivel a hatszög szemköztes szögei egyenlők és így három, nem szomszédos szög összege (az ábrán

$\alpha+\beta+\gamma$

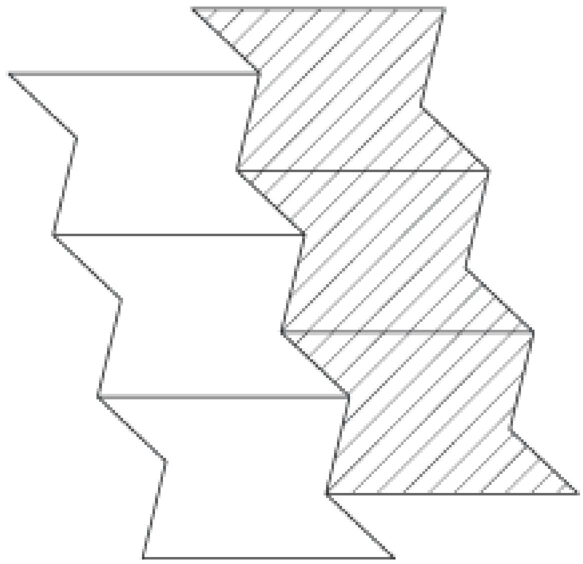
) éppen a hatszög szögösszegének fele, azaz

360□

. Megjegyezhetjük, hogy itt a hatszög konvex voltát sem kell kikötni (11. ábra).

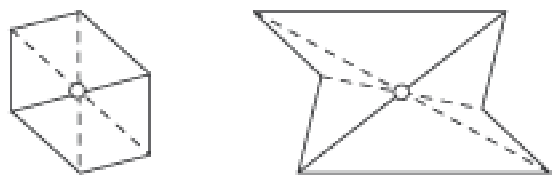


10. ábra



11. ábra

A paralelogramma-parkett és a háromszögparkett közötti igen egyszerű kapcsolat a paralelogrammák középpontos szimmetriájából következik. Ugyanilyen kapcsolatot tudunk felfedezni a négyszögek és a középpontosan szimmetrikus hatszögek között is. Egy tetszőleges négyszög bármelyik oldalfelező pontjára vonatkozó tükörképével együtt középpontosan szimmetrikus hatszöget alkot (itt sem szükséges a konvex kikötés, 12. ábra).



12. ábra

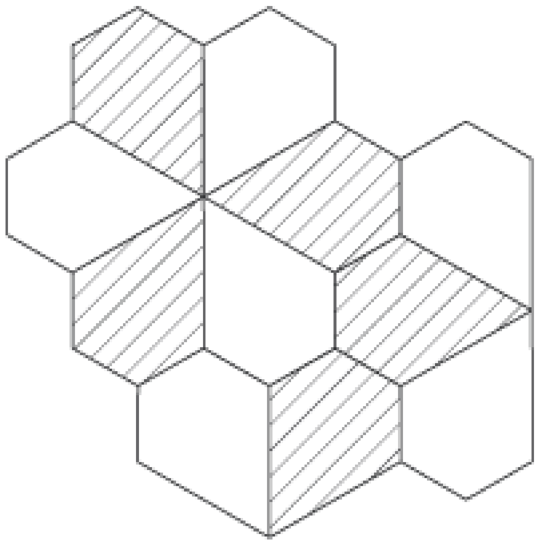
Ezekből a hatszögekből viszont – és így a négyszögekből is – készíthető parkett.

Eddigi eredményeinket összefoglalva kimondhatjuk: *minden háromszögből és minden négyszögből készíthető parkettázás*.

Logikusan vetődik fel a kérdés: milyen ötszögekkel fedhető le a sík? Itt már korántsem várható az, hogy a három-, ill. négyszögekre nyert tételünk érvényben marad; pl. szabályos ötszögekkel nem fedhető le a sík; ezt könnyen beláthatjuk, ha meggondoljuk, hogy annak egyik szöge

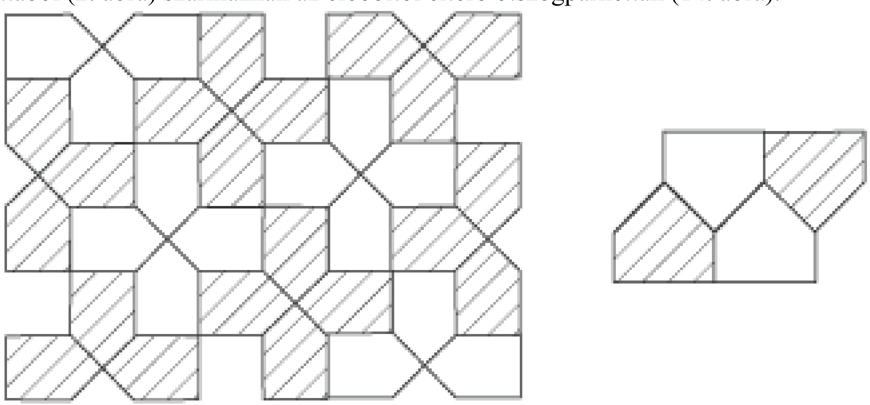
108□

-os. Egyszerűen megadhatunk azonban másféle ötszögekből készített parkettát. Egy ilyen a 3. ábra szabályos hatszögeiből készült (13. ábra).



13. ábra

A négyzetparkettából (2. ábra) származnak az előbbitől eltérő ötszögparketták (14. ábra).



14. ábra

Megfigyelhetjük, hogy az itt szereplő ötszögek mindegyikének van legalább egy,

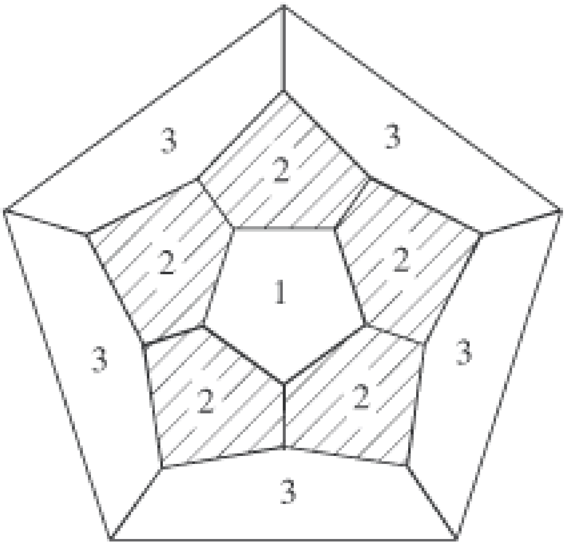
90□

-nál nem nagyobb szöge. Megmutatjuk, hogy ez szükségszerű: *ha egy ötszög minden szöge tompaszög, akkor nem készíthető belőle parketta*. (Tehát speciálisan a szabályos ötszögből sem.) Nevezzük csomópontnak az elképzelt parkettán azokat a pontokat, amelyekben élek végpontjai találhatók! Mivel két tompaszög összege nagyobb

180□

-nál, ezért egy csomópontban legfeljebb három sokszög, tehát legfeljebb három él találkozhat. Mivel legalább három élnek kell is találkoznia, ezért a szóban forgó ötszögparketta minden csomópontja három élű.

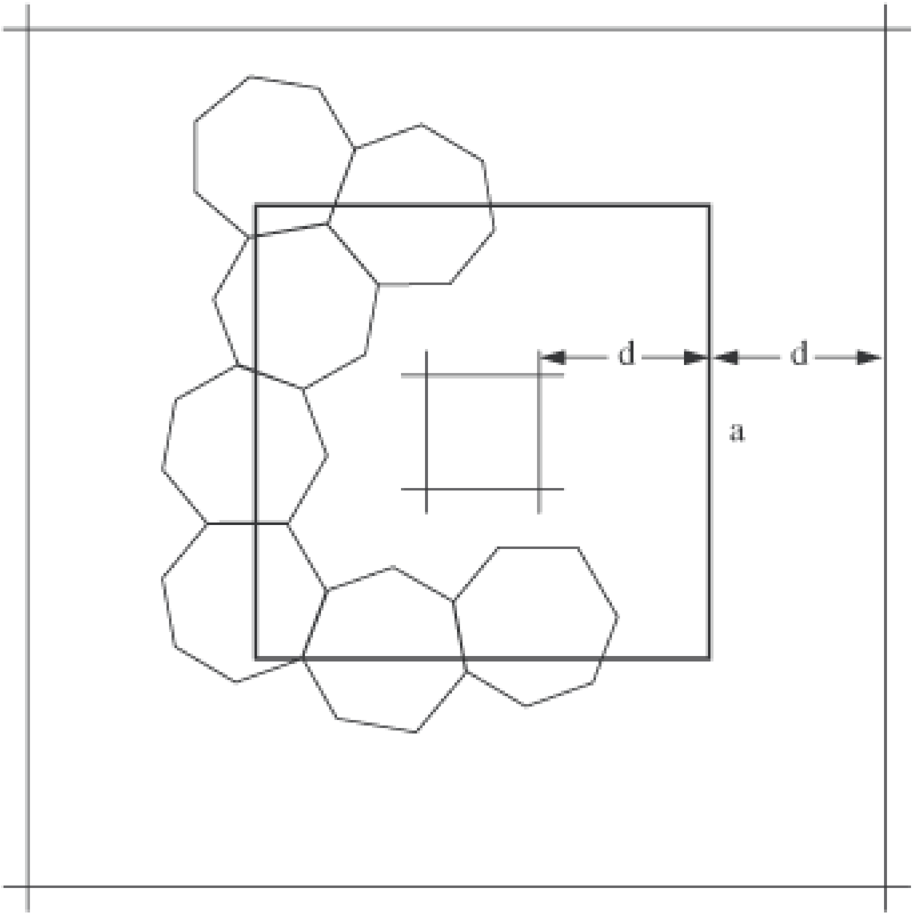
Állításunknál többet is igazolunk: megmutatjuk, hogy még nem egybevágó ötszögekből sem lehet készíteni olyan parkettát, amelynek minden csomópontja három élű. A 15. ábrán egy tetszőleges, 1-gyel jelölt ötszögből kiindulva kíséreljük meg a parketta felépítését. A kiindulási ötszöget a sötétebb árnyalatú, 2-es típusú ötszögek fogják közre, ezekhez pedig a 3-asok köre csatlakozik. Itt azonban már minden csomópontban pontosan három él van, a parkettázás tehát nem folytatható tovább, nem létezik ilyen parkett.



15. ábra

Nagyon nehéz kérdésnek bizonyult, hogy melyek azok az ötszögek, ill. hatszögek, amelyekből parkett készíthető, és a kérdés teljes egészében még ma sem tisztázott. Mindebből hajlamosak lennénk levonni azt a következtetést, hogy a hatnál nagyobb oldalszámú konvex sokszögekre a feladat még nehezebb, még áttekinthetetlenebb. Meglepő azonban, hogy ez nincs így; ez a kérdés röviden elintézhető: *hatnál nagyobb oldalszámú konvex sokszögből nem készíthető parkett.*

Ezt az állítást *indirekt* módon bizonyítjuk be. Feltesszük, hogy az állítással ellentétben létezik olyan parkett, amelynek építőelemei 6-nál nagyobb oldalszámú egybevágó konvex sokszögek, s megmutatjuk, hogy ebből a feltevésből kiindulva képtelen következtetésre jutunk, ami a kiindulás lehetetlenségét bizonyítja. A feltételezett parkettel kapcsolatosan vezessünk most be néhány jelölést! Legyen a parkettázás alapsokszögének oldalszáma n, területe t. Nyilvánvalóan meg tudunk adni olyan kört, amellyel lefedhető ez a sokszög; legyen ennek átmérője d. Ez azt jelenti, hogy az alapsokszög tetszőleges két pontjának távolsága nem lehet nagyobb d-nél. (A d értéke nyilván nincs egyértelműen meghatározva, egy határon túl tetszőlegesen választható.) A rövidség kedvéért tételezzük fel, hogy parkettünk csomópontjai nem belső pontjai egyetlen élnek sem, tehát pl. nem olyanok, mint az 1. ábra csomói. (E feltevés nélkül is hasonlóan, de valamivel hosszabban menne a bizonyítás.) Borítsunk most rá parkettánkra egy a oldalú négyzetet (16. ábra)!



16. ábra

Mivel a feltevés szerint készíthető az egész síkra kiterjedő parkett,

a
értéke *tetszőleges nagy* lehet. Jelöljük

c
-vel a négyzetbe vagy határára eső csomópontok számát! Ez a négyzet teljesen lefed

x
számú sokszöget, a teljesen és részben lefedett sokszögek együttes száma pedig legyen

y
. Húzzunk most a négyzet oldalaival, tőlük

d
távolságra, párhuzamosokat; ezek egy

a+2*d*
, ill.

a–2*d*
oldalú négyzetet zárnak közre.

Az előbb megnevezett

y

számú sokszöget teljesen lefedi az

$a+2d$
oldalú négyzet, mivel ezeknek a sokszögeknek van közös pontjuk az

a
oldalú négyzettel; ezért az

y
számú sokszög területének összege kisebb az

$a+2d$
oldalú négyzet területénél:

$yt<(a+2d)2$,
azaz

$$y<((a+2d)2/t). \tag{1}$$

Hasonlóan: az

$a–2d$
oldalú négyzetet teljesen lefedi az

a
oldalú négyzetben teljesen benne lévő

x
számú sokszög, ezért ezek területének összege nagyobb az

$a–2d$
oldalú négyzet területénél:

$xt>(a–2d)2$,
azaz

$$x>((a–2d)2/t). \tag{2}$$

Becsüljük most meg az

a
oldalú négyzetben lévő sokszögcsúcsok számát! E négyzet belsejében vagy határán fekvő csomópontok képzésében azok a sokszögek vesznek részt, amelyek vagy benne vannak a négyzetben, vagy oda benyúlnak; jelölésünk szerint ezek száma

y
. Mivel minden sokszögnek

n
csúcsa van, ezért a szóban forgó csomópontokban lévő csúcsok száma legfeljebb

yn

. Mivel viszont egy csomóban a sokszögek konvex volta miatt legalább három csúcsnak kell egyesülnie, a csomókban elhelyezkedő csúcsok száma legalább

$3c$
, így

$$3c \le yn. \qquad (3)$$

Ha most (1) mindkét oldalát megszorozzuk

n
-nel, akkor azt kapjuk, hogy

$yn<((a+2d)2n/t)$.
Ebből (3)-ra tekintettel

$3c<((a+2d)2n/t)$,
azaz

$$c<((a+2d)2n/3t). \qquad (4)$$

Végezzünk most becslést az

a
oldalú négyzetben fekvő sokszögek szögeinek összegére! Ez az összeg nyilván

$(n-2)180^\circ \cdot x$,
mivel

x
számú sokszögről van szó és mindegyik sokszög szögeinek összege

$(n-2)180^\circ$
. Másrészt: az egy csomóhoz tartozó valamennyi szög összege

360°
, ezért az összes csomóban lévő, számításba jövő szögek összege nem lehet nagyobb

$c \cdot 360^\circ$
-nál (lehet, hogy kevesebb, mert ha a csomó a négyzet határán van, nem minden szöge tartozik az összeghez), ezért

$c \cdot 360^\circ \ge (n-2)180^\circ \cdot x$.

Ezt az egyenlőtlenséget

360°
-kal osztva kapjuk:

$c \ge ((n-2)x/2)$,
vagy még (2)-t is figyelembe véve

$$c>(n-2/2)\square((a-2d)2/t). \tag{5}$$

Hasonlítsuk most össze (4)-et és (5)-öt:

$((n-2)(a-2d)2/2t)<((a+2d)2n/3t)$.

$6t$-vel megszorozva az egyenlőtlenség mindkét oldalát, majd azonos algebrai átalakításokat végezve és rendezve kapjuk, hogy

$(n-6)(a2+4d2)<ad(20n-24)$.

Mivel $n>6$, $(n-6)$-tal eloszthatjuk az egyenlőtlenség mindkét oldalát

$$a2+4d2<(20n-24/n-6)ad. \tag{6}$$

A bal oldalt csökkentjük, ha $4d2$-et elhagyjuk; majd mindkét oldalt a-val osztjuk:

$a<(20n-24/n-6)d$.

Vizsgáljuk most meg d együtthatóját, mekkora lehet maximális értéke?

$(20n-24/n-6)=20+(96/n-6)$.

Mivel $n-6$ minimális értéke 1, $(96/n-6)$ maximális értéke 96, ezért

$(20n-24/n-6)\leq 20+96=116$.
Tehát

$a<116d$.

A szóban forgó parkett létezéséből tehát arra a következtetésre jutottunk, hogy az

a
négyzetoldalnak kisebbnek kell lennie egy rögzített értéknél,

$116d$
-nél. Ez azonban ellentmond annak a feltevésnek, hogy a parkettre tetszőlegesen nagy négyzet borítható, azaz a *teljes síkot lefedő parkett nem létezhet*. Ezzel bebizonyítottuk állításunkat.

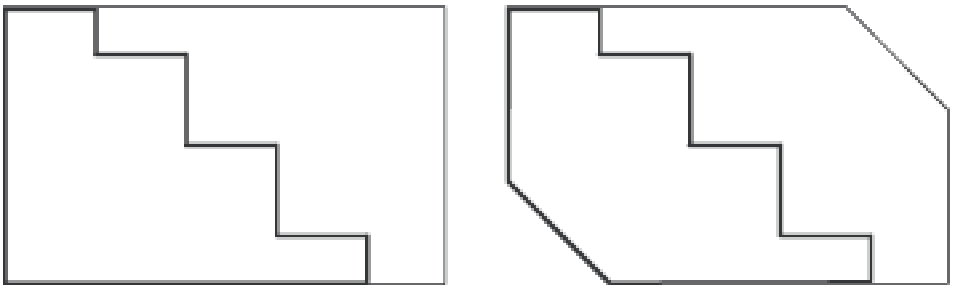
Érdemes még egyszer áttekinteni a fenti bizonyítást abból a szempontból, hogy hol használtuk ki az egyes feltételeket. Az

$n \geq 7$
feltételre a számolás (6)-tal jelölt és utolsó lépésénél volt szükség. A sokszög konvexitásából következett, hogy egy csomóban legalább három csúcs találkozik. Nem használtuk ki azt, hogy a sokszögek egybevágók, hanem csupán azt, hogy területeik egyenlők és mindegyikük lefedhető egy

d
átmérőjű körrel.

Egybevágó konkáv sokszögekből készíthető parkett

$n \geq 7$
esetén is, sőt tetszőleges oldalszámú sokszöget meg lehet adni, amelyből síkparkettázás készíthető. Erre mutat példát a 17. ábra páros, ill. páratlan oldalszámra; két-két sokszög téglalapba, ill. középpontosan szimmetrikus hatszögbe foglalható.



17. ábra

A parkett gyakorlati felhasználásával művészi, esztétikai szempontok is jelentkeztek. Az egybevágó elemekből készített parkettnél szebb mintákat szolgáltatnak a különböző elemekből összetett parketták; mutatósságuk még fokozható, ha az azonos elemeket azonos színűre is festik. A különböző elemekből készítendő parkettázás lehetőségének vizsgálata ugyancsak bonyolult kérdés; válasszunk ki itt is olyan speciális esetet, amely a gyakorlati alkalmazások szempontjából is elsődleges. Kössük ki, hogy

1. a parkettlapok csak szabályos sokszögek lehetnek;
2. a parkettázás csomópontjai nem belső pontjai egyetlen élnek sem;
3. bármely két különböző csomópontban azonos számú hatszög, azonos számú négyszög, azonos számú ötszög stb. szerepel, de a különböző fajta sokszögek számának egy csomón belül nem kell megegyeznie.

A fenti feltételeknek eleget tevő parkettákat *homogéneknek* nevezzük.

Az összes homogén parketta meghatározása matematikailag nem jelent nehéz feladatot, legfeljebb eléggé hosszadalmas. Néhány konkrét eset meghatározásával megmutatjuk azt az utat, amelyen megtalálható az összes megoldás.

Mivel a szabályos sokszögek konvexek, a homogén parkett egy csomópontjában legalább három sokszög találkozik. A szabályos sokszög egy-egy szöge

$((n-2)\square 180\square/n)=180\square-(360\square/n)$,
ahol

n
az oldalszámot jelenti. Ez a képlet egyébként azt mutatja, hogy az

n
növekedésével a szabályos sokszög szögei is nőnek; a legkisebb szögei tehát a szabályos háromszögnek vannak, ezek

$60\square$
-osak. Ebből már következik, hogy egy csomóban legfeljebb hat csúcs találkozhat, hiszen egy csomóban a szögösszeg

$360\square$
. Feladatunkat tehát négy részre vághatjuk szét aszerint, hogy egy csomó 3, 4, 5 vagy 6 csúcsot tartalmaz. Nézzük meg például, hogyan lehetne meghatározni a második esethez tartozó parkettákat. (Itt egy csomóban négy csúcs van.) Legyen a csomót alkotó szabályos sokszögek oldalszáma rendre

$n1$
,

$n2$
,

$n3$
és

$n4$
. Ezek egy-egy szöge

$180\square-(360\square/n1)$, $180\square-(360\square/n2)$, $180\square-(360\square/n3)$, $180\square-(360\square/n4)$.

Az egy csomó képzésében részt vevő négy szög – tehát a fenti négy szög – összege

$360\square$
:

$4\square 180\square-360\square((1/n1)+(1/n2)+(1/n3)+(1/n4))=360\square$.
Ezt átrendezve kapjuk, hogy

$(1/n1)+(1/n2)+(1/n3)+(1/n4)=1$.
Feladatunk most megtalálni ennek az egyenletnek pozitív egész megoldásait. Az

n
-ek értékei természetesen 2-nél nagyobbak. Tegyük fel, hogy nagysági sorrendjük

$n1\leq n2\leq n3\leq n4$,
azaz

$(1/n_1)\geq(1/n_2)\geq(1/n_3)\geq(1/n_4)$.

Vizsgáljuk meg először a háromszögek számát! Egy csomóban nem lehet 3 háromszög, mert ha pl.

$n_1=n_2=n_3=3$
, akkor

$3\cdot(1/3)+(1/n_4)=1$,

$(1/n_4)=0$
következik, ami lehetetlen. Tegyük fel, hogy két háromszög van, tehát

$n_1=n_2=3$
, ekkor:

$(1/n_3)+(1/n_4)=(1/3)$, $n_3=3+(9/n_4-3)$.
Mivel

n_3
egész; kell, hogy

(n_4-3)
pozitív egész osztója legyen 9-nek, tehát a lehetséges értékek

n_4-3	1	3	9
n_4	4	6	12
n_3	12	6	4

A nagyságrendi megkötések miatt ebből csak az

$n_3=6$
,

$n_4=6$
és az

$n_3=4$
,

$n_4=12$
megoldás lehetséges. A két háromszöget tartalmazó csomókhoz tehát a lehetséges megoldások:

A:(3, 3, 6, 6) és B:(3, 3, 4, 12).
Tegyük most fel, hogy a csomóban csak egy háromszög van

$(n_1=3)$
; ekkor

n_2
legkisebb értéke 4 lehet. Ebben az esetben

$(1/3)+(1/4)+(1/n_3)+(1/n_4)=1$,
ebből az előbbiekhez hasonlóan

$(1/n_3)+(1/n_4)=(5/12)$.
Mivel

$(1/n_3)\geq(1/n_4)$,

$(5/12)=(1/n3)+(1/n4)\leq(2/n3)$,

$n3\leq(24/5)<5$,
tehát

$n3$
csak 4 lehet; az egyenletbe való helyettesítéssel látható, hogy ekkor

$n4=6$
. Kaptuk tehát a

C:(3, 4, 4, 6)
megoldást.

$n2$
nem lehet 4-nél nagyobb, mert ebben az esetben

$n3$
és

$n4$
is nagyobb 4-nél és így

$1=(1/3)+(1/n2)+(1/n3)+(1/n4)\leq(1/3)+(1/5)+(1/5)+(1/5)=(14/15)<1$
lenne, ami lehetetlen. Ha tehát a csomóban van háromszög, csak az

A
,

B
,

C
esetek léphetnek fel.

Ha nincs a parkettának háromszög építőeleme,

$n1$
legkisebb értéke 4. De ekkor

$n2$
,

$n3$
,

$n4$
közül egyik sem lehet 4-nél nagyobb, mert akkor reciprokjaik összege kisebb lenne 1-nél, tehát nem elégítenék ki az egyenletet; a nagyságrendi megkötések miatt viszont mindegyiknek legalább 4-nek kell lennie, így ebben az esetben az egyedül lehetséges megoldás

D:(4, 4, 4, 4).

Hasonló okból nem lehet

$n1$

4-nél nagyobb szám. Ezzel egyenletünknek mind a négy megoldását megtaláltuk. Kérdéses azonban, hogy az egyenlet megoldásainak megfelel-e mindig egy, az egész síkot lefedő parketta. Ebben nem lehetünk eleve biztosak, hiszen mi a parketta létezésének feltételét csak egyetlen csomóban vizsgáltuk, és nem biztos, hogy ez az egész síkra kiterjeszthető. Vizsgáljuk meg ebből a szempontból az

A

,

B

,

C

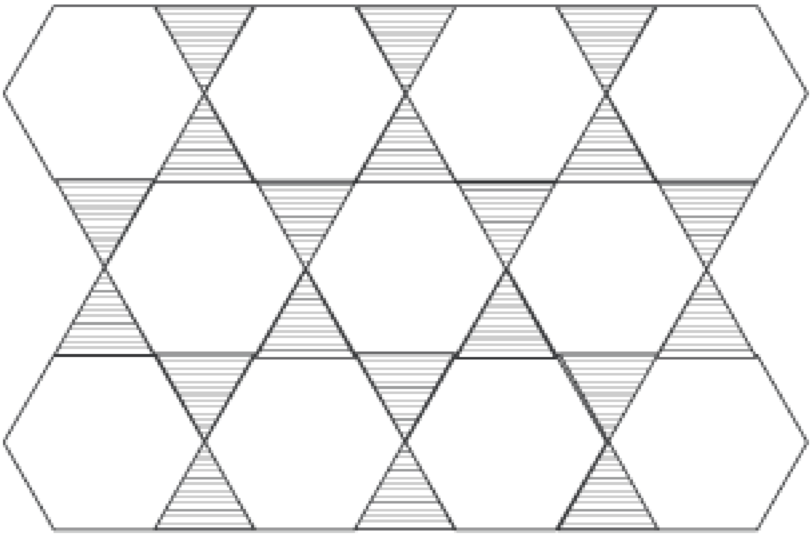
,

D

megoldásokat! Az

A

megoldás könnyen megvalósítható (18. ábra).

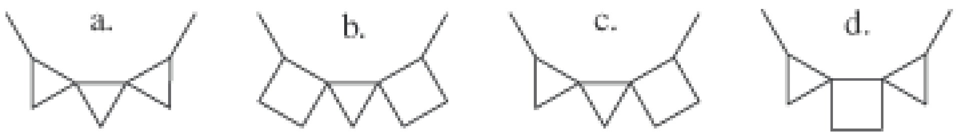


18. ábra

Megmutatjuk, hogy a

B

megoldáshoz nem tartozik megvalósítható parketta. Válasszunk ki ui. egy tizenkétszöget! Ennek két szomszédos oldalához nem csatlakozhat négyzet, mert akkor a közös csúcsban két négyzet lenne, ellentétben a feltevéssel. A tizenkétszög három szomszédos oldalához csatlakozó négyzetek, ill. háromszögek a 19. ábrán

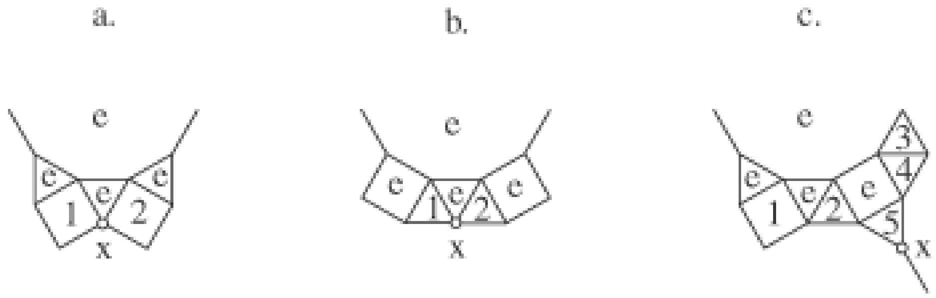


19. ábra

látható négy különböző módon helyezkedhetnek el. Kíséreljük meg ezeknek a kiterjesztését az egész síkra! Az eredeti sokszögeket jelöljük az ábrán

e

-vel! A további építés minden esetben egyértelmű, az építés sorrendjét a számozás mutatja (20. ábra).

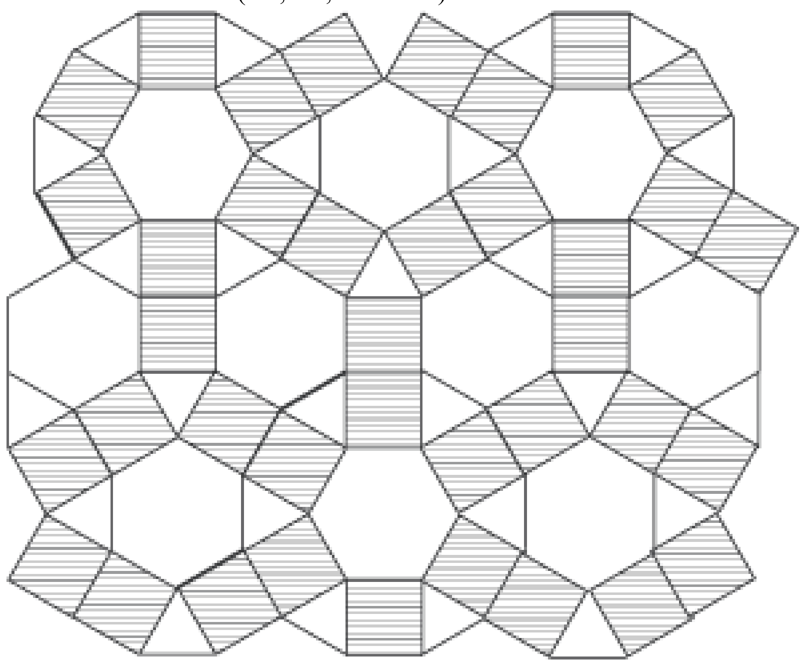


20. ábra

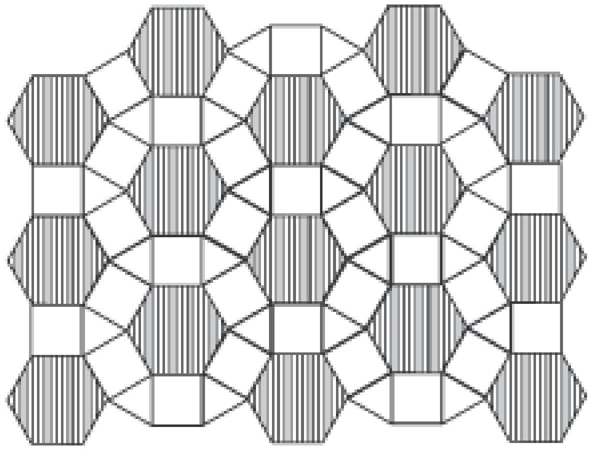
Az *a)*, *b)*, *c)* ábrákon a továbbépítés lehetetlensége az

X
-szel jelölt csomóban mutatkozik meg, mert ide szükségképpen olyan sokszögek kerültek össze, amelyeket feltételünk nem enged meg. A *d)* eset kiterjesztésének lehetetlensége abból következik, hogy szükségképpen visszavezetődik a *b)*, ill. *c)* esetre aszerint, amint az egyik háromszöggel szomszédos tizenkétszög-oldalra négyzetet vagy háromszöget illesztünk. A

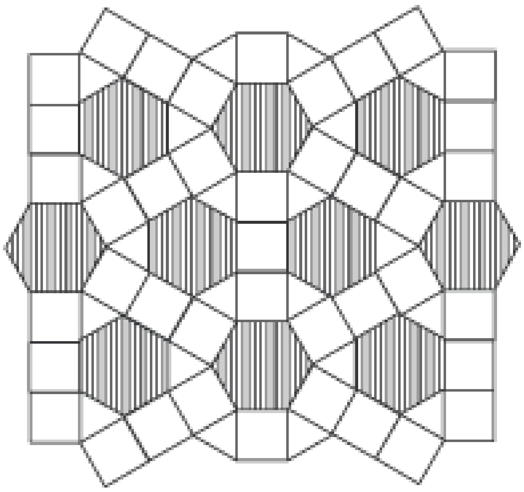
C
megoldás többféle módon is realizálható (21., 22., 23. ábrák).



21. ábra



22. ábra

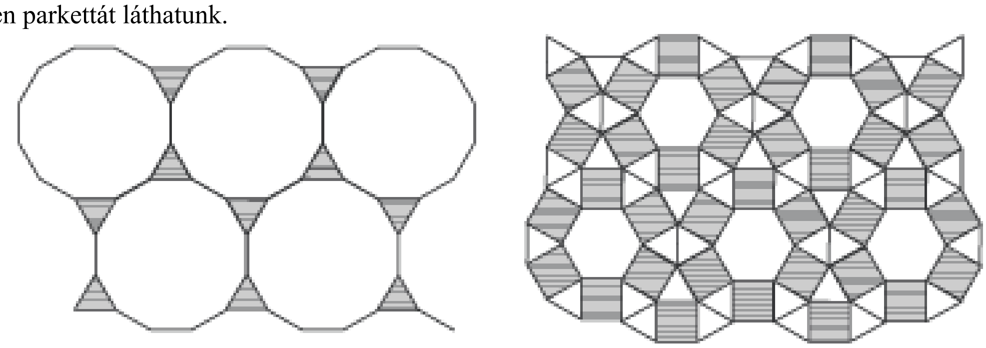


23. ábra

A

D
megoldást már a 2. ábráról ismerjük.

Ha a homogén parketta definíciójában szereplő 3. feltételt elhagyjuk, ún. *inhomogén parkettákat* kapunk. A 24. ábrán szabályos tizenkétszögből és háromszögből felépített homogén parkettát és egy abból származtatott inhomogén parkettát láthatunk.



24. ábra

A parkettázás kapcsán felmerülő geometriai problémák számtalan további gondolatot, kérdést vetnek fel.

Megemlítünk ezek közül néhányat. Szó volt róla, hogy pl. körökkel nem rakható ki a sík hézagtalanul. Ha ennek ellenére egybevágó körökkel szeretnénk kirakni egy padlózatot, felvetődik a kérdés, hogyan lehet a köröket a *legsűrűbben* elhelyezni, azaz úgy, hogy a le nem fedett és lefedett részek területének aránya a lehető legkisebb legyen. Bebizonyítható, hogy ez akkor következik be, ha a körök középpontjai egy szabályos háromszögekből készített parkett csomópontjai. Általában igen nehéz megtalálni a legsűrűbb elhelyezkedést abban az esetben, ha körök helyett más síkalakzatokat veszünk.

Az itt tárgyalt kérdések *térbeli* megfelelői is természetszerűen merülnek fel: milyen testekkel lehet hézagtalanul és egyrétűen kitölteni a teret? A paralelepipedonokon (paralelogramma alapú hasábok, pl. a téglatest, kocka) kívül az első pillanatban aligha tudnánk erre alkalmas testet mondani, bár vannak ilyenek; itt is azonban még számos kérdés megoldatlan. Ugyancsak megoldatlan az egybevágó gömbök legsűrűbb elhelyezésének kérdése is. Érdekes különbség a sík és tér szerkezete között, hogy míg tetszőleges háromszöggel nagyon egyszerűen tudunk parkettát készíteni, a háromszög térbeli megfelelőjével, a tetraéderrel (háromszög alapú gúla) nem tölthető ki a tér egyrétegűen és hézagtalanul. Azonos élhosszúságú tetraéderrel és oktaéderrel azonban már lehetséges a tér egy homogén kitöltése. Az ehhez hasonló jellegű kérdések gyakorlati alkalmazást is nyernek pl. a kristályszerkezetek vizsgálatánál. A matematika e problémaköre a geometria egy, az utóbbi évtizedekben kialakult ágához, az ún. *diszkrét geometriához* tartozik, amelynek fejlődését magyar matematikusok kiváló eredményekkel segítették elő.

ÉRDEKES SZÁMOK

SURÁNYI JÁNOS

1. Az embernek tevékenységei során időtlen idők óta szüksége volt tárgyak megszámlálására. A számok nevét eleinte tárgyakhoz kötötte, de igen hamar kezdett tudatára ébredni annak is, hogy itt nem valami tapinthatóról, érzékelhetőről, hanem elvont fogalomról van szó. Nehezen tudott azonban belenyugodni a számok teljes elvontságába, ezért igyekezett felruházni azokat különböző tulajdonságokkal. Ma is számon tartanak pl. szerencsés és szerencsétlen számokat, régebben pedig ezt a számmisztikát numerológia néven külön tudománynak tekintették. Az ilyen belemagyarázott tulajdonságok mellett azonban a számok tényleges tulajdonságait is kezdték felfedezni és azokat hozták kapcsolatba emberi tulajdonságokkal. Egyik példa erre a tökéletes számok csoportja.

A görögök a számok osztóit a szám részeinek tekintették. A mindennapi életben ma is szoktunk beszélni egy mennyiség valamilyen „hányadrész"-éről. Egy szám osztójának segítségével fel lehet bontani a számot ennyi egységből álló egyenlő részekre. Érthető, hogy ilyen módon magát a számot nem tekintették a szám osztójának. Rakjuk most össze a szám *különböző* „részeit". (Mai szóhasználattal adjuk össze a valódi – azaz a számnál kisebb – osztóit.) Pl. a 4 „részei" 1 és 2, összegük

$1+2=3$

; a 2 vagy 5 egyetlen „része" az 1; a 6 „részei" 1, 2, 3, összegük

$1+2+3=6$

; a 12 „részei" 1, 2, 3, 4, 6, összegük 16. A görögök az olyan számokat, amelyek „összetehetők" különböző „részeikből" – mint pl. a 6 – egy igen magas fokú tökéletesség jelképének tekintették és *tökéletes számoknak* nevezték. További példák tökéletes számokra: 28, osztói 1, 2, 4, 7, 14, összegük 28; 496; 8128. Eddig 3710 tökéletes szám ismeretes.

2. Két olyan számot, amelyek bármelyikének részeit összeadva a másik számot kapjuk, a barátság legmagasabb fokú kifejezésének tekintették, és *barátságos számoknak* nevezték. Pythagorasz állítólag a legmagasabb elismerésként mondta két tanítványáról, ezek olyan barátok, mint a 220 és a 284. Valóban,

220 valódi osztói:

1, 2, 4, 5, 10, 11, 20, 22, 44, 55, 110,
összegük:

$1+2+4+5+10+11+20+22+44+55+110=284$

284 valódi osztói:

1, 2, 4, 71, 142,
összegük:

$1+2+4+71+142=220.$

Érdekes, hogy az ókori forrásokban csak ez a barátságos számpár szerepel. Az újkor elején viszont Fermat, Mersenne, Descartes és mások sorozatban állítják elő az ilyen számpárokat, 6-, 8-jegyűeket is. Egy másik érdekesség, hogy a következő legkisebb barátságos számpárt viszont csak a múlt század közepén egy Niccolò Paganini nevű 16 éves olasz fiú adja meg, ez az 1184 és 1210.

1184 valódi osztói:

1041 (2004. május óta)

1, 2, 4, 8, 16, 32, 37, 74, 148, 296, 592,

összegük:

1+2+4+8+16+32+37+74+148+296+592=1210;

1210 valódi osztói:

1, 2, 5, 10, 11, 22, 55, 110, 121, 242, 605,

összegük:

1+2+5+10+11+22+55+110+121+242+605=1184.

3. Lényegesen újabb keletű számérdekességet nyújtanak a *főnixszámok*. Legismertebb példájuk a

142 857
. Ennek néhány többszörösét az alább látható táblázat tünteti fel.

$$142\,857 \cdot 1 = 142\,857$$
$$142\,857 \cdot 2 = 285\,714$$
$$142\,857 \cdot 3 = 428\,571$$
$$142\,857 \cdot 4 = 571\,428$$
$$142\,857 \cdot 5 = 714\,285$$
$$142\,857 \cdot 6 = 857\,142$$

Figyeljük meg, hogy ezek a többszörösök a kiinduló szám számjegyeiből épülnek fel ugyanabban a sorrendben is, mint ahogyan az eredeti számban következnek, csak más jegynél kezdve, és ha a szám végére értünk, az elejét kell utána írni folytatólag. A főnixszámok a képzelet alkotta főnix madárról kapták nevüket. Ez – a monda szerint – ha elégetik is, újra feléled hamvaiból.

4. Egyre világosabb lett minden gondolkodó ember előtt, hogy a számok nem rendelkeznek a nekik tulajdonított csodálatos tulajdonságokkal. Azok azonban, akik szívesen foglalkoznak számokkal, továbbra is felfigyeltek a számok különböző érdekességeire.

Igen egyszerű szabályosság szerint épülnek fel a csupa egyező jeggyel írt számok. Pl. az 1, 11, 111, 1111, …. Szorozzuk meg mindegyiket önmagával. Az eredményt a lent látható táblázat tünteti fel.

$$1 \cdot 1 = 1$$
$$11 \cdot 11 = 121$$
$$111 \cdot 111 = 12\,321$$
$$1111 \cdot 1111 = 1\,234\,321$$

$$3 \cdot 3 = 9 \qquad 9 \cdot 9 = 81$$
$$33 \cdot 33 = 1\,089 \qquad 99 \cdot 99 = 9\,801$$
$$333 \cdot 333 = 110\,889 \qquad 999 \cdot 999 = 998\,001$$
$$3333 \cdot 3333 = 11\,108\,889 \qquad 9999 \cdot 9999 = 99\,980\,001$$

Járjunk el hasonlóképpen a csupa 3-assal és a csupa 9-essel írt számokkal! Ezeknek az eredményeit is láthatjuk.

5. Az érdekességek további sorolása helyett vegyük most közelebbről szemügyre, hogy min is alapszanak a felsorolt jelenségek. Hogyan találhatunk tökéletes, ill. barátságos számokat, folytatódnak-e az utoljára felírt táblázatokban talált szabályszerűségek? Tudunk-e további hasonló táblázatokat készíteni?

Kezdjük az utolsó példával! Vajon igaz-e, hogy az öt 9-essel írt

99 999
számot megszorozva önmagával

9 999 800 001
számot kapjuk? A

99 999
egy híján

100 000
. Vele tehát úgy is szorozhatunk, hogy a szorzandó

100 000
-szereséből egyszer levonjuk a szorzandót. Ha ezt a

99 999
szorzandóval végezzük, könnyen láthatóan a várt eredményre jutunk:

$$\begin{aligned} 99\,999 \cdot 99\,999 &= 99\,999 \cdot 100\,000 - 99\,999 = \\ &= 9\,999\,900\,000 - \\ &\quad - 99\,999 = \\ &\overline{\quad 9\,999\,800\,001.} \end{aligned}$$

Az is világos, hogy ez a szabályosság nem múlott azon, hogy éppen öt jeggyel írt számot szoroztunk meg önmagával. Hasonló lesz az eredmény, akárhány 9-essel írt számmal végezzük is el ugyanezt.

Hasonlóan adhatunk magyarázatot a csupa 3-assal írt számok szorzatára is. Egy szorzatot tekinthetünk olyan összegnek, amelynek minden tagja a szorzandó, a tagok száma pedig a szorzóval egyenlő. Ha a szorzandó pl.

33 333
, akkora minden tagot három

11 111
-es tag összegére bonthatunk, és így a tagok száma háromszor annyi lesz, mint előzőleg, azaz esetünkben

99 999
. Az imént láttuk azonban, hogyan szorozhatunk egyszerűen ezzel a számmal. Ugyanazzal a meggondolással most a következő számításhoz jutunk.

$$\begin{aligned} 33\,333 \cdot 33\,333 &= 11\,111 \cdot 99\,999 = \\ &= \ 1\,111\,100\,000 - \\ &\quad - 11\,111 \\ &\overline{\quad 1\,111\,088\,889.} \end{aligned}$$

Ismét világos, hogy hasonlóan járhatunk el, akárhány hármassal írt számból indulunk is ki és hasonló eredményre is jutunk.

Rendkívül egyszerű az 1-esekkel írt számok önmagukkal való szorzásakor talált szabályszerűség magyarázata. Válasszuk példának ismét az ötjegyű

11 111

számot! A szorzást a szokásos módon végezzük, minden részletszorzat öt 1-esből áll, rendre egy-egy jeggyel jobbra írva. Így az utolsó részletszorzat első jegye az első részletszorzat – maga a szorzandó – utolsó jegye alá kerül. Ebben az oszlopban tehát annyi 1-es van, ahány 1-es magában a szorzandóban (ill. a szorzóban); ettől jobbra, balra pedig rendre eggyel csökken az 1-esek száma.

$$\begin{array}{ccccccccc}
1&1&1&1&1&\cdot\;1\;1\;1\;1\;1&&& \\
&1&1&1&1&1&&& \\
&&1&1&1&1&1&& \\
&&&1&1&1&1&1& \\
&&&&1&1&1&1&1 \\
\hline
1&2&3&4&5&4&3&2&1
\end{array}$$

Ez a szabályszerűség már nem marad meg akárhány jegyű számokra, mert ha valamelyik oszlopban 9-nél több 1-es gyűlik össze, akkor ebből átvitel adódik, és ez megbontja az eredmény jegyeinek szabályszerűségét. Ha 10 helyett egy nagyobb alapszámú számrendszerben végezzük a számítást[11] akkor többjegyű számokig marad fenn hasonló szabályszerűség, azonban mindig csak kevesebb 1-essel írt számokra lesz érvényes, mint amekkora a számrendszer alapszáma. Azt is érdemes megnézni, hogy az előbbi szabályosságok még milyen más alapú számrendszerben jelentkeznek. Ajánljuk az olvasónak, hogy próbáljon más számrendszerekben is keresni hasonló tulajdonságot mutató számokat.

6. Visszatérve a 10-es számrendszerhez, szorozzunk meg önmagukkal csupa 6-ossal írt számokat. A jelentkező szabályosságot nem nehéz az előbbiek mintájára igazolni.

$$\begin{array}{rcrcr}
6 & \cdot & 6 & = & 36 \\
66 & \cdot & 66 & = & 4\,356 \\
666 & \cdot & 666 & = & 443\,556 \\
6666 & \cdot & 6666 & = & 44\,435\,556
\end{array}$$

Még egyszerűbb szabályosság adódik, ha megnöveljük egygyel a csupa hármassal, ill. a csupa hatossal írt számokat és ezeket szorozzuk önmagukkal. Az eredményt táblázatban foglaltuk össze:

$$\begin{array}{rcrcrcrcrcr}
4 & \cdot & 4 & = & 16 & \qquad & 7 & \cdot & 7 & = & 49 \\
34 & \cdot & 34 & = & 1\,156 & & 67 & \cdot & 67 & = & 4\,489 \\
334 & \cdot & 334 & = & 111\,556 & & 667 & \cdot & 667 & = & 444\,889 \\
3334 & \cdot & 3334 & = & 11\,115\,556 & & 6667 & \cdot & 6667 & = & 44\,448\,889
\end{array}$$

Megint nem túl nehéz a fellépő szabályszerűségek igazolása a korábban talált eredmények segítségül vételével. Csak azt kell meggondolni, mennyivel növekszik a szorzat, ha mindkét tényezőjét 1-gyel növeljük. Pl.

3334□3334=3334□3333+3334=3333□3333+3333+3334= =11 108 889+1111+5556=11 115 556.

7. Nézzük meg most a főnixszámokat! Milyen szabályszerűséget találhatunk a

142 857

számban? Feltűnhet pl., hogy a középső két jegyből alkotott 28-as szám az előtte álló 14 kétszerese, az utána

[11] Lásd Erdős P. és Surányi J.: *Válogatott fejezetek a számelméletből*, Polygon, 1996., Szeged, 7-9. old.

álló 57 pedig a 28 kétszeresénél 1-gyel nagyobb. Vajon a hasonló szabályszerűség alapján képezett

234 693
főnixszám lesz-e?

234 693□1 =234 693 234 693□2 =469 386 234 693□3 =704 079.

Amint látjuk, a szám kétszeresénél még kezdődik valami hasonló szabályszerűség, az utolsó jegyekben azonban már elromlik, a háromszorosnál pedig már semmi szabályosságot sem fedezhetünk fel. Az okot tehát mélyebben kell keresnünk. Számítsuk ki főnixszámunk következő többszörösét:

142 857□7=999 999.

Így a hat 9-essel írott számhoz jutunk. Ez már mutat valami érdekességet. Ha ehhez egyet hozzáadunk,

1 000 000
-t kapunk. Ez más szóval azt jelenti, hogy az egymillió 7-tel való osztásakor a hatodik osztás után újra 1 a maradék. Egészen hasonló lesz a helyzet, ha nem

1 000 000
-t, hanem 1-et osztunk 7-tel, mindössze annyi különbséggel, hogy az osztást egy 0 leírásával és utána tizedesvessző kitételével kell kezdeni. Valóban akár 1, akár

1 000 000
az osztandó, az eljárás ugyanaz; minden egyes maradék után 0-t írunk és így osztunk tovább. A hatodik lépés után újra 1-et kapunk maradékul, ekkor azonban a továbbiakban a korábbi maradékok ismétlődnek ugyanabban a sorrendben, és így a hányados jegyei is periodikusan ismétlődnek: periodikus (szakaszos) tizedestörthöz jutunk, amelynek szakasza éppen főnixszámunk.

Nem minden egész szám reciprokának tizedestört alakja szolgáltat azonban főnixszámot, hiszen az

(1/2)=0,5
az 5 számot adja, és ez nem főnixszám. Igaz ugyan, hogy az

(1/2)
esetében nem jutottunk periodikus tizedestörtre, de ugyancsak nem főnixszám az

(1/3)=0,3333…
számból adódó 3-as szám, valamint a

(10/99)
szakaszából adódó 10-es szám sem.

1 : 13 = 0,076 923
10
 90
 120
  30
   40
    1

Tegyünk próbát az

(1/13)
szakaszát alkotó

076 923
számmal! Első 13 többszörösét alább láthatjuk.

076 923 · 1 = 076 923
076 923 · 2 = 153 846
076 923 · 3 = 230 769
076 923 · 4 = 307 692
076 923 · 5 = 384 615
076 923 · 6 = 461 538
076 923 · 7 = 538 461
076 923 · 8 = 615 384
076 923 · 9 = 692 307
076 923 · 10 = 769 230
076 923 · 11 = 846 153
076 923 · 12 = 923 076
076 923 · 13 = 999 999

Bár ismét nem kaptunk főnixszámot, a felírt táblázat mégis mutat valami hasonló érdekes jelenséget. A 12-szeresig haladva a szorzatok fele mutatja a főnixszámoknál talált szabályosságot. (Ezek a táblázatban kövéren vannak szedve.) A másik részük azonban a kétszeresnél először fellépő szorzat jegyeiből keletkezik hasonló szabályosság szerint, mint a többi szorzat az eredeti számból.

(A szám 13-szorosa itt is a hat 9-essel írt szám, hiszen a hatodik részletmaradék 1, vagyis 13-nak a

76 923
-szorosa 1-gyel kisebb az 1-es után hat 0-val írt számnál, a milliónál.) Nem is várható azonban, hogy mindegyik szorzat az eredeti számból keletkezzék úgy, hogy a jegyek egymás közötti sorrendjét nem változtatjuk, csak más számjegynél kezdjük írni; ehhez ugyanis 12 jegyre volna szükség, számunk pedig csak hatjegyű.

8. Tehát egy szám reciprokának tizedestört alakja akkor ad főnixszámot, ha ez a tizedestört periodikus és periódusa 1-gyel kevesebb jegyből áll, mint az a szám, amelynek reciprokát vettük. A 7 után először a 17 rendelkezik ezzel a tulajdonsággal: reciproka periodikus tizedestört és szakasza 16 jegyű. Ez a szakasz valóban főnixszámot is szolgáltat, ha a szakasz elején álló 0-t is hozzászámítjuk a számhoz, amint azt már az

(1/13)
esetében is tettük:

10 : 17 = 0,058 823 529 411 764 7
100
150
140
40
60
90
50
160
70
20
30
130
110
80
120
1

0 588 235 294 117 647□2 =1 176 470 588 235 294 0 588 235 294 117 647□3 =1 764 705 882 352 941 0 588 235 294 117 647□4 =2 352 941 176 470 588.

Csak az első néhány többszöröst írtuk fel, azonban a többszörösök kiszámítása nélkül is – ami igen hosszadalmas munka volna – meggyőződhetünk arról, hogy így valóban főnixszámot kaptunk. A szakasz ugyanis ott fog befejeződni, ahol először kapunk az osztás folyamán olyan maradékot, amely már korábban is előfordult. Ha azonban 17-tel osztunk és az osztás maradék nélkül nem végezhető el, akkor csak az 1, 2, 3, …, 14, 15 és végül 16 léphet fel maradékul, tehát összesen 16 szám. Ha tehát 16 jegyű szakaszt kaptunk, akkor 16 különböző maradéknak kellett fellépnie, tehát minden lehetséges maradéknak. (A fenti osztásban a „levett” 0-kat kisebb jeggyel írtuk, és a hányados 0 jegyéhez tartozó 10 maradékot is újra leírtuk, hogy világosan látható legyen a 16 különböző maradék.)

Ha most az

(1/17)
helyett pl. a

(11/17)
-et vesszük, ez egyrészt az

(1/17)
11-szerese. Másrészt, ha ezt tizedestörtté alakítjuk, ugyanazt a számítást végezzük, amelyet az

(1/17)
tizedestörtté alakításakor végeztünk az abban is fellépő 11 maradéktól kezdve. Azok a jegyek következnek a hányadosban, és azok a maradékok, amelyek az

(1/17)
átalakításakor a 11 maradéktól kezdve léptek fel, egészen addig, amíg az 1 maradékhoz nem jutunk, ezután pedig azok a maradékok és így a hányadosnak is azok a jegyei következnek, amelyek az

(1/17)
osztás elején léptek fel, amíg csak meg nem ismétlődik a 11 maradék. Ismét periodikus tizedestört lesz tehát az eredmény, és ennek a szakasza ugyanazokból a jegyekből áll ugyanabban a sorrendben, mint az

$(1/17)$
szakasza, csak más számjeggyel kezdődik. Az így kapott szám, mint már említettük, az

$(1/17)$
szakaszából adódó szám 11-szerese.

Hasonló meggondolással láthatjuk be, hogy a szám 16-szorosáig mindegyik többszörös ugyanazokból a jegyekből áll ugyanabban a sorrendben, mint az eredeti szám, csak más jeggyel kezdve, és ha a szám végére értünk, akkor az elején álló jegyeket kell utána írni ismét az eredeti sorrendben. Ezzel beláttuk, hogy az

$(1/17)$
szakasza valóban főnixszámot szolgáltat.

Valószínűnek látszik, hogy végtelen sok főnixszám létezik, ezt azonban nem sikerült ez ideig sem bebizonyítani, sem megcáfolni.

9. Térjünk most vissza a tökéletes és barátságos számokra! Az már Euklidésznél megtalálható, hogy ha

p
olyan természetes szám, amelyre teljesül, hogy 2-nek a

p
-edik hatványából egyet levonva prímszámot kapunk, akkor ennek a prímszámnak a

$2p-1$
-szerese tökéletes számot szolgáltat. Ha

$p=2$
, 3, 5, 7, akkor éppen a felsorolt tökéletes számokat kapjuk. Más alakúak nem is ismertek, ennek ellenére csak 2000 évvel később Eulernak sikerült bebizonyítania, hogy az összes páros tökéletes szám ilyen alakú.

Páratlan tökéletes számot eddig nem találtak, és nem is látszik valószínűnek, hogy léteznék. Ezt azonban eddig nem sikerült bebizonyítani. Mindenesetre Euler és utána sokan mások különböző feltételeket találtak, amelyeket a páratlan tökéletes számoknak ki kell elégíteniük, ha vannak ilyenek; például egy páratlan tökéletes számnak legalább 6 különböző prímosztója kell hogy legyen, és a legnagyobb kitevő, amelyre egy prímosztót hatványozva még a hatvány is osztója a számnak, egy prímosztó esetében páratlan, a többi esetében páros (tehát legalább 2).

A legkisebb szám is, ami az eddig ismert feltételek mindegyikét kielégíti, már csillagászati nagyságú. Ez is arra utal, hogy páratlan tökéletes szám létezése valószínűtlen.

Az elmondott tények bizonyítása nem igényel ugyan nagyon mély matematikai ismereteket, de egy keveset mégis, legalábbis ismerni kell a bizonyításokhoz a számelmélet alaptételét.[12] Ez azt mondja ki, hogy minden, 1-nél nagyobb egész szám vagy prím, vagy felbontható véges sok prímszám szorzatára, és bármilyen módon jutunk el egy ilyen felbontáshoz, mindig ugyanazok a prímek fognak benne szerepelni, mégpedig mindegyik ugyanannyiszor. Ez a tétel módot ad pl. esetünkben egy szám összes osztójának az áttekintésére, és ennek alapján az osztók összegének meghatározására. A bizonyítás további részleteibe azonban, amelyek nem egyszer hosszadalmas számolással is járnak, e helyen nem bocsátkozunk bele.

10. Azt mondhatjuk, hogy még kevesebbet tudunk a barátságos számokról. Ismét rég ismeretes az, hogy ha képezünk egy számsorozatot az 5-tel mint első taggal, majd minden további szám az előző kétszeresénél 1-gyel

[12] Lásd. pl. Erdős P. és Surányi J. idézett könyve I. fejezet 15–17. old., továbbá tökéletes számokkal kapcsolatban VIII. fejezet 262–264-ig és 267–270-ig; vagy Faragó L.: *A számelmélet elemei* (szakköri füzet), Tankönyvkiadó, 1954., Budapest. III. fejezet, 28–33. old.

nagyobb, ezután egy második sorozatot a 17-ből indulva, és minden szám után annak 4-szeresénél 3-mal nagyobb számot írva, akkor a következő módon találhatunk barátságos számpárt. Ha az első sorozatban az

$n–1$
-edik és

n
-edik szám prím, és a második sorozatban is az

n
-edik prím, akkor az első sorozat talált két prímszámának a szorzatát

$2n$
-nel szorozva, és a második sorozatban talált prímszámnak ugyancsak a

$2n$
-szeresét véve, barátságos számpárt kapunk:

5,	11,	23,	47,	95,	191,	383
17,	71,	287,	1151,	4607,	18 431,	73 727

$4 \cdot 5 \cdot 11 = 220$ $4 \cdot 71 = 284$

$16 \cdot 23 \cdot 47 = 17\,296$ $16 \cdot 1151 = 18\,416$

$128 \cdot 191 \cdot 383 = 9\,363\,584$ $128 \cdot 73\,727 = 9\,437\,056$.

A prímszámokat aláhúzással jelöltük meg. Az első sor következő 6 száma közül csak a negyedik prím, a másodikban viszont a 15. szám osztható 7-tel, így legfeljebb az első sorozat 15. és 16. számából kiindulva kaphatunk újra barátságos számokat; ezek már felül vannak az ezer billión.

Itt már láthatóan nem áll az, hogy az ismert barátságos számpárok mind az említett módon keletkeztek volna. A korábban felsoroltak között is találunk ezektől eltérő szerkezetűt. Viszont másfelől azt sem tudjuk, hogy az előbb leírt két sorozatban előfordul-e végtelen sok prímszám, s ennek megfelelően azt sem, hogy van-e végtelen sok, az előbb leírt módon származtatható barátságos számpár. Azt sem tudjuk továbbá, hogy van-e végtelen sok barátságos számpár; nem tudjuk, alkothat-e egy páros és egy páratlan szám barátságos számpárt, még kevésbé azt, hogy állhat-e egy ilyen számpár két olyan számból, amelyeknek nincs 1-nél nagyobb közös osztója.

Hasonlóképpen az sem ismeretes, hogy van-e végtelen sok

$2p–1$
alakú prímszám. Így azt sem tudjuk, hogy végtelen sok páros tökéletes szám van-e vagy csak véges számú. A

$2p–1$
alakú prímeket Mersenne-féle prímeknek szokás nevezni. Azt nem nehéz belátni, hogy egy ilyen alakú szám csak akkor lehet prím, ha a

p
kitevő is prímszám. Viszont az már nem igaz, hogy ilyen módon minden

p
prím kitevőre prímszámot kapnánk. Pl. a

$p=11$
esetben

211–1=2047=23□89.

Különösen az elektronikus számítógépek elterjedése óta sikerült eldönteni sok csillagászatian nagy Mersenne-számról, hogy prím-e vagy sem. Az eddig ismert legnagyobb Mersenne-féle prím (1998 áprilisi adat):

23 021 377–1
. Ez több, mint

909 500
jegyű. Az ezzel képezhető tökéletes számmal együtt jelenleg 37 tökéletes szám ismeretes.[13] Mindez azonban nem ad további felvilágosítást abban a tekintetben, hogy végtelen sok Mersenne-féle prím van-e vagy csak véges számú. Természetesen ilyen módon azt sem tudjuk, hogy van-e végtelen sok páros tökéletes szám vagy sem.

Kézenfekvő ezek után felvetni azt a kérdést, hogy akadnak-e prímek a

2*n*+1
alakú számok között, és melyek, ill. mennyien. Itt viszont azt lehet megállapítani az

n
kitevőről, hogy annak is 2 hatványának kell lennie. *Fermat* úgy sejtette, hogy 2 bármilyen hatványát választva kitevőnek, prímszámhoz jutunk. Ez a sejtés nem bizonyult helyesnek, mert *Euler* megmutatta, hogy ha a

32=25
kitevőt vesszük, akkor már összetett számhoz jutunk:

225+1=4 294 967 297=641□6 700 417.

Megvizsgálták azóta 2 számos magasabb hatványát és az derült ki, hogy a hozzájuk tartozó Fermat-számok nem prímek. Nem tudjuk azonban, hogy van-e további Fermat-féle prím, noha az sincs kizárva, hogy végtelen sok legyen.

A Fermat-féle prímek érdekességét nagyban megnövelte az, hogy – másfél évszázaddal későbben – *Gauss* munkássága nyomán kiderült, hogy prím oldalszámú szabályos sokszöget csak akkor szerkeszthetünk körzővel és vonalzóval, ha az oldalszám Fermat-prím. Ebből már könnyen tárgyalható az összetett oldalszámú sokszögek szerkeszthetősége.

11. Már az eddigiekben is láttuk, hogy egy szám sok különböző szempontból lehet érdekes. Azt gondolhatná az ember, hogy minden számnak van valamilyen érdekessége. Ezt mutatja a következő történet is. Egy *Ramanuyan* nevű hindu matematikust meglátogatott egyszer egy barátja és érkezésekor megjegyezte, hogy az A 1729 rendszámú taxin érkezett oda. „Úgy gondolom – tette hozzá –, hogy ennek a számnak nincs semmi különösebb érdekessége.” *Ramanuyan*, aki nagy ismerője volt a számok tulajdonságainak, azonnal azt felelte: „Tévedsz, ez a legkisebb olyan szám, amelyik két különböző módon írható föl két harmadik hatvány összegeként.” Valóban:

1729=1728+1=123+13=1000+729=103+93.

Felírva az első 12 köbszámot, és sorra hozzáadva mindegyiket az előzőkhöz és önmagához, láthatjuk, hogy a kisebb összegek mindegyike csak egyféleképpen jön létre. Nem is téved, aki úgy véli, hogy minden szám érdekes szám. Könnyen beláthatjuk matematikai szigorúsággal, hogy ez nem lehet másképp.

[13] 2004. május óta 41 tökéletes számot ismerünk, a legnagyobb

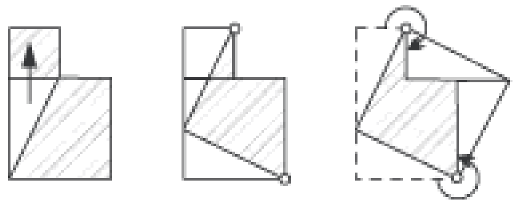
224 036 583–1
, ez már több, mint 7 millió jegyű.

Előrebocsátjuk, hogy amikor érdekes számokról beszéltünk, mindig pozitív egész számokra gondoltunk. Ha nem volna minden pozitív egész szám érdekes, akkor a nem érdekes számok között volna egy legkisebb. Ámde a *legkisebb* nem érdekes számnak lenni, ez már magában nem csekély érdekesség. Így viszont máris ellentmondásba jutottunk azzal, hogy ez a szám nem volna érdekes szám.

Az okoskodás egészen logikus, mégis kifogás merül fel ellene. Akkor volna elfogadható bizonyításnak, ha egyértelműen meghatározható volna az, hogy mikor érdekes egy szám. Bizonyára akad olyan ember is, akinek szám eleve nem lehet érdekes. De ezen túl is, különböző emberek véleménye messzemenően eltér abban, hogy ki mikor tart egy számot érdekesnek, mikor kevésbé érdekesnek. De nem is úgy vetődik fel a kérdés, hogy ez a szám érdekes, ez már nem érdekes, hanem az egyik nagyon érdekes, egy másik nem annyira, egy harmadik már egyáltalán nem. Így a fenti gondolatmenetet sem tekinthetjük bizonyításnak, hiszen az a tulajdonság, amiről be akartunk látni valamit, maga is nagymértékben határozatlan.

Annyit esetleg kiolvashatunk az okoskodásból, hogy minden számban található valami érdekesség, ez viszont olyan határozatlan állítás, ami bizonyítás nélkül is nyugodtan elfogadható.

12. Nem akarom egyre erőszakoltabb tulajdonságokra rásütni az „érdekes" jelzőt, így mindössze két számféleséget említek még meg. Gyakran hallani pythagoraszi számhármasokról. Ismeretes, hogy egy derékszögű háromszög befogóira rajzolt négyzetek területének összege az átfogó fölé rajzolt négyzet területével egyenlő. Ez az itt következő ábrákból is könnyen leolvasható.



1. ábra 2. ábra 3. ábra

Az első ábrán a derékszögű háromszög két befogójára emelt négyzetből álló idomban a háromszöget feltoljuk és alatta még egy háromszöget helyezünk el 90°-kal elforgatva (2. ábra). Ezután a háromszögeket a megjelölt pontok körül 270°-kal elforgatva az idomot az átfogó fölé rajzolt négyzetté alakíthatjuk át (3. ábra).

Már a görögöket érdekelte az, hogy van-e olyan derékszögű háromszög, amelyik oldalainak mérőszáma egész szám. Az ilyen háromszög oldalainak hosszát kifejező számhármast nevezik pythagoraszi számhármasnak. Kiderült, hogy végtelen sok ilyen van, megadható három formula: az $x=r\cdot 2uv$, $y=r\cdot(u^2-v^2)$, $z=r\cdot(u^2+v^2)$ számok minden pythagoraszi számhármast előállítanak, ha r tetszés szerinti természetes szám, u és

v
pedig olyan természetes számok, amelyek közül az első a nagyobbik, nincs 1-nél nagyobb közös osztójuk és az egyik páros, a másik páratlan. Az alábbi táblázat néhány, az

$r=1$
esetnek megfelelő, ún. pythagoraszi alaphármast tartalmaz:

u	2	3	4	4	5	5	6	6	7	7	7	8	8	8	8	9	9	9
v	1	2	1	3	2	4	1	5	2	4	6	1	3	5	7	2	4	8
x	4	12	8	24	20	40	12	60	28	56	84	16	48	80	112	36	72	144
y	3	5	15	7	21	9	35	11	45	33	13	63	55	39	15	77	65	17
z	5	13	17	25	29	41	37	61	53	65	85	65	73	89	113	85	97	145

Megfigyelhetjük, hogy bármelyik hármas számainak szorzata osztható 60-nal. Belátható, hogy ez teljesül bármelyik pythagoraszi számhármasra.

Visszatérve formuláinkra, azt könnyű kiszámolni, hogy az első két kifejezés négyzetének összege a harmadik kifejezés négyzetét adja, akármilyen számokat jelentenek is a bennük levő betűk. Az is világos, hogy ha a betűk egész számot jelentenek, akkor a kifejezések értéke is egész lesz. Annak bizonyítása, hogy már akkor is megkapunk minden pythagoraszi számhármast, ha a betűk az előbb mondott feltételeknek eleget tevő egészeken futnak végig, ismét a számelmélet alaptételére építhető fel; ennek részleteivel azonban nem kívánok itt foglalkozni.

13. Ismeretes, hogy az egész középkoron át a matematika nem fejlődött lényegesen. A matematikai tevékenység főként egyes számolóművészek mutatványaira szorítkozott. Ezeknek az ismereteiről viszont keveset tudunk, hiszen ismereteiket igyekeztek titokban tartani, nehogy vetélytársaik ellessék azokat, és ennek révén fölébük kerekedjenek. A feladatokból, amelyekkel foglalkoztak és megoldásaikból azonban nyilvánvaló, hogy ismerniök kellett valamilyen formában a pythagoraszi számhármasokat, mert gyakran használták azokat meggondolásaikban.

Ennek a hosszú korszaknak egyik kimagasló tehetségű matematikusa volt *Fibonacci*, más néven *Leonardo Pisano* (a pisai Lénárd, Bonaccio fia); mint számolóművész is legendás hírre tett szert. Fiatal korában apjával Keleten járva megismerkedett a helyiérték-rendszeren alapuló arab számírással és hazatérve annak első propagátora volt Európában. Neve különösen a Fibonacci-féle számsorozatról közismert. Ennek első és második száma 1, a továbbiakat pedig úgy számítjuk ki, hogy az előző két szám összegét vesszük. Ha tehát a sorozat

n
-edik számát

u_n
-nel jelöljük, akkor a sorozat képzési szabályát az

$u_{n+2}=u_n+u_{n+1},\ u_1=u_2=1$
formula fejezi ki. A sorozat első néhány száma:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
u_n	1	1	2	3	5	8	13	21	34	55	89	144	233	377	610	987	1597	2584
$u_1+\ldots+u_n$	1	2	4	7	12	20	33	54	88	143	232	376	609	986	1596	2583	4180	6764

A számsorozat felírt tagjai alá odaírtuk az addig terjedő számok összegét is. Ha összehasonlítjuk a két sorozatot, észrevesszük, hogy az alsó sor csupa 1-gyel kisebb számból áll, mint a felső sor a harmadik számtól kezdve. Ha megfigyelésünk helyes, akkor ez azt jelenti, hogy a sorozat első

n
számának összege 1-gyel kisebb a sorozat

(n+2)
-edik számánál. Ennek az észrevételnek a helyességét nem nehéz belátni. Fel kell csak írni a sorozat egymás utáni számait a harmadiktól az

(n+2)
-edikig, mindegyiket a képzési szabály szerint az előző két taggal kifejezve. Ha számainkat összeadjuk, a jobb oldal első oszlopából éppen

$$\begin{array}{l}
u_3 = u_1 + u_2\\
u_4 = u_2 + u_3\\
u_5 = u_3 + u_4\\
\hdashline
u_{n+1} = u_{n-1} + u_n\\
u_{n+2} = u_n + u_{n+1}\\
\hline
u_{n+2} - u_2 = u_{n+2} - 1 = u_1 + u_2 + \ldots + u_{n-1} + u_n.
\end{array}$$

az első

n
Fibonacci-szám összegét kapjuk. A jobb oldal második oszlopa a bal oldali oszloptól csak annyiban különbözik, hogy egy taggal előbb kezdődik és egy taggal előbb fejeződik be. Így a jobb oldali első oszlop összege a bal oldalon fellépő utolsó tagnál a jobb oldal második oszlopában álló első taggal, vagyis a Fibonacci-sorozat második tagjával kisebb, ez a tag azonban 1. Ezzel be is láttuk az észrevett szabályosság helyes voltát.

A Fibonacci-számoknak sok további érdekessége állapítható meg. Belátható pl. hogy az egyjegyű Fibonacci-számok kivételével mindig vagy négy, vagy öt egymás utáni Fibonacci-szám áll ugyanannyi jegyből. Ha megnézzük a felírt Fibonacci-számokat, azok között egy négyzetszámot találunk, a 144-et. Több nem is fordul elő a sorozatban, ennek bizonyítása azonban igen nehéz. Azt viszont nem nehéz belátni – megfelelő számelméleti alapismeretekkel –, hogy bármely két Fibonacci-szám legnagyobb közös osztója szintén Fibonacci-szám, mégpedig az, amelynek sorszáma a két Fibonacci-szám sorszámának legnagyobb közös osztója. Azt is be lehet bizonyítani, hogy bármely egész számnak van többszöröse a Fibonacci-számok között. A legkisebb ilyen Fibonacci-szám sorszámára azonban csak felső korlátot sikerült megadni, általában sikerül találni jóval kisebb sorszámú Fibonacci-számot is, amelyik az adott számnak többszöröse.

Lehetne még folytatni a felsorolt számok tulajdonságait, még inkább az érdekes számok sorozatát, annyi azonban talán az elmondottakból is látszik már, hogy sok szempontból vizsgálhatjuk a számok érdekességét, és eközben tetszetős összefüggések mellett számos, eddig még megoldatlan problémára is bukkanunk.

HÁNY SZÍN KELL A TÉRKÉP SZÍNEZÉSÉHEZ?

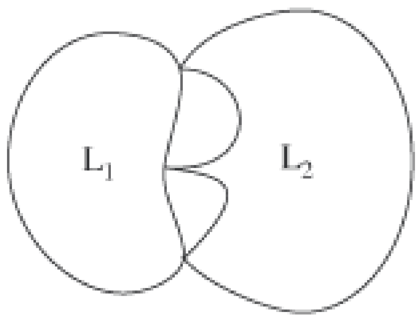
ANDRÁSFAI BÉLA

1. 1. A NÉGYSZÍN PROBLÉMA

A térképeken színezéssel szokás áttekinthetővé tenni az országok rendszerét, mégpedig úgy, hogy egy ország minden részét ugyanolyan színűre, a különböző országokat pedig különböző színűre festik be. Az áttekintést nem zavarja, ha nem szomszédos országok ugyanazt a színt kapják. A színezésnél akkor kell *két országot szomszédosnak* tekintenünk, ha határvonaluknak van közös szakasza; tehát az 1. ábrán látható

$L1$
és

$L2$
nem szomszédos országok.



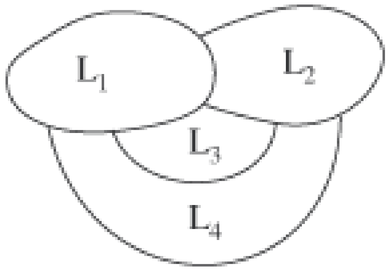
1. ábra

Egy térképet

p
színnel jól színezhetőnek mondunk, ha

p
színnel úgy színezhetők az országai, hogy egy ország színezéséhez a

p
szín közül csak egyet használunk, és a szomszédos országok különböző színt kapnak. A térképek elkészítéséhez célszerű minél kevesebb színt használni. Ha egy térképen pl. 100 ország van, akkor 100 színnel biztosan jól színezhető. De szükséges-e ilyen sok szín? Ha az országaink olyanok, hogy mindegyiknek van egy-egy része mindegyikben, akkor igen, hiszen valamennyi lehet valahol szomszédos. Talán az országok feldaraboltsága miatt van szükségünk ilyen sok színre? Zárjuk most ki ezt a lehetőséget! Nevezzünk egy térképet *normál térképnek*, ami azt jelenti, hogy bármely országának két tetszőleges pontja összeköthető az országon belül haladó útvonallal. Ilyen országokat *összefüggőknek* mondunk. Több mint 100 éve *Cayley* vetette fel a problémát: vajon hány szín elegendő bármilyen normál térkép jó színezéséhez? A 2. ábrán látható



2. ábra

normál térkép négy országának jó színezéséhez 4 szín szükséges, hiszen a négy ország közül bármely kettőnek van közös határa, azaz a négy ország *páronként szomszédos*. A kérdéses minimális színszám tehát legalább 4. Az eddig felrajzolt normál térképek mindegyikét sikerült 4 színnel jól színezni, de a mai napig senki sem tudta bizonyítani, hogy 4 szín minden normál térkép jó színezéséhez elegendő. Ez a híres *négyszín probléma*.[14]

A probléma gyakorlatilag nem túl érdekes, hiszen egyrészt a valóságban nem csupán normál térképek fordulnak elő, másrészt az bebizonyított tény, hogy 5 színnel már bármilyen normál térkép jól színezhető. (Ez az ún. ötszín probléma. Megoldásával a későbbiekben foglalkozunk.) Ennek ellenére a probléma mélyen behatolt a matematika területére, és sokan megkísérelték, hogy a benne rejlő bizonytalanságot feloldják. Az ilyen kísérletek hasznossá válhatnak, mert általuk olyan módszerek birtokába juthatunk, amelyek más, gyakorlatilag is fontos problémák megoldásaihoz vezetnek.

A kutatások eddigi eredményeinek részletes áttekintése messzire vezetne. Itt főleg e tárgykör érdekességeire szeretnénk a figyelmet irányítani.

2. 2. A TÉRKÉP LÉNYEGTELEN MÓDOSÍTÁSA. EULER TÉTELE

A térképek színezésénél a méretek természetesen nem játszanak szerepet. Ha pl. egy jól színezett térképet gumilapra rajzolunk, a gumit szakítás nélkül tetszés szerint nyújthatjuk, közben térképünk jól színezett marad, és továbbra is ugyanazok az országok lesznek egymással szomszédosak, mint eredetileg. Nem teszünk különbséget gömbre és síkra rajzolható térkép között, hiszen az említett nyújtást megengedve, minden síkra rajzolható térkép gömbre is rajzolható, és bármely gömbre rajzolt térkép síkra teríthető. Az utóbbinál nehézséget okozhat, ha a gömbfelületet az országok teljesen lefedik. Ekkor ideiglenesen kiemelünk egy

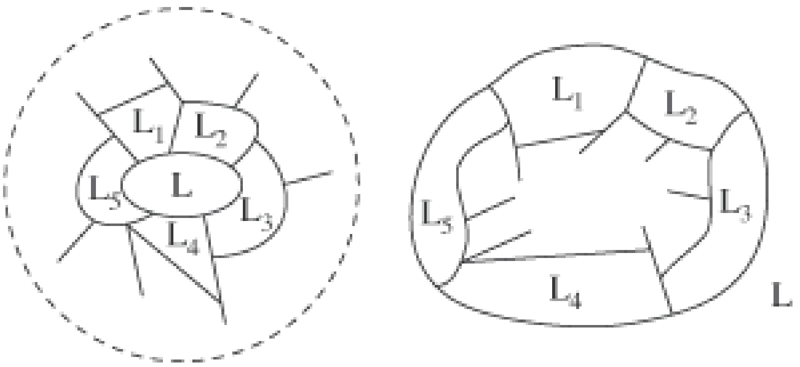
L

országot, a megmaradt térképet kiterítjük a síkra, és a sík le nem fedett részét tekintjük az

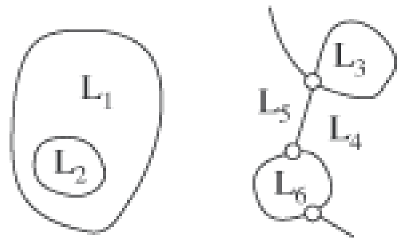
L

országnak (3. ábra).

[14]Ennek a cikknek megírása és első megjelenése után 1976-ban Kenneth Appel és Wolfgang Haken, az Illinois-i egyetem matematikusai elektronikus számítógép igénybevételével bebizonyították a négyszín-tételt. Érdekességképpen megjegyezzük, hogy a bizonyítás leírásához kb. 800 gépelt oldalra volt szükség, elkészítéséhez pedig mintegy 1200 órányi gépidőt használt fel az említett két matematikus. Ún. klasszikus bizonyítást még nem sikerült találni a négyszín-tételre, ezért változatlan az érdeklődés egy ilyen jellegű bizonyítás iránt.



3. ábra



4. ábra

Egy térképen *csúcspontnak* olyan pontot nevezünk, amelyből legalább háromfelé indul határvonal. A határvonalnak azt a szakaszát, amelynek végpontjai csúcspontok, de belsejében nem tartalmaz csúcspontot, *élnek* nevezzük. A határvonalak nem mindig rakhatók össze élekből. A 4. ábra térképén pl. egyetlen élt sem tartalmaz

*L*1
és

*L*2
határvonala,

*L*3
határvonala egy,

*L*4
határvonala pedig két élből áll. *Elemi felületnek* mondunk egy országot akkor, ha az az említett nyújtással körlappá alakítható. Ha a gömbfelületet teljesen lefedjük elemi felületekkel, átfedés nélkül, akkor az elemi felületek

l
, a csúcspontok

c
és az élek

$é$
száma között fennáll a következő, ún. Euler-féle összefüggés:

$l+c=é+2$

Pl. a kocka felülete megengedett nyújtással gömbbé fújható fel; ezen a gömbön a lapok (az elemi felületek) száma 6, a csúcspontok száma 8 és az élek száma 12; tehát ekkor valóban

$l+c=é+2$

Az összefüggés bizonyításához képzeljük el, hogy gömbünk egy égitest, minden élre egy gátat építünk az él teljes hosszában (így bármelyik csúcspontból bármelyikbe eljuthatunk a gátakon haladva), és egyetlen elemi felületet víz borít. Szeretnénk az égitest összes elemi felületét egy-egy gát megnyitásával sorjában elárasztani vízzel. Olyan gátat felesleges nyitnunk, amelynek már mindkét oldalát víz mossa, és azzal sem tudnánk újabb elemi felületet vízzel elárasztani, ha olyan gátat nyitnánk, amelynek mindkét oldala száraz. Tehát minden lépésben olyan gátat nyitunk, amelynek egyik oldala száraz, a másik pedig már elárasztott. Így egy-egy gát megnyitása egy-egy elemi felületet áraszt el vízzel, tehát

$l–1$
gátat kell megnyitnunk.

Vizsgáljuk most az érintetlen gátak rendszerét! Azt állítjuk, hogy még ezen haladva is eljuthatunk bármelyik csúcspontból bármelyikbe. Hiszen két olyan csúcspont között, amelyek közé eső gátat megnyitottuk, az ezáltal elárasztott elemi felület határának megmaradt részén az összeköttetés még fennáll.

Érintetlen gátakon haladva bármelyik csúcspontból bármelyik másikba csak egy úton juthatunk el, ha közben nem szabad visszafordulni. Ha ugyanis két út volna két csúcspont között, ezeken haladva körsétát tehetnénk. Ez olyan zárt gátvonalat jelentene, amelyen víz nem folyhatott át, tehát létezne el nem árasztott elemi felület.

Most jelöljünk ki egy csúcspontot, legyen ez

A
. Minden más csúcspontba állítsunk egy-egy őrt! Induljanak el az őrök az egyetlen lehetséges úton

A
felé, de álljanak meg, mielőtt újabb csúcspontba érnének. Így minden érintetlen gáton pontosan egy őr fog állni. Ugyanis két őr ugyanarra a gátra csak szembe haladva juthatna, így útjukat

A
-ig folytatva, e két út együttvéve körsétát tenne lehetővé. Ha pedig egy gátra egyetlen őr sem jutna, akkor ez a gát és a végpontjairól

A
felé induló utak együttvéve ismét körsétát tennének lehetővé.

Az érintetlen gátak száma tehát egyenlő az őrök számával. Az őrök száma pedig

$c–1$
, mert

A
kivételével minden csúcspontban állt egy őr. A megnyitott és az érintetlen gátak együttes száma okoskodásunk szerint

$é=(l–1)+(c–1)$,
ez pedig az Euler-féle összefüggést bizonyítja.

3. 3. A PÁRONKÉNT SZOMSZÉDOS ORSZÁGOK MAXIMÁLIS SZÁMA

A 2. ábra szerint lehetséges a gömbön olyan normál térkép, amelynek van négy, páronként szomszédos országa. Olyan normál térkép viszont nem lehetséges, amelyen négynél több, páronként szomszédos ország van.

Ennek az állításnak az igazolásához tételezzük fel az ellenkezőjét; azt, hogy a gömb egy normál térképén az $L1$, $L2$, $L3$, $L4$ és $L5$ országok páronként szomszédosak. Jelöljünk ki mindegyikük belsejében egy-egy pontot! Ezek a $P1$, $P2$, $P3$, $P4$ és $P5$ pontok. Kössük össze mindegyiket mindegyikkel egy-egy, a gömbön haladó vonallal a következőképpen: két pontot összekötő vonal csupán abban a két országban haladjon, amelyek belsejében kijelölt pontokat köt össze, és az öt ország mindegyikében a befutó négy vonalnak csak a kijelölt Pi pont legyen közös pontja. Ekkor a Pi pontokat összekötő vonalak nem metszik egymást. (Az 5. ábrán négy országhoz végrehajtottuk az utasítást. A Pi pontokat összekötő vonalak szaggatottak.) A $P1$, $P2$ és $P3$ pontokat egymás közt összekötő vonalak a gömböt két elemi felületre bontják. A $P4$

-et

$P5$
-tel összekötő vonal a többit nem metszi, tehát

$P4$
és

$P5$
a két elemi felület egyikében van. Legyen ez

E
. Vegyük fel

E
-ben a

$P4$
pontot, és húzzuk meg a

$P1$
,

$P2$
és

$P3$
pontokat

$P4$
-gyel összekötő vonalakat! Ezek

E
-t három elemi felületre – az

$E1$
-re,

$E2$
-re és

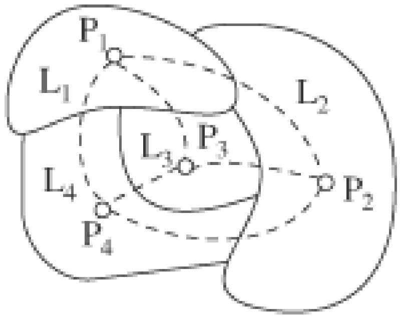
$E3$
-ra – bontják, és

$P5$
valamelyikben – mondjuk

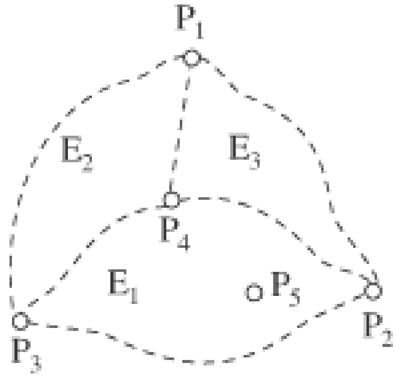
$E1$
-ben – benne van. (6. ábra). De ekkor nem létezhet a

$P5$
-öt

$P1$
-gyel összekötő olyan vonal, amely a többi vonal valamelyikét ne metszené.



5. ábra

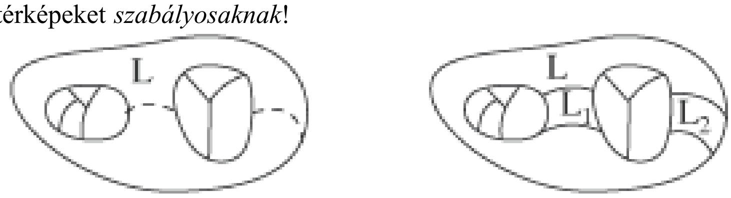


6. ábra

Ezzel ellentmondásra jutottunk, tehát kiinduló feltételezésünk nem lehet igaz, azaz *nem lehet a gömbre olyan normál térképet rajzolni, amelyen van öt, páronként szomszédos ország*. Ezt akartuk bizonyítani.

4. 4. AZ ÖTSZÍN PROBLÉMA

Most bebizonyítjuk a már említett ötszín tételt, amely szerint: *Öt színnel már bármilyen normál térkép jól színezhető a gömbön*. Megmutatjuk, hogy ezt az állítást elegendő olyan normál térképekre bizonyítani, amelyeken minden ország elemi felület, és minden csúcspontból pontosan háromfelé indul határvonal. Nevezzük el az ilyen térképeket *szabályosaknak*!



a. b.

7. ábra

I. Legyen normál térképünkön

L

egy nem elemi felületország, mint pl. a 7/a. ábrán. Az

L

belsejében levő országcsoportokat kössük össze az

L

-et körülvevő határvonalhoz vezető útszakaszokkal (szaggatott vonalak). Ezeket a szakaszokat bővítsük ki egy-egy

L

-ben fekvő országgá (7/b. ábra). Így

L

-ből elemi felület lesz. A felvett új

*L*1

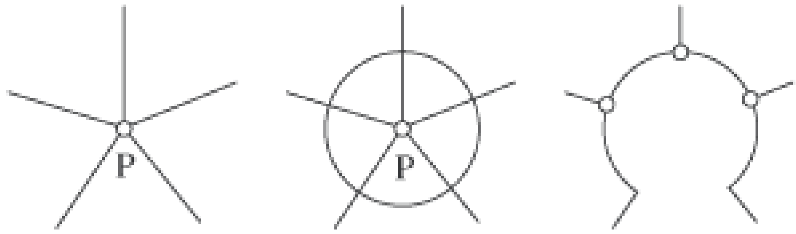
és

*L*2

ország a többi ország között eredetileg fennálló szomszédosságokat nem szüntette meg, ha tehát az így módosított térkép 5 színnel jól színezhető, akkor az eredeti is.

II. Ha van a normál térképen olyan

P
csúcspont, amelyből háromnál többfelé indul határvonal (8/a. ábra),



a. b. c.

8. ábra

akkor rajzoljunk köréje olyan kört, amelybe a

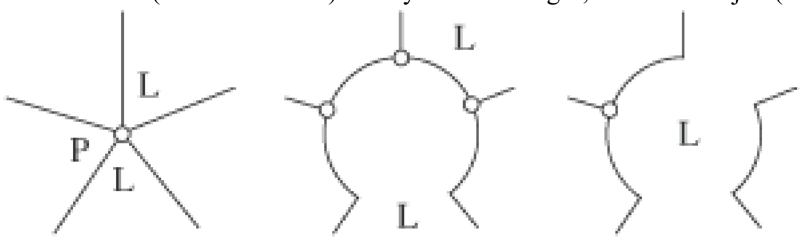
P
csúcsponton kívül más csúcspont nem esik, és amely a

P
-ből induló határvonalakat csak egy-egy pontban metszi (8/b. ábra). E kör belsejét csatoljuk az egyik

P
körüli országhoz (8/c. ábra). Ezzel szomszédosságot nem szüntettünk meg, esetleg létrehoztunk. Ha tehát az így módosított térkép 5 színnel jól színezhető, akkor az eredeti is. Ez után a módosítás után előfordulhat, hogy egy új él egy országon belül halad; ha pl. egy

L
ország a

P
pont körül többször is felbukkan (9/a. és 9/b. ábra). Az ilyen él fölösleges, és ezért töröljük (9/c. ábra).



a. b. c.

9. ábra

Az **I.** és **II.** lépések ismételt alkalmazásával minden normál térkép szabályossá alakítható, és – amint láttuk – az eredeti normál térkép jó színezéséhez sem kell több szín, mint az **I.** és **II.** lépések ismételt alkalmazásával nyert szabályos térkép jó színezéséhez.

Megemlítjük, hogy ha a gömb egy szabályos térképén kettőnél több ország van, akkor közöttük nincs olyan, amelynek határvonalán csúcspont ne volna. Ekkor tehát minden határvonal élekből összerakható. Az is igaz, hogy szabályos térképen nincs csupán egy éllel határolt ország. Mindkét állítás belátását az olvasóra bízzuk.

Ezek után azt fogjuk bizonyítani, hogy egy tetszőleges, a gömböt befedő T szabályos térkép 5 színnel jól színezhető. Jelöljük T országainak számát l-lel, éleinek számát $é$-vel és csúcspontjainak számát c-vel!

T minden országa elemi felület, tehát Euler tétele szerint $l+c=é+2$.

Számoljuk össze a csúcspontok segítségével az éleket! T minden csúcspontjához 3 különböző él illeszkedik. Minden él pontosan két csúcsponthoz illeszkedik, ti. a két végpontjához. Ezért igaz, hogy $3c=2é$.

Ha az Euler-egyenlőség mindkét oldalát 6-tal szorozzuk, és $6c$ helyébe $4é$-t írunk, akkor a következőre jutunk:

$$6l=2é+12. \qquad (\square)$$

Ebből viszont az következik, hogy T-nek van 6-nál kevesebb éllel határolt országa. Tegyük fel ugyanis az ellenkezőt! Számozzuk meg az országokat az 1, 2, …, l számokkal, és jelöljük h_i-vel az i-edik országot határoló élek számát! Minthogy minden él két különböző országnak közös határszakasza,

$h_1+h_2+\ldots+h_l=2é$

. Feltevésünk szerint minden

i
-re

$h_i \geq 6$
, tehát a bal oldalt nem növeljük, ha minden

h_i
helyett 6-ot írunk. Ekkor azt kapjuk, hogy

$6l \leq 2é$
. Ez pedig lehetetlen, mert

(□)
szerint

$2é$
12-vel kevesebb, mint

$6l$
. Tehát

T
tartalmaz 2, 3, 4 vagy 5 éllel határolt országot.

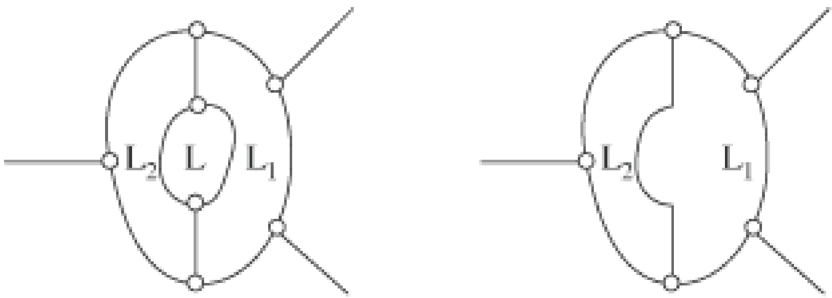
Ha

T
tartalmaz egy kétélű

L
országot, akkor azzal pontosan két másik ország szomszédos (10/a. ábra). Csatoljuk

L
-et valamelyik szomszédjához, pl.

L_1
-hez (10/b. ábra)!



a. *b.*

10. ábra

Így

T
-ből egy szabályos

T'
térképet állítottunk elő. Elég kimutatni, hogy

T'
5 színnel jól színezhető, Ebből ugyanis következik, hogy

T
5 színnel jól színezhető, mert

T'
színezésekor

*L*1
-hez és

*L*2
-höz az 5 színből kettőt használunk fel. A többi 3 szín bármelyikével színezhettük az újból visszaállított

L
országot, és máris jól színeztük

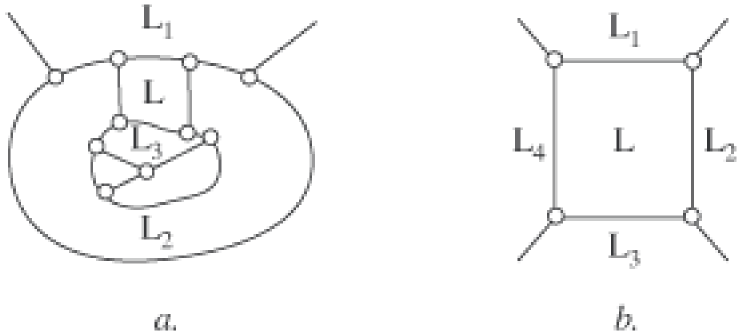
T
-t 5 színnel.

Ha

T
tartalmaz egy háromélű

L
országot, akkor azzal pontosan 3 ország szomszédos. Az előbbi utat követve,

L
visszaállításakor még mindig 2 szín közül választhatunk.



11. ábra

Ha

T
tartalmaz egy négyélű

L
országot, akkor azzal 3 vagy 4 ország lehet szomszédos, amint azt a 11/a. és 11/b. ábra mutatja. Az előbbi utat most is követhetjük, de ahhoz, hogy

T'

ismét szabályos legyen, a 11/a. ábra esetén

L
-et nem szabad

L_2
-höz csatolnunk. A visszaállításkor

L
számára még mindig marad 2, ill. 1 szín.

Ha

T
tartalmaz egy ötélű

L
országot, akkor a vele szomszédos országok között biztosan van kettő, amelyek nem szomszédosak. Mert ha pl.

L
szomszédai

L_1
,

L_2
,

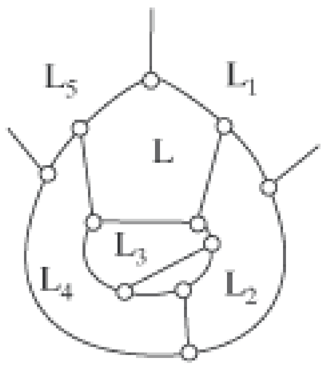
L_3
,

L_4
,

L_5
, és

L_2
szomszédos

L_4
-gyel, akkor a 12. ábrából



12. ábra

látható, hogy

$L3$
nem lehet szomszédos sem

$L1$
-gyel, sem

$L5$
-tel.

T'
-t most

L
két élének törlésével állítjuk elő:

$L1$
és

$L3$
felé törlünk egy-egy élt. Ha

T'
jól színezhető 5 színnel, akkor

T
is, mert

T'
színezésekor

L
szomszédaihoz mindössze 4 színt használunk fel (

$L1$
és

$L3$
ugyanazt a színt kapja), és így

L
számára a visszaállításkor marad még egy szín.

Okoskodásunk szerint elegendő azt bizonyítanunk, hogy a felsorolt redukciókkal nyert

T'
jól színezhető 5 színnel. De

T'
ismét szabályos térkép, tehát érvényes rá a

(□)
összefüggés, amiből viszont következik, hogy

T'
szintén tartalmaz 2, 3, 4 vagy 5 élű országot. Ennélfogva redukciós eljárásunkat addig folytathatjuk, amíg csupán 5 ország nem marad. Azt 5 színnel tudjuk jól színezni, és így eljárásunkból következik, hogy

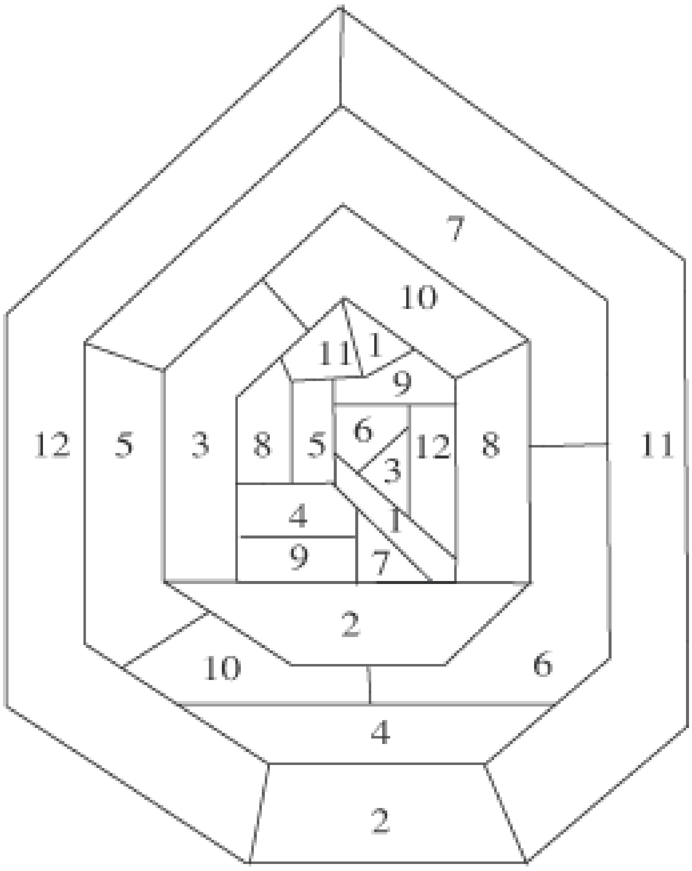
T
is jól színezhető 5 színnel. Az **I.** és **II.** figyelembevételével pedig azt nyertük, hogy bármilyen normál térkép 5 színnel jól színezhető.

5. 5. TETSZŐLEGES TÉRKÉPEK SZÍNEZÉSE

A térképek színezésének problémáját a normál térképek színezésére korlátoztuk. Amint láttuk, ez lényeges megszorítást jelent. A minimális színszükséglet függ attól, hogy az országok különálló részekből állnak-e vagy sem. Felmerülhet a kérdés: vajon akkor is akármilyen sok szín szükséges a térképek jó színezéséhez, ha az országok nem akármilyen sok különálló részből állnak?

A 13. ábrán olyan térképet láthatunk, amelyen 12 ország van, minden ország két különálló részből áll (az azonos országokhoz tartozó részeket azonos számmal láttuk el), és az országok mind páronként szomszédosak. Ennek a térképnek a jó színezéséhez nyilván 12 szín szükséges. Másrészt sikerült bebizonyítani, hogy *12 színnel minden olyan térkép jól színezhető, amelyen minden ország legfeljebb két különálló részből áll*.

Érdekes, hogy e bonyolultabb feladatot sikerült teljesen megoldani. Megoldatlan még az a kérdés, amelynél megengedünk a térképen kettőnél több különálló részből álló országot is.



13. ábra

6. 6. KÉT GÖMB SZÍNEZÉSI PROBLÉMÁJA

A technika mai fejlettségi fokán elképzelhető, hogy Földünk lakói eljutnak más bolygóra is, pl. a Marsra. Ha a Mars felületét részekre bontják, és ezeket a részeket szétosztják a Föld országai között, előfordulhat, hogy a Földön nem szomszédos országok a Marson szomszédosak lesznek. A Föld–Mars térkép elkészítésénél is kívánatos, hogy ugyanahhoz az országhoz tartozó részek azonos színűek legyenek. Ha az országok mindkét gömbön egy-egy összefüggő részből állnak, akkor hány szín elegendő két gömb térképeinek jó színezéséhez? A 14. ábrán 8 ország látható két gömbfelületen. Ezek az országok páronként szomszédosak, tehát a minimális színszám legalább 8. Azt azonban nem tudjuk, hogy 8 szín minden esetben elegendő-e.

14. ábra

Az **I.** és **II.** alatti okoskodásunkból látható, hogy a szabályos térképek

p

színnel való jó színezhetőségéből következik bármilyen normál térkép

p

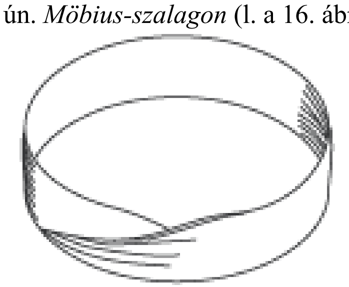
színnel való jó színezhetősége. Ezért a normál térképek vizsgálata helyett a következőkben csak szabályos térképeket vizsgálunk.

7. 7. A MÖBIUS-SZALAG ÉS A TÓRUSZ SZÍNEZÉSE

Nézzük meg a 15. ábrán látható szalagra rajzolt térképet!

15. ábra

Számokkal megkülönböztetett hat országot láthatunk. Az 1-gyel, 2-vel és 3-mal jelölt országok két-két különálló részből állnak, és mind a hat ország páronként szomszédos. Térképünk nem szabályos, mert országai nem összefüggők. Ha viszont térképünk két oldalát a szalag egy csavarása után összeragasztjuk, akkor a különálló országrészek egyesülnek. Az így kapott ún. *Möbius-szalagon* (l. a 16. ábrán!)



16. ábra

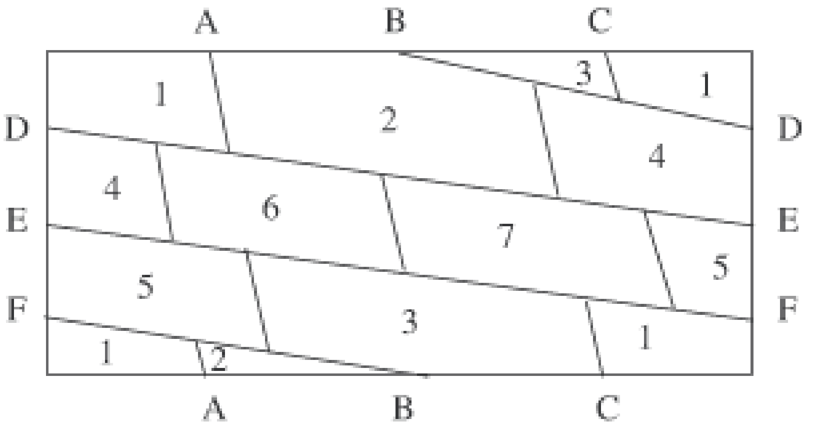
már mind a hat ország elemi felület, és mind páronként szomszédosak. Nincs ez ellentmondásban azzal, amit a 3. pontban bizonyítottunk? Az csak olyan normál térképekre vonatkozik, amelyek szakítás nélküli nyújtással gömbre teríthetők. Ezt nem tehetjük meg a Möbius-szalaggal.

Általában, ha két felület a megengedett nyújtással egymásra teríthető, akkor azt mondjuk, hogy a két felület *homeomorf.* (Itt mindig csak véges kiterjedésű felületekre gondolunk.) Tehát a gömb (ill. a sík) és a Möbius-szalag nem homeomorfak. Minthogy a Möbius-szalagra lehet olyan szabályos térképet rajzolni, amelynek hat országa páronként szomszédos, szabályos térképeinek jó színezéséhez legalább 6 színre van szükség. Több szín használata már fölösleges, ugyanis bizonyított tény a következő:

A Möbius-szalagra rajzolt bármely szabályos térkép 6 színnel jól színezhető.

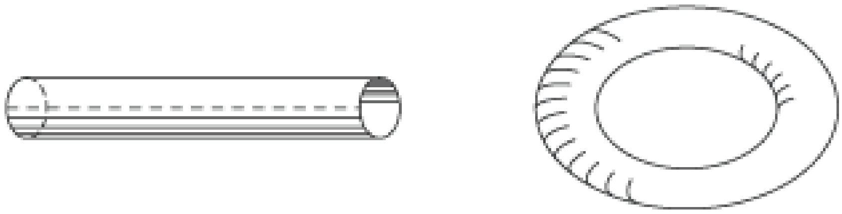
Az eredmény meglepő, hiszen az „egyszerűbb” síkfelület szabályos térképeinek jó színezéséhez szükséges minimális színszámot csak később sikerült pontosan meghatározni.

A Möbius-szalagnak több érdekes tulajdonsága van15, itt csak egyre hívjuk fel a figyelmet: E felületnek nincs két különböző oldala, azaz bármely pontjából kiindulva be tudjuk járni egész felületét anélkül, hogy a szalag szélét átlépnénk. Talán ez okozta, hogy hat, páronként szomszédos országot tartalmazó szabályos térképet is rajzolhattunk rá, és a szabályos térképeinek jó színezéséhez szükséges minimális színszámot meg lehetett határozni? Látni fogjuk, hogy nem. Vegyük ugyanis szemügyre a 17. ábrán



17. ábra

látható szalagra rajzolt térképet! Ezen hét országot találunk, amelyek mind páronként szomszédosak. Ha a szalagot gumiból készítjük el, összeilleszthetjük olyan felületté, mint amilyen a mentőöv felülete: Ragasszuk össze először a szalag vízszintes éleit, majd az így kapott cső két végét illesszük egymáshoz (l. a 18. ábrát)!



18. ábra

A ragasztásnál ügyeljünk arra, hogy az azonos betűvel jelölt pontok egybeessenek! Az így kapott ún. *tórusz* felületén (gyűrűfelületen) az előre kijelölt hét ország mindegyike elemi felület, és páronként mind szomszédosak. (E térkép szabályos a tóruszon.) Ahhoz tehát, hogy a tóruszon bármilyen szabályos térképet jól színezhessünk, legalább 7 színre van szükségünk. Ismét meglepő, hogy a gömbnél „bonyolultabb” tóruszra sikerült bizonyítani a következőt:

A tóruszra rajzolt bármilyen szabályos térkép 7 színnel jól színezhető.

A tórusznak, ugyanúgy, mint a gömbnek, két oldala van (külső és belső). A gömb, a Möbius-szalag és a tórusz

15Lásd még e kötet Csalafinta felületek című cikkét!

közül semelyik kettő sem homeomorf.

8. 8. TETSZŐLEGES FELÜLETRE RAJZOLT TÉRKÉPEK SZÍNEZÉSE

Jelentse egy

F
felülethez

$lmax$
azt a legnagyobb számot, ahány páronként szomszédos, összefüggő ország

F
-re rajzolható, és

$smin$
azt a legkisebb számot, ahány színnel már bármilyen,

F
-re rajzolható normál térkép jól színezhető. Természetesen minden felületre

$lmax \leq smin$.

Eddig azt láttuk, hogy gömbre (vagy síkra)

$4 = lmax \leq smin \leq 5$,
Möbius-szalagra

$lmax = smin = 6$
és tóruszra

$lmax = smin = 7$.
Felmerülhetnek bennünk a következő kérdések: Vannak-e még másfajta felületek is? Ha igen, mit tudunk a hozzájuk rendelt

$lmax$
és

$smin$
számokról? Bármilyen nagy

k
számhoz található-e olyan felület, amelyhez tartozó

$lmax$
nagyobb, mint

k
? Mindezekre a kérdésekre pontos választ tudunk adni a következőkben:

lmax
is van.

Minden felületre

lmax=smin
.

Az

lmax
és

smin
számokat minden egyoldalú felületre sikerült pontosan meghatározni, kétoldalú felületekre viszont sok esetben csak olyan közelítő értéket sikerült megadni, amely a pontos értéktől nem tér el többet két egységnél.

9. 9. A NÉGYSZÍN TÉTEL ÁTFOGALMAZÁSAI

A térképszínezési problémák közül máig is a legtöbbet emlegetett probléma a gömbön a kérdéses négyszín tétel. Több olyan tételt ismerünk, amelyet eddig szintén nem sikerült bizonyítani, de kimutatható, hogy megoldásuk bizonyítaná a gömb tetszőleges normál térképének jól színezhetőségét 4 színnel. Három ilyen tételt megemlítünk:

a.) A sík mindazon

T
térképei jól színezhetők 4 színnel, amelyeket a következőképpen lehet előállítani: Egy

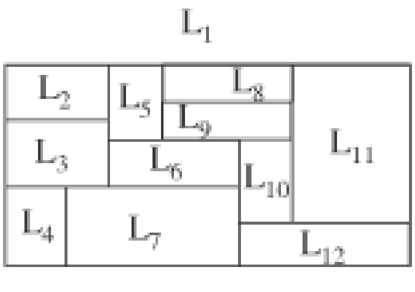
T□
téglalapot feldarabolunk téglalapokra.

T
országai a

T□
-on belüli téglalapok és a

T□
-ot körülvevő rész. Egy ilyen

T
térkép látható a 19. ábrán.



19. ábra

b.) A gömbre rajzolt bármely szabályos térkép élei színezhetők 3 *színnel úgy, hogy egy él egy színt kapjon, és minden csúcspontban különböző színű élek találkozzanak.*

c.) A gömbre rajzolt bármely szabályos térkép csúcspontjai a

+1
és

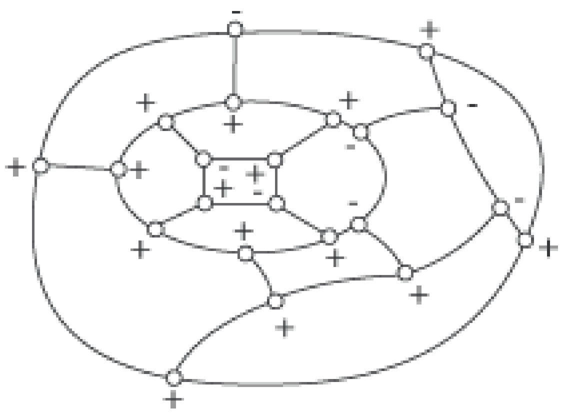
–1
számokkal számozhatók úgy, hogy minden ország határvonalán a számok összege 3*-mal osztható legyen.* Ilyen számozás látható a 20. ábrán.

+1
, ill.

–1
helyett röviden a

+
, ill.

–
jeleket írtuk.



20. ábra

10. 10. FELADATOK

E témakör sok szórakoztató feladata közül megemlítünk néhányat. A feladatokat úgy válogattuk, hogy megoldásaikkal ezen a területen végzett kutatások néhány módszerét is megmutathassuk.

(1) Rajzoljunk a síkra tetszés szerint egyeneseket! Ezzel a síkot összefüggő „országokra” bontottuk. Bizonyítsuk be, hogy az így kapott térkép 2 színnel jól színezhető!

(2) A sík egy

T
térképén minden országot 3 él határol.

T
csúcspontjai megszámozhatók az 1, 2 és 3 számokkal úgy, hogy minden él különböző számokkal jelölt

csúcspontokat köt össze. Mutassuk meg, hogy ekkor

T
jól színezhető 2 színnel!

(3) Bizonyítsuk be, hogy ha egy, a gömböt teljesen lefedő térkép 2 színnel jól színezhető, akkor minden csúcsponthoz páros számú él illeszkedik!

(4) Rajzoljunk a síkra köröket, és minden körben húzzunk meg egy-egy húrt! Ezzel a síkot összefüggő országokra bontottuk. Igazoljuk, hogy ha bármely két húrnak legfeljebb egy közös pontja van, akkor térképünk 3 színnel jól színezhető!

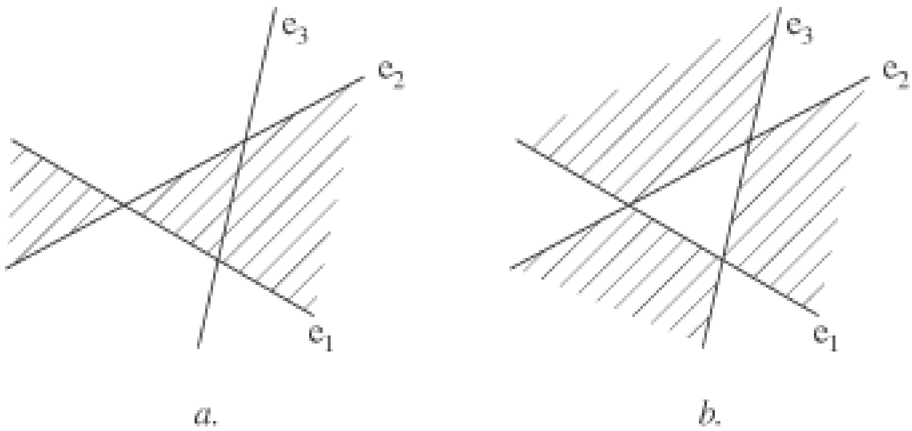
11. 11. MEGOLDÁSOK

(1) Fektessük le egyenként az egyeneseket a síkra! Az első,

$e1$
-gyel jelölt egyenes a síkot két félsíkra osztja. Színezzük az egyiket fehérrel, a másikat feketével! A második egyenes,

$e2$
a síkot szintén két félsíkra osztja. Ezek egyikén cseréljük fel a színeket: ami fehér volt, legyen fekete és viszont. Fektessük le a harmadik egyenest,

$e3$
-at (21/a. ábra)!



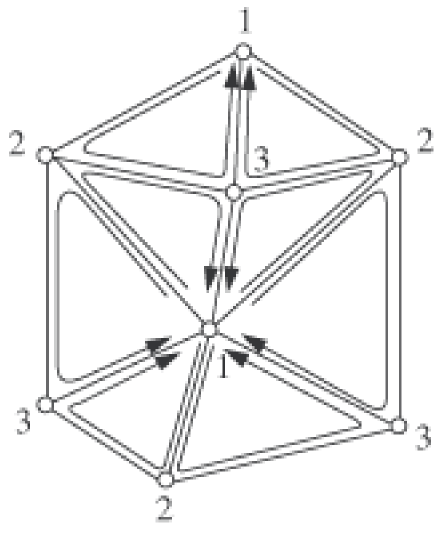
21. ábra

Az

$e3$
által kettéosztott sík egyik felén ismét cseréljük fel a színeket (21/b. ábra)! Ezt az eljárást akármeddig folytathatjuk. Minden lépés után térképünk 2 színnel – fehérrel és feketével – jól színezett lesz.

(2) Tekintsük

T
csúcspontjainak egy említett számozását (22. ábra)!



22. ábra

A csúcspontokhoz rendelt számok sorrendjének megfelelően minden ország körüljárható vagy az óramutató járásával megegyező, vagy azzal ellentétes irányban. Az egyik irányban körüljárt országokat fehérre, a másik irányban körüljártakat feketére színezzük. Ekkor a szomszédos országok mindig különböző színt kapnak, hiszen egy él két oldalára eső országok bejárási iránya ellentétes.

(3) Ha a 2 színnel jól színezett térkép bármely

P
csúcspontját körüljárjuk, akkor minden élen áthaladva színt váltunk, és így a

P
körüli országok közül minden második ugyanolyan színű. Ebből következik, hogy

P
-hez páros számú élnek kell illeszkednie.

(4) Egy konfiguráció (a körök közül egy és a belerajzolt húr) három részre osztja a síkot. Számozzuk e három részt a 0, 1, és 2 számokkal! Minden országba írjuk azt a számot, amelyikkel megjelölt részbe esik a három közül. Ezt a számozást minden konfigurációra hajtsuk végre! Minden lépésnél egy-egy újabb szám kerül minden országba. Az egy országon belüli számokat adjuk össze, osszuk el az összeget 3-mal, és írjuk az országba a maradékot! Ezután aszerint színezünk egy országot fehérre, feketére vagy pirosra, hogy a beleírt maradék 0, 1 vagy 2. Be kell látnunk, hogy ekkor két, tetszés szerinti szomszédos ország különböző színű (azaz különbözők a maradékaik). Legyen két szomszédos ország

$L1$
és

$L2$
, határaik közös éle pedig

h
. Vegyük el a térképről a

h
élt tartalmazó

K

konfigurációt. Ekkor

$L1$
és

$L2$
között nincs határ. Tehát

$L1$
és

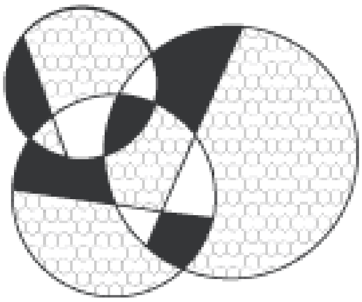
$L2$
a

K
-tól különböző konfigurációk mindegyikének ugyanabba a részébe esik, a

K
által felosztott 3 rész közül pedig különbözőbe. Ezért

$L1$
és

$L2$
maradékai nem egyezhetnek meg. Így tehát térképünk 3 színnel – fehérrel, feketével és pirossal – jól színezett.



23. ábra

A SZERENCSEJÁTÉKOK ÉS A VALÓSZÍNŰSÉGSZÁMÍTÁS

RÉNYI ALFRÉD

1. 1. BEVEZETÉS

E cikknek az a célja, hogy a szerencsejátékokra – különösen a kártyajátékokra – vonatkozó közérthető és érdekes feladatokon keresztül a valószínűségszámítás bizonyos fogalmait és módszereit ismertesse.[16] Bár nem törekedtem a példaként szereplő szerencsejátékok szabályainak teljes részletességgel való ismertetésére; azonban igyekeztem a példákat úgy megfogalmazni, hogy az is megértse őket, aki az illető játékot nem ismeri alaposan. Természetesen azért annak, aki bridzsel, a bridzsre vonatkozó példák többet mondanak, mint annak, aki nem ismeri a játékot stb.

Felmerülhet az olvasóban a kérdés: érdemes-e a kártyajátékokkal és általában a szerencsejátékokkal tudományos alapon foglalkozni? Erre a kérdésre véleményem szerint feltétlenül igennel kell válaszolni. Nemcsak azért érdemes, mert ez segít a kombinatorika és a valószínűségszámítás megértéséhez, hanem azért is, mert e problémák tudománytörténeti érdekességgel bírnak, hiszen a valószínűségszámítás kialakulásában a szerencsejátékokra vonatkozó feladatok közismerten igen nagy szerepet játszottak. Jó példa erre a kártyakeverés fogalma, amely összefügg nemcsak a kémiai technológiában használatos keverési eljárásokkal, hanem a termodinamika alapvető fogalmaival is.

2. 2. A KÁRTYAKEVERÉSRŐL

Amikor a kártyák eloszlására vonatkozó kérdéseket a valószínűségszámítás alapján válaszoljuk meg, hallgatólagosan mindig feltesszük, hogy az a csomag kártya, amelyből a lapokat kiosztják, „jól össze van keverve”. E kifejezést a kártyások gyakran használják, de mivel nem szokták precízen definiálni, nem árt először röviden foglalkozni e kérdéssel.

Valószínűségszámítási szempontból egy csomag kártyát akkor nevezünk jól megkevertnek, ha a keverés után a lapok összes lehetséges sorrendje, permutációja ugyanakkora valószínűséggel bír;

n

lap esetében

$n!$

permutáció lehetséges, tehát jól összekevert az a kártyacsomag, amelynél feltehetjük, hogy minden egyes sorrendnek a valószínűsége

$(1/n!)$

. Egy tetszőleges, a kártyák sorrendjétől függő

A

esemény valószínűsége tehát

$(k/n!)$

[16]Az egyes szerencsejátékok valószínűségszámítási vizsgálatába nem mehetek itt bele részletesen, azonban az egyes játékokkal kapcsolatban utalok olyan könyvekre, ahol részletesebb tárgyalást találhat az érdeklődő olvasó.

, ahol k jelenti azoknak a sorrendeknek a számát, amelyek mellett az A esemény bekövetkezik.

Például, ha egy 52 lapos kártyát megkeverünk, annak a valószínűsége, hogy a legfelül fekvő lap ász legyen, $(1/13)$-dal egyenlő; ugyanis az 52! sorrend közül egy meghatározott ásszal kezdődő sorrendek száma 51! (hiszen ha pl. a legfelső lap a kőr ász, az alatta levő 51 lap 51! sorrendben helyezkedhet el), tehát mivel a csomagban 4 ász van, $k=4\cdot 51!$, és így a keresett valószínűség

$$(k/52!)=(4\cdot 51!/52!)=(4/52)=(1/13).$$

A valóságban a keverés úgy történik, hogy az egyik játékos (vagy a kártya keverését végző gép) 10–20-szor végez el egy bizonyos mozdulatot. Minden egyes mozdulat a kártyacsomag egy átrendezését, vagyis egy permutáció alkalmazását jelenti a kártyák sorrendjére. Egy számsorozat permutációi tudvalevőleg csoportot alkotnak.

Két permutáció, pl. P és Q szorzatán (amit PQ-val jelölünk) azt értjük, hogy először elvégezzük a P átrendezést, és az eredményül kapott sorrenden végrehajtjuk a Q átrendezést. Például, ha $n=32$, vagyis 32 lapos (magyar) kártyáról van szó, és P a

$$\begin{pmatrix} 1 & \dots & \dots & 16 & 17 & \dots & \dots & 32 \\ 17 & \dots & \dots & 32 & 1 & \dots & \dots & 16 \end{pmatrix}$$

permutáció, tehát első helyre az eredetileg 17-ik helyen levő lap kerül, második helyre az eredetileg 18-ik helyen levő lap, és így tovább, az utolsó helyre az eredetileg a 16-ik helyen álló lap (ez úgy hajtható végre, hogy a 32 lapos kártyacsomagról együtt leemelem a felső 16 lapot, ezt leteszem az asztalra, és erre ráteszem az alsó tizenhat lapot), és $Q=P$, vagyis másodszorra is ugyanazt a műveletet végzem, akkor

$PQ=P^2=I$

az azonos permutáció, tehát a két művelet végeredményeként újból az eredeti sorrend áll elő.

A keverés folyamatára mármost két kézenfekvő matematikai modellt (a tényleges folyamat leegyszerűsített képét) állíthatunk fel. Az első modellt determinisztikus modellnek, a másodikat sztochasztikus modellnek fogjuk nevezni. Először tegyük fel, hogy a keverő minden egyes mozdulatnál *pontosan* ugyanazt az átrendezést végzi a kártyacsomagon. Ha ezt a permutációt

P

-vel jelöljük, akkor e mozdulat

k

-szor való elvégzése egyenértékű az eredeti sorrend egyetlen átrendezésével, mégpedig a

P

permutáció

k

-adik hatványának megfelelő átrendezésével. Az ilyen keverési eljárás elvi szempontból nem kielégítő, hiszen (elvben) pontosan kiszámítható, hogy mi lesz a végeredmény, de gyakorlatilag sem megfelelő, mégpedig annál kevésbé, minél kisebb a

P

permutáció rendje, vagyis az a legkisebb

r

pozitív egész szám, amelyre

$P^r=I$

, ahol

I

az azonos permutációt jelenti, amelynél minden lap a helyén marad. (A fent példaként említett

P

permutációnak a rendje

$r=2$

.) Ugyanis, ha a

P

permutáció rendje

r

, akkor ez azt jelenti, hogy a

P

,

P^2

, …,

P^r

permutációk mind különbözők (

P

minden magasabb hatványa viszont már megegyezik ezek egyikével), tehát akárhányszor is ismétlünk egy r-ed rendű P permutációt, ezáltal elvileg nem tudunk létrehozni r-nél több különböző sorrendet. Persze, ha létezne olyan P permutáció, amelynek rendje $n!$, ennek ismétlésével minden sorrendet létrehozhatnánk; ilyen permutáció azonban nem létezik, ha $n \geq 3$; ugyanis az n-ed rendű szimmetrikus csoport nem ciklikus (sőt nem is Abel-csoport, míg a ciklikus csoportok azok). Tisztán matematikai szempontból igen érdekes kérdés a permutációk rend szerinti eloszlása; e kérdéssel foglalkozik *Erdős Pál* és *Turán Pál* egy nemrégiben megjelent dolgozata [1]; mivel azonban a keverésnél valóban soha nem pontosan ugyanazt a műveletet ismételjük meg (még akkor sem, ha gép végzi a keverést), a kérdéssel itt nem foglalkozunk részletesen.[17] A keverés másik – a valósághoz sokkal közelebb álló – modellje a következő:

Feltesszük, hogy az egyes keverő mozdulatok a véletlentől is függnek, és egy mozdulatnál bizonyos valószínűséggel minden permutáció felléphet. Feltesszük továbbá, hogy az egyes mozdulatok egymástól függetlenek. Pontosabban ez azt jelenti, hogy ha n elem $n!$

[17] Erdős és Turán [1]-ben bebizonyítják, hogy az $n!$ permutáció közül a legtöbbnek a rendje az $e^{(1/2-\varepsilon)\log^2 n}$ és $e^{(1/2+\varepsilon)\log^2 n}$ határok közé esik, ahol $\varepsilon > 0$ tetszőleges kicsiny számnak választható, ha n elég nagy (itt log n az n szám természetes logaritmusát jelöli). A kártyakeverésre vonatkoztatva ez azt jelenti, hogy 52 kártya legtöbb permutációja esetében kevesebb, mint 200 000-szer kell a keverő mozdulatot megismételni, hogy a kártyák újból az eredeti sorrendbe kerüljenek. Összehasonlításul vegyük figyelembe, hogy 52! egy 68 jegyű számóriás.

lehetséges permutációját valahogy megszámozzuk és

Q_j
jelöli azt a permutációt, amely a

j
-edik sorszámot kapja, akkor a keverés minden egyes mozdulatával a keverő a

Q_j
permutációt

q_j
valószínűséggel hajtja végre

$(j=1,2,\ldots,n!)$
. (Természetesen

$\sum_{j=1}^{n!} g_j=1$
.) Ha tehát a keverő

i
-edik mozdulatánál a

Π_i
permutációt hajtja végre, akkor

Π_i
véletlen permutáció, amelynek eloszlása

$P(\Pi_i=Q_j)=q_j \ (j=1,2,\ldots,n! \ i=1,2,\ldots),$
(azaz a

Π_i
véletlen permutáció

q_j
valószínűséggel lesz egyenlő egy előírt

Q_j
permutációval), és

Π_i
eloszlása független a

$\Pi_1,\ldots,\Pi_{i-1}$
permutációktól.

Ez esetben

k
mozdulat után (ha a lapok eredetileg az

$1,2,\ldots,n$
sorrendben voltak) a

$\Pi_1\Pi_2\ldots\Pi_k=\Pi^{(k)}$
permutáció jön létre. A

$\Pi(k)$
permutációk ún. Markov-láncot alkotnak. Mármost ismeretes, hogy ha a

$\{q_j\}$
eloszlás olyan, hogy az összes permutáció csoportjának tetszőleges valódi

G
részcsoportjára

(2.1)

$$\sum_{Q_j \in G} g_j < 1$$

(vagyis az eloszlás nincs egy valódi részcsoportra koncentrálva), akkor nagy

k
esetében

$\Pi(k)$
eloszlása közel egyenletes lesz, vagyis minden permutáció körülbelül ugyanolyan

(közel $(1/n!)$)
valószínűséggel fog létrejönni nagyszámú keverőművelet után; pontosabban ez esetben[18]

(2.2)

$$\lim_{k\to\infty} P(\Pi(k)=Q_j)=(1/n!) \quad (j=1,2,\ldots,n!)$$

[A (2.1) feltétel teljesül például, ha

j
minden egyes értékére

$q_j>0$
.]

Gyakorlatilag a mondottakból az következik, hogy ha a keverés minden egyes mozdulata véletlenszerű, elvben minden átrendezést létrehozhat, és ha elég nagyszámú keverő mozdulatot teszünk, akkor indokolt az a feltevés, hogy a lap „jól össze van keverve". Arra a kérdésre, hogy mi értendő „elég" nagyszámú mozdulaton, itt nem térünk ki.

Mindenesetre érdemes volt a keverés folyamatával ilyen részletesen foglalkoznunk, hiszen tudvalevőleg a hamiskártyások leggyakrabban alkalmazott fogása éppen a nem kielégítő keverés. (Lásd az 5-öt is!)

3. 3. A KÁRTYÁK ELOSZLÁSÁRA VONATKOZÓ FELADATOK

Vizsgáljnuk meg néhány egyszerű példát!

[18] E tétel speciális esete a valószínűségszámítás centrális határeloszlás-tétele topologikus csoportokra vonatkozó általánosításának, amely a Haar-mértékhez való konvergenciára ad feltételt. L. pl. [2] és [3].

A. *Póker.* A pókernél minden játékosnak 5 lapot osztanak ki az 52 lapból. A játéknál a következő figurákat különböztetik meg:

1. *Royal Flush*: ha az 5 lap ugyanabból a színből van, és nagyság szerint sorban következnek (az ász egyaránt tekinthető 1-esnek vagy a király után következőnek), pl. kőr 9, 10, Bubi, Dáma, Király.
2. *Póker*: ha az 5 lap között 4 egyforma van (pl. 4 király stb.), az ötödik emellett akármilyen lehet.
3. *Full Hand* (vagy *Full House*): ha az 5 lap közül 3 egyforma és a másik kettő is egymással megegyező (pl. három tízes és két király).
4. *Szín (Couleur)*: ha az 5 lap egyszínű (pl. mind az öt treff-lap).
5. *Sorozat* (szekvencia): ha az öt lap nagyság szerint sorban következik, de nem mind ugyanolyan színű: pl. treff 5, treff 6, pikk 7, kőr 8, káró 9.
6. *Hármas*: három megegyező lap (pl. három 7-es), a másik kettő tetszőleges (de nem megegyező, hiszen az Full Hand-et jelent).
7. *Két pár*: két-két megegyező lap (pl. két ász, két 6-os), az ötödik tetszőleges (de a két pártól különböző).
8. *Egy pár*: két megegyező lap (pl. két dáma), a másik 3 tetszőleges, egymástól és a pártól különböző lap.

E figurák valószínűségeit könnyen kiszámíthatjuk úgy, hogy összeszámoljuk, hogy egy figura hányféleképpen valósulhat meg, és ezt a számot elosztjuk 5 lapnak az 52 lapból való összes lehetséges kiválasztásának számával. Pl. a póker

$13 \cdot 48$
-szor valósulhat meg, míg 52 lapból 5 lapot

$\binom{52}{5}$
-féleképpen választhatunk, tehát a póker valószínűsége

$$\frac{13 \cdot 48}{\binom{52}{5}}=\frac{13 \cdot 48 \cdot 120}{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}=\frac{1}{4165}=0{,}000240.$$

Itt látszólag másként számoltunk, mint az előző pontban, mert a lapok sorrendjét nem vettük figyelembe. Ez az eredményt nem befolyásolja, hiszen öt lapot

$5!=120$
sorrendben kaphatunk meg, és ha ezt a faktort mind a számlálóból, mind a nevezőből elhagyjuk, a tört értéke nem változik. Az előző pontban említett eljárást követve a póker valószínűségét a következőképpen számíthatjuk ki. Ha pl. 4 játékos játszik, és körben osztanak, az első helyen ülő játékos kapja az első, az ötödik, a kilencedik, a tizenharmadik és tizenhetedik lapot. Azoknak a sorrendeknek a száma, amelyeknél ezen az 5 helyen 5 meghatározott lap áll megadott sorrendben, nyilván 47!, hiszen ennyiféleképpen lehet a többi 47 lapot elrendezni a többi 47 helyen. Tehát annak valószínűsége, hogy a játékosnak pókert osztanak

$$\frac{13 \cdot 48 \cdot 5!\,47!}{52!}=\frac{13 \cdot 48}{\binom{52}{5}}=\frac{1}{4165},$$
azaz így is ugyanazt az eredményt kapjuk, mint az előbb. Az összes figura valószínűségét az alábbi táblázat adja meg (6 tizedesjegy pontosságig):

Royal Flush	0,000 014

Póker	0,000 240
Full Hand	0,001 385
Szín	0,001 967
Sorozat	0,003 532
Hármas	0,021 055
Két pár	0,047 373
Egy pár	0,422 570
Semmilyen figura	0,501 864
Összesen:	1,000 000.

A pókernél tehát annál értékesebb egy figura, minél kisebb a valószínűsége. (Megváltozhat azonban a valószínűségi sorrend, ha nem 52 lappal játszunk, vagy ha a csomagban egy vagy több joker is van.)

A póker szabályai szerint minden játékosnak, miután lapját megnézte (és befizette a bankba a megfelelő összeget), joga van lapjai közül egyeseket eldobni és helyettük újakat kérni. Míg tehát az első leosztásnál

$1/2$

-nél valamivel nagyobb a valószínűsége annak, hogy a játékos nem kap semmilyen figurát; ha ilyen esetben 5 új lapot kér, annak valószínűsége, hogy másodszorra sem kap figurát, megint kereken

$1/2$

. De mivel a két esemény majdnem független, annak valószínűsége, hogy a második osztás után sincs egy játékosnak semmilyen figurája, már csak kb.

$1/4$

. Ha tehát 3 másik játékossal játszom pókert, akkor annak valószínűsége, hogy ezek közül egyiknek sincs még egy párja sem, körülbelül[19]

$1/64$

; ez azt jelenti, hogy ha csak 1 párom van, szinte biztosra vehetem, hogy a másik három játékos közül legalább az egyiknek szintén van egy párja vagy annál jobb lapja.

A lapok felmutatása esetén tehát 1 párral igen csekély a nyerés valószínűsége; más kérdés persze, hogy a pókerben blöffölni is lehet.

Utalunk itt *Jordan K.* könyvére [4], amely számos további, a pókerjátékkal kapcsolatos valószínűségszámítási feladatot tárgyal.

B. *Bridzs.*[20] A bridzsnél 52 lapot osztanak szét 4 játékos között, akik közül a két-két szemben ülő koalícióban van. A játék két részből áll: a licitálásból és a tényleges lejátszásból. A licitálásra nézve számos rendszer ismeretes. A Culbertson-rendszer szerint (lásd [5]) a játékosok először értékelik a saját lapjukat és ennek

[19] Ezek az események nem teljesen függetlenek, de közel azok; így nem követünk el nagy hibát, ha összeszorozzuk a valószínűségeket.

[20] A bridzs tulajdonképpen nem szerencsejáték a szó szoros értelmében, hiszen a játékban a játékosok tudása sokkal nagyobb szerepet játszik, mint a véletlen. A lapok eloszlása azonban itt is a véletlentől függ, és így a briddzsel kapcsolatban is számos valószínűségszámítási feladat merül fel.

alapján döntik el, hogy mit licitáljanak. Az értékelés az ún. „trick”-ek összeszámlálásából áll, a következő szabályok szerint.

Ász (ugyanolyan színű király és dáma nélkül):	1	trick
Ász és király (egy színből, az uo. színű dáma nélkül):	2	trick
Ász és dáma (egy színből, az uo. színű király nélkül):	1,5	trick
Ász, király és dáma egy színből:	2,5	trick
Király (ugyanolyan színű ász és dáma nélkül):	0,5	trick
Király és dáma egy színből (uo. színű ász nélkül):	1	trick

A teljes lap értékét a benne levő trickek összege adja meg. Pl. a következő lap:

PIKK:	KŐR:	KÁRÓ:	TREFF:
ÁSZ, DÁMA,	KIRÁLY,	DÁMA, 7	8
10, 8, 7	DÁMA, 4, 3, 2		

értéke

$1{,}5+1=2{,}5$
trick.

A kiosztott 13 lap értéke a véletlentől függő szám, tehát valószínűségi változó.

Vizsgáljuk most meg, hogy mi a *várható értéke* egy játékosnak kiosztott lapoknak.[21]

Ehhez tulajdonképpen ki kellene számítanunk a lap értékének eloszlását, tehát, hogy mekkora valószínűséggel lesz az egy játékos kezében lévő 13 lap összértéke egyenlő a 0; 0,5; 1; 1,5; 2; 2,5; 3; 3,5; 4; 4,5; 5; 5,5; 6; 6,5; 7; 7,5; 8; 8,5; 9; 9,5; 10 számok mindegyikével. (Könnyen belátható, hogy ezek a lehetséges trick-számok.) Bár ez az út sem túl nehéz, itt egy egyszerűbb eljárást fogunk követni, amely a várható érték egy jól ismert tulajdonságán alapszik, mégpedig azon, hogy valószínűségi változók összegének várható értéke egyenlő a tagok várható értékeinek összegével. Jelöljék a négy játékos lapjainak teljes értékét rendre

[21]Egy valószínűségi változó várható értékét úgy számítjuk ki, hogy veszszük a változó lehetséges értékeinek a megfelelő valószínűségekkel mint súlyokkal képezett súlyozott középértékét; ha tehát a

ξ
valószínűségi változó az

$x_1, x_2, \ldots x_n$
értékeket rendre

$p_1, p_2, \ldots p_n$
valószínűséggel veszi fel, akkor várható értéke

$M(\xi)=p_1x_1+p_2x_2+\ldots+p_nx_n$.
A várható érték jelentőségét a nagy számok törvénye mutatja, amely szerint, ha a valószínűségi változó értékét nagy számú független kísérletnél megfigyeljük, akkor a megfigyelt értékek számtani közepe szinte bizonyosan igen közel lesz a valószínűségi változó várható értékéhez.

□1, □2, □3, □4. Nyilvánvaló, hogy (jól kevert lap esetében) mind a négy játékos lapjának várható értéke ugyanakkora. Mármost az első játékos lapjának trickértéke maga is négy tag összege:

□1=□11+□12+□13+□14,

ahol □11, □12, □13, □14 azt jelentik, hogy hány trickje van az első játékosnak rendre pikkből, kőrből, káróból és treffből.

Hasonlóképpen bontható fel 4 tag összegére a másik 3 játékos lapjának teljes értéke is:

□2 =□21+□22+□23+□24, □3 =□31+□32+□33+□34, □4 =□41+□42+□43+□44.

Mármost nyilvánvaló, hogy az □*ij* (*i*=1,2,3,4; *j*=1,2,3,4) valószínűségi változók várható értékei mind egyenlők, tehát ha közös értéküket *m*-mel jelöljük, azaz

M(□*ij*)=*m* (*i*,*j*=1,2,3,4),

akkor a várható érték additivitása folytán

M(□1)=4*m*=*M*(□11+□21+□31+□41).

Mármost

□11+□21+□31+□41

a pikk színből a négy játékosnál lévő trickek összege; vagyis azt láttuk be, hogy egy játékos kezében lévő lap várható értéke egyenlő a pikk színből a négy játékos kezében lévő trickek összegének várható értékével.

Ez utóbbi mennyiség kiszámításához csak azt kell megvizsgálnunk, hogy a pikk ász, király és dáma hogyan oszlik el a 4 játékos között. Ha e 3 lap mindegyike más játékosnál van, akkor

□11+□21+□31+□41=1,5
. Ha a pikk ász és a pikk király egy kézben van, de a pikk dáma másnál; vagy ha a pikk ász és a pikk dáma egy kézben van, de a pikk király másnál van; végül ha a pikk király és a pikk dáma egy kézben van, a pikk ász másnál, akkor

□11+□21+□31+□41=2
. Végül, ha mindhárom lap egy kézben van, akkor

□11+□21+□31+□41=2,5
.

Annak valószínűségét, hogy a pikk ász, király és dáma más-más játékosnál van, a következőképpen számíthatjuk ki. Ha a pikk ászt az

A
játékos kapta, ő még további 12 lapot kap, a többi 3 játékos pedig 39 lapot 51 lap közül, tehát annak valószínűsége, hogy ne

A
kapja a pikk királyt,

(39/51)
. Ha a pikk ászt, ill. királyt az

A
, ill.

B
játékos kapta, akkor annak valószínűsége, hogy a pikk dámát

C
vagy

D
kapja, hasonló meggondolással

(26/50)
.

Így a keresett valószínűség

(39□26/51□50)=(169/425)=0,398.

Hasonlóképpen látható be, hogy annak valószínűsége, hogy a szóban forgó lap közül kettő egy kézben legyen, de a harmadik másnál legyen

3□(12/51)□(39/50)=(234/425)=0,550.
Végül annak valószínűsége, hogy a pikk ász, király és dáma egy kézben legyen:

(4(133)/(523))=(22/425)=0,052.

A három valószínűség összege természetesen 1-gyel egyenlő:

0,398+0,550+0,052=1

. Egy játékos lapjának várható értéke ily módon

$$1{,}5\cdot 0{,}398+2\cdot 0{,}550+2{,}5\cdot 0{,}052=0{,}597+1{,}100+0{,}130=1{,}827$$

, vagyis kerítve 1,8. Ez az eredmény alátámasztja a Culbertson-féle licitrendszernek azt a szabályát, hogy induláshoz legalább 2,5 trick szükséges – hiszen 2 trick még alig jobb az átlagosnál –, míg ha a partner indult, az emeléshez 1,5 trick is elegendő, vagyis már egy, az átlagot megközelítő lap elégséges; hiszen a két ellenfél lapjának várható összértéke 3,6 trick, és mivel

$$2{,}5+1{,}5=4$$

, az induló partnere már 1,5 trick esetében számíthat arra, hogy ő és partnere együtt erősebbek az ellenfeleiknél.

Egy másik tanulságos, bridzsre vonatkozó kérdés a következő: ha egy játékosnak a kezében 2 ász van, mi a valószínűsége, hogy a partnerénél van a két hiányzó ász, illetve, hogy ezek közül csak egy, vagy egyik sincs nála? Nyilván a szóban forgó 2 ász a másik 3 játékos kezében lévő 39 lap közt van, és így eloszlásukra

$$\binom{39}{2}$$

egyformán valószínű lehetőség van; ezek közül az első kérdés szempontjából

$$\binom{13}{2}$$

eset kedvező, tehát annak valószínűsége, hogy mind a két hiányzó ász a partnernél van

$$\binom{13}{2}\Big/\binom{39}{2}=\frac{6}{57}.$$

Annak valószínűsége, hogy a 2 hiányzó ász közül az egyik van a partnernél

$$\frac{13\cdot 26}{\binom{39}{2}}=\frac{26}{57},$$

míg annak valószínűsége, hogy egyik sincs a partnernél

$$\frac{\binom{26}{2}}{\binom{39}{2}}=\frac{25}{57}.$$

A 3 valószínűség összege természetesen eggyel egyenlő:

$$\frac{6+26+25}{57}=1.$$

A bridzs-játék részletes és közérthető valószínűségszámítási tárgyalása megtalálható *É. Borel* és *A. Chéron* könyvében [6].

4. 4. JÁTÉKSTRATÉGIÁK

Az egyszerűség kedvéért ebben a részben a következő leegyszerűsített szerencsejátékra szorítkozunk. Egy játékos (nevezzük őt Péternek) játszik a bank ellen; a játék játszmák sorozatából áll. Minden egyes játszmában Péternek jogában áll eldönteni, hogy mekkora összeget kockáztat; ezt az összeget Péter *tétjének* nevezzük. Péter köteles tétjét előre letenni az asztalra, tehát soha nem tehet nagyobb tétet, mint amennyi pénz van nála összesen. Ezek után egy véletlen kísérletet hajtanak végre, amelynek két lehetséges kimenetele van, az

A

és az

$\bar{A}$

esemény, melyek valószínűségei

p

és

q

$(p+q=1, 0<p<1)$
. Ha a kísérlet eredményeképpen az

A
esemény következik be, Péter megtartja a tétjét és ezen kívül a bank annyit fizet Péternek, mint amennyi Péter tétje volt. Ha a kísérlet eredményeképpen az

$\bar{A}$
esemény következik be, Péter tétjét megkapja a bank.

E típusba tartozik pl. a fej vagy írás játék, amelynél (szabályos érme esetében)

$p=(1/2)$
, továbbá ide tartozik a rulett, feltéve, hogy Péter mindig a pirosra tesz. Ez esetben, mivel a rulettkorongon 18 piros és 18 fekete pozitív szám van, továbbá egy zérus, és ha zérus jön ki, akkor a bank nyer,

$p=(18/37)$
.

Közismert, hogy ha

$p\leq(1/2)$
, nem létezhet olyan játékrendszer, amely biztos nyereséget biztosítana Péter részére. Egyszerűség kedvéért szorítkozzunk a

$p=(1/2)$
esetre! Legyen

$\square k=+1$
, ha a

k
-adik játszmánál Péter nyer (tehát az

A
esemény következik be) és

$\square k=-1$
, ha veszít, (tehát az

$\bar{A}$
esemény következik be). Jelölje

Sk
Péter tétjét a

k
-adik játszmában!

Sk
nyilván függhet

$\square 1,\square 2,\ldots,\square k-1$
-től:

$Sk=Sk(\square 1,\ldots,\square k-1)$

.

S_k értéke csak nemnegatív szám lehet;

$S_k=0$ azt jelenti, hogy Péter nem vesz részt a k-adik játszmában.

$S_n\neq 0$ és $S_j=0$, ha $j>n$, azt jelenti, hogy Péter az n-edik játszma után abbahagyja a játékot.

Ha Péter N forinttal a zsebében ül le játszani, Péter egy lehetséges stratégiáján az $S_k(\square_1,\dots,\square_{k-1})\ (k=1,2,\dots)$ nemnegatív függvények egy tetszőleges sorozatát értjük (S_1 egy állandó, ahol az $\square_i$ változók mindegyike a ± 1 értékeket veheti fel, és ezek a függvények eleget tesznek az $N+\sum_{k=1}^{n}\square_k S_k(\square_1,\dots,\square_{k-1})\geq 0$ feltételeknek $(n=1,2,\dots)$). Legyen $\xi_0=N$.

A

$$\xi_n=N+\sum_{k=1}^{n}\square_k S_k(\square_1,\dots,\square_{k-1})$$

$(n=1,2,\dots)$ összeg nyilván megadja, hogy mennyi pénze van Péternek az n-edik játszma után. A

$\xi_n (n=0,1,\ldots)$
valószínűségi változók ún. *martingált* (lásd [7]) alkotnak: ez azt jelenti, hogy

ξ_n
várható értéke amellett a feltétel mellett, hogy

$\xi_1,\xi_2,\ldots,\xi_{n-1}$
értéke adott, mindig egyenlő

ξ_{n-1}
-gyel.

Könnyen belátható, hogy ha

$p=(1/2)$
, akkor

$M(\xi_n)=N$

$(n=1,2,\ldots)$
, tehát semmilyen játékrendszer sem garantál Péter számára biztos nyereséget. Érdemes röviden foglalkozni a következő – szerencsejátékosok között a valószínűségszámítás nem ismerése folytán népszerű – hibás „játékrendszerrel”, amely szerint Péternek addig kell mindig 1 forintot megtennie, amíg először nem kerül nyerésbe: ekkor (1 forint nyereséggel) abba kell hagynia a játékot. Valóban úgy látszik, mintha e rendszer Péternek 1 forint biztos nyereséget garantálna, hiszen (1 valószínűséggel) előbb vagy utóbb Péter nyerésbe kerül. Valójában azonban ez a játékrendszer nem nyújt biztos nyereséget. Nyilvánvaló ugyanis, hogy ez a játékrendszer a fenti definíció szerint nem megengedett, hiszen ha például Péter az első

N
játszmában veszít, vagy az első

$N+2M$
játszma során összesen

$N+M$
-szer veszít és

M
-szer nyer, úgy, hogy közben soha sincs nyerésben, akkor nem tudja folytatni a játékot és így pozitív valószínűséggel elveszti teljes pénzét. Valójában Péter várható nyeresége e játékban 0, amit következőképpen bizonyíthatunk be: Jelölje

$f_k(N)$
annak a valószínűségét, hogy Péter előbb veszíti el mind az

N
forintját, mintsem

$k-N$
forint nyereségre tenne szert. Ez esetben, ha

$N\geq 2$

(4.1)	$f_k(N)=(1/2)f_k(N+1)+(1/2)f_k(N-1).$

Ugyanis a szóban forgó esemény kétféleképpen következhet be: úgy, hogy Péter az első játszmában veszít, és az ezután következő játék során előbb veszti el maradék $N–1$ forintját, mintsem $k–N+1$ forintot nyerne; vagy úgy, hogy Péter az első játszmában nyer, és ezután előbb veszti el $N+1$ forintját, mintsem $k–N–1$ forintot nyerne.

Könnyen belátható,22 hogy a (4.1) differencia-egyenlet összes lehetséges megoldása $fk(N)=AN+B$ alakú. Mivel $fk(0)=1$ (hiszen ha Péternek semmi pénze sincs, nem játszhat és így nem is nyerhet) és $fk(k)=0$ (hiszen ha a játék kezdetén Péternek már k forintja van, akkor nem is kell játszania), tehát

22Ezt a következőképpen bizonyíthatjuk be: Ha $fk(N)$ eleget tesz (4.1)-nek, akkor $g(N)=fk(N)–(N/k)(fk(k)–fk(0))–fk(0)$ is eleget tesz (4.1)-nek és $g(0)=g(k)=0$. Legyen max $g(N)=G=g(N1)$, akkor indukcióval belátható, hogy $g(N1+j)=G(j=1,2,…,k–N1)$ és hasonlóképpen $g(N1–j)=G(j=1,2,…,N1)$, tehát $g(N)=0,$ ha $N=0, 1, …, k$ és így $fk(N)=AN+B$, ahol $A=(1/k)(fk(k)–fk(0))$ és $B=fk(0)$.

$fk(N)=1-(N/k)$
. A minket érdeklő esetben

$k=N+1$
, tehát

$(1/N+1)$
annak a valószínűsége, hogy Péter előbb veszti el mind az

N
forintját, minthogy 1 forint nyereségre tenne szert. Eszerint Péter játékrendszere mellett nyereségének várható értéke

$1(1-(1/N+1))-N$□$(1/N+1)=0$.

Az elmondottak alapján úgy gondolhatná az olvasó, hogy a valószínűségszámítás a szerencsejátékot játszót csak arról győzheti meg, hogyha csupán azért játszik, mert nyereségre törekszik (és nem azért is, mert a játék szórakoztatja), akkor jobban teszi, ha nem is játszik. Ez azonban nincs így: ha a játékos azt kéri a matematikustól, hogy dolgozzon ki számára biztos nyerést garantáló játékrendszert, akkor lehetetlent kíván, és a matematikus nem segíthet rajta. Ha azonban a játékos elérhető célt tűz ki maga elé, a matematikus választ adhat arra a kérdésre, hogy e cél elérésére mi a legjobb út.

Vizsgáljuk először a következő kérdést! Péter fej vagy írást játszik; a játék kezdetén

N
forintja van, és elhatározza, hogy addig játszik, ameddig vagy pénze felnövekszik

$M>N$
forintra, vagy minden pénzét elveszti. Milyen játékrendszer mellett lesz annak valószínűsége, hogy nyer, maximális? Ha

$w=w(N,M)$
jelöli annak valószínűségét, hogy Péter

M
forinttal hagyja abba a játékot, mivel Péter

N
forintnál többet nem veszthet és nyereségének várható értéke 0 kell, hogy legyen, tehát

$w(M-N)-(1-w)N=wM-N\leq 0$
, vagyis

$w\leq(N/M)$
kell, hogy legyen. Kérdés, milyen játékrendszer mellett lehet elérni, hogy a nyerés valószínűsége

$w=(N/M)$
legyen?

Nevezzük „merész” stratégiának azt, amikor Péter mindaddig egész pénzét egyszerre megteszi tétként, amíg pénze

$\leq(M/2)$
, míg ha pénze

$x>(M/2)$

, de

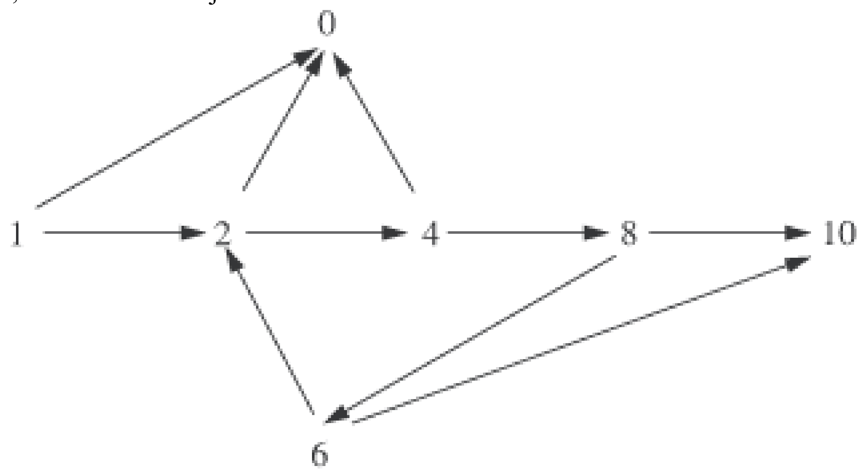
$x<M$
, akkor csak

$(M–x)$
-et tesz meg, vagyis pontosan annyit, hogy ha a következő játszmában nyer, akkor éppen elérje a célul kitűzött

M
forintot. Például ha

$N=1$
és

$M=10$
, akkor Péter a következőképpen játszik: az első játszmában 1 forintot tesz meg: ha veszít, kénytelen abbahagyni a játékot; ha nyer, most már 2 forintja lesz.



1. ábra

Ez esetben a második játszmában 2 forintot tesz meg: ha veszít, bánatosan távozik; míg ha nyer, akkor a következő játszmában újból egész pénzét (tehát most 4 forintot) tesz meg. Ha veszít, üres zsebbel hazamegy; ha nyer, akkor már 8 forintja van; most már csak 2 forintot tesz meg, így ha nyer, máris elérte a 10 forintot és így abbahagyja a játékot; de ha veszít, akkor is marad 6 forintja és így még tudja folytatni a játékot: megtesz 4 forintot, ha nyer, megvan a 10 forintja, és így örömmel távozik; ha veszít, még marad 2 forintja, és azt a következő játszmában megteheti, és így tovább. E számpéldában Péter pénzének alakulását a következő irányított gráf mutatja, amelyben minden pontból 2 él vezet ki és mindkét élen való továbbhaladás valószínűsége

$(1/2)$
. Ha

p_i
jelöli annak valószínűségét, hogy az

i
pontból

$(i=1,2,4,6,8)$
a játékos a

10.
pontba jut, nyilvánvalóan fennállnak a következő egyenletek:

$p_8 = (1/2)+(1/2)\cdot p_6$ $p_6 = (1/2)+(1/2)\cdot p_2$ $p_4 = (1/2)p_8$ $p_2 = (1/2)p_4$ $p_1 = (1/2)p_2$.

Ez az 5 egyenletből álló lineáris egyenletrendszer megoldható (ugyanis a determinánsa nem 0). Behelyettesítéssel nyerjük, hogy

$p_2=2p_1$
,

$p_4=4p_1$
,

$p_8=8p_1$
, továbbá

$16p_1-p_6=1$
,

$p_6-p_1=(1/2)$
; ebből

$p_1=(1/10)$
és így

$p_i=(i/10)$

$(i=1,2,4,6,8)$
.

Tehát

$w(1,10)=(1/10)$
. Hasonlóképpen számítható ki

$w(N,M)$
, ha

N
és

M
tetszőleges pozitív számok,

$N<M$
és

(M/N)
racionális. Ha azonban

N
és

$M>N$
igen nagy számok, ez a módszer nem célravezető, mert igen sok egyenletből álló egyenletrendszerre vezet. Ezért az általános esetben más bizonyítási módszert célszerű alkalmaznunk.

Általában igaz, hogy ha

N és M tetszőleges pozitív (nem feltétlenül egész) számok és $N<M$, akkor $w(N,M)=(N/M)$. Ezt a következőképpen láthatjuk be. Nyilván feltehetjük, hogy $M=1$ és $0<N<1$, hiszen választhatjuk M-et a pénz egységeként. Legyen $w(N,1)=f(N)$ $(0\le N\le 1)$! Nyilván fennáll a következő egyenlet

(4.2)

$$f(x)=\begin{cases}(1/2)f(2x), & \text{ha } 0\le x\le(1/2)\\ (1/2)+(1/2)f(2x-1), & \text{ha } (1/2)\le x\le 1\end{cases}$$

Ezt a függvényegyenletet hasonló módszerrel oldjuk meg, mint az előbb (4.1)-et. Legyen $g(x)=f(x)-x$, akkor $g(x)$ nyilván eleget tesz a

(4.3)

$$g(x)=\begin{cases}(1/2)g(2x), & \text{ha } 0\le x\le(1/2)\\ (1/2)g(2x-1), & \text{ha } (1/2)\le x\le 1\end{cases}$$

egyenletnek.

Mármost $g(x)$ korlátos, $-1\le(x)\le 1$, hiszen $f(x)$ valószínűség és így

$0 \le f(x) \le 1$
. Legyen

$G = \sup_{0 \le x \le 1} g(x)$
és

x_n
egy olyan számsorozat, amelyre

$\lim_{n\to\infty} g(x_n) = G$
.

Az

x_n
(korlátos) sorozatból kiválasztható egy konvergens részsorozat: jelöljük ezt

y_n
-nel, akkor tehát

$\lim_{n\to\infty} y_n = y$ és $\lim_{n\to\infty} g(y_n) = G$.

Ha

$0 \le y_n \le (1/2)$
végtelen sok

n
-re, akkor (4.3) szerint

$G = \lim_{n\to\infty} g(y_n) \le (1/2) \lim_{n\to\infty} \sup g(2y_n) \le (G/2),$

ha viszont

$(1/2) \le y_n \le 1$
végtelen sok

n
-re, akkor

$G = \lim_{n\to\infty} g(y_n) \le (1/2) \lim_{n\to\infty} \sup g(2y_n - 1) \le (G/2),$
tehát mindenképpen

$G \le (G/2)$
, azaz

$G \le 0$
.

Legyen most

$g = \inf_{0 \le x \le 1} g(x)$
. Hasonló meggondolással belátható, hogy

$g \ge 0$
; de ez azt jelenti, hogy

$g=G=0$
, vagyis

$g(x)\equiv 0$
és így

$f(x)\equiv x$
, amit bizonyítani akartunk.

Megjegyzendő, hogy a fej vagy írás játék esetében minden olyan stratégiánál, amelynél Péter 1 valószínűséggel véges sok lépésben vagy elveszti az összes pénzét, vagy eléri a célul kitűzött nyereséget, ugyanannyi a nyerés valószínűsége, mint a fent tárgyalt „merész” stratégiánál.

Vizsgáljunk azonban most egy olyan játékot, amelynél minden egyes játszmában Péter

p
valószínűséggel megnyeri a tétjét és

$q=1-p$
valószínűséggel elveszti, ahol

$0<p<(1/2)$
. (Ilyen játék pl. a rulett, ha Péter mindig a pirosra tesz, amely esetben, mint láttuk,

$p=(18/37)$
). Ebben az esetben már nem mindegy, hogy Péter milyen stratégiát alkalmaz, és a „merész” stratégia valóban optimális. Tegyük fel megint, hogy Péternek a játék kezdetén

x
pénze van, ahol

$0<x<1$
, és célja az, hogy pénze 1-re növekedjék fel. Jelölje

$g(x,p)$
annak a valószínűségét, hogy Péter eléri a célját, ha a „merész” stratégiát alkalmazza. Ugyanazzal a meggondolással, amely a

$p=(1/2)$
esetében a (4.2) függvényegyenletre vezetett, azt nyerjük, hogy

$g(x,p)$
az alábbi függvényegyenletnek tesz eleget:

(4.4)

$$g(x,p)=\{pg(2x,p),\quad \text{ha } 0\le x\le(1/2)\ /\ p+(1-p)g(2x-1,p)\quad \text{ha } (1/2)\le x\le 1,$$

továbbá eleget tesz a

$g(0,p)=0$
és

$g(1,p)=1$
feltételeknek. Az a meggondolás, amellyel beláttuk, hogy a (4.2)-nek eleget tevő

$f(x)$ függvény azonos x-szel, (4.4)-re alkalmazva arra az eredményre vezet, hogy a (4.4) függvényegyenletnek csak egy, a $g(0,p)=0$, $g(1,p)=1$ mellékfeltételeknek eleget tevő megoldása van. (Ezt egyébként először *G. de Rham* bizonyította be, l. [10].)

Mármost a (4.4) függvényegyenlet egy megoldását a következőképpen konstruálhatjuk meg: legyenek $\xi_1,\xi_2,\ldots$ független valószínűségi változók, amelyek a 0 és 1 értékeket p és $1-p$ valószínűséggel veszik fel. Legyen $\eta=\sum_{n=1}^{\infty}(\xi_n/2^n)$ és jelölje $F_p(x)$ az η valószínűségi változó eloszlásfüggvényét! Akkor egyszerűen belátható, hogy $F_p(x)$ eleget tesz az

$$F_p(x)=\begin{cases} pF_p(2x), & \text{ha } 0\le x\le(1/2) \\ p+(1-p)F_p(2x-1), & \text{ha } (1/2)\le x\le 1 \end{cases}$$

függvényegyenletnek és az $F_p(0)=0$, $F_p(1)=1$ feltételeknek. Így tehát $F_p(x)=g(x,p)$. Azt a $\mu_p(A)$ mértéket a $(0,1)$ intervallum Borel-halmazain, amelyre $\mu_p(I_{a,b})=F_p(b)-F_p(a)$, ha $0\le a<b\le 1$, ahol

Ia,b
az

$a \leq x < b$
intervallumot jelöli, a következőképpen is jellemezhetjük: a

$(0,(1/2))$
intervallum mértéke

p
, az

$((1/2),1)$
intervallumé

$1-p$
. A

$(0,(1/2))$
intervallumra jutó

p
mértéket

$p:(1-p)$
arányban osztjuk el a

$(0,(1/4))$
és

$((1/4),(1/2))$
részintervallumokra, hasonlóképpen járunk el az

$((1/2),1)$
intervallummal és így tovább. Így tehát a

$((k/2n),(k+1/2n))$

$(k=0,1,\ldots,2n-1)$
intervallumok közül

(nl)
számú intervallumnak a mértéke

$pl(1-p)n-l$
lesz

$(l=0,1,\ldots,n)$
. Az

$Fp(x)$
függvényről könnyen ki lehet mutatni, hogy szigorúan monoton növekvő, folytonos és szinguláris függvény, tehát deriváltja majdnem mindenütt 0. A

$\mu p1$
és

μ_{p_2} mértékek ortogonálisak, ha $p_1 \neq p_2$. Nyilván $\mu_{1/2}$ a közönséges Lebesgue-féle mértékkel azonos, mivel $F_{1/2}(x) = g\left(x, \frac{1}{2}\right) = x$ $(0 \leq x \leq 1)$.

Azt, hogy a $p < \frac{1}{2}$ esetben már nem mindegy, hogy milyen stratégiát alkalmaz valaki, és hogy a „merész” stratégia optimális, nem fogjuk itt általánosan bebizonyítani, csak egy számpéldával illusztráljuk. Tegyük fel, hogy Péter 25 Ft-tal ül le rulettezni $\left(p = \frac{18}{37}\right)$ és célja az, hogy 100 Ft-ra tegyen szert, és a merész stratégiát alkalmazza. Ez esetben akkor és csak akkor fogja elérni célját, ha megnyeri az első két játszmát, és ennek a valószínűsége $p^2 = 0{,}2366\ldots$. Mármost nézzük meg, mi Péter nyerési esélye, ha azt az óvatosabb stratégiát alkalmazza, és mindig csak 25 Ft-ot tesz meg. Könnyű belátni, hogy ez esetben Péter csak $\left(\frac{p^3}{1-2p+2p^2}\right)$ valószínűséggel éri el célját, és $\left(\frac{p^3}{1-2p+2p^2}\right) < p^2$, ha $p < \frac{1}{2}$; hiszen $p^2 - \left(\frac{p^3}{1-2p+2p^2}\right) = \left(\frac{p^2(1-p)(1-2p)}{p^2+(1-p)^2}\right) > 0$.

Speciálisan, ha $p = \frac{18}{37}$, akkor $\left(\frac{p^3}{1-2p+2p^2}\right) = 0{,}2301\ldots$, tehát Péter nyerési esélye a „merész” stratégia mellett több, mint $23{,}5\%$, míg az „óvatos” stratégia mellett kevesebb.

A most tárgyalt problémához hasonló általánosabb kérdések vizsgálatával foglalkozik *L. E. Dubbins* és *L. J. Savage* könyve [8].

5. 5. EGY MATEMATIKUS HARCA A JÁTÉKKASZINÓK ELLEN

Befejezésül egy igen érdekes esetről számolunk be, amely jól mutatja, hogy mire képes a szerencsejátékok matematikai elmélete és mire nem. *Edward O. Thorp* amerikai matematikus évekkel ezelőtt, amikor a Los Angeles-i egyetemen tanított, a téli szünidőben néhány napot töltött Las Vegasban és ennek során ellátogatott az egyik játékkaszinóba, ahol huszonegyest játszott – és természetesen vesztett. Ezen bosszankodva gondolkodni kezdett azon, hogy a huszonegyes játéknál (úgy, ahogy azt a nevadai játékkaszinókban játsszák23) mi a legjobb stratégia egy játékos részére.

Tudvalevőleg a huszonegyesnél az „osztó” (a kaszinó alkalmazottja) minden játékosnak egy jól megkevert 52 lapos kártyából 2–2 lapot oszt ki: a játékosok lapjaikat nem mutatják meg az osztónak. Az „osztó” önmagának is oszt 2 lapot, de ezek közül az elsőt köteles megmutatni a játékosoknak. A lapokat a következőképpen értékelik: minden figura (bub, dáma, király) 10 pontot ér, a többi lap az ász kivételével annyit, amennyi rá van írva (tehát pl. a hetes hét pontot). Az ászt minden játékos szabad elhatározással értékelheti 1-esnek vagy 11-esnek. A játékban az nyer, akinek lapjai pontértékeinek összege legjobban megközelíti a 21-et anélkül, hogy túllépné azt.

Minden játékosnak joga van lapjainak megnézése után anynyi újabb lapot kérni, ahányat csak akar; de ha lapjainak pontösszege meghaladja a 21-et, köteles lapjait felmutatni és kiesik a játékból. Az „osztó” önmagának is oszthat újabb lapokat. A játékosok tétjeiket megadott alsó és felső korlát között szabadon választhatják meg. Minden játékos külön játszik az „osztó” ellen. Ha a játékos lapja jobb, mint az osztóé, akkor a játékos ugyanannyit nyer, mint amennyi a tétje volt; ha rosszabb, elveszti tétjét; ha viszont egyformák (pl. mindkettőé 21), pénz nem cserél gazdát.

Az osztó nagy előnye abban áll, hogy a játékos minden esetben köteles felmutatni lapjait, és így ha azok összértéke meghaladja a 21-et, tétjét mindenképpen elveszti; akkor is, ha az osztó lapjainak összértéke is meghaladja a 21-et; ez ugyanis esetleg ki sem derül, mivel ha minden játékos eldobta a lapjait, az osztónak nem kell megmutatnia a sajátját, csak beseperi a tételeket.

Thorp megfigyelte, hogy a kaszinók egész szigorú szabályokkal előírják alkalmazottaiknak, hogy milyen rendszer szerint kell játszaniuk,24 így pl. előírják, hogy ha az osztó lapjainak összértéke eléri vagy meghaladja a 17-et, nem szabad újabb lapot osztania magának. *Thorp* úgy gondolta, hogy az a tény, hogy a játékost semmilyen merev szabály nem köti és – ellentétben az osztóval – nem köteles első lapját felmutatni; továbbá, hogy a tét összegét ő választhatja meg, elvileg lehetővé teszi nyerő rendszer kidolgozását az osztó fent említett előnye ellenére is. Ezt főként arra alapozta, hogy akkoriban – a játék meggyorsítása érdekében – a nevadai kaszinókban az volt a szokás, hogy az osztó nem keverte meg a lapot minden játszma után, hanem addig használt egy csomag kártyát, amíg az el nem fogyott, és csak azután szedte össze és keverte meg az elhasznált lapokat. Ilyen módon az a játékos, aki megjegyzi, hogy már milyen lapok „mentek ki”, és ennek alapján rugalmasan változtatja stratégiáját, megnövelheti nyerési valószínűségét; feltéve, hogy tudja, hogy a rendelkezésre álló információt hogyan használja fel. Ehhez a játékosnak nyilván az egyes lapok kihúzásának feltételes valószínűségeit kell figyelembe vennie nem teljes kártyacsomag esetén, és ezek alapján tudja kidolgozni legelőnyösebb stratégiáját. Persze egyszerű és könnyen megjegyezhető szabályokra van szükség, hiszen pillanatok alatt kell döntenie, hogy kér-e még további lapot vagy sem. *Thorp* nem sajnálta a fáradságot és – a Massachusetts Institute of Technology IBM 704-es elektronikus számítógépének segítségével – kidolgozott25 egy könnyen megjegyezhető stratégiát, amely a játékosnak néhány százalék előnyt biztosít a kaszinóval szemben. Az Amerikai Matematikai Társulat 1960-ban Washingtonban tartott találkozóján előadást tartott számításairól. Az előadás nagy feltűnést keltett és néhány nap múlva egy üzletembertől levelet kapott, aki

23Az Amerikai Egyesült Államok többi tagállamában játékkaszinók nem működhetnek legálisan. Nevadában azonban ezt megengedik a törvények.

24A merev szabályok előírásával próbálja a kaszinó megakadályozni, hogy alkalmazottai összejátszanak a játékosokkal és – nyereségrészesedés ellenében – a kaszinó rovására szándékosan veszítsenek.

25A számítás a gép 3 óráját vette igénybe.

100 000
dollárt ajánlott fel neki arra a célra, hogy rendszerét a gyakorlatban kipróbálja. Thorp elfogadta az ajánlatot és – rendszerét betanulva – elutazott Nevadába, hogy kísérletet tegyen. A kísérlet fényesen sikerült: 2 óra leforgása alatt

17 000
dollárt nyert. A kaszinó tulajdonosa persze távolról sem lelkesedett annyira a tudományos kísérlet sikeréért, mint Thorp és társa, és másnap különböző kifogásokkal megakadályozták őket abban, hogy újra játszanak. Később Thorp más kaszinókban is próbálkozott, de a híre mindenütt előtte járt és minden kaszinó bezárta előtte ajtóit. Néhányszor álszakállal vagy kínainak maszkírozva sikerült a játékasztalhoz jutnia; de akárhogy is változtatta el az arcát, elárulta őt az, hogy állandóan nyert. Így tehát rendszerének gyakorlati felhasználását abba kellett hagynia. Azzal bosszulta meg magát az őt kiutasító kaszinókon, hogy rendszerét könyv alakban megírta és kiadta (lásd [9]). E könyv alapján olyan sokan tanulták meg a nyerő stratégiát, hogy a nevadai kaszinók kénytelenek voltak gyökeresen megváltoztatni a játékszabályokat. A változtatás többek között abban állt, hogy ma már minden játszma után újra kevernek; ezzel kihúzták a talajt a stratégia lába alól. Ily módon visszajutottunk oda, ahonnan kiindultunk: a kártyák keverésének jelentőségéhez.

IRODALOM

[1] [1] Erdős P. – Turán P.: On some problems of a statistical group theory, I.; *Zeithschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **4** (1965), 175–186.

[2] [2] Prékopa A. – Rényi A. – K. Urbanik: O predeljnom raszpredelenii dlja szum nezaviszimih szlucsajnih velicsin na bikompaktnih kommutativnih topologicseszkih gruppah, *Acta Math. Ac. Sci. Hung.* **7** (1956), 11–16.

[3] [3] U. Grenander: *Probabilities on algebraic structures*, Wiley, New York, 1963.

[4] [4] Jordan K.: *Fejezetek a klasszikus valószínűségszámításból*, Akadémiai Kiadó, 1955.

[5] [5] E. Culbertson: *New and Complete Summary of Contract Bridge*, John C. Winston, Philadelphia, 1935.

[6] [6] É. Borel – A. Chéron: *Théorie mathématique des bridges, a la portée de tous*, Gauthier-Villars, Paris, 1955. 1–424.

[7] [7] J. L. Doob: *Stochastic processes*, Wiley, New York, 1953.

[8] [8] L. E. Dubbins – L. J. Savage: *How to gamble if you must*, McGraw–Hill, New York, 1965.

[9] [9] E. O. Thorp: *Beat the dealer, A winning strategy for the game of twenty one*; Blaisdell, New York, 1962.

[10] [10] G. de Rham: Sur quelques courbes definies par des équations fonctionelles, *Rendiconti del Seminario Matem. Univ. Torino* **16** (1956–57), 101–113.

BŰVÖS NÉGYZETEK

BAKOS TIBOR

A bűvös négyzet érdekes, sok szabályszerűséget mutató számelrendezés. Egy négyzetet az oldalaival párhuzamos egyenesekkel bizonyos számú sorra és ugyanannyi oszlopra osztunk fel, majd a következő feladatot tűzzük ki: töltsük meg ezt az ábrát egymás utáni természetes számokkal úgy, hogy minden kis mezőbe egy szám jusson, és a számok összege minden sorban, minden oszlopban és mindkét átló mentén ugyanannyi legyen. Ebben a mai ember számára már csak a magas fokú rend az érdekesség: a különböző vonalakon más-más számok állnak együtt, és az összeg mégis mindig ugyanaz. Tetszik ez, és a siker reménye a megoldásra ösztönöz bennünket. A régiekben azonban más érzelmeket is kiváltott egy-egy bűvös négyzet szemlélete: tiszteletet parancsolt, félelmet keltett, bűvészetnek látszott. Innen ered az elnevezés is. Némelyek például hasznot húztak a bűvös négyzettel díszített, a „bajoktól megvédő” kis amulettek árusításából. Egyeseket boszorkánysággal vádoltak és fogságba is vetettek ilyen számösszeállítások készítéséért.

Pedig nincs a bűvös négyzetekben semmi bűvészet, csak érteni kell a szabályszerűségekhez. És ha ez már bűvészet, akkor bárki megtanulhatja. Valóban, az úgynevezett bűvészek is mindig csupán kevésbé ismert természeti törvények alapján végzett mutatványokkal keltettek csodálatot nézőikben. Ezért nincs szükség a nem egészen szerencsés „bűvös négyzet” elnevezés megváltoztatására.

Az egyre több és több számot felhasználó bűvös négyzetek tanulmányozásához csak ezzel az elhatározással foghatunk hozzá: nem akarunk a terjedelmük miatt már fárasztó összeadásokban gyönyörködni, hanem mindig igyekszünk lehetőleg kevés munkával eredményre jutni, miután a kevesebb számot felhasználó eredményeknek magyarázatát adtuk.

Számos régi, a matematikával szórakoztatni kívánó könyv csak kész bűvös négyzeteket, képzési „recepteket” közölt. Ez is érdekes; de különb szórakozás érteni is, mit, miért csinálunk! Kis részben mi is a közlésre kényszerülünk, de kérjük az olvasót, ne ugorja át a magyarázó, bizonyító részleteket se, vagy kellő pihenő után, napok múltával térjen vissza hozzájuk, amikor már hozzászokott az érdekességekhez. Ugyanis az elmélyülések gyakran újabb sikerek előkészítői. Kívánjuk, hogy az ilyen részletek még több szórakozást nyújtsanak számára, és újabb problémák felvetését jelentsék. A problémakör szinte kifogyhatatlan. Itt a régi eredményekből is csak ízelítőt adhatunk, és már ezekhez is sok további, könnyen felfedezhető érdekesség kapcsolódik. Az sem baj, ha felfedezésünkről kiderül, hogy már mások is rátaláltak. Jó tudni viszont, hogy újabb és újabb efféle problémák mai matematikusokat is foglalkoztatnak, és hogy a legutóbbi időkben is számos érdekes eredmény született.

A sorok és oszlopok egyező számát a bűvös négyzet *rendszámának* szokás nevezni. Látni fogjuk – vagy legalábbis sejteni –, hogy 3-tól kezdve minden rendszámhoz lehet bűvös négyzetet képezni. (Másodrendű bűvös négyzet nyilvánvalóan lehetetlen.) A képzés lényeges különbségeket mutat aszerint, hogy a rendszám páros vagy páratlan. Elsőnek mindkét fajtából a legkisebb rendszámút vesszük, a harmad- és a negyedrendű bűvös négyzeteket, majd nagyobb rendszámú bűvös négyzet képzésére adunk útmutatást. Az összes megoldást csak a 3-as rendszámra adhatjuk meg, mert már negyedrendű négyzet is 7040 van, az ötödrendűektől kezdve pedig számuk ez ideig még nem is ismeretes.

A legegyszerűbb bűvös négyzet 3 sorral és 3 oszloppal az

1,2,3,4,5,6,7,8,9
számokból épül fel. Összegük 45, így mindegyik soron, oszlopon és átlón – röviden: mindegyik *vonalon* – 15-öt kell összegként kapnunk, ez lesz négyzetünk ún. *bűvös állandója.*

$$\begin{matrix} x & y & z \\ u & v & w \\ r & s & t \end{matrix}$$

1. ábra

Ezt tudva bizonyára többen gondolnak az egyes számoknak egyenletrendszerrel való meghatározására. Azonban az 1. ábra szerint jelölt 9 ismeretlenre csak 8 egyenletet írhatunk fel, a 3 sor, a 3 oszlop és a 2 átló követelményét:

$$(1)\ x+y+z=15, \qquad (4)\ x+u+r=15, \qquad (7)\ x+v+t=15,$$
$$(2)\ u+v+w=15, \qquad (5)\ y+v+s=15, \qquad (8)\ z+v+r=15.$$
$$(3)\ r+s+t=15, \qquad (6)\ z+w+t=15,$$

Sőt az oszlopok és sorok összegét előíró (1)–(6) egyenletek közül egyet el is kell vetnünk, mint ami nem mond újat. Ugyanis pl. a (6) egyenlet következik (1)–(5)-ből, hiszen az (1)–(3) egyenleteket összeadva visszakapjuk, hogy a 9 ismeretlen összege 45, és ebből elvéve (4) és (5) összegét, visszamarad a (6) egyenlet. (6)-ot elhagyva a többi 7 egyenlet független rendszert alkot, egyikből sem következik a másik. Csakhogy ennyi kevés 9 ismeretlen meghatározására, amint ezt a 2–6. ábrák példái is mutatják. Ezek minden előírt vonalán 15 az összeg, de több ábrán egyenlő számok lépnek fel; ahol pedig csupa különböző számot látunk, ott nem az előírt számok szerepelnek. A felírt *összeg-követelményeket* az előírtaktól különböző számok is teljesíthetik. Az előírt számok *szerepeltetésének követelményét* viszont nem lehet egyenletben felírni. Nem is a beírandó számok az ismeretlenek, hanem a helyzetük.

$$\begin{matrix} 5 & 5 & 5 \\ 5 & 5 & 5 \\ 5 & 5 & 5 \end{matrix} \qquad \begin{matrix} 4 & 6 & 5 \\ 6 & 5 & 4 \\ 5 & 4 & 6 \end{matrix} \qquad \begin{matrix} 5 & 1{,}5 & 8{,}5 \\ 8{,}5 & 5 & 1{,}5 \\ 1{,}5 & 8{,}5 & 5 \end{matrix}$$

2. ábra *3. ábra* *4. ábra*

$$\begin{matrix} 5{,}6 & 6{,}2 & 3{,}2 \\ 2{,}6 & 5 & 7{,}4 \\ 6{,}8 & 3{,}8 & 4{,}4 \end{matrix} \qquad \begin{matrix} 5{,}6 & 5{,}9 & 3{,}5 \\ 2{,}9 & 5 & 7{,}1 \\ 6{,}5 & 4{,}1 & 4{,}4 \end{matrix}$$

5. ábra *6. ábra*

Valami mégis kihozható egyenletrendszerünkből: az, hogy a középső szám csak

$v=5$

lehet. Ugyanis a középső mezőn átmenő 4 vonal összege a (2), (5), (7) és (8) egyenletekből:

$(u+w)+(y+s)+(x+t)+(z+r)+4v=60,$

másrészt a 3 sor összege az (1), (2) és (3) egyenletekből:

$x+y+z+u+v+w+r+s+t=45,$

és ezt az előbbi összegből kivonva

$3v=15$

,

$v=5$

. (Az 5-ös az előírt számok nagyság szerinti sorrendjében is középen áll.)

Célszerű a hátralévő 8 számon egyelőre csak páros vagy páratlan voltukat nézni. Mindegyik fajtából 4 van: 2, 4, 6 és 8, illetőleg 1, 3, 7, 9. Próbálkozzunk a páratlanok elhelyezésével! Az elsőnek elhelyezett páratlan számot az 1. ábrán

x

helyére téve, az innen kiinduló átló – az ún. *főátló* – másik sarokmezejére ugyancsak páratlan számot kell tennünk, mert

x

és 5 összege páros, és ezt 15-re, páratlan összegre csak páratlan szám egészítheti ki. Most azonban a harmadik páratlan számot akármelyik mezőre próbálva hasonlóan adódik, hogy minden mezőre páratlan számot kell írnunk. Pl.

s

helyére téve a harmadik páratlant, az alsó sor miatt

r

, a középső oszlop miatt

y

is páratlan, és máris 5 páratlant helyeztünk volna el.

Eszerint, ha egyáltalán van harmadrendű bűvös négyzet az 1–9 számokból, annak sarkán nem állhat páratlan szám. Más szóval: sarkon csak páros szám állhat. Legyen pl.

$x=8$

, így

$t=2$

, ekkor a 4 és 6 páros számok elhelyezésére a másik átló – az ún. *mellékátló* – végein két lehetőség van; legyen pl.

$r=4$

, így

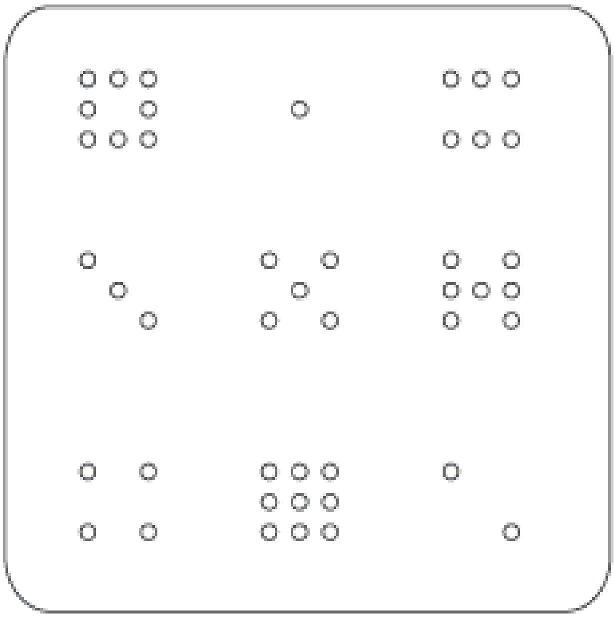
$z=6$

. Ezzel a szélső vonalak egy-egy mező híján be vannak töltve, üres mezejük száma az összeg-követelményből kiszámítható. Ezekre éppen az előírt négy páratlan szám kerül (7. ábra), és egy csapásra a középső sor és a középső oszlop összege is 15.

8	1	6
3	5	7
4	9	2

7. ábra

Kínában találtak olyan, fémlemezből készült régi amulettet, amelyre ezt a bűvös négyzetet nem számjegyekkel vésték rá, hanem a mezőkbe megfelelő számú lyukat fúrtak (8. ábra).



8. ábra

Így a lemez bármelyik élére állítva olvasható, akár hátulról is, mintha tükörben szemlélnénk. Vagyis egyetlen bűvös négyzetből a négyzetlap szimmetriái révén 8 bűvös négyzetet kapunk.

A sarokmezők más betöltésével próbálkozva már nem kapunk kilencedik bűvös négyzetet. Hiszen pl. megtartva

x=8
és

t=2
értékét és

r=6
-ot,

z=4
-et véve ugyanazt kapjuk, amit lemezünkről a főátló körüli megfordítással is leolvashatunk.

x
megválasztására a 2, 4, 6, 8 számokból 4 lehetőség van, ebből mindig

t
is megkapható,

r
-re pedig a hátralevő két páros számból 2 lehetőség van, tehát a számítás is

4□2=8
bűvös négyzetet ad.

(Ha már tudjuk, hogy páratlan számaink csak oldalközépen állhatnak, ügyetlenség lenne ragaszkodni ahhoz, hogy előbb ezeket helyezzük el. Akkor ugyanis minden szélső vonalon csak egy ismert számunk lenne, és a sarkokra kerülő számokat egyenletrendszerrel kellene kiszámítanunk.)

A kész négyzeteket szemlélve az is feltűnik, hogy a szemben fekvő csúcs-párokon, oldalközép-párokon álló két-két szám összege mindig 10, és az ilyen számpárok már az eredeti felsorolásban is az 5-ösre mint középpontra nézve tükrösen helyezkednek el, pl. 2 és 8. Az összeg értéke természetes:

$15–v=15–5=10$

; hozzátehetjük viszont, hogy a számpárok tükrös elhelyezkedése feltűnhet már a 9. ábrán,

1	2	3
4	5	6
7	8	9

9. ábra

számaink ún. *alapállásában* is. Szokás az eredeti felsorolás elölről és hátulról számított ugyanannyiadik tagját – pl. a 2-t és 8-at – egymás számtanilag, *aritmetikailag* tükrös párjának nevezni. Így azt mondhatjuk: az előírt számok aritmetikailag tükrös párjai a bűvös négyzetben a középmezőre nézve *geometriailag* tükrösen helyezkednek el, a pár nélkül álló 5-ös pedig a tükrös pár nélküli mezőn, a középmezőn.

Ezek után már kissé fanyarul fogadjuk az innen adódó felfedezést: harmadrendű bűvös négyzet minden száma helyére aritmetikailag tükrös párját írva ismét bűvös négyzetet kapunk – hiszen már tudjuk, hogy nincs több harmadrendű bűvös négyzet. – Ez igaz, nagyobb rendszám mellett azonban majd ez a fogás is adhat egy bűvös négyzetből újat.

Két más elindulás a harmadrendű bűvös négyzetek felé. Vegyük figyelembe egyidejűen a szerepeltetési és az összeg-követelményeket! Állítsuk össze előírt számainkból az összes 3 tagú, 15-ös összegű kombinációkat! A bűvös négyzet mindegyik vonalán egy ilyennek kell állnia, maga a bűvös négyzet ezeknek mintegy az összeszövése. Könnyű belátni, hogy csak a következő 8 kombináció megfelelő:

I 1+5+9,	IV 2+5+8,	VII 3+5+7,
II 1+6+8,	V 2+6+7,	VIII 4+5+6.
III 2+4+9,	VI 3+4+8,	

Éppen 8 bűvös vonalat kell kialakítanunk, tehát mindegyik kombináció fellép egy vonalon.

Vizsgáljuk meg, kombinációink mely párjaiban fordul elő közös szám. Nyilvánvaló, hogy egymást csak olyan két kombináció keresztezheti, amelyekben van közös szám; másrészt párhuzamosan csak olyanok állhatnak, amelyekben nincs ilyen. A közös számot – ha van – táblázatunk tünteti fel

	I	II	III	IV	V	VI	VII	VIII
I	x	1	9	5	—	—	5	5
II	1	x	—	8	6	8	—	6
III	9	—	x	2	2	4	—	4
IV	5	8	2	x	2	8	5	5
V	—	6	2	2	x	—	7	6
VI	—	8	4	8	—	x	3	4
VII	5	—	—	5	7	3	x	5
VIII	5	6	4	5	6	4	5	x

(egynél több közös szám persze nem lehet). A IV és VIII kombinációnak minden másikkal van közös száma, ezért csak olyan vonalra tehetők, amellyel párhuzamosan nem fut vonal, vagyis az átlókra. Közös számuk az 5-ös, csak az állhat középen. (A befejezést ebből az állapotból már fentebb láttuk.)

Egy kis statisztikával is megsejthetjük a megoldást. A sarokmezők 3 követelményben szerepelnek: egy sorban, egy oszlopban és egy átlóban; a középmező 4, az oldalközepek 2 követelményben. Fenti kombinációinkban viszont az 5-ös szám négyszer fordul elő, a páros számok háromszor, a többi páratlan kétszer. Ebből adódik, hogy csak úgy érdemes próbálkoznunk, hogy középre az 5-öst állítjuk és a sarkokra a páros számokat.

Negyedrendű bűvös négyzetben az 1-től 16-ig terjedő számokkal minden vonalon 34-et kell kapnunk, ennyi az összegük negyedrésze. Fentebbi eljárásaink átvétele azonban nehézségeket ígér: az első eljárás mintájára 16 ismeretlenünk, 10 bűvös vonalunk, 9 független egyenletünk lenne, és nem lenne az 5-ösre emlékeztető fix szám, hiszen – ha van megoldás – a megfelelő amulettlemez elforgatása minden számot elmozdít. A kombináló eljárások mintájára gondolva számaink 34-es összegű, négytagú kombinációinak száma (az előbbi 8 helyén) 86. Mégse csüggedjünk, itt 880 különböző lemez-amulettet lehetne kifúrni. (Ezek tarthatók 7040 állásban.)

Számaink alapállását szemlélve (10. ábra)

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

10. ábra

azt mondhatnánk: induljunk ki abból, hogy mindkét átló összege máris 34. – De jók voltak a harmadrendű alapállás átlói is (9. ábra), sőt még középső sora és oszlopa is, mégsem maradt a helyén, csak az 5-ös! Nézzük hát, hol és mennyit kell elvenni, hozzátenni. Az 1. sor összege 10, tehát hiánya 24, a 2. sor hiánya 8, tovább már többletek vannak: ismét 8 és ismét 24. Cseréljük hát a 2-est az alsó sorban alatta álló 14-essel; különbségük 12, ennyivel csökken a cserével a hiány is, a többlet is. A mondott hiányok és többletek egyszerre eltűnnek, ha még az egymás fölött álló 3, 15, az 5, 9 és a 8, 12 számpárokat is megcseréljük. Hasonlóan az egymás utáni oszlopokban 6 és 2 egységnyi hiány van, majd meg 2 és 6 egységnyi töblet, és ez mind eltűnik a következő négy páros cserével: 9, 12; 5, 8; 14, 15; 2, 3; melyeknek tagjai – az előbbi cserék után is – ugyanabban a sorban állnak, tehát cseréik az előbb elért helyes sorösszegeket nem rontják el. Kész!

A kapott 11. ábra a legegyszerűbb szerkezetű negyedrendű bűvös négyzetek egyike. Képzését – mindjárt a végeredményt tekintve – így egyszerűbb megjegyezni: a nyolc átlós szám a helyén marad, a többi nyolc pedig a négyzet középpontjára tükrös helyzetű párokban helyet cserél. Ez az előbbi eljárással szemben csak 4 páros csere, ugyanis mindegyik cserélt szám két cserében vett részt, és a második cserék után mindegyik szám annak a számnak a helyére került,

1	15	14	4
12	6	7	9
8	10	11	5
13	3	2	16

11. ábra

amely viszont az ő eredeti helyét foglalta el. A most megcserélt számok egyszersmind aritmetikailag is tükrösek, összegük páronként 17.

5	11	14	4
16	6	3	9
1	15	10	8
12	2	7	13

12. ábra

1	15	10	8
14	4	5	11
7	9	16	2
12	6	3	13

13. ábra

A 12-13. ábrák még két bűvös négyzetet adnak. Már e három megoldáson is számos további érdekességet mutathatunk be. Képezzük a 12. ábra aritmetikai tükörképét, vagyis írjuk minden száma helyére az 1, 2, …, 15, 16 felsorolás hátulról számított ugyanannyiadik számát, más szóval azt, amely őt 17-re egészíti ki. Így bármely bűvös négyzetből ismét bűvös négyzetet kapunk; ugyanis ha tetszés szerinti számokkal

$x+y+z+v=34$
, akkor egyszersmind

$(17-x)+(17-y)+(17-z)+(17-v)=68-34=34$
is teljesül. A kapott 14. ábra

12	6	3	13
1	11	14	8
16	2	7	9
5	15	10	4

14. ábra

a 12. ábrából képzett amulett-lemezről semmilyen forgatás útján sem olvasható le, új bűvös négyzet. (A szélső oszlopokban megfordult a sorrend, és a belső oszlopokban is van valami hasonló.)

Új megoldást ad így a 13. ábra is; viszont a 11. ábra aritmetikai tükörképét a középpont körüli

180°
-os elforgatással – geometriai tükrözéssel – is megkapjuk, ez nem új megoldás.

Írjuk fel számainkat egy-egy gyufásdoboz címkéjére, így írás és radírozás helyett rakosgatással kísérletezgethetünk. Cseréljük fel így egymással a 13. ábra két belső oszlopát mint egészet (15. ábra).

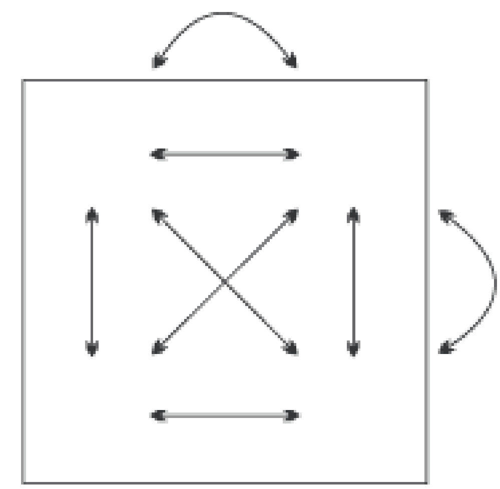
1	10	15	8
14	5	4	11
7	16	9	2
12	3	6	13

15. ábra

1	10	15	8
7	16	9	2
14	5	4	11
12	3	6	13

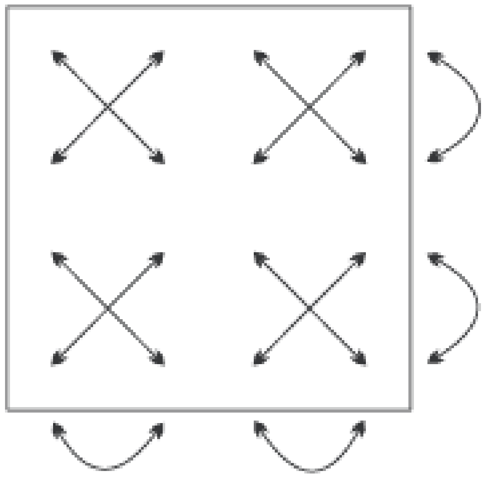
16. ábra

Ezzel csak az átlók bűvösségét rontottuk el, a középső négy szám most rosszul áll. Eredeti átlójukba úgy is visszajuttathatjuk őket – és új bűvös négyzetet kapunk –, ha az átmeneti 15. ábra két belső sorát felcseréljük (16. ábra).



17. ábra

A 13. és 16. ábrák abban térnek el, hogy a 17. ábra egy-egy kéttollú nyila végeinél álló számok felcserélődtek. Ezt az eljárást bármely más negyedrendű bűvös négyzetre is alkalmazhatjuk – sőt kis módosítással magasabb rendűekre is –, megérdemli hát, hogy külön elnevezést kapjon: nevezzük *első bűvös transzformációnak* (átalakításnak). Ezzel az ábrázolásmóddal mutatjuk be a 18. ábrán a *második bűvös transzformációt*: minden szélső vonal cserél a szomszédjával; így minden szám elmozdul, kétszer is cserél. A végeredmény 8 páros cserével is leírható.



18. ábra

Így készült az imént kapott 16. ábrából a 19. ábra. Ezt „ősével”, a 13. ábrával közvetlenül összehasonlítva azt látjuk, hogy az alábbi számnégyesek tagjai körben egymás helyére léptek (ezért a rajzbeli feltüntetés bonyolultabb):

1, 4, 13, 16;
8, 5, 12, 9;
15, 11, 3, 7;
10, 14, 6, 2.

16	7	2	9
10	1	8	15
3	12	13	6
5	14	11	4

19. ábra

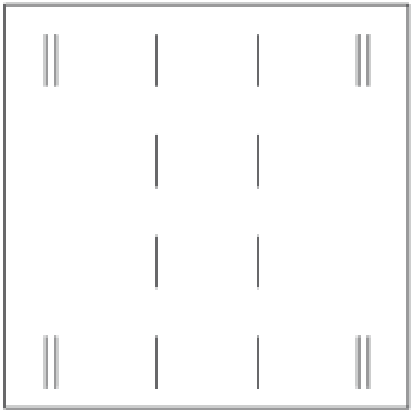
Próbálja ki az olvasó, új bűvös négyzetet kap-e, ha a 13. ábrára előbb a második transzformációt alkalmazza, majd az eredményre az elsőt.

A 11. ábra aritmetikai tükörképéből akkor is új bűvös négyzetet kapunk, ha csak a középső oszlopokat cseréljük, mert mindkét átló középső két számának összege ugyanannyi, 17. Ez a bűvös négyzet híressé vált, mert *A. Dürer* festőművész *Melancolia* című rézmetszetén is szerepel; egymás mellett álló 15-ös és 14-es száma összeolvasva a kép keletkezésének évét adja.

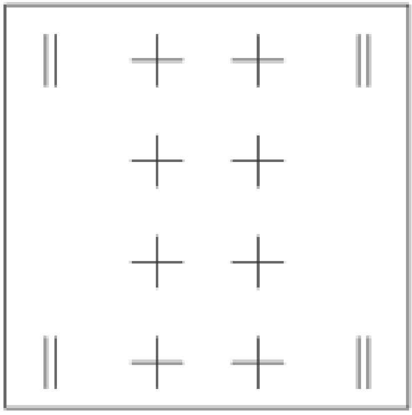
Vizsgálhatnánk ábráinkon az említett 86 bűvös négytagú kombináció elhelyezkedéseit is. Sok érdekesség közt azt is tapasztalhatnók, hogy egy kombináció mindig a négy sarokmezőn „ül”, egy másik a négy középmezőn, azaz minden eddigi negyedrendű bűvös négyzeten mutatkoznak az előírást meghaladó „rejtett” bűvösségek is. Nevezzük a mondott mezőnégyeseket *bűvös kereteknek*. Pl. a 12. és 13. ábrán:

5+4+13+12=6+3+10+15=1+8+13+12=4+5+16+9=34.
Könnyű belátni, hogy ennek minden negyedrendű bűvös négyzeten így kell lennie. Vegyük egy tetszés szerinti négyzet 1. sorának, 4. sorának és a két átlójának számait, összegük

4□34
. A 20. ábrán minden figyelembe vett szám mezejére egy kis álló vonalat írtunk; ebből látjuk, hogy eddigi összegünkben a sarokmezők számai 2-szer szerepelnek, a 2. és 3. oszlop számai pedig 1-szer.



20. ábra



21. ábra

Hagyjuk most el a 2. és 3. oszlopot, írjunk az elhagyott számok helyére vízszintes vonalat (21. ábra)! Ahol kétféle jelünk egymáson keresztben áll, a megfelelő szám további meggondolásunkban már nem szerepel.

Az elhagyott számok összege

2□34
, így a visszamaradottaké is

2□34
, ennyi tehát a sarokszámok összegének 2-szerese, vagyis összegük 34. Most már a két átlóból a sarokszámokat elhagyva a középső keret számai maradnak vissza, ezek összege ismét

2□34–34=34
. – Két további bűvös keret is adódik: a szélső sorok belső (azaz nem sarki) számainak összege is a bűvös állandó, pl. a 13. ábrán

15+10+6+3
, és a szélső oszlopok belső számaié is:

14+7+11+2
. Az olvasó az előbbi meggondolást folytatva beláthatja, hogy ez is mindig érvényes.

A 13. ábrának van egy további érdekessége, ez a tulajdonság a 11-12. ábrákban nincs meg. Írjuk a 13. ábrát sok példányban négyzet alakú, egybevágó kőlapokra és kövezzük ki velük a sík egy részét a „kockás papír” mintájára; a lapok egyformán álljanak (22. ábra)!

1	15	10	8	1	15	10	8	1 . . .
14	4	5	11	14	4	5	11	14 . . .
7	9	16	2	7	9	16	2	7 . . .
12	6	3	13	12	6	3	13	12 . . .
1	15	10	8	1	15	10	8	1 . . .
14	4	5	11	14	4	5	11	14 . . .
7	9	16	2	7	9	16	2	7 . . .
.	.	.	.	.	.	.	.	. . . .

22. ábra

Bármelyik bűvös négyzetünkből képzünk ilyen ún. *bűvös kövezetet*, abból bárhol kiragadva egy 4 sorból és 4 oszlopból álló, rendes állású négyzetet (vagyis a határvonal lehet nem kőlap határán is), mindig ún. *félig bűvös négyzetet* kapunk. Ezen azt szokás érteni, hogy a sorok és az oszlopok összege a bűvös állandó – ami a kövek ismétlődése miatt magától értetődik –, az átlós összegek azonban általában különböznek a bűvös állandótól. Nos, a 13. ábrából képzett bűvös kövezetből bárhogyan jelölve ki egy

4×4

-es négyzetet, mindig helyes átlós összegeket, vagyis bűvös négyzetet kapunk. Egy ilyen áll a 22. ábra jobbra lefelé elcsúsztatott keretezésében.

A négyzet bal felső sarokszámának a 16 szám bármelyikét választhatjuk, ezzel az észrevétellel tehát a 13. ábrából 15 új bűvös négyzetet kapunk, nyilvánvalóan csupa olyat, amit a kifúrt lemez szimmetriáival nem kaphatunk meg. Az ilyen kövezetet adó bűvös négyzetet *minden átlóján bűvösnek* mondjuk, idegen szóval *pándiagonálisnak*. (Ha meg akarjuk vizsgálni, hogy egy bűvös négyzetünk pándiagonális-e, nem szükséges belőle kövezetet felírnunk, bár ennek vizsgálatához elég egy

7×7

mezős része. Ugyanis a keretezés máshová csúsztatásával előálló négyzetek átlós számnégyeseit a 13. ábrán is összekereshetjük. Pl. bal felső sarokszámnak az 5-öst véve, tőle jobbra lefelé találjuk a 2-t, ez lesz a főátló második száma. Még egy ilyen lépéssel átlépnénk a jobb szomszéd kőre, amelyet most elég, ha odagondolunk, hiszen azon ismét a 13. ábra szerinti az elrendezés. Így a főátló 3. száma a következő sor kezdő száma a 12, 4. száma pedig az alsó szomszéd kőre gondolva a felső sor 2. száma, a 15. A mellékátló számait hasonlóan kereshetjük össze, kiindulva az 5-ös bal szomszédjából, a 4-esből, és balra lefelé lépegetve. Egy találó név segít a megjegyzésben: az ilyen mezőnégyeseket *megtört átlóknak* szokás nevezni.)

A 13. ábrából fentebb származtatott 16. és 19. ábrák nem örökölték a pándiagonális tulajdonságot. Be lehet bizonyítani, hogy negyedrendű pándiagonális bűvös kövezeten bármelyik számtól bármelyik átlós irányban indulva a második helyen az aritmetikailag tükrös párját találjuk, pl. a 22. ábrán mindegyik 2-től jobbra föl, jobbra le, balra le és balra föl két lépésre a 15-ös számot. (Vegyük pl. a főátlót 2-szer, az 1. és 3. oszlopot, a bal felső sarokból balra lefelé induló megtört átlót, majd vegyük el a 2. és a 4. sort, valamint az 1. sor 3. mezejéből balra lefelé induló megtört átlót; mindennek a 20-21. ábrák módján való egyszerű nyilvántartása mutatja, hogy a főátló 1. és 3. száma összegének 4-szeresét kapjuk, és ez a bűvös állandó 2-szeresével egyenlő, mert az állandót előbb 5-ször vettük, majd 3-szor elhagytuk.)

Apróságok. A negyedrendű bűvös négyzetekben nem tapasztaltuk a páros és páratlan számok olyan különválását, mint a harmadrendűben. Sőt itt minden aritmetikai tükrös számpár tagjai „ellentétes párosságúak”, pl. 3 és 14. Hogy helyes képet kapjunk, megjegyezzük, hogy a harmadrendű négyzetekben is csak addig áll a „párosakat a sarokra” elv, amíg az elrendezendő számokat 1-nél vagy más páratlan számnál kezdjük. Ha a 0-t vesszük a legkisebb számnak – vagyis a sorozat eleje is, vége is páros –, akkor a 7. ábra minden száma helyére az 1-gyel kisebb szám lép (23. ábra), és minden párosság ellentétesre fordul.

7	0	5
2	4	6
3	8	1

23. ábra

Valóban, nemcsak 1-nél szokás kezdeni az elrendezendő számokat, és nemcsak egymás utáni egész számokból szokás bűvös négyzeteket képezni, lehet haladni nagyobb lépésekkel, néhol kihagyásokkal. Ilyenek az 5–6. ábrák is. (Hogyan készültek a 7. ábrából?) Persze csak akkor tetszetős az eredmény, ha a beírt számok valami jól észrevehető szabályszerűséget mutatnak. Képeztek azonban bűvös négyzeteket pl. egymás utáni törzsszámokból is.

A 23. ábrát kissé agyafúrt módon így is értelmezhetjük: a 7. ábrához hozzáadtuk azt a harmadrendű bűvös négyzetet, amelynek minden száma –1. A hozzáadáson azt értettük, hogy az ugyanazon helyen álló számpárokat adtuk össze, pl. a 2. sor 3. helyén álló 7-hez a másik négyzet 2. sora 3. helyén álló –1-et, így jött létre a 23. ábra 2. sorának 3. mezején a 6-os szám. – Ezt a gondolatot hamarosan hasznosítani fogjuk.

Bűvös négyzetek belső szerkezete. A 23. ábrán a 0–8 számok éppen a 3-as rendszámhoz hozzáillő 3-as számrendszer öszszes, két jeggyel leírható számai, ugyanis a

0,1,2,3,4,5,6,7,8
szám 3-as számrendszerbeli alakja

00,01,02,10,11,12,20,21,22
(az első hármat egy elöl álló 0-val írtuk, mert célszerűbb, ha mindegyik számunk kétjegyű). Pl.

$7=2\cdot 3+1$
, ebből a 3-as alapszámot és a szorzás, összeadás jelét a szokásos rövidítésben nem írjuk ki, ahogyan a „huszonegy"-ben is

$2\cdot 10+1$
-et gondolunk. Itt „kettő-egy"-nek olvassuk a 21-et. A 3-as számrendszer csak a 0, 1, 2 számjegyeket használja, a „három" írása már „egy–nulla", azaz 1 nagyobb egység, „egyes" pedig nincs.

Írjuk új számaink új alakját a 23. ábra helyére, azután mindjárt válasszuk szét az első és második – más szóval hármas és egyes helyiértékű – jegyeket két egyforma négyzet megfelelő mezőire (24-26. ábrák)!

21	00	12	‖	2	0	1	1	0	2
02	11	20	‖	0	1	2	2	1	0
10	22	01	‖	1	2	0	0	2	1

24. ábra *25. ábra* *26. ábra*

Mindkét rész-négyzet minden bűvös vonalán az összeg ugyanaz: 3. Csakhogy ebben nincs semmi érdekes, minden vonalon egy 0, egy 1-es és egy 2-es áll, csupán az egyik-egyik átlón látunk

$1+1+1$
-et. A részábrák azt mutatják, hogy ha volna 3 Ft-os papírpénz, és a 23. ábrát ilyenekből és 1 Ft-osokból raknánk ki úgy, hogy a forintosokat – mihelyt 3 van belőlük – papírpénzre kell váltanunk, akkor az ábra minden vonalán papírpénzben is, 1 Ft-osokban is ugyanannyi pénz állana. – Ez hát a bűvösség „titka".

Nem is lehet máshogy 3-féle különböző tárgyból 3–3 darabot úgy elrendezni 3 sorban és 3 oszlopban, hogy mindegyik sorban és mindegyik oszlopban csupa különböző tárgy álljon, csak a 25-26. ábrák módján. Ekkor persze nem beszélünk összegről. Ez a két ábra viszont egymásból is előállítható, hiszen ilyen amulett is készíthető.

Többször is találkozunk majd hasonló szabályszerűségekkel, ezért érdemes megállapodni a következő beszédmódban: a 24. ábra a 25-26. ábrákból mint *összetevőkből*, *egyesítéssel* adódik.

A két összetevő felcserélésével ismét bűvös négyzetet kapunk, a 24. ábrának a függőleges tengelyre való tükörképét. Ha viszont a 25. ábrát önmagával egyesítenénk, nem kapnánk igazi bűvös négyzetet, mert nem minden szám lenne különböző; ugyanígy egyedül a 26. ábrából sem. A helyes egyesítésekben azért áll elő csupa különböző szám, mert a 25. ábra bármelyik 3 egyező számjegye helyén a 26. ábrában csupa különböző számjegy áll, és persze megfordítva is.

Hasonló szabályszerűségeket látunk, ha a 11-13. ábrák 4-ed-rendű bűvös négyzeteiben is a rendre 1-gyel kisebb számokat vesszük, éspedig most a 4-es számrendszerbeli alakjukban, pl. 11 helyére

11=2□4+3
-at, rövidítve 23-at (olvasd kettő –három), majd ezeket is felbontjuk 4-es (4 Ft-os) és 1-es (1 Ft-os) helyiértékű számjegyeik külön négyzeteire (27-29. ábrák).

00	32	31	03	0	3	3	0	0	2	1	3
23	11	12	20	2	1	1	2	3	1	2	0
13	21	22	10	1	2	2	1	3	1	2	0
30	02	01	33	3	0	0	3	0	2	1	3

27. ábra

10	22	31	03	1	2	3	0	0	2	1	3
33	11	02	20	3	1	0	2	3	1	2	0
00	32	21	13	0	3	2	1	0	2	1	3
23	01	12	30	2	0	1	3	3	1	2	0

28. ábra

00	32	21	13	0	3	2	1	0	2	1	3
31	03	10	22	3	0	1	2	1	3	0	2
12	20	33	01	1	2	3	0	2	0	3	1
23	11	02	30	2	1	0	3	3	1	2	0

29. ábra

A 27. ábra két rész-négyzete is átvihető egymásba

90□
-os forgatással; ugyanez áll a 29. ábrából képezhető bűvös kövezet összetevőire, de magukra a négyzet összetevőire nem. A 28. ábra összetevőinek átlóiban 7 és 5, illetőleg 2 és 10 az összeg, ezeket „fordított helyiértékkel” egyesítve csak félig bűvös négyzeteket kapnánk. Ezek kevésbé szabályos alakzatok, de szintén érdekesek. Ha viszont a 27. és 29. ábra összetevőit egyesítenők a helyiértékek felcserélésével, bűvös négyzeteket kapnánk, helyes átlókkal. (Kérdés, újak-e ezek számunkra?)

Könnyű képezni az összes olyan négy-négy 0, 1, 2, 3 számjeggyel kitöltött

4×4
mezős négyzetet, amelyben mindegyik sor, oszlop és átló csupa különböző jegyet tartalmaz, és ezért az összegek is mindig egyenlők. Ezeket előállítva az eddigi elemzésből áttérhetünk új négyzetek képzésére. Jegyek helyett az

első sorba az A, B, C, D betűket írva B alá vagy C-t, vagy D-t kell írnunk, mert az átló miatt A-t már nem írhatunk (30-31. ábra). Mindegyik változatból egy-egy befejezést kapunk, mert ennyiből mindegyikben egymás után adódik a főátló kitöltése, a 2. oszlopé, majd a mellékátlóé, a 4. soré, a 3. oszlopé, végül a még hátralevő mezőké. E két négyzet egyesíthető csupa különböző betűpárt tartalmazó négyzetté (a 32. ábrán a 31. ábra betűit minden mezőn másodiknak írtuk, és megkülönböztetésül A, B, C, D helyett X, Y, Z, V-t használtunk).

A	B	C	D
D	C	B	A
B	A	D	C
C	D	A	B

30. ábra

A	B	C	D
C	D	A	B
D	C	B	A
B	A	D	C

31. ábra

AX	BY	CZ	DV
DZ	CV	BX	AY
BV	AZ	DY	CX
CY	DX	AV	BZ

32. ábra

Azért használtunk betűket, mert így a 32. ábra 1152 bűvös négyzet közös „képlete”. Ugyanis az A, B, C, D betűk helyére a 0, 1, 2, 3 számjegyeket $4\cdot3\cdot2=24$-féleképpen írhatjuk vissza, ugyanígy X, Y, Z, V helyére is, végül 2-féleképpen választhatjuk meg, hogy melyik betűnégyes adja meg a négyes helyi értékű számjegyeket, és ez a $24\cdot24\cdot2=1152$ behelyettesítés csupa különböző bűvös négyzet. Tessék megpróbálni! 8-asával 1-1 amulett-lemezzel volnának megadhatók.

Az önmagukban is érdekes 30-31. ábrák (és magasabb rendű megfelelőik) más problémákban is fellépnek, pl. bástya-probléma a sakktáblán. *L. Euler* híres matematikus (1707–1783) is foglalkozott velük, *latin négyzeteknek* nevezte őket. Néha az átlók követelményét nem teljesítő efféle elrendezéseket is így nevezik (25-26. ábra). A 32. ábrát (és magasabb rendű megfelelőit) pedig *Euler-négyzeteknek* nevezik. (Euler sokat kutatott ilyen 6-odrendű négyzetek után; csak 1900-ban bizonyították be, hogy 6-odrendű Euler-négyzet nem létezik, de tizedrendűt találtak már (*E. T. Parker* 1957; *L. Weisner* 1959).)

Át lehetne írni betűkre a 27. és 29. ábrák összetevőit is, ezekben a számjegy-visszahelyettesítési lehetőségek száma kisebb. A 4-edrendű bűvös négyzetek kérdését jelentősen megkönnyítette *Fitting* német matematikusnak az az észrevétele (1931), hogy mivel a 4 maga is négyzetszám: 2^2, ezért a négyes és egyes helyiértékű összetevő-négyzetek tovább felbonthatók a kettes számrendszerben. Pl. a 27. ábrán már felbontott 11. ábra további bontása a 33. ábra négy összetevője, helyiértékük sorra

23=8
, 4, 2, 1. Bizonyára máris látja az olvasó, hogy ezek más sorrendben is bűvös négyzetet adnak, továbbá hogy bármelyik összetevőben minden 0 helyett 1-et írva és minden 1-es helyett 0-t: újabb és újabb 4-edrendű négyzeteket írhatunk fel. Vannak köztük számunkra újak is. Ilyen átalakításban az egyesíthetőség feltételei bonyolultabbak, mégis sokkal szebb az összes bűvös négyzet számának, szerkezetének megállapítása, és főleg sokkal kevesebb munkával jár, mint ahogyan ezt régen sok lélekölő próbálgatással – és mégsem megnyugtató módon – végezték.

0	1	1	0	0	1	1	0	0	1	0	1	0	0	1	1
1	0	0	1	0	1	1	0	1	0	1	0	1	1	0	0
0	1	1	0	1	0	0	1	1	0	1	0	1	1	0	0
1	0	0	1	1	0	0	1	0	1	0	1	0	0	1	1

33. ábra

Eljárások bűvös négyzet képzésére, bármely páratlan rendszámhoz. Készítsünk bűvös kövezetet a 7. ábrából (34. ábra)!

3	5	7	3	5	7	3	5	7
4	9	2	4	9	2	4	9	2
8	1	6	8	1	6	8	$\boxed{1}$	6
3	5	7	3	5	$\boxed{7}$	3	5	7
4	9	2	$\boxed{4}$	9	2	4	9	2
8	1	6	8	1	6	8	1	6
3	5	7	3	5	7	3	5	7

34. ábra

Erről leolvashatunk néhány olyan szabályszerűséget, amelynek általánosításaképpen már régen kimondtak néhány „szabályt” bármely páratlan rendszámhoz bűvös négyzet képzésére.

Megjelöltük a kövezet egyik 1-esét, majd a hozzá legközelebbi 2-est, az ehhez legközelebbi 3-ast, az ehhez legközelebbi 4-est és így tovább 9-ig. Majdnem mindig 1-et jobbra és 1-et fölfelé lépve találjuk meg az 1-gyel nagyobb szám legközelebbi előfordulását, kivétel csak a 4-es és a 7-es, ezeket a 3-asból, illetőleg a 6-osból 1-et lefelé lépve találjuk, vagyis egymással ismét megegyező elv szerint. Nem is lehet a 3-asból és a 6-osból jobbra fölfelé lépni, mert ott már a kisebb 1-es, illetőleg 4-es szám egy másik előfordulása áll.

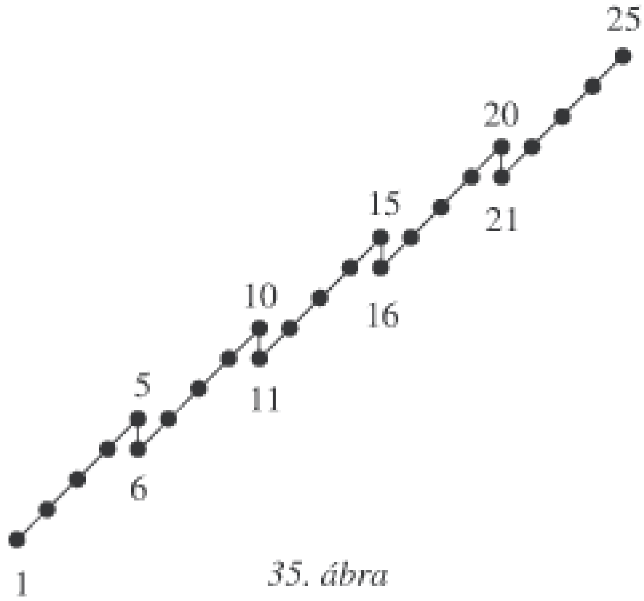
Kipróbálhatjuk, hogy bármely páratlan rendszámhoz bűvös kövezetet kapunk a következő előírások alapján:

I. *Számainkat növekvő sorrendben, egymás után írjuk be a kövezetbe*; ezt úgy értjük, hogy egyidejűen a kövezet minden kőlapjára beírni gondoljuk a soron következő számot, vagyis pl. az első beírt 1-est még a 2-esek beírása előtt megismételjük jobbra és balra haladva annyi mezőnként, ahányad rendű kövezetet képezünk, majd minden bejegyzésünkből kiindulva fölfelé és lefelé is.

II. *Az egymás után következő számok első példányát a közvetlen előző szám első példányához képest mindig ugyanolyan elmozdulás – az ún. lépés – után írjuk be, ha ott még üres mezőt találunk.* A lépés: 1 mezővel jobbra és egyidejűen 1 mezővel fölfelé.

III. *Ha a lépés már foglalt mezőre vinne, akkor a soron következő számot egy másfajta elmozdulás – az ún. ugrás – után helyezzük el; majd ismét lépésekkel haladunk tovább mindaddig, míg ismét ugrásra nem* *kényszerülünk.* Az ugrás: 1 mezővel lefelé.

A 35. ábrán az így képezett 5-ödrendű bűvös kövezetből csak a számok első bejegyzéseinek helyeit látjuk kis pontokkal és a menetvonallal.



35. ábra

Ha úgy választjuk meg a kövezeten a négyzet határát, hogy a beírandó számok felsorolásának középső tagja – pl. 5-ödrendű négyzetben az

1+52

összeg fele, a 13-as szám – a négyzet középső mezején álljon, akkor az átlókon is a bűvös állandót kapjuk összegként, tehát bűvös négyzetet vágtunk ki a kövezetből.

Régebben ezt az eljárást (és a továbbiakat is) csak egyetlen négyzet kitöltésére szorítkozva mondták ki, ezért elsőnek az 1-es szám helyzetét adták meg, így: az 1-est a felső sor középső mezejére írjuk. A fenti II. és III. előíráshoz kapcsolódó „takarékossági" utasítás így szólt (lásd 5-ös rendszám esetére a 35. ábrát!): ha a lépés átvitt a négyzet jobb oldali határvonalán, akkor az így elért külső mezőről úgy térünk vissza a négyzetbe, hogy a rendszámmal egyenlő számú mezővel balra lépünk (pl. a 4-es esetében); hasonlóan ha fönt futottunk ki (pl. a 2-essel), akkor az illető oszlopban lent juthatunk vissza a keretbe. Látjuk, hogy ezek éppen a kövezetképzés elveinek a visszafordításai. A régiek így papírt takarítottak meg, nekünk viszont szabadabban mozog a kövezeten a szemünk és a képzeletünk, és több szabályszerűséget vehetünk észre.

A fenti, talán 1000 évesnél is régebbi eljárást *sziámi* vagy *indus*, vagy *kínai szabálynak* nevezik.

A 34. ábra kövezetét persze más lépéssel, más ugrással is képezhettük volna. Bemutatunk két, ugyancsak régen felismert hasonló eljárást, amelyek kissé merészebb lépést vagy ugrást használnak; mindkettő leolvasható a 34. ábráról, 5 és magasabb rendszám esetén viszont már a 36. ábrától, illetőleg a megfelelő indus négyzettől különböző bűvös négyzetet állítanak elő.

9	11	18	25	2	9	11
15	17	24	1	8	15	17
16	23	5	7	14	16	23
22	4	6	13	20	22	4
3	10	12	19	21	3	10
9	11	18	25	2	9	11
15	17	24	1	8	15	17

36. ábra

Az ún. *„sakktábla-szabály”* szintén az indus lépést használja, ugrása: 2-t fölfelé, az 1-es helye pedig a négyzet középső mezeje fölötti mező (ez 3-adrendű négyzetben egyszersmind a felső sor középső mezeje is). *Bachet* francia matematikus írta le az eljárást kissé más módon 1600 körül.

A XIV. század közepe óta ismert *„lóugrás-szabály”* lépése: 1-et jobbra, 2-t lefelé, ugrása: 4-et lefelé, kiindulás itt is a felső sor középső mezejéről (*Moszkhopulósz* görög kutatótól). Ez 5-öd- és 7-edrendben pándiagonális bűvös kövezetet állít elő, 9-edrendben viszont nem.

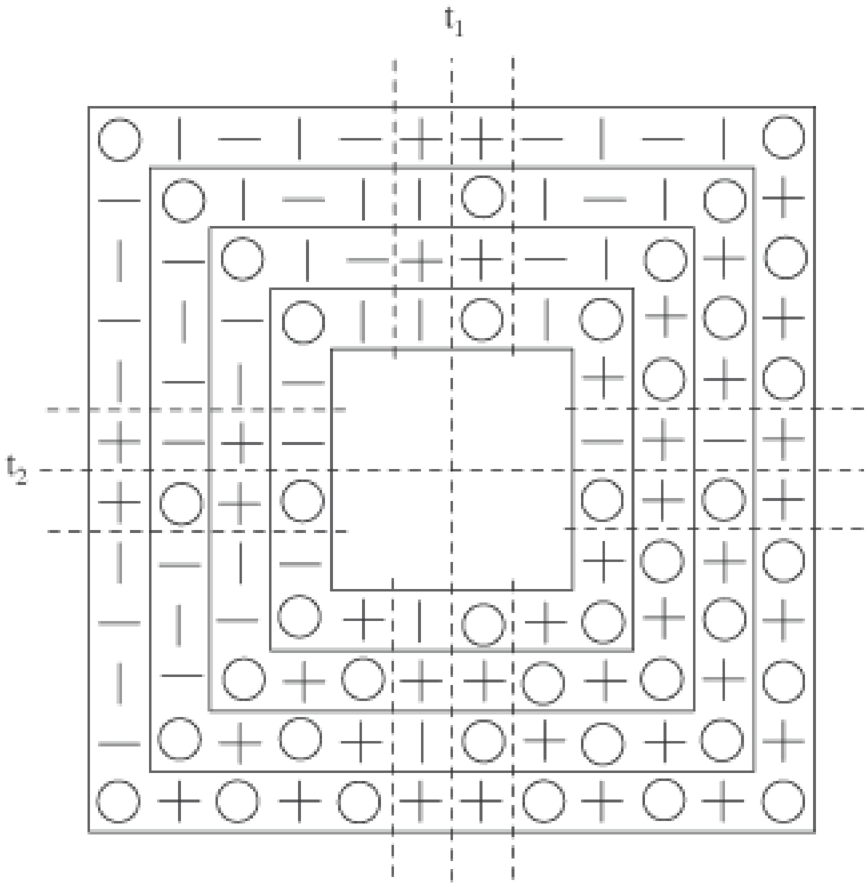
Végtelen sok ilyen lépés–ugráspárt lehet megadni, de ezek csak akkor adnak bármely páratlan rendszámhoz bűvös négyzetet, ha a lépés vízszintes és függőleges része, valamint az ugrás is

$2n$
mezőnyi, ahol

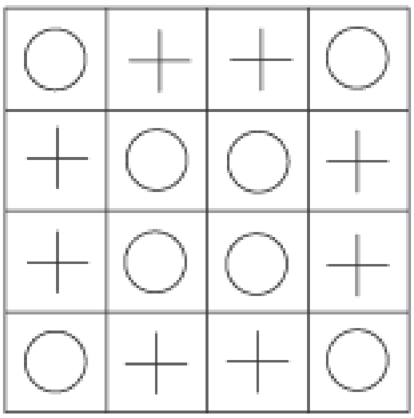
n
pozitív egész szám vagy 0. Messze vezetne ennek bizonyítása, ezért csupán megemlítjük, hogy mindig két latin négyzetet kapunk, ha az elrendezés összes számát 1-gyel csökkentjük, majd átírjuk a rendszámmal egyenlő alapú számrendszerbe – amivel mindig éppen az illető számrendszer összes, két jeggyel írható számát kapjuk –, végül a számjegyeket helyi értékenként két négyzetbe különválasztjuk. Az egyik átlón azonban rendszerint csupa egyenlő számjegy áll, a nagyságra nézve középső számjegy. Egy másik könnyen kimondható eleme a bizonyításnak, hogy a képzésben ugrás mindig a rendszám többszörösei után válik szükségessé (a 36. ábrán 5, 10 stb. után), és hogy az utolsó bejegyzés utáni ugrás – amire persze már nincs szükség – mindig 1-essel elfoglalt mezőbe vinne.

Páros rendszámú bűvös négyzetek képzése. Minden 4-nél nagyobb, páros rendszámhoz képezhetünk egy bűvös négyzetet a 37–39. ábrákon bemutatott *felcserélési táblázatok* alapján. (*Devedec*-eljárás). A 37. ábrán a 2-esével növekvő rendszámokhoz (6-tól kezdve) egy-egy keretet látunk, ez lesz a megfelelő bűvös négyzet határa, a 38–39. ábrákon pedig egy-egy

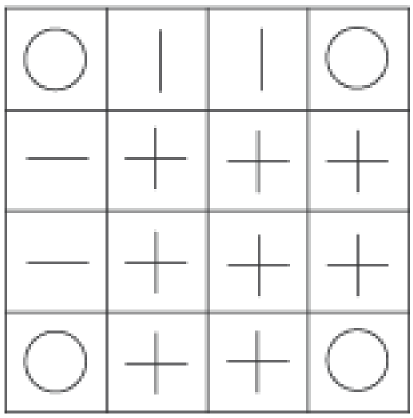
4×4
mezős négyzetet. Ha a rendszám osztható 4-gyel, akkor a 37. ábra közepébe, *magjába* a 38. ábrát gondoljuk bemásolva, ha pedig a rendszám 4-gyel osztva maradékul 2-t ad, akkor a 39. ábrát. Az ábrákon levő jelek bizonyos számhelyettesítéseket jelölnek.



37. ábra

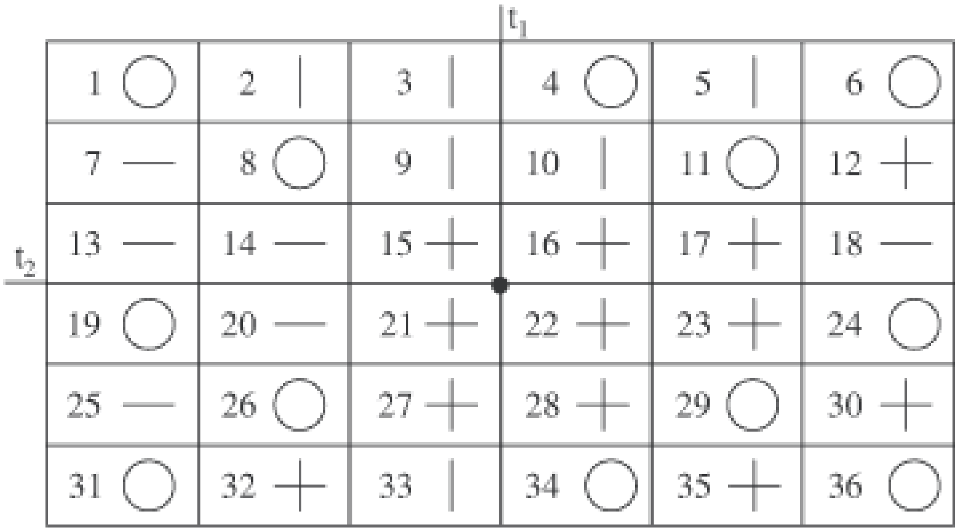


38. ábra



39. ábra

A kitöltendő négyzet mezőibe először a 9-10. ábrák mintájára beírjuk az elrendezendő számok alapállását. Ezután minden egyes szám mellé átmásoljuk a 37. ábra megfelelő keretén belül a megfelelő helyen található jelet, a 38., illetőleg a 39. ábra jeleit is. A jelek arra adnak utasítást, hogy a mellettük álló szám helyére mely más mezőről hozzuk át az ott található számot. Az összes ilyen mozgatást elvégezve bűvös négyzetet kapunk. Az eljárást 6-odrendű négyzetre előkészítve a 40. ábrán mutatjuk be, az eredményt pedig a 41. ábrán.



40. ábra

1	32	33	4	35	6
12	8	27	28	11	25
18	17	22	21	20	13
19	23	16	15	14	24
30	26	10	9	29	7
31	5	3	34	2	36

41. ábra

A jelek jelentése: a kis kört tartalmazó mezők számai a helyükön maradnak. Az „–” jelű mezőkre arról a mezőről hozzuk át az ott talált számot, amely mezőnknek tükrös párja a négyzet függőleges

t_1

tengelyére, pl. 7 helyére a 12-t. A „

□

” jelű mezőket hasonlóan a vízszintes

t_2

tengelyre tükrös párjukról töltjük be, pl. 2 helyére a 32-t írjuk. Végül a „

+

” jelű mezőkre a négyzet középpontjára nézve tükrös száma lép, pl. 12 helyére 25, ami egyébként mindig egyszersmind az aritmetikai tükörképe is. A példában minden szám vagy a helyén maradt, vagy páros cserében vett részt, pl. 13 és 18, vagy pedig beletartozik egy négytagú csoportba, amelynek tagjai körben egymás helyére lépnek, pl. 2 a 35 helyére, 35 az 5 helyére, 5 a 32-ére, ez pedig a 2 helyére.

12-nél nagyobb rendszámhoz a 37. ábrát ki kell terjesztenünk a könnyen felismerhető szabályszerűségek alapján: az átlók mentén csupa kör áll, a

t_1

,

t_2

tengelyeket közrefogó sorok, oszlopok jelei minden második keretben ismétlődnek, a mellékátlótól balra fölfelé levő többi mezőn a „–” és „

□
” jelek váltakoznak sakktáblaszerűen, a jobbra lefelé levő mezőkön pedig körök és

+
jelek.

Bűvös transzformációk 4-nél magasabb rendszámú bűvös négyzetekre. A 4-edrendű négyzeteknél megismert bűvös transzformációkhoz hasonlóan minden nagyobb rendszámú bűvös négyzetből újabbakat képezhetünk a következők szerint:

I. Az első bűvös transzformáció mintájára felcserélünk olyan két sort, amelyek egymás tükrös párjai a négyzet vízszintes tengelyére, majd felcseréljük az ugyanilyen sorszámú oszlopokat is.

II. A második bűvös transzformáció mintájára felcserélünk egymással két nem tükrös helyzetű sort és egyidejűleg a tükörképeiket is egymással, és ugyanezt végezzük az ugyanolyan sorszámú oszlopokkal is. Páratlan rendszám esetén a négyzetnek van középső sora és középső oszlopa, ez a mondott cserékben nem vehet részt, mert a tükörképe önmaga.

Kitekintés további problémákra. Korántse higgyük, hogy mindent láttunk a bűvös négyzetekről. Megemlítünk két olyan eljárást, amely kész bűvös négyzetek felhasználásával képez magasabb rendűeket. Az egyik, az ún. szorzási eljárás, két tetszés szerinti rendszámú bűvös négyzet felhasználásával olyat állít elő, amelynek rendszáma amazok rendszámainak szorzata, pl. egy 3-ad- és egy 4-edrendű négyzet felhasználásával egy

3□4=12
-edrendű bűvös négyzetet. Ennek az elvnek a további elemzésével rájöttek, hogy lehet a rendszámot 2-vel is megszorozni, bár 2-odrendű bűvös négyzet nem létezik (*Aubry*, francia, 1930 körül).

A másik eljárás egy bűvös négyzetből „csak” 2-vel magasabb rendszámút képez úgy, hogy minden oldalán 1 mezőnyi széles *szegélyt* készít köréje részben az újonnan belépő számokból, részben az elhelyezendő sorozat legkisebb számaiból, a *magban* levő eredeti bűvös négyzet számait pedig megfelelően ugyanannyival emeli (*M. Stifel*, német, 1550 körül).

L. Bieberbach német matematikus 1954-ben általános eljárást adott a magasabb rendszámú szegélyek egyszerű kitöltésére, de már azóta is adódtak ezen a téren egyszerűsítések.

Képeztek olyan bűvös négyzeteket, amelyekben a számok helyére a négyzetüket vagy a köbüket írva ismét bűvös négyzet adódik. (Egész számokat és négyzeteiket használva az eddig ismert ilyen négyzetek legkisebbike 8-adrendű.)

A bűvös négyzetek térbeli megfelelőiként bűvös kockák is képezhetők. Ezek a hasonló problémák újabb légióit vetik fel.

Az első kiadás óta eltelt 30 év szinte kötelezi a még élő szerzőket, hogy tegyenek hozzá valamit korábbi cikkükhöz, figyelembe véve annak témakörében az időközben elért újabb eredményeket.

Ilyenek pl. a „nagy szabályok” nyomán készült, vagy véletlenül észrevett mikrováltozásai (vadhajtásai? nemesítései?) a bűvös négyzeteknek, ahogyan a 42. ábra bűvös négyzete többféleképpen is átmehet egy-egy másikba négy-négy szám páronkénti helycseréjével.

21	8	19	12	5
9	15	1	23	17
11	24	7	20	3
10	16	13	4	22
14	2	25	6	18

42. ábra

A 8 és 2 helycseréje „eltorzít” két sorösszeget, de a 19-25 csere korrigálja. Tetszetősebben: a $8+19$ „dominó” cserél a $(2+25)$-tel. Vagy a $20+6$ a $(24+2)$-vel. A sort lehetne még folytatni. Ez a „nyitás” az 5×5-ös méretben „indul” igazán; kilátásai áttekinthetetlenek (talán 800 ezer példány), a 6×6-osban pedig kb. félmilliárd! a (lényegesen) különböző bűvös négyzetek száma.

Befejezésül még egy különlegességét említjük meg az olyan 4×4-es bűvös négyzeteknek, amelyeknek „sarkaiban” is 34 az ott levő négy szám összege. Ilyen négyzet látható a 43. ábrán: $7+10+16+1=4+13+11+6=9+8+2+15=14+3+5+12$.

7	10	4	13
16	1	11	6
9	8	14	3
2	15	5	12

43. ábra

Jelölje a a táblázatban szereplő valamelyik számot $(a=1,2,\ldots,16)$! Hajtsuk végre a következő transzformációt:

(□)	$a\to2a$, ha $a\le8$,

	$a \to 2a - 17$, ha $a > 8$.

Így a következő négyzethez jutunk (44. ábra):

14	3	8	9
15	2	5	12
1	16	11	6
4	13	10	7

44. ábra

Könnyű ellenőrizni, hogy ez is olyan bűvös négyzet, amelynek „sarkaiban” is 34 az ott levő négy szám összege. Ha ez utóbbi bűvös négyzet számaira alkalmazzuk a (

□

) transzformációt, szintén hasonló tulajdonságú bűvös négyzetet kapunk. Végül: a (

□

) transzformáció nyolcadik végrehajtása után visszakapjuk az eredeti, a 43. ábrán feltüntetett bűvös négyzetet.

IRODALOMJEGYZÉK

[1] [1] Berger György: *Bűvös négyzetek*, Dacia Könyvkiadó, Kolozsvár-Napoca, 1986.

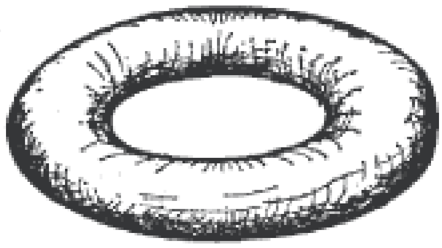
[2] [2] Bakos Tibor: *Ki tud többet a bűvös négyetekről?* I. füzet, Eötvös Loránd Fizikai Társulat, Budapest, 1998.

CSALAFINTA FELÜLETEK

BOGNÁR MÁTYÁS

1. BEVEZETÉS

A köznapi szóhasználat felületen egy térrészt (testet) határoló alakzatot ért. Egy golyónak a felülete gömbfelület, a mentőöve tóruszfelület (lásd az 1. ábrát),



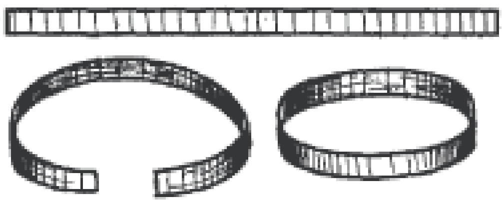
1. ábra

a hengert határoló felület a henger palástjából és két körlapból áll; a gúla felülete pedig a gúla palástjából és egy sokszöglapból tevődik össze. Az ilyen típusú felületeket zárt vagy nem határolt felületeknek szokás nevezni.

Felületnek hívjuk azonban gyakran a hengerpalástot vagy a félgömböt is. Ezek a felületek ugyan nem határolnak térrészt, mégis zárt felületekből származtathatók. Zárt felületnek egy vagy több zárt görbe által határolt összefüggő részét alkotják. Ezért határolt felület a nevük.

2. A MÖBIUS-SZALAG

Téglalap alakú szalagot (például ragasztószalagot) hajlítsunk meg és ragasszuk össze két szemközti oldalát (lásd a 2. ábrát)!

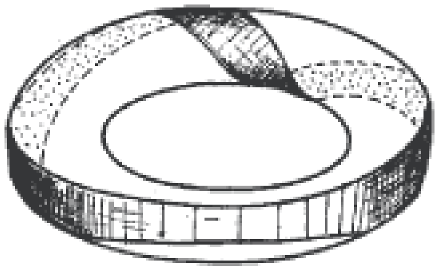


2. ábra

Ily módon határolt felületet nyertünk, hiszen a szalag nem más, mint egy henger palástja. Csavarjuk meg most kétszer a szalagot és ezután ragasszuk össze a két szélét (lásd a 3. ábrát)!

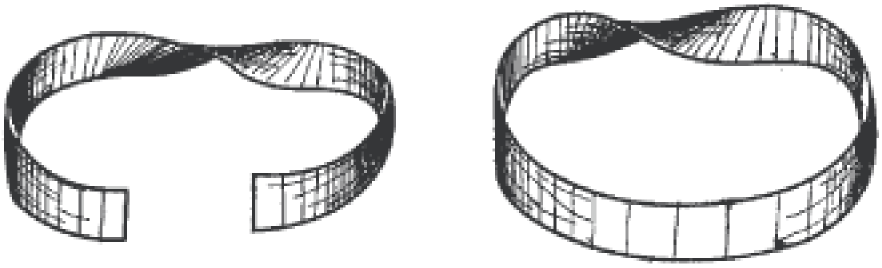
3. ábra

Újra határolt felületet kaptunk, hiszen az így nyert, kétszer megcsavart szalag a tóruszfelület határolt részeként fogható fel (lásd a 4. ábrát)!



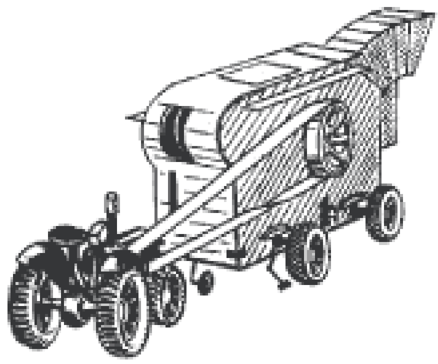
4. ábra

Csavarjuk meg ezek után egyszer a szalagot és így ragasszuk össze a két szélét (lásd az 5. ábrát)!



5. ábra

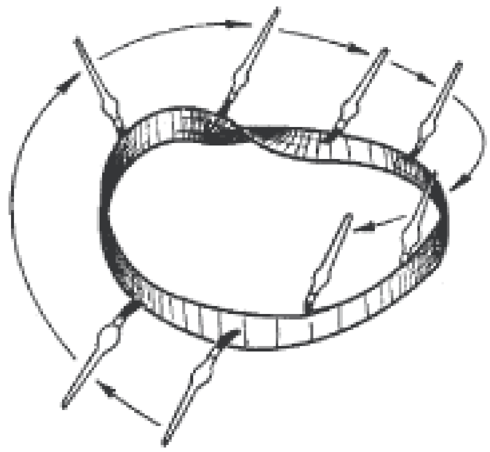
Az így nyert alakzatot először Möbius és Listing német matematikusok írták le 1862-ben, illetve 1865-ben. Felfedezőjéről Möbius-szalagnak hívjuk. Ilyen szalagot használnak például időnként cséplőgép meghajtásánál (lásd a 6. ábrát)!



6. ábra

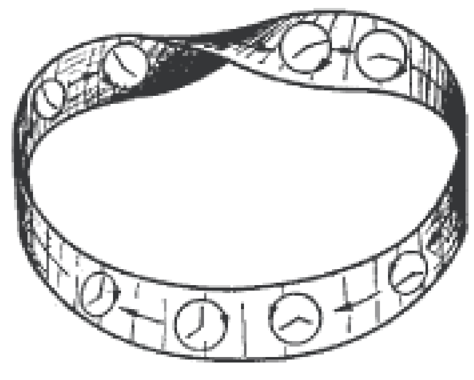
Felvetődik a kérdés, vajon a Möbius-szalag is határolt felület-e? Más szóval van-e olyan zárt felület a térben, amelyből egy vagy több zárt görbe Möbius-szalagot metszhet ki?

Belátjuk, hogy nincs ilyen zárt felület. Ha ugyanis volna, akkor ennek az általa határolt térrész felé eső oldalát pirossal, a külső oldalát pedig zölddel befestve, magán a Möbius-szalagon is két oldalt: egy piros és egy zöld oldalt különböztethetnénk meg. Könnyen meggyőződhetünk azonban arról, hogy ha a Möbius-szalagot egy helyen elkezdjük befesteni pl. piros színnel és a festést tovább folytatjuk; akkor anélkül, hogy a határvonalat átlépnénk; egy idő múlva kiindulási pontunk túlsó oldala is piros lesz; sőt ha tovább megyünk, akkor semmi sem marad befestetlenül (lásd a 7. ábrát)!



7. ábra

Más módszerrel is meggyőződhetünk arról, hogy a Möbius-szalag nem lehet része valamely testet határoló zárt felületnek. Ellenkező esetben ugyanis az átlátszó számlapú és burkolatú órát tetszés szerinti helyzetből elmozgatva, és valamilyen út megtétele után eredeti helyére visszahozva, az óra mutatói az út megtétele után mindig ugyanabban az irányban forognának, mint az út megtétele előtt. Márpedig, ha az órát a Möbius-szalag középvonala mentén toljuk végig, akkor eredeti helyére visszatérve éppen ellenkező irányban forognak az óramutatók, mint eredetileg (lásd a 8. ábrát)!

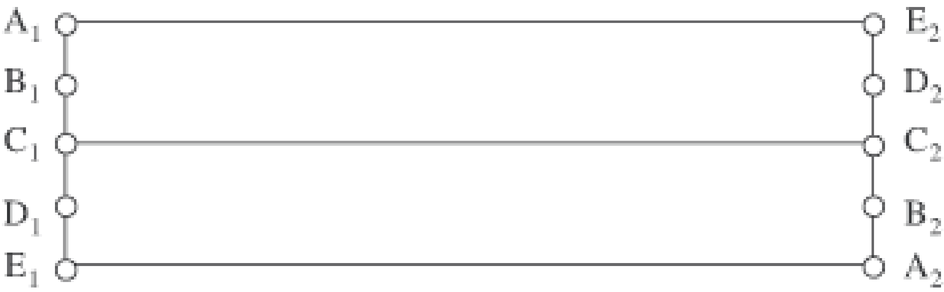


8. ábra

(Az óra járásának megfordulása természetesen csak látszólagos, hiszen míg előbb az óra számlapja nézett az ábra szemlélője felé, az út megtétele után az óra hátlapja van vele szemben; feltéve, hogy a szemlélő a szokásos módon néz az ábrára, azaz olyan irányból, amelyből annak feliratát egyenes állásúnak látja.)

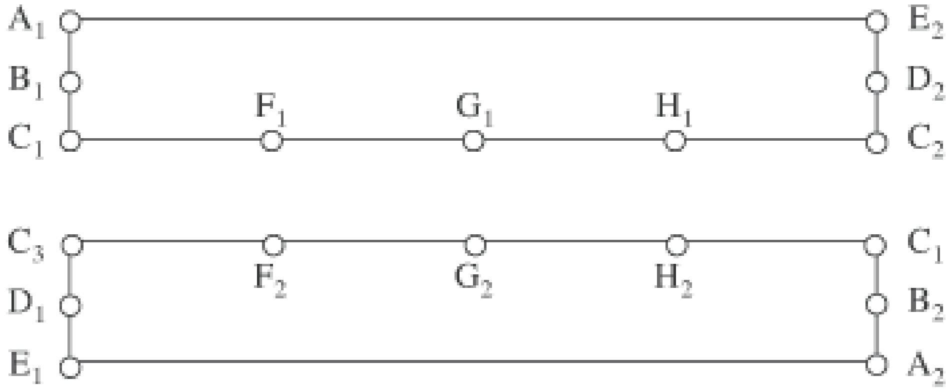
A Möbius-szalag tehát nem sorolható a bevezetőben leírt felületek közé. Ezért ki kell terjesztenünk a felület fogalmát. Előbb azonban magán a Möbius-szalagon végzünk el bizonyos átalakításokat.

Ha a Möbius-szalagot a ragasztás helyén szétvágjuk, ismét téglalapot nyerünk. Jelöljünk meg ezen a téglalapon néhány olyan pontpárt, amelynek elemei a Möbius-szalagon össze voltak ragadva! Ilyen például a 9. ábrán



9. ábra

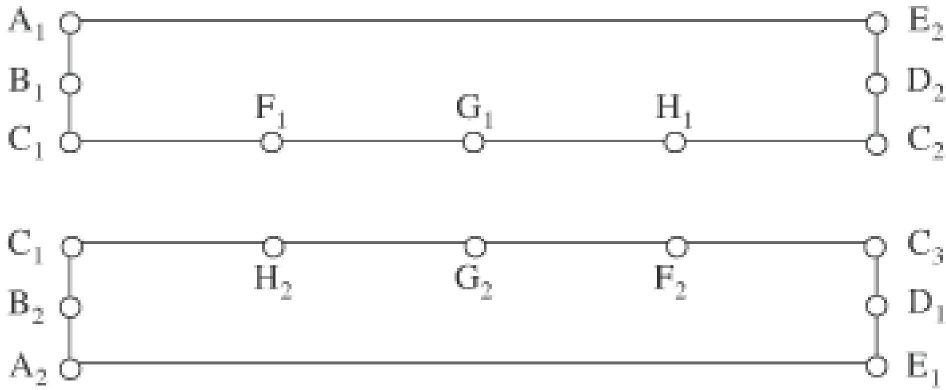
A_1 és A_2, B_1 és B_2, C_1 és C_2, D_1 és D_2, E_1 és E_2. Vágjuk el a téglalapot $\overline{C_1C_2}$ középvonala mentén és a „kettévágott” pontokat jelöljük ismét azonos betűkkel (lásd a 10. ábrát)!



10. ábra

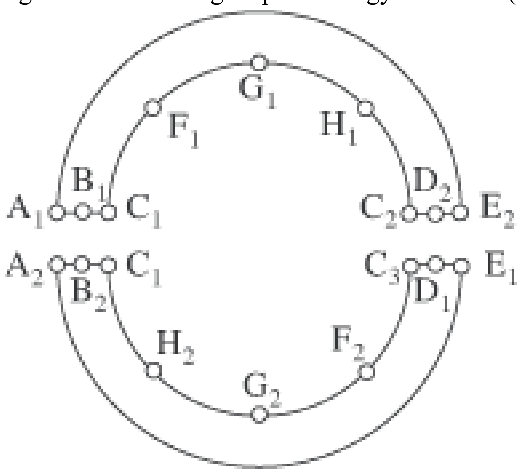
E két téglalapból az összetartozó pontok összeragasztásával nyerjük a Möbius-szalagot. Ezt az összeragasztást most más sorrendben fogjuk elvégezni, mint előzőleg.

Forgassuk át a 10. ábra alsó téglalapját a rövidebbik középvonala körül (lásd a 11. ábrát)!

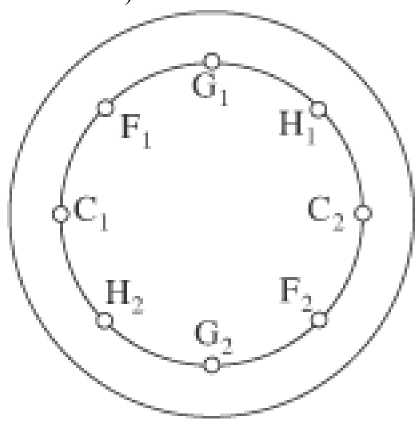


11. ábra

Hajlítsuk meg most mindkét téglalapot félkörgyűrű alakra (lásd a 12. ábrát)!



12. ábra



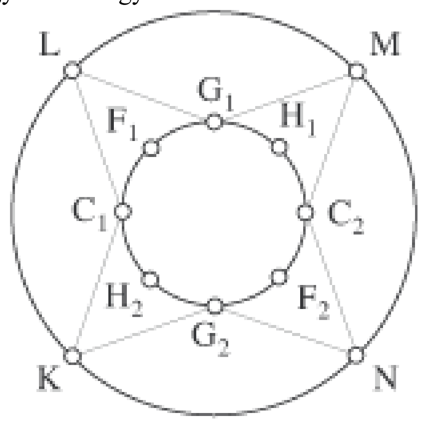
13. ábra

(Persze ehhez a téglalapoknak nemcsak hajlékonynak, hanem nyúlékonynak is kell lenniük. Anyaguk például plasztilin vagy rétestészta lehet.) Ha a két félkörgyűrűt összetoljuk, akkor éppen az összeragasztandó pontok esnek egybe, és így az alakzatot körgyűrűvé ragaszthatjuk össze (lásd a 13. ábrát).

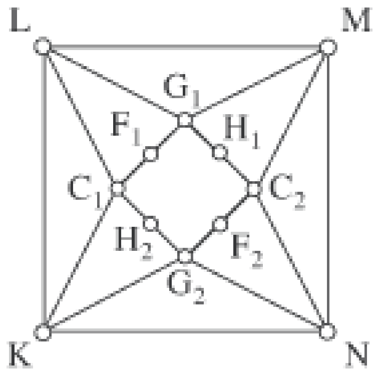
Ezek után már csak a körgyűrű belső körén kell összeragasztani minden pontot az átellenesével, hogy eljussunk a Möbius-szalag egy átalakított formájához.

Ezt az összeragasztást a háromdimenziós térben lehetetlen elvégezni. Ha azonban megengedjük, hogy az összeragasztás után nyert felület metszhesse önmagát, akkor már végrehajtható a ragasztás.

Osszuk fel ugyanis a körgyűrűt a 14. ábrán



14. ábra



15. ábra

látható módon (részben görbe vonalú) háromszögekre és egyenesítsük ki ezeket a háromszögeket (lásd a 15. ábrát)! Emeljük ki a síkból a

C_1
és

C_2
pontokat és közelítsük őket egymáshoz! Közelítsük egymáshoz a

G_1
és

G_2
pontokat is! A

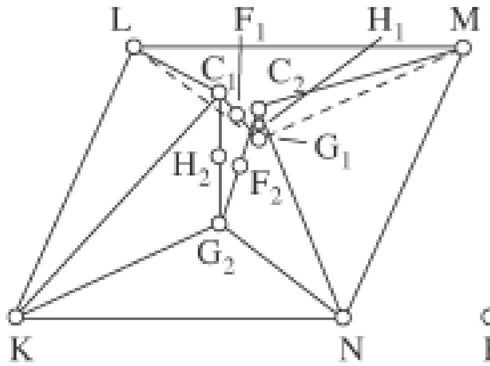
$C_1G_1C_2G_2$
négyszögből torz négyszög lesz (lásd 16. ábrát)! Ha végül is összeragasztjuk a

C_1
és

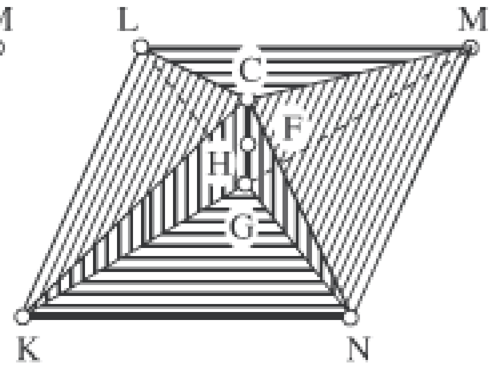
C_2
, valamint a

G_1
és

G_2
pontokat, és ezzel együtt az általunk kifeszített szakaszokat is (lásd a 17. ábrát),



16. ábra



17. ábra

akkor a 13. ábra belső körének átellenes pontjai egy helyre kerülnek, de feleslegesen is összeragadnak bizonyos pontok. Így például

F
és

H
ugyanaz a pont lesz. Ezen úgy segíthetünk, hogy a 17. ábra

$\overline{CG}$
szakaszának pontjait kettős pontoknak tekintjük. Így például az

F
pontot úgy fogjuk fel, hogy az rajta van az

LNC
háromszögön, de nincs rajta a

KMC
háromszögön, a

H
pont pedig pontja a

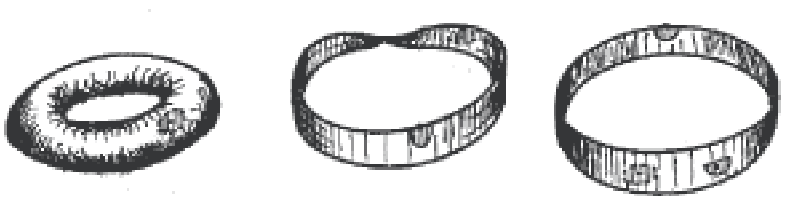
KMC
háromszögnek, de nem pontja az

LNC
háromszögnek.

3. A FELÜLET FOGALMA

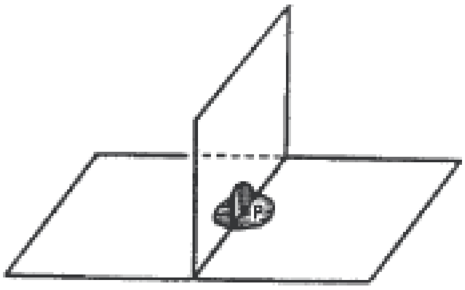
A továbbiakban a térrészt határoló zárt felületeket, és azoknak egy vagy több (ugyancsak zárt) görbe által határolt részeit kétoldalú felületeknek hívjuk.

Ennek megfelelően a felületeknek olyan definícióját adjuk meg, amely a kétoldalú felületeket is és a Möbius-szalagot is a felületek közé sorolja. Vizsgáljuk meg ezért először is, hogy mi a közös a kétoldalú felületek és a Möbius-szalag struktúrájában. Ha megnézzük azokat az alakzatokat, amelyek körülveszik egy kétoldalú felület vagy egy Möbius-szalag valamely pontját, akkor azt tapasztaljuk, hogy amennyiben a pont nincs a felület határvonalán, akkor az azt körülvevő alakzatok között lesznek olyanok, amelyek erős hasonlatosságot mutatnak a nyílt (határvonalától megfosztott) körlappal (lásd a 18. ábrát)!

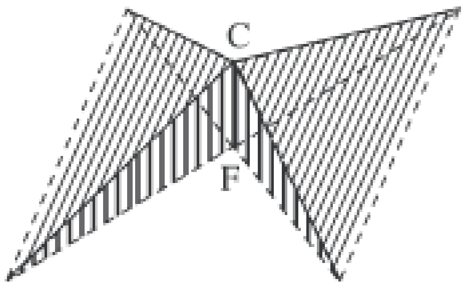


18. ábra

(A határvonalától megfosztott körlap olyan, mint a héjától megfosztott párizsi szelet.) Ha a pont a határvonalon fekszik, akkor ilyen nyílt körlapszerű alakzat nem veheti körül, de zárt (határvonalával együtt vett) félkörlaphoz hasonlító alakzat már igen. Persze zárt félkörlaphoz hasonlító alakzatok belső (nem határon fekvő) pontot is körülvehetnek (lásd a 18. ábrát)!



19. ábra



20. ábra

Tekintsük most a 19. ábra nyitott könyvét! Itt a

P
pontot körülvevő alakzatok egyike sem lesz a nyílt körlappal vagy a zárt félkörlappal azonos felépítésű. Mindegyik elágazik a

P
pontnál. Ezért ezt az alakzatot nem is soroljuk a felületek közé.

A felület szemléletes definíciója tehát a következő: A felületek olyan összefüggő alakzatok, amelyeken minden pontot körülvesz egy zárt félkörlappal azonos felépítésű részalakzat. A felület belső pontjai azok a pontok, amelyeket nyílt körlappal azonos felépítésű részalakzat is körülvesz. Ha a felület valamely pontja nem belső pont, akkor az határpontja a felületnek.

Valószínű, hogy a felületek most megfogalmazott meghatározásával kapcsolatosan még a kevésbé éles szemű kritikusban is számos ellenvetés támad.

Először is nem tisztázott kérdés, hogy mit is értsünk alakzaton. Elég, ha arra gondolunk, hogy a 13. ábrán két pont (például

$G1$
és

$G2$
) számított egynek, a 17. ábrán pedig egy pont kettőnek. Tisztázatlan az is, mikor állítható egy alakzatról, hogy a nyílt körlappal vagy a zárt félkörlappal azonos felépítésű.

Pontos megfogalmazást kívánhat a szigorú kritikus arra vonatkozóan is, hogy az alakzat egy része egy pontot körülvesz-e vagy nem.

Mindezek a kérdések csakugyan tisztázásra szorulnak, hiszen különben tág tere nyílna a szubjektivizmusnak, a szemlélőn múlna például az, hogy a kúp palástját azonos felépítésűnek tartja-e a körlappal, vagy sem.

Vannak olyan esetek is, ahol a szemlélet bizonytalanná válik; gondoljunk például a 17. ábra

C

pontjára és egy azt körülvevő kis sapkára (lásd a 20. ábrát)! A

CF

él pontjait itt – a

C

pont kivételével – természetesen kettős pontoknak kell tekintenünk. Azonos-e az így nyert alakzat felépítése a körlapéval vagy sem? A térben jól látó szemlélő sem ad rögtön igenlő választ erre a kérdésre. Ha azonban meggondoljuk, hogy az alakzat szerkezeti felépítése éppen olyan, mint egy négyzet alapú gúla palástjáé a határoló négyzetvonal nélkül, akkor már könnyebb igennel válaszolnunk.

Az imént felvetett kérdésekre csak a matematika egyik igen elvont ágának, a topológiának eszközeivel adhatunk megnyugtató választ.

Nem lehet célunk, hogy teljes egészében ismertessük itt a topológia apparátusát, mert nagyon is eltérnénk eredeti témánktól. Ha azonban figyelmünket egyelőre csak azokra a felületekre fordítjuk, ahol nincsenek kettős pontok vagy pontpárok, akkor már könnyebben válaszolhatunk.

Alakzaton tehát most a tér pontjainak bizonyos összességét, más szóval a tér pontjaiból álló halmaznak egy részhalmazát értjük.

Valamely

P

pontnak egy térbeli környezetén olyan halmazt értünk, amely tartalmaz egy

P

középpontú golyót; más szóval

VP

akkor környezete

P

-nek, ha van olyan pozitív

□

-szám, amelynél a térnek

P

-től

□

-nál kisebb mérőszámú távolságra lévő pontjai (és így maga

P

is) valamennyien

VP

-hez tartoznak. Egy pontnak természetesen végtelen sok környezete van.

Legyen most

M

a tér pontjainak tetszőleges részhalmaza, és

P

az

M

halmaz egy pontja! Ekkor a

P

pontnak egy

M

-beli környezetén

P

pont egy térbeli környezetének

M

-hez tartozó pontjait értjük.

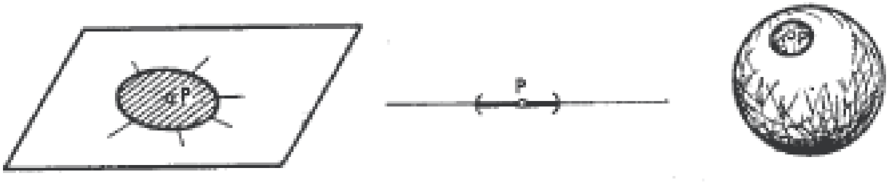
Így például egy sík, egyenes, illetve gömbfelület valamely

P

pontjának környezete olyan halmaz a síkon, egyenesen, illetve gömbfelületen, amely tartalmaz

P

középpontú körlapot, egyenesszakaszt, illetve gömbsüveget (lásd a 21. ábrát)!



21. ábra

A tér valamely

G

részhalmaza nyílt, ha maga a

G

halmaz térbeli környezete minden

G

-hez tartozó pontnak. Egy

F

halmaz zárt, ha a kiegészítő halmaz, vagyis az a halmaz, amely a térben

F

pontjainak elhagyásakor megmarad, nyílt halmaz. Így például egy gömb belső pontjai és külső pontjai is nyílt halmazt alkotnak (külön–külön és együttesen is!), egy gömbfelület pedig zárt halmaz a térben.

Egy

M

halmaz korlátos, ha van olyan gömb, amely tartalmazza azt.

Legyen

M
a tér pontjainak valamely részhalmaza,

N
egy nyílt, illetve zárt halmaz a térben és

N'
az

N
halmaz

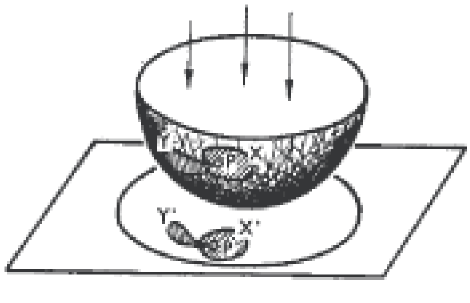
M
-hez tartozó pontjainak az összessége! Ekkor

N'
-t

M
-ben nyílt, illetve zárt halmaznak hívjuk. Ily módon a nyílt körlap a síkban nyílt halmaz, a zárt félkörlap pedig a síkban (de a térben is) zárt.

Felhívjuk itt az olvasó figyelmét arra, hogy a nyílt és zárt kifejezéseket többféle értelemben használtuk. Például egy nyílt körlap a térben nem lesz nyílt halmaz, egy nem zárt (hanem határolt) felület pedig zárt halmaznak számít a térben.

Tekintsünk most egy nyílt körlapot, és tekintsük egy vele azonos sugarú gömbfelület nyílt (határkörétől megfosztott) félgömbjét! Helyezzük el őket a 22. ábrán jelzett módon!



22. ábra

Ha a félgömböt párhuzamos sugarakkal felülről megvilágítjuk, akkor a félgömb minden pontjának árnyéka a körlapon fekszik. A különböző pontok árnyékai is különbözők lesznek, és minden körlapon fekvő pont árnyéka lesz a félgömb valamelyik pontjának. Azt mondjuk, hogy kölcsönösen egyértelmű megfeleltetést létesítettünk a félgömb és a körlap pontjai között.

Általánosságban akkor beszélünk két alakzat pontjai közötti kölcsönösen egyértelmű megfeleltetésről vagy leképezésről, ha az egyik alakzat minden pontjához hozzá van rendelve a másik alakzat egy pontja úgy, hogy különböző pontokhoz különböző pontok legyenek hozzárendelve, és a második alakzat minden pontja megfeleljen az eredeti valamelyik pontjának.

A félgömbnek a körlapra történő vetítése mindkét irányban folytonos megfeleltetés. Ezen a következőt értjük:

Szemeljük ki a félgömb valamely tetszőleges

P

pontját, egy tetszés szerinti olyan

X

részhalmazát, amely környezete

P

-nek a félgömbön, végül olyan tetszőleges

Y

részhalmazát, amely nem környezete

P

-nek a félgömbön (lásd a 22. ábrát)! Legyen

P'

a

P

pont,

X'

az

X

halmaz,

Y'

pedig az

Y

halmaz képe (árnyéka). Ekkor

X'

környezete

P'

-nek,

Y'

pedig nem környezete

P'

-nek a körlapon.

Általában valamely

M

térbeli alakzatnak egy másik

M'

térbeli alakzatra történő kölcsönösen egyértelmű leképezését mindkét irányban folytonos vagy topologikus vagy homeomorf leképezésnek hívjuk, ha az

M

alakzat bármely

P
pontjára és

X
részhalmazára nézve igaz a következő:

X
akkor és csak akkor környezete a

P
pontnak

M
-ben, ha

X
képe környezete

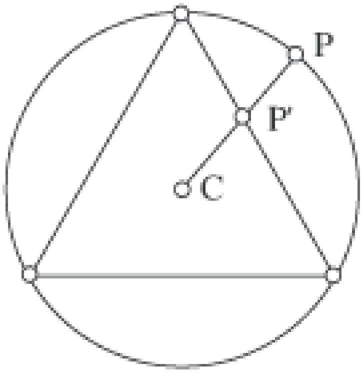
P
képének

M'
-ben.

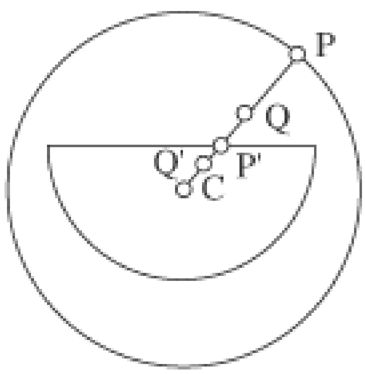
Két alakzatot akkor hívunk homeomorfnak, illetve akkor mondjuk, hogy az egyik a másik topologikus képe, ha létesíthető köztük homeomorf leképezés.

Lássunk néhány példát homeomorf leképezésekre:

A nyílt félgömb és a nyílt körlap homeomorf voltát az imént már láttuk, ha a bizonyítást nem is hajtottuk végre teljes részletességgel. Ugyanúgy látható be, hogy egy körvonal és a beleírt szabályos háromszög kerülete (határvonala) homeomorfak. Hiszen a kör középpontjából való vetítés homeomorf leképezést létesít a két alakzat között (lásd a 23. ábrát)!



23. ábra



24. ábra

Homeomorf egymással a zárt körlap és a zárt háromszöglap, valamint a nyílt körlap és a nyílt háromszöglap is. A homeomorfiát az a leképezés létesíti, amely a körvonal tetszőleges

P
pontja esetén arányosan húzza össze a

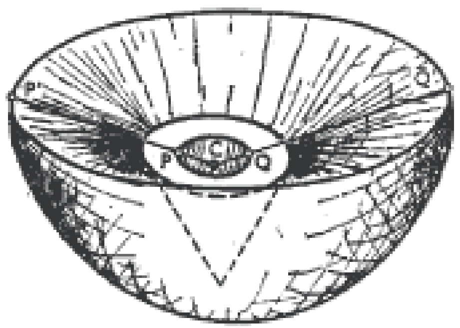
$\overline{CP}$
szakaszt a

$CP^{\ulcorner}$
szakaszra (lásd a 23. ábrát)!

Ugyanúgy látható be egy zárt félkörlap és egy zárt körlap homeomorf volta (lásd 24. ábrát)!

Homeomorf továbbá például egy gömbfelület egy kocka felületével. Ha ugyanis egybe ejtjük a kocka és a gömb középpontját, akkor ebből a közös középpontból történő vetítés homeomorf leképezést létesít a két alakzat között.

Homeomorf egy zárt félgömb egy körgyűrű és annak külső körére emelt kúppalást egyesítésével. A homeomorfiát itt ugyancsak egy vetítés szolgáltatja. Ez a homeomorf leképezés a körgyűrű belső körvonalát a félgömb határkörébe viszi át, méghozzá úgy, hogy átellenes pontok képei átellenes pontok lesznek (lásd a 25. ábrát)!

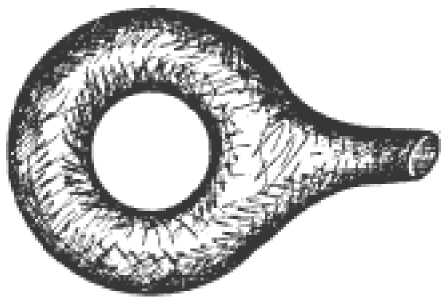


25. ábra

Szemléletesen azt mondhatjuk tehát, hogy két alakzat akkor homeomorf, ha igen rugalmas gumiból készítve el azokat, bármelyik átalakítható a másikba szakítás és ragasztás nélkül, természetesen húzást, nyújtást, zsugorítást megengedve. Így például, ha egy gumiból készült kockafelületet jól felfújunk, akkor anélkül, hogy elszakítanánk vagy valahol összeragasztanánk, előbb–utóbb gömbalakot vagy legalábbis gömbhöz közeli alakot vesz fel.

A homeomorf alakzatoknak ez a „gumigeometriai” meghatározása azonban nem pontos, mert például a 2. és 3. ábrák egyszerű, illetve kétszer csavart szalagjai homeomorfak egymással, de gumiból elkészítve az egyiket, szétvágás és ragasztás nélkül az semmilyen húzással, nyújtással vagy zsugorítással nem vihető át a másikba.

Persze a ragasztás- és szakításmentes „gumigeometriai” leképezések is időnként meglepő átalakításokat tesznek lehetővé. Nézzük például a 26. ábra

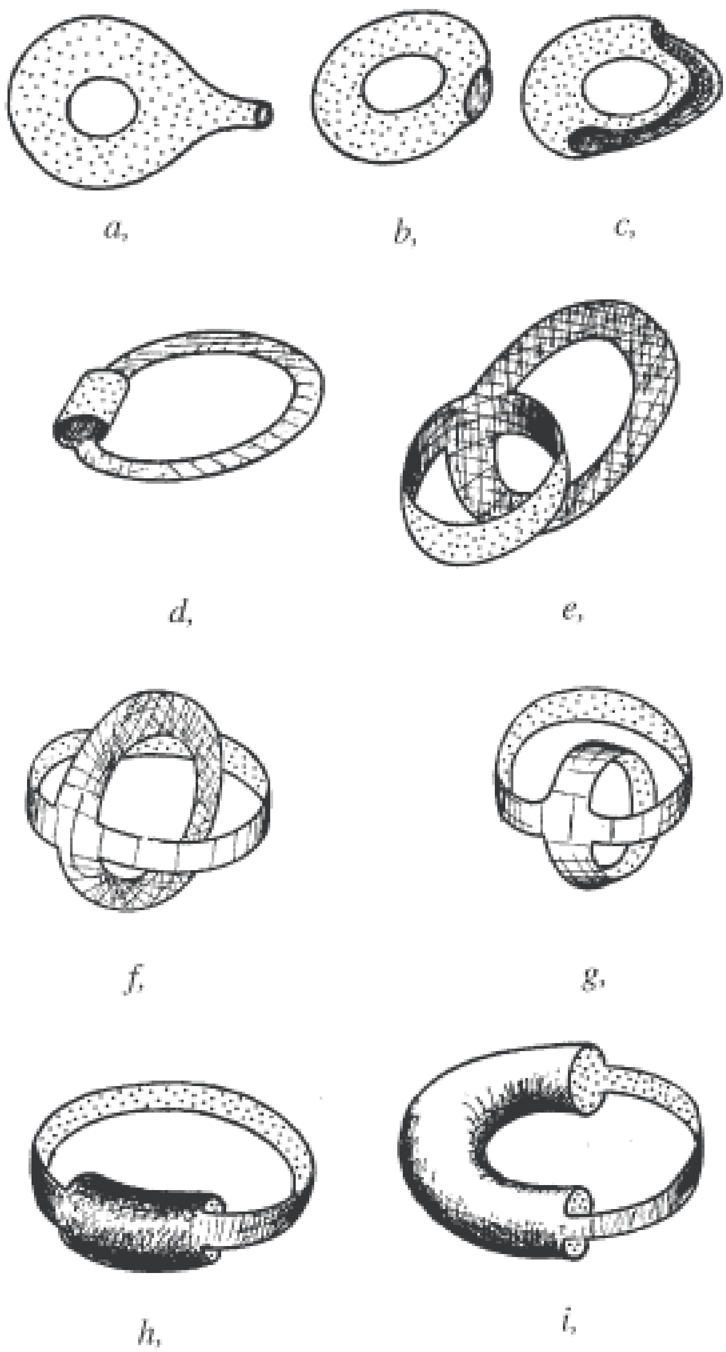


26. ábra

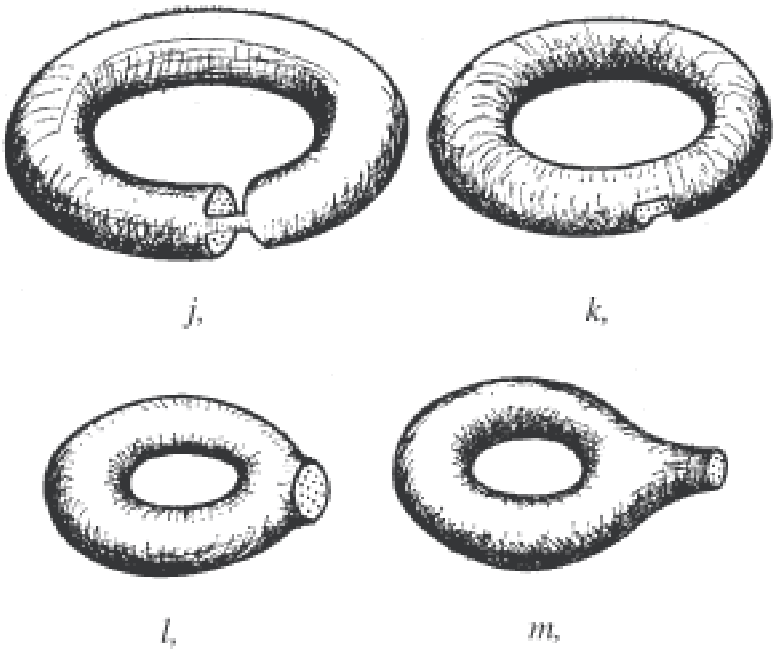
úgynevezett fogantyúfelületét! Képzeljük el, hogy a felület külseje pirosra, belseje pedig feketére van festve. A feladat a felület olyan átalakítása (kifordítása), hogy kívül legyen a fekete és belül a piros oldal. Mint azt a 27. ábrasorozat mutatja, ez az átalakítás elvégezhető szakítás és ragasztás nélkül.

Egyenesszakasszal homeomorf alakzatokat egyszerű íveknek hívunk (lásd a 28. ábrát)! Az ív két végpontja a szakasz két végpontjának a homeomorf leképezés által nyert képe lesz.

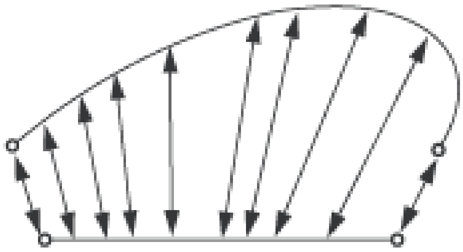
Az egyszerű ív összeköti (saját) két végpontját. Egy alakzatot ívszerűen összefüggőnek hívunk, ha bármely két pontja összeköthető az alakzatban haladó egyszerű ívvel.



27. ábra



A 27. ábra folytatása



28. ábra

Egyszerű zárt görbének hívjuk a körvonal topologikus képét. Az előkészületek megtétele után most már visszatérhetünk eredeti kérdésünk megválaszolására, tudniillik arra, hogy végül is mit értsünk felületen. Egyelőre – mint arról már szóltunk – csak az önmagukat nem metsző térbeli alakzatokkal foglalkozunk. Nem foglalkozunk a végtelenbe nyúló felületekkel sem. A (térbeli) felület matematikailag pontos definíciója a következő: A tér valamely

F

részhalmazát felületnek hívjuk, ha az ívszerűen összefüggő, korlátos és zárt, azonkívül bármely pontjának van zárt körlappal homeomorf

F

-beli környezete.

Ez természetesen eléggé elvont definíció, de valójában a felület szemléletes meghatározásának matematikailag szabatosan megfogalmazott pontos mása.

Ha minden

F

-beli pontnak van nyílt körlappal homeomorf

F

-beli környezete is, akkor a felület zárt vagy nem határolt, egyébként határolt. A felület határpontjai azok a pontok, amelyeknek nincs nyílt körlappal homeomorf környezetük. A határpontok összessége a felület határa, amely egy vagy több, de mindig véges sok egyszerű zárt görbéből áll.

4. A FELÜLETEK OSZTÁLYOZÁSA

Vizsgáljuk meg most, hogy milyen típusú felületek vannak! Ennek megállapításához először azt a kérdést kell eldöntenünk, hogy mikor tekintünk két felületet azonos, és mikor különböző típusúnak.

Az elemi geometria a kocka és a gömb felületét különböző típusúnak tekinti, hiszen nem hasonlók egymáshoz. Ha azonban akkor sorolunk két felületet egy osztályba, ha azok homeomorfak, abban az esetben a kocka és a gömb felülete már ugyanabba az osztályba kerül. Ezt a topológiai osztályozást vesszük majd alapul.

A geometrián belül a topológiai osztályozás adja a legáttekinthetőbb rendszerezést. Bár ebben az osztályozásban azonos osztályba kerül az egyszerű szalag és a kétszer megcsavart szalag is, vagy a lyukas tórusz és a 26. ábra fogantyúfelülete, de például a gömbfelület és a tóruszfelület, az egyszerű szalag és a Möbius-szalag itt is különböző osztályokhoz tartozik.

A felületek topológiai osztályozását a múlt század második felében *Möbius*, *Dyck* és *Petersen* végezték el.

Az osztályokat úgy tekinthetjük át legcélszerűbben, ha mindegyik osztályból mintát veszünk, más szóval minden egyes osztályt valamely beletartozó felülettel képviseltetünk. Bármely kettőt szemeljük is ki az így nyert minták közül, ezek sohasem lesznek homeomorfak, de bármely térbeli felület homeomorf lesz a minták egyikével.

Lássunk tehát egy ilyen mintakollekciót:

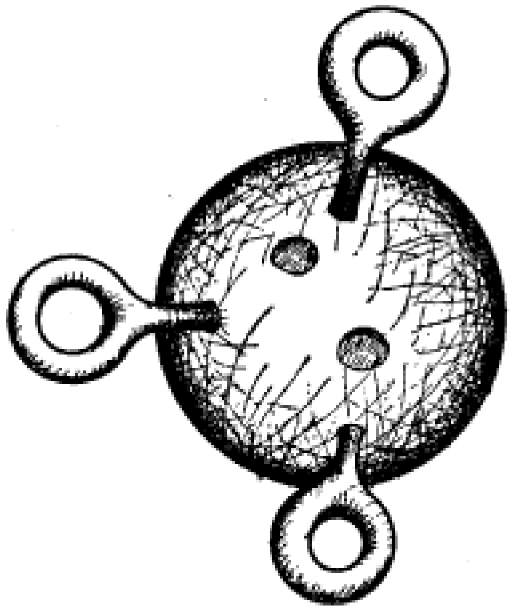
A. Kétoldalú felületek.

Akkor kétoldalú a felület, ha egy órát a felületen fekvő bármely zárt görbén végigvezetve az eredeti helyére való visszatérésekor az ugyanúgy jár, mint amikor elindítottuk.

Vágjunk ki a gömbfelületből

$p+q$
darab gömbsüveget (ablakot)! Közülük

p
darabot ragasszunk be fogantyúval (lásd a 29. ábrát)!



29. ábra

Ekkor q számú határgörbével rendelkező kétoldalú felületet nyertünk. Ha két ilyen mintán a p és q számok nem azonosak, akkor a felületek nem homeomorfak. Másrészt viszont bármely kétoldalú felület valamely p fogantyús és q ablakos gömbbel lesz homeomorf.

Magán a gömbfelületen $p=0$, $q=0$, a tóruszfelületen $p=1$, $q=0$, a fogantyún $p=1$,

$q=1$
, az egyszerű szalagon

$p=0$
,

$q=2$
, a körlapon

$p=0$
,

$q=1$
stb.

B. Egyoldalú felületek.

Akkor egyoldalú a felület, ha van rajta olyan zárt görbe, amelyen végigvezetve az órát, annak járása „megfordul”.

Ragasszunk hozzá egy körlap határvonalához

p
darab (egyszer) csavart és

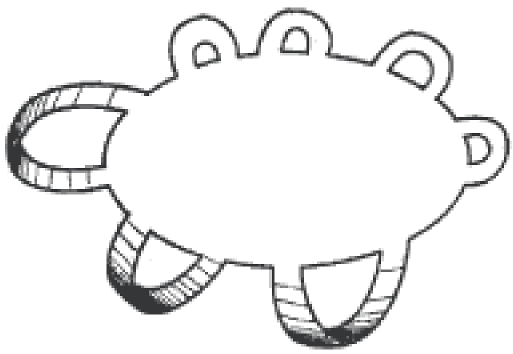
q
darab egyszerű szalagot (

$p \geq 1$
,

$q \geq 0$
)! A 30. ábrán például

$p=3$
és

$q=4$
.



30. ábra

Ha két mintán a

p

és

q
számok nem egyeznek meg, akkor nem homeomorfak a felületek, de bármely (térbeli) egyoldalú felülethez tartozik olyan

p
és

q
, amelynél az homeomorf a

p
csavart és

q
egyszerű szalaggal ellátott körlappal.

Így például a Möbius-szalagon

p=1
és

q=0
.

5. EGYOLDALÚ ZÁRT FELÜLETEK

A kétoldalú felületek között vannak olyanok (

q=0
típusúak), amelyeknek egyetlen határvonaluk sincs, míg az egyoldalú felületek mindegyikének van legalább egy határvonala. Ha azonban az egy határvonallal bíró (

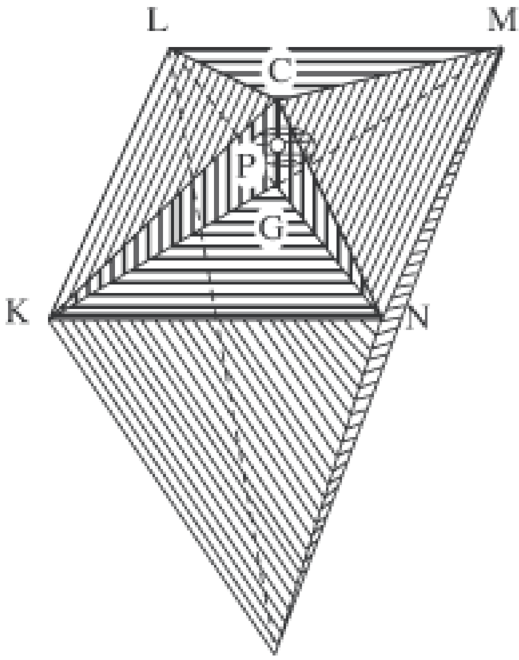
q=0
) egyoldalú felületek határvonalát összeragasztjuk egy

L
körlap

K
határoló körvonalával, más szóval körlapot ragasztunk a felülethez, akkor már zárt egyoldalú felületet nyerünk. Csakhogy ez a beragasztás sohasem vihető véghez anélkül, hogy ne kapnánk kettős pontokat is.

A Möbius-szalagnak a 17. ábrán megadott átalakításához a határvonala mentén egy gúla palástját ragasztva hozzá, egyoldalú zárt

F1
felületet nyerünk (lásd a 31. ábrát),



31. ábra

de a

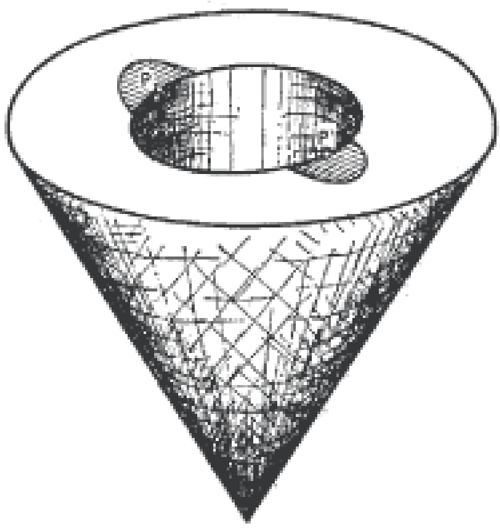
$\overline{CG}$
szakasz pontjai

C
és

G
kivételével itt természetesen kettős pontok lesznek.

A 13. ábra Möbius-szalagjából egy kúppalást hozzáragasztásával szintén egyoldalú zárt

$F2$
felületet kapunk, csakhogy itt a körgyűrű belső körén az átellenes pontpárok mindegyike egy pontnak számít (lásd a 32. ábrát)!



32. ábra

A Möbius-szalag körlappal történő beragasztását úgy is értelmezhetjük, hogy egy

f
topologikus leképezést létesítünk az

L
körlapot határoló

K
körvonal és az

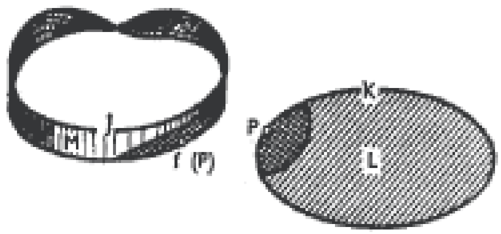
M
Möbius-szalag

J
határvonala között. Egy pontnak tekintjük a

$[P,f(P)]$
párokat, ahol

P
végigfut a

K
körvonalon (lásd a 33. ábrát)!



33. ábra

Ezzel az azonosítással az

M
Möbius-szalag és az

L
körlap együtt ismét zárt (nem határolt) egyoldalú

$F3$
felületet alkot.

Felvetődik a kérdés: vajon az így meghatározott alakzatok valóban felületek-e, és ha igen, akkor homeomorfak-e vagy sem?

Először meg kell mondanunk, hogy mit értsünk ezeken a furcsa alakzatokon egy elem környezetén.

Tekintsük az

$F3$
alakzatot!

A tér pontjainak egy

N
összességét

$F3$
-mal összeférő halmaznak hívjuk, ha a

K
körvonal minden

N
-hez tartozó

P
pontjával együtt

$f(P)$
is

N
-hez tartozik, és ezenkívül

$f(P)$

N
-hez tartozása maga után vonja azt, hogy

P
is

N
-hez tartozzék. Az

$F3$
alakzat egy közönséges

Q

pontjának valamely környezetét a

Q
térbeli környezetét alkotó,

$F3$
-mal összeférő

N
halmazok jelölik ki úgy, hogy a környezet elemei az

N
halmaz

M
-hez és

L
-hez tartozó, de

J
-n és

K
-n nem fekvő pontjaiból álljanak, valamint azokból a

$[P,f(P)]$
pontpárokból, ahol

P
az

N
-hez és

K
-hoz is hozzátartozik. Ha

$[P,f(P)]$
az

$F3$
valamely pontpárja, akkor környezetének képezésekor olyan

N
összeférő halmazból kell kiindulnunk, amely

P
-nek is és

$f(P)$
-nek is környezete. Az eljárás további menete ugyanaz, mint előbb volt (lásd a 33. ábrát)!

Hasonlóan járhatunk el

$F2$
esetében is, csak ott az összeférő halmazok azok az

N
halmazok lesznek, amelyek a körgyűrű belső körvonalának minden

P
pontjával együtt annak a körvonalon fekvő átellenesét is tartalmazzák.

Az

$F1$
alakzatnál más utat kell követnünk.

$F1$
elemei az

$F1$
-hez tartozó pontok a

$\overline{CG}$
szakasz belső pontjai nélkül, továbbá a

$(P,KCM\square)$
és

$(P,LCN\square)$
párok, ahol

P
a

CG
szakasz belső pontja. Ilyen módon a kettős pontokat mintegy felbontottuk két különálló elemre. Egy elem környezetének megállapításakor itt három típusú elemet kell megkülönböztetnünk. Ha

P
nem tartozik a

$\overline{CG}$
szakaszhoz, akkor

$F1$
-nek

X
részhalmaza környezete

P
-nek, ha van olyan

P
középpontú golyó, amelynek

$F1$
-hez tartozó közönséges pontjai

X
-hez is hozzátartoznak. Ha

$P \equiv C$
vagy

$P \equiv G$
, akkor még azt is meg kell követelnünk, hogy azok a

$(Q, KCM\square)$
és

$(Q, LCN\square)$
párok, amelyeknél

Q
a golyóhoz tartozik, szintén az

X
halmaz elemei legyenek. Ha végül

$(P, KCM\square)$
típusú elemet veszünk, akkor a

P
középpontú golyónak a

$KCM\square$
-höz tartozó pontjaitól és azoktól a

$(Q, KCM\square)$
pároktól, ahol

Q
a golyóhoz tartozik, kell megkövetelnünk, hogy

X
elemei legyenek (lásd a 31. ábrát)! Ha

$(P, LCN\square)$
típusú elemről van szó, akkor ugyanúgy járunk el, csak

$KCM\square$
helyébe mindenütt

$LCN\square$
kerül.

Ilyen módon, bár az

$F2$
és

$F3$
alakzatokban pontokon kívül pontpárok is szerepelnek az elemek között,

$F1$
-ben pedig kettős pont is van, mégis értelmezhető volt rajtuk a környezetfogalom, és így beszélhetünk ezeknek az alakzatoknak homeomorf vagy nem homeomorf voltáról is.

Belátható, hogy

*F*1
,

*F*2
és

*F*3
homeomorfak, továbbá az is, hogy

*F*1
(és így

*F*2
vagy

*F*3
) minden elemének – legyen az akár közönséges pont, akár pontpár, akár egy kettős pont egyik fele – van nyílt körlappal homeomorf környezete. Az is belátható, hogy

*F*1
bármely két pontját össze lehet kötni egyszerű ívvel. Gyanítjuk tehát, hogy

*F*1
-et,

*F*2
-t és

*F*3
-at is a zárt felületek közé sorolhatjuk, ha nem is homeomorfak valamely kétoldalú felülettel, sőt egyetlen olyan felülettel sem, amelyről az előző fejezetekben beszéltünk.

A mondottak bizonyítása érdekében még tovább kellene terjesztenünk a felület fogalmát úgy, hogy abba a régi felületek és az

*F*1
,

*F*2
és

*F*3
alakzatok egyaránt beleférjenek. Mi most itt nem térhetünk ki erre, mivel ez túlságosan messzire, az absztrakt topologikus terekhez vezetne. Pusztán annyit jegyzünk meg, hogy az egyoldalú, egyetlen határvonalú felületek mindegyike beragasztható egy körlappal ugyanúgy, ahogyan ezt a Möbius-szalagból képzett

*F*3
felületnél láttuk. Ilyen módon nyerjük az egyoldalú zárt felületeket. Ezek a zárt felületek akkor és csak akkor lesznek homeomorfak, ha a beragasztandó egyoldalú határolt felületek homeomorfak voltak.

Befejezésül néhány példát mutatunk egyoldalú zárt felületekre.

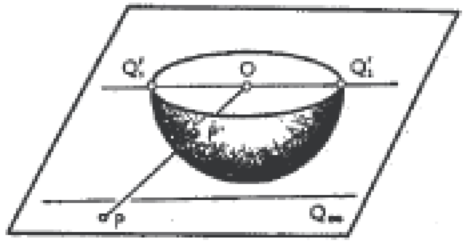
A. A projektív sík.

Egészítsük ki a sík minden egyenesét a közönséges pontokon kívül egy úgynevezett végtelen távoli ponttal úgy, hogy a párhuzamos egyenesekhez ugyanaz a végtelen távoli pont tartozzék, a nem párhuzamosokhoz pedig különböző végtelen távoli pontok tartozzanak! A végtelen távoli pontokkal ilyen módon kiegészített síkot projektív síknak hívjuk.

Belátjuk, hogy ez a projektív sík homeomorf a Möbius-szalag körlappal való beragasztása révén nyert egyoldalú zárt

*F*3
felülettel.

Állítsunk a síkra egy félgömböt úgy, hogy a sík a félgömböt éppen peremétől legtávolabb eső pontjában érintse (lásd a 34. ábrát)!



34. ábra

A gömb

O
középpontjából történő vetítés a közönséges sík és a nyílt félgömb között homeomorf leképezést létesít. Végtelen távoli pont megfelelőjét úgy kapjuk meg, hogy

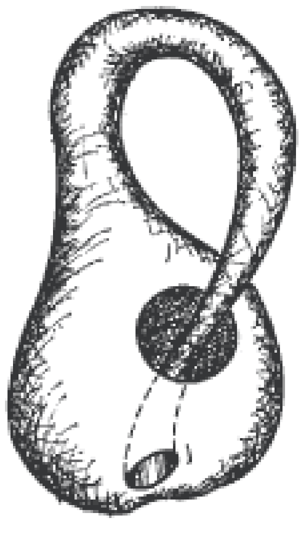
O
-ból párhuzamost húzunk a végtelen távoli pontot tartalmazó valamelyik egyenessel. Ez a vetítősugár a félgömb peremén egy átellenes pontpárt fog kimetszeni. A projektív sík tehát homeomorf a félgömbből a határkör átellenes pontpárjainak összeragasztása után nyert felülettel, vagyis

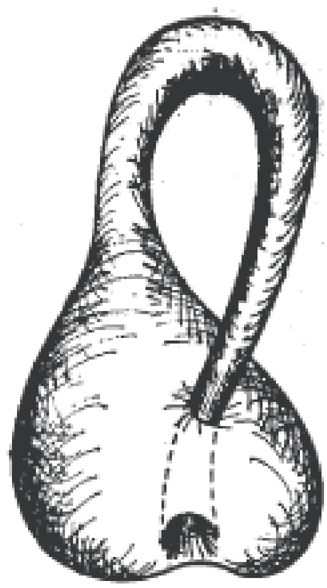
*F*2
-vel és így

*F*3
-mal is.

B. A Klein-féle cső.

Tekintsük a 35. ábra egy határvonalú, egyoldalú felületét! Ragasszuk be a körvonalat egy körlappal! Ekkor egy önmagát metsző, egyoldalú zárt felületet nyerünk (lásd a 36. ábrát)!





35. ábra 36. ábra

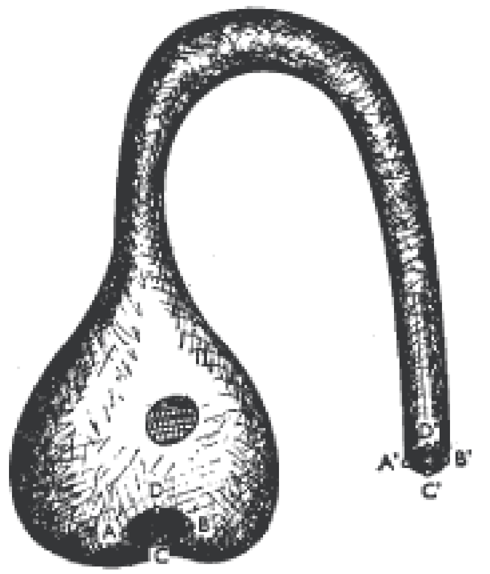
Ez a Klein-féle cső. A Klein-féle cső homeomorf a két (egyszer) csavart szalaggal ellátott körlap

(p=2,q=0)
körlappal való beragasztásával.

Hogy erről meggyőződhessünk, a 35. ábrán látható felületről azt kell kimutatnunk, hogy homeomorf a két (egyszer) csavart szalaggal ellátott körlappal.

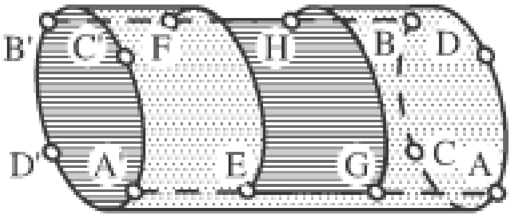
Vágjuk szét a felületet a

J
körvonal mentén, és húzzuk ki a cső elvékonyodó részét a lyukból (lásd a 37. ábrát)!



37. ábra

Jelöljük meg azonos betűkkel a szétvágott pontokat! Kissé átalakítva a 37. ábrát, a 38. ábra alakzatához jutunk.



38. ábra

Vágjuk szét ezt az ábrát az

A'E

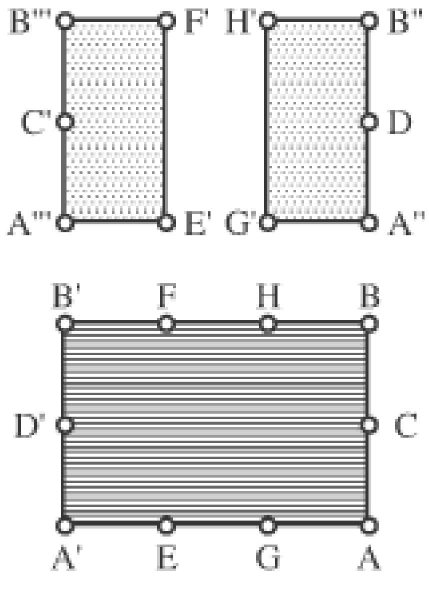
,

B'F

,

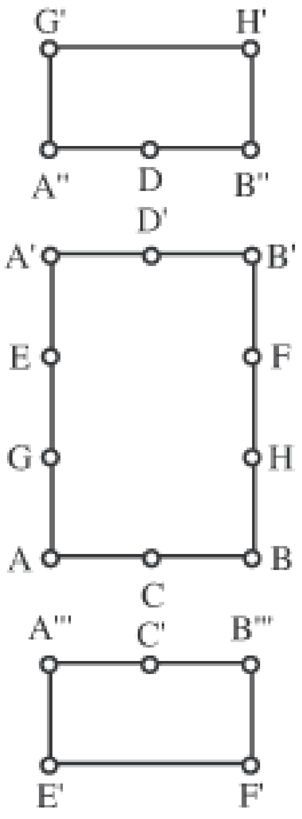
GA
és

HB
szakaszok mentén, és a szétvágott pontokat újra jelöljük azonos betűkkel! Terítsük ki a síkba az így nyert alakzatokat (lásd a 39. ábrát)!

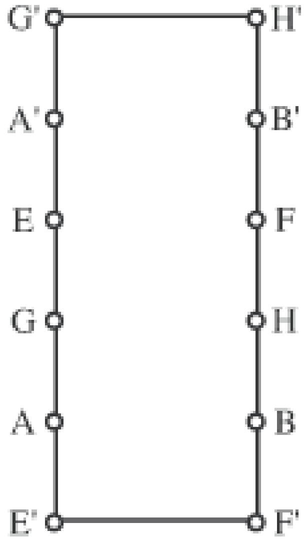


39. ábra

Toljuk el a 39. ábra alsó kis téglalapját a nagy fölé, a felsőt pedig a nagy téglalap alá (lásd a 40. ábrát)!



40. ábra



41. ábra

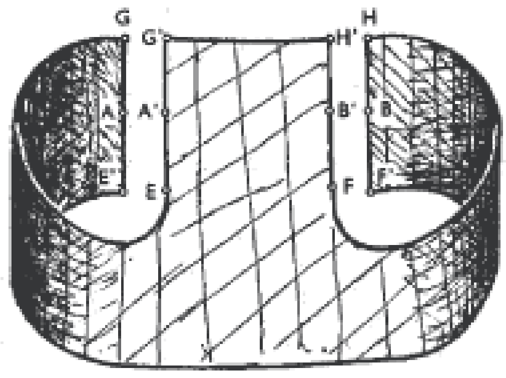
A szétvágott pontok kerültek egymás fölé, és így ezeket összeragaszthatjuk (lásd a 41. ábrát)! A

GAE'

szakaszt kihúzzuk szalagszerűen és megcsavarjuk. Ugyanezt tesszük a

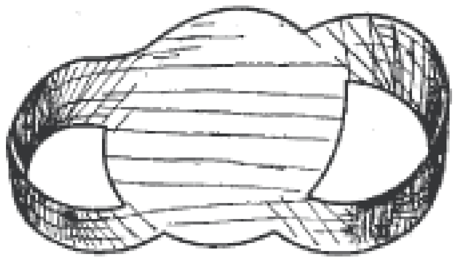
HBF'

szakasszal is (lásd a 42. ábrát)!



42 ábra

Ismét egymás mellé kerültek a szétvágott pontok. Az azonos jelzésű pontok összeragasztása és az ábra némi kiigazítása után két (egyszer) csavart szalaggal ellátott körlapot kapunk (lásd a 43. ábrát)!



43. ábra

A lyukas Klein-féle cső tehát csakugyan homeomorf ezzel az alakzattal, maga a Klein-féle cső pedig a 43. ábra felületének körlappal történő beragasztottjával lesz homeomorf.

□ □

□

Nincs szándékunkban tovább szaporítani a példákat. Azoknak, akik a témával behatóbban kívánnak foglalkozni, Kőnig Dénes: „Az analysis situs elemei” (Akadémiai Kiadó, 1918) vagy Boltjanszkij–Jefremovics magyar nyelven megjelent „Szemléletes topológia” (Tankönyvkiadó, 1965) című könyveit ajánlhatjuk.

EGY IGAZÁN CSUDÁLATOS BIZONYÍTÁS[26]

RÓNYAI LAJOS

1993. június 23-án Cambridge-ben a Sir Isaac Newton Intézetben Andrew Wiles angol matematikus bejelentette, hogy bebizonyította a *Fermat-sejtést*. A felfedezés híre igazi szenzációvá vált, ami azelőtt még sohasem fordult elő matematikai eseménnyel. A világ vezető napilapjai nagy terjedelemben írtak a dologról. Néhány hónappal később kiderült, hogy a nyári szemináriumon vázolt gondolatmenetben egy jelentős hiányosság van.

Már megint a szokásos történet – mondták sokan a Fermat-sejtésre adott számos korábbi kérészéletű „megoldásra” gondolva. A szakértők másként vélekedtek rámutatva, hogy Wiles munkája óriási előrelépést hozott a terület egyik központi problémájával kapcsolatban. Néhányan közülük úgy látták, hogy a bizonyításban levő lyuk is befoltozható pár hónap, esetleg pár év alatt. Hamarosan kiderült, hogy igazuk volt. Már 1994 nyarán voltak olyan hírek, hogy létezhet egy kerülőút, amely mentén teljessé tehető a bizonyítás. 1994. október 24-én Wiles két kéziratot adott közre, melyek igen alapos és gondos bírálat után megjelentek az *Annals of Mathematics* című folyóiratban. Az egyik dolgozat annak az elliptikus görbékre vonatkozó eredménynek a bizonyítását tartalmazza, amelyből következik a Fermat-sejtés. A második – a szintén brit Richard Taylorral közös – munka foglalkozik a kerülőutat jelentő algebrai állítással.

A következő néhány oldalon összefoglaljuk a sejtés történetét. Közben érintünk néhány kapcsolódó matematikai fogalmat. Eközben ejtünk szót régi és mai sötétben botorkálásról is.

Mi lehet az a matematikai kérdés, amiről többhasábos cikkek jelennek meg a nagy napilapokban? Nos, a Fermat-sejtés ártatlanul egyszerűen hangzik:

Ha

$n>2$
egész szám, akkor nincsenek olyan nullától különböző

x,y,z
egészek, melyekre

$x^n+y^n=z^n$
teljesül.

Egyszerű állítás, nem kell sok előismeret a megfogalmazásához. Valósággal csábítja az embert, hogy kezdjen el számolgatni, gondolkodni rajta. Hosszú története során sokakat megejtett különös varázsával. Többek számára jelentette azt a meghatározó élményt, ami a matematikusi pálya választásához vezette őket. Személyes ismerőseim között is van ilyen kolléga. Wiles maga is már kisiskolásként találkozott a kérdéssel.

A sejtés eredetét kutatva i.e. 250-ig tekinthetünk vissza. Ezidőtájt született az alexandriai Diofantosz nagyhatású munkája, az *Aritmetika*, amely – ismereteink szerint – először adott közre valamelyes rendszerbe foglalva számelméleti és algebrai eredményeket. Íme egy jellegzetes feladat a II. Könyvből: *osszunk fel egy adott négyzetet két négyzetre!* A probléma és a Diofantosz által közölt megoldás a következő pontosabb „modern” megfogalmazást sugallja: *keressük az*

$x^2+y^2=z^2$

[26]Köszönetet mondok Bródy Ferencnek, Ivanyos Gábornak, Krámli Andrásnak és Szabó Rékának értékes észrevételeikért.

egyenlet egész megoldásait, szokásos nevükön a pitagoraszi számhármasokat.

Nem nehéz meghatározni az összes ilyen hármast. Az általánosság különösebb sérelme nélkül szorítkozhatunk arra az esetre, amikor mindhárom szám pozitív, semelyik kettőnek nincs közös prímosztója, és

x
páros (nevezzük ezeket primitív hármasoknak). Ekkor az

$(x^2/4)=(1/2)(z+y)\cdot(1/2)(z-y)$
összefüggést használva kapjuk, hogy

$(1/2)(z+y)$
és

$(1/2)(z-y)$
egészek, sőt négyzetszámok:

$(1/2)(z+y)=u^2$, $(1/2)(z-y)=v^2$.

A gondolatmenetet folytatva könnyen adódik, hogy a primitív hármasok mind megkaphatók

$x=2uv$, $y=u^2-v^2$, $z=u^2+v^2$
alakban, ahol

$u>v$
pozitív, relatív prím egészek, és az egyikük páros.

Most ugorjunk egyet térben és időben, a hellenisztikus Egyiptom virágkorából XIII. Lajos és Richelieu bíboros Franciaországába! Pierre de Fermat (1601–1665) hivatását tekintve jogász volt. Egyetemi tanulmányainak befejezése után pályájának végéig a közigazgatás és igazságszolgáltatás területén dolgozott Toulouse-ban. Úgy tűnik azonban, hogy igazi szenvedélye a matematika volt, ezen belül is az egész számok tulajdonságainak vizsgálata – a *számelmélet*, ahogy ma mondanánk. Fermat tehát pusztán kedvtelésből foglalkozott matematikával. Mégis olyan sok és fontos eredmény fűződik a nevéhez, hogy méltán tartják kora egyik legjelentősebb matematikusának. Akkoriban a tudomány korántsem volt olyan jól szervezett gépezet, mint manapság. Nem voltak például szakfolyóiratok. A gondolatok elsősorban levelezés útján terjedtek. Fermat is kedvelte ezt a műfajt. Olyan hírneves levelezőtársai voltak, mint René Descartes, Marin Mersenne, Christian Huygens vagy Blaise Pascal. Leveleiben adta közre seregnyi, nagy horderejű felfedezését többek között a matematikai analízis, a valószínűségszámítás, az optika és a számelmélet problémáiról. Heves indulatokat kavart azzal, hogy feladványok formájában csak a következtetéseinek végeredményét írta meg. Kesztyűt dobott ezzel a tudományos világnak, hogy találják meg mások is a megoldáshoz vezető – sokszor rendkívül nehéz – utat. Fermat emellett csípős hangú kritikusa volt kortársai munkáinak. Meglehetősen fagyos viszonyba keveredett például a nyugati gondolkodás atyjával, Descartes-tal, miután *sötétben botorkálásnak* minősítette utóbbinak a fénytörés természetével[27] foglalkozó érvelését. (Felettébb fullánkos megjegyzés egy a *fény* viselkedését *megvilágítani* szándékozó fejtegetésről.) Ez a kritika vezetett el aztán az optika egyik alaptörvényének, a *Fermat-elvnek*[28] a felfedezéséhez.

A jelen írás központi témája is egy Fermattól származó tudományos kihívás. Ezt vélhetően 1637 táján jegyezte fel az *Aritmetika* latin fordításának margójára. A probléma csak sokkal később, 1670-ben került napvilágra,

[27] Arról az összefüggésről van szó, amelyet ma Snellius–Descartes-törvénynek nevezünk: sík közeghatáron a megtört fénysugár a beesési síkban marad, és a beesési, valamint a törési szög szinuszának hányadosa a két közeg minőségére jellemző állandó.

[28] A Fermat-elv szerint a fénysugár két pont között mindig azon az úton halad, amelyre vonatkozóan az optikai úthossz a legkisebb vagy – ritkábban – legnagyobb. Az elvből könnyen következnek a geometriai optika további nevezetes összefüggései, így a Snellius–Descartes-törvény is.

amikor fia, Samuel közzétette apja feljegyzéseinek egy részét.

Fermat nagy figyelemmel olvasta az Aritmetikát, amire számos, a könyvbe írt megjegyzéséből következtethetünk. A négyzet felosztásával kapcsolatban imént idézett részhez az alábbi széljegyzetet fűzte:

Cubum autem in duos cubos, aut quadrato-quadratum in duos quadrato-quadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos ejusdem nominis fas est dividere; cujus rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caparet.

Magyarul – Bródy Ferenc fordításában – mindez így hangzik:

Nincsen mód viszont felosztani köböt két köbre, sem négyzetes négyzetet két négyzetes négyzetre, és általában a négyzeten túl a végtelenig semmiféle hatványt két ugyanolyan nevezetűre; mely dolognak igazán csudálatos bizonyítását találtam. Szűkebb a margó, semhogy befogadná.

Ez a Fermat-sejtés eredeti megfogalmazása. Olyan állításként írta le, amelyet bizonyítani tud. Azóta eltelt kb. 350 év, és egészen napjainkig senkinek sem sikerült bizonyítást találnia. Ezért általános a vélemény, hogy Fermat valószínűleg tévedett, elnézett valamit, és nem volt *igazán csudálatos bizonyítása*. Az 1800-as évek elejére minden más (levelekben és jegyzetekben fennmaradt) állítását sikerült tisztázni, csak ez az egy állt ellen makacsul minden kísérletnek. Innen származik a sejtés másik gyakran használt elnevezése: Fermat Utolsó Tétele.

Pedig erőfeszítésben nem volt hiány. Az

$n=3$
és

$n=4$
esetekről később Fermat maga is írt. Sokak szemében ez is amellett szól, hogy nem volt minden

n
-re érvényes bizonyítása. Az

$n=4$
kitevőre adott gondolatmenete fennmaradt – szintén lapszéli jegyzetek formájában. Valamivel általánosabban azt igazolta, hogy az

$x^4+y^4=z^2$
egyenletnek nincs csupa nem nulla egészekből álló megoldása.

Röviden bemutatjuk Fermat bizonyítását. Egy tanulságos módszerről van szó, mely több más problémára is alkalmazható, és amelynek a ma használt igen kifinomult változatait *Fermat-leszállásnak* nevezik.

Ha van az

$x^4+y^4=z^2$
egyenletnek csupa pozitív egészekből álló megoldása, akkor van olyan is, amelyben

z
a lehető legkisebb. Ekkor (

x
és

y
esetleges felcserélése után)

x^2, y^2 és z primitív pitagoraszi hármast alkot. Vannak tehát olyan u,v pozitív egészek, melyekre

$$x^2=2uv,\ y^2=u^2-v^2,\ z=u^2+v^2. \qquad (\square)$$

A $v^2+y^2=u^2$ egyenlőség szerint v,y,u is pitagoraszi hármas, ami primitív is, mert különben x^2,y^2,z sem volna primitív. A v páros, mert feltevésünk szerint y páratlan. A ($\square$)-beli első egyenlőségből ezután arra következtethetünk, hogy u négyzetszám ($u=c^2$, $c>0$), v pedig egy négyzetszám kétszerese. Alkalmazzuk ismét a primitív pitagoraszi hármasok leírását, mégpedig ezúttal a v,y,u hármasra! Alkalmas t és s

pozitív egészekkel érvényesek a következők:

(†) $v=2ts,\ y=t2-s2,\ u=t2+s2.$

Az első egyenletből és a

v
-re vonatkozó előbbi megállapításunkból közvetlenül látszik, hogy

t
és

s
is négyzetszám. Legyen tehát

$t=a2$
és

$s=b2$
, ahol

a
és

b
pozitív egészek. A (

†
) utolsó egyenlete szerint

$a4+b4=c2$
. Másfelől nyilvánvaló, hogy

$z=u2+v2>u2=c4\geq c.$

Foglaljuk össze, amit eddig kaptunk: feltettük, hogy van az egyenletnek csupa pozitív egészekből (vagy ami ezzel egyenértékű: nem nulla egészekből) álló megoldása. Ezek után az egyenlet egy olyan megoldásából indultunk ki, amelyben a jobb oldali szám a lehető legkisebb. Ez volt az

x,y,z
hármas. Ebből nyertünk egy olyan megoldást – az

a,b,c
hármast –, amelyben a jobb oldalon álló

$c>0$
egész kisebb, mint

z
. Ez nyilván képtelenség. Kiinduló feltevésünk ellentmondáshoz vezetett, az egyenletnek tehát nincs pozitív egészekből álló megoldása.

Fermatnak az Utolsó Tétel bizonyításával kapcsolatos tévedéséről több elképzelés is van. Egy népszerű nézet szerint feltehetőleg arra gondolt, hogy a leszállás módszere könnyen átvihető tetszőleges

n
kitevőre. Ilyen általánosítást azonban senki sem tudott kidolgozni.

Az

$n=4$
kitevő kizárása után elég a kérdést azokban az esetekben nézni, amikor

$n>2$
prímszám. Ez azért igaz, mert a

$q\cdot r$
kitevőhöz tartozó, ellenpéldát adó megoldásból (

q,r
pozitív egészek) ellenpéldát kaphatunk a

q
kitevőhöz is. Az

$n=3$
esetet L. Euler oldotta meg 1753-ban. L. Dirichlet és A. M. Legendre találtak bizonyítást

$n=5$
-re 1825-ben. Ugyancsak Dirichlet nevéhez fűződik az

$n=14$
kitevő (1832), az

$n=7$
esetet pedig G. Lamé tudta kezelni 1839-ben.

1847 jeles esztendő a sejtés történetében. Ekkor született az első nevezetes hibás bizonyítás, méghozzá olyan híres matematikusok műveként, mint A. Cauchy és G. Lamé (avagy nem csak az iskolai matekdolgozatokban találhatók gyógyíthatatlan hibák). A másik – sokkal fontosabb – fejleményt E. E. Kummer eredményei jelentették. Kummer egy egészen más jellegű számelméleti probléma vizsgálata során dolgozott ki olyan eszközöket, melyekkel az Utolsó Tétel több kitevő esetére is igazolható. Módszere működik például a 37, 59 és 67 kivételével minden 100-nál kisebb

n
prímre. Kummer munkája tekinthető az *algebrai számelmélet* hajnalának. Kummer eredményeire építve H. S. Vandiver (1920 körül) igazolta a sejtést

$n<100$
-ra. A nevek és az évszámok mutatják, hogy kiváló matematikusok foglalkoztak a problémával, de ennek ellenére a haladás meglehetősen lassú volt. Komoly előrelépést jelentettek az olyan feltételek, amelyek révén számítógépek segítségével lehet vizsgálni a problémát nagyobb kitevőkre. Ezek segítségével 1992-re

$n\le 4\,000\,000$
-ig igazolták a sejtést (J. Buhler, R. Crandall, R. Ernvall, T. Metsänkylä).

Fermat problémája az évek során rendkívül népszerűvé vált a matematikával hivatásszerűen foglalkozók és az amatőrök között egyaránt. Több értékes díjat tűztek ki a megoldó jutalmazására; ezek közül a legismertebb az 1908-ban alapított Wolfskehl Díj. Ennek a népszerűségnek az árnyoldala, hogy elképesztő számú hibás bizonyítást tettek közzé. Csak az 1908–1912 időszakban ezer felett volt a hibás kísérletek száma.

Fontos mérföldkő Gerd Faltings (1983) egy általános, sok egyenletre érvényes végességi tétele, melyből következik, hogy egy adott

$n>2$
-re az

$x^n+y^n=z^n$
egyenletnek csak véges sok primitív egész megoldása lehet.

A Fermat-sejtés jelenkori történetében főszerep jutott az *elliptikus görbéknek*. Mielőtt az eseményekről beszélnénk, vessünk egy pillantást ezekre a rendkívül érdekes, gazdag struktúrájú objektumokra! Elliptikus görbén egy

$y^2=f(x)$
alakú egyenlettel definiált síkbeli görbét értünk, ahol

$f(x)$
egy

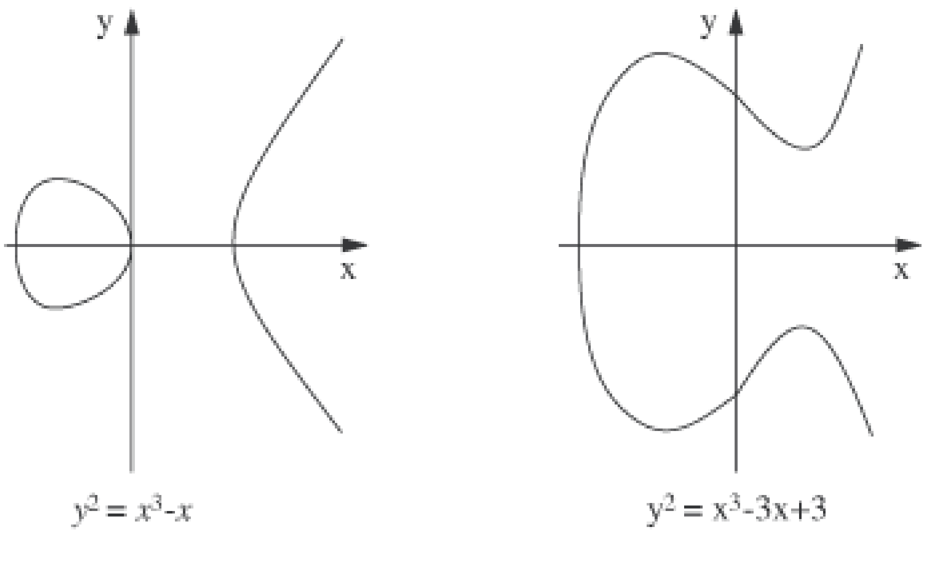
ax^3+bx^2+cx+d
alakú polinom, melynek három különböző gyöke van. A görbe *diszkriminánsa* a

$\Delta=(a_1-a_2)^2(a_1-a_3)^2(a_2-a_3)^2$
mennyiség, ahol

a_1,a_2,a_3
az

$f(x)$
polinom gyökei. A gyökök különbözősége egyenértékű a

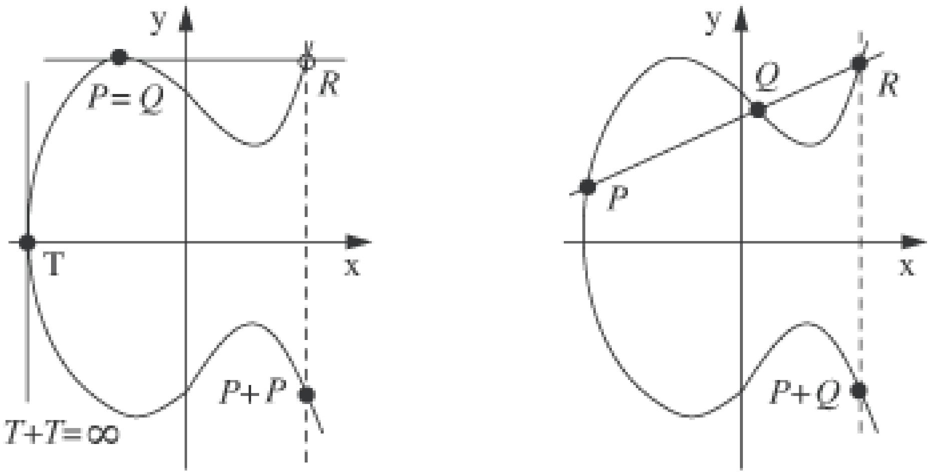
$\Delta\neq0$
feltétellel. A következő ábra két elliptikus görbe grafikonját mutatja.



Elliptikus görbék

Az elliptikus görbéknek sok szép geometriai és számelméleti tulajdonsága van. Egy ilyen érdekes tulajdonság, hogy lehet egy az összeadásra emlékeztető műveletet definiálni a görbe pontjain. Ebből a célból még vegyünk a görbéhez egy

∞-nel jelölt „pontot”! Erről a mágikus pontról feltételezzük, hogy rajta van minden függőleges (az y-tengellyel párhuzamos) egyenesen, és hogy az x-tengelyre vonatkozó tükörképe önmaga. Ezután a görbe P és Q pontjainak $P \square Q$ összegét a következő recepttel határozhatjuk meg: legyen a P, Q pontokon átmenő egyenes és a görbe harmadik metszéspontja R; ennek az x-tengelyre való S tükörképe (ami szintén pontja a görbének) a $P \square Q$ összeg. Ha $P=Q$, akkor az összekötő egyenesükön a görbe P-beli érintőjét kell érteni. Előfordulhat az is, hogy R megegyezik a P, Q pontok valamelyikével; ekkor az egyenes R-ben érinti a görbét. Végül legyen $\infty \square \infty = \infty$.



Összeadás elliptikus görbe pontjain

A

□
műveletre teljesülnek az összeadás szokásos azonosságai, és

∞
játssza a 0 szerepét: a görbe tetszőleges

P,*Q*,*R*
pontjaira

P□*Q*=*Q*□*P*
,

∞□*P*=*P*
,

(*P*□*Q*)□*R*=*P*□(*Q*□*R*)
. Ezek a tulajdonságok, kivéve az utolsót, az asszociatív szabályt, könnyen igazolhatók. Szintén egyszerű belátni, hogy a görbe tetszőleges

P
pontjához a

P
-nek az

x
-tengelyre való

R
tükörképe az egyetlen olyan pont, amelyre

P□*R*=∞
teljesül. Az összeadásnál megszokott értelemben használhatjuk tehát az

R=□*P*
jelölést. Jelöljük

E-vel az $y^2=x^3-2x$ görbét!

Példa. Tekintsük az E görbe $P_1=(0,0)$, $P_2=(\sqrt{2},0)$ és $P_3=(2,2)$ pontjait! Az □ művelet definícióját használva könnyen adódik, hogy $P_1 □ P_1=\infty$, $P_1 □ P_2=(-\sqrt{2},0)$ és $P_1 □ P_3=(-1,1)$. Nézzük meg közelebbről ezek közül az utolsót: a P_1 és P_3 pontokon átmenő egyenes egyenlete $y=x$. Ennek az E-vel való harmadik metszéspontja $(-1,-1)$. A $(-1,-1)$ pontnak az x-tengelyre vonatkozó tükörképe pedig $(-1,1)$.

Két pont összegének koordinátái kifejezhetők az összeadandók koordinátáival és az elliptikus görbe

a,b,c,d együtthatóival, mégpedig csak a $+,-,\square,/$ műveletek segítségével. Ebből két fontos következtetés vonható le. Egyik, hogy a $\square$ művelet definiálható algebrai úton, koordinátákkal. Másfelől, ha az a,b,c,d együtthatók mind racionális számok, akkor racionális koordinátájú pontok összege is racionális pont lesz (a ∞ -t racionális pontnak tekintjük). Szemléltetésül nézzük, hogyan számíthatók ki az E görbe egy $P=(x,y)$ pontjára a $P\square P$ pont u, v koordinátái:

$$u=-2x+((3x^2-2/2y))^2,\quad v=-y+(3x^2-2/2y)(x-u). \qquad (\square)$$

A kifejezések nem értelmesek, ha $y=0$. Ez azt a tényt tükrözi, hogy ekkor $P\square P=\infty$; másképpen fogalmazva a görbe P -beli érintője függőleges.

Ha a görbe együtthatóit jelentő a,b,c,d számok egészek, akkor tetszőleges p prímszámra tekinthetjük az $y^2\equiv ax^3+bx^2+cx+d \pmod p$ kongruenciát. Ennek egy megoldásán egy egész számokból álló

(u,v)
párt értünk, melyre

$v^2-au^3-bu^2-cu-d$
osztható

p
-vel. Vegyük észre, hogy csak az

u,v
számok

p
-vel való osztási maradékán múlik, hogy az

(u,v)
pár megoldás-e. Ezért a megoldásokról teljes képünk marad, ha kikötjük még a

$0\le u,v<p$
egyenlőtlenségeket. Szokás még ezeket a megoldásokat a görbe modulo

p
pontjainak is nevezni. Az

E
görbe modulo 5 pontjait rövid számolás után megkaphatjuk, észrevéve, hogy

v^2
maradéka csak 0, 1 vagy 4 lehet:

$(0,0)$
,

$(1,2)$
,

$(1,3)$
,

$(2,2)$
,

$(2,3)$
,

$(3,1)$
,

$(3,4)$
,

$(4,1)$
,

$(4,4)$
.

Egy egészegyütthatós

F
elliptikus görbe és egy

p
prímszám esetén jelölje

$mp(F)$
az

F
görbe modulo

p
pontjainak számát! Ezzel a jelöléssel az előbbiek alapján

$m5(E)=9$
.

Néhány, a görbéhez képest rosszul viselkedő prímszámot kivéve – ezek a prímek mind osztói a

Δ
diszkriminánsnak – értelmezhető a

□
összeadás az egészegyütthatós görbék modulo

p
pontjain is. Itt is szükség van a

∞
pontra, és két pont összegét ugyanazokkal az algebrai kifejezésekkel számíthatjuk ki, amelyek a valós pontokra megadják az összeg koordinátáit. Például az

E
görbe modulo 5 pontjaira használhatók a (

□
) formulák, ha a

$P□P$
összeget akarjuk meghatározni. Megemlítjük még, hogy tetszőleges racionális

a,b,c,d
együtthatók esetén is értelmezhetők a modulo

p
pontok és az

mp
mennyiségek mindazokra a

p
prímekre, amelyek nem osztói egyik együttható nevezőjének sem.

A racionális együtthatós elliptikus görbék számelméleti tulajdonságainak vizsgálata (egész és racionális koordinátájú pontok, modulo p pontok, a $\square$ művelet hatása ezeken) ebben a században vált intenzívvé. Érdekességként azért megemlítjük, hogy az egyik első ilyen jellegű állítást szintén Fermat margószéli feljegyzései között találták. Arról a tényről van szó, hogy az $y^2=x^3-2$ egyenletnek csak két egész megoldása van, nevezetesen $(3,\pm5)$. Igen szép és fontos eredmény L. J. Mordell tétele (1921): egy racionális együtthatós görbének van véges sok racionális koordinátájú pontja, melyekből az összes racionális pont megkapható a $\square$ művelet alkalmazásával.

A modulo p pontokkal kapcsolatban Emil Artin a doktori disszertációjában (1924) fogalmazta meg sejtésként a $|p-m_p(F)|\le 2\sqrt{p}$ egyenlőtlenséget. Itt F tetszőleges egész együtthatós elliptikus görbe, p pedig a görbe diszkriminánsát nem osztó prímszám. Ennek a sejtésnek és általánosításainak vizsgálata egy egészen új tudományterület, az *aritmetikai geometria* kialakulásához, majd pedig fantasztikus sikereihez vezetett. Ezt mutatja egyebek között, hogy művelői közül már ketten is elnyerték a matematikai Nobel-díjnak számító *Fields-érmet*: Pierre Deligne és Gerd Faltings. Utóbbi lényegében a már említett végességi tételével érdemelte ki a kitüntetést. Artin sejtését Helmut Hasse igazolta 1934-ben.

1955-ben egy fiatal japán matematikus, Taniyama Yutaka igen merészen hangzó sejtést tett közzé a racionális együtthatós elliptikus görbékről. Később ennek Shimura Goro japán és André Weil francia kutatók finomabb megfogalmazását adták. Ezt a hipotézist Taniyama–Shimura–Weil-sejtésnek (röviden TSW-sejtés) fogjuk nevezni. Az ebben foglalt állítások kimondásához nagyon sok előkészület kellene, amire itt nem vállalkozhatunk. Érzékeltetésül csak annyit, hogy a TSW-sejtés szerint tetszőleges racionális együtthatós $y^2=ax^3+bx^2+cx+d$ görbéhez vannak olyan nem állandó $f(z)$ és $g(z)$ *moduláris függvények*, melyekre $f(z)^2=ag(z)^3+bg(z)^2+cg(z)+d$ teljesül. A moduláris függvényekről most csak annyit jegyzünk meg, hogy komplex változós komplex értékű függvények, és hogy elképesztően gazdag szimmetriákkal[29] rendelkeznek. Az első moduláris függvényeket Henri Poincaré fedezte fel a múlt század végén. A sok szimmetriát annyira valószerűtlennek tartotta, hogy hetekig méregette bizalmatlanul, hibára vadászva a számításait, míg végre el tudta fogadni, hogy a mesés tulajdonságú függvények tényleg léteznek.

A TSW-sejtésnek van egy (nem kevésbé bonyolult) számelméleti megfogalmazása is. Ennek különlegessége, hogy az állítás csupán a görbe modulo

p

tulajdonságain múlik; valójában pusztán az

mp

(

$p=2,3,5,\ldots$

) számsorozat tulajdonságaként is kifejezhető. Ezen a ponton érdemes rámutatni a sejtés egy sajátos vonására. Nevezetesen arra, hogy három – egymástól látszólag távoleső – terület között létesít kapcsolatot. Ezek: a komplex függvények világa (moduláris függvények), a geometria (elliptikus görbék) és a számelmélet (kongruenciák megoldásai). A TSW-sejtéssel kapcsolatos első – és sokáig utolsó – jelentős eredményt Shimura érte el 1971-ben, megmutatva, hogy teljesül görbék egy (véges) családjára.

A következő fontos esemény már összefűzi a két szálat, az Utolsó Tételt és az elliptikus görbéket. 1985-ben Gerhard Frey német matematikus a Fermat-egyenlet tanulmányozása során meglepő kapcsolatot talált. Tegyük fel, hogy

$n>3$

prím, és a páronként relatív prím

a,b,c

egészekre teljesül, hogy

$an+bn=cn$

,

b

páros és

$a+1$

osztható 4-gyel. Könnyű meggondolni, hogy ha az Utolsó Tétel nem igaz, akkor ilyen

n,a,b,c

négyes létezik. Egy ilyen „megoldáshoz” Frey a következő egészegyütthatós elliptikus görbét rendelte (ún. Frey-görbe):

$y2=x(x–an)(x+bn)$.

A görbe diszkriminánsa:

$\Delta=a2nb2nc2n$

szép szimmetrikusan függ az

a,b,c

[29] Szimmetrián – némi egyszerűsítéssel – itt a függvény értelmezési tartományának olyan transzformációját értjük, amely a függvény értékét nem változtatja meg. Például a jól ismert

$\sin(x+2\pi)=\sin(x)$

azonosságra gondolva elmondhatjuk, hogy az

x□$x+2\pi$

eltolás a szinuszfüggvény egy szimmetriája.

számoktól. A német kutatót vizsgálatai ahhoz a meggyőződéshez vezették, hogy *a Frey-görbékre nem teljesülhet a TSW-sejtés*. Volt elképzelése a bizonyításról is, de egy jelentős részkérdést nem tudott kezelni. Néhány hónappal később Frey vázlatos gondolatmenetéből a francia Jean-Pierre Serre pontos hipotézist formált. Egy levelében úgy fogalmazott, hogy

TSW+□

-ból következik a Fermat-sejtés. Serre hipotézise innen kapta az

□

-sejtés becenevet. Nem kellett sokáig várni a következő áttörésre: Kenneth A. Ribet amerikai matematikus 1986-ban virtuóz érveléssel igazolta az

□

-sejtést. Ribet munkája nyomán tehát világossá vált, hogy a TSW-sejtésből következik a Fermat-sejtés. Az eredmény óriási port vert fel – egyelőre még csak a matematikusok körében. Sok helyen taglalták izgalommal és csodálkozva, hogy az

n

-edfokú Fermat-egyenlet megoldhatóságának kérdése *harmadfokú* egyenletekhez köthető.

Alighanem elmondhatjuk, hogy az Utolsó Tétel ettől fogva Andrew Wiles asztalára került. Wiles pályája kezdete óta foglalkozik elliptikus görbék aritmetikájával. Első messzire szóló – tanárával, John Coates-szal közösen elért – eredménye e tárgyban 24 esztendős korában, 1977-ben jelent meg. A princetoni egyetem professzoraként már régóta a témakör egyik vezető tekintélyének számított. Ribet eredménye nyomán úgy vélte, hogy gyermekkori álmát, a Fermat-sejtés megoldását, a hazai pálya előnyével kísérelheti meg.

Hét esztendeig dolgozott a TSW-sejtésen teljesen egyedül és titokban. Még legközelebbi kollégái, ismerősei sem tudtak nagy munkájáról. Nem nehéz elgondolni, miért döntött így. Az első időkben nyilván szerette volna elkerülni, hogy afféle reménytelen vállalkozásba bonyolódott La Mancha lovagjának tartsák. Később meg, amikor már egyre közelebb került a célhoz, nem akarta másnak átengedni a befejezés dicsőségét. Ezért nem hozta nyilvánosságra az időközben elért, önmagukban is nagy jelentőségű részeredményeit.

Szinte a semmiből indulva kellett óriási utat megtennie. Ennek érzékeltetésére talán elég, ha megjegyezzük, hogy Shimura már említett 1971-es munkája óta nem volt érdemi előrelépés a TSW-sejtéssel kapcsolatban. Wiles később úgy beszélt erről az időszakról, mint valami ismeretlen, sötét kastély bejárásáról, felfedezéséről:

„Először belépsz a kastély első termébe, amely teljesen sötét. Botorkálsz körbe, beleütközve a bútorokba, de aztán fokról fokra feltérképezed a tárgyak helyét. Végül, mondjuk hat hónap múlva, megtalálod a villanykapcsolót. Felkapcsolod a villanyt és egyszerre minden világos lesz. Pontosan látod, merre jártál. Ezután a következő terembe lépsz, és további hat hónapot töltesz a sötétben. Szóval ezek az áttörések, amelyek néha egy pillanat, máskor egy-két nap művei, ezek a többhónapos sötétben botorkálásokból öltenek alakot, amelyek megelőzik őket; nélkülük nem lennének."

Matematikai eredmények és módszerek félelmetes mennyiségű és mélységű arzenálját kellett bevetnie. Bizonyos kérdések kezelésére neki magának kellett teljesen új eszközöket kidolgoznia. Hónapról hónapra épült a nagy bizonyítás. Végül aztán 1993 tavaszán úgy érezte, hogy legyőzte a sárkányt.

Wiles az 1993 júniusában Cambridge-ben rendezett számelméleti konferencián terítette ki kártyáit, mégpedig nem kis dramaturgiai érzékkel. Három egymást követő napon tartott előadást a témáról. A második előadás végére kiderült, hogy *végtelen sok elliptikus görbére igazolta a TSW-sejtést*. A TSW-sejtés tehát felébredt húszesztendős Csipkerózsika-álmából. A harmadik előadás előtt a hallgatóság izgatottan találgatta, vajon kijön-e a sejtés *elég sok görbére* ahhoz, hogy abból már következzék az Utolsó Tétel. 1993. június 23-án a harmadik előadáson kiderült, hogy *igen*: a TSW-sejtést bizonyítani tudja elliptikus görbék egy nagy osztályára, melybe, ha léteznének, beletartoznának a Frey-görbék is. Ez az osztály a *félig stabil* görbékből áll. A félig stabilitás feltétele a modulo

p
redukált görbékre vonatkozó kikötés, ahol

p□Δ
. A feltételt nem nehéz megérteni

$y2=(x–A)(x–B)(x–C)$
alakú görbék esetén, ha

A,B,C
egészek. A modulo

p
redukált görbe akkor tesz eleget a kikötésnek, ha az

A,B,C
számok nem mind adják ugyanazt a maradékot

p
-vel osztva (

$p=2,3$
-ra valamivel bonyolultabb a helyzet, ezt nem részletezzük). Felhasználva, hogy

a,b,c
páronként relatív prímek, a Frey-görbére

$p>3$
esetén egyszerű ellenőrizni a félig stabilitás feltételét

$(A=0, B=an, C=–bn)$
.

Az Utolsó Tétel bizonyítását hallatlan lelkesedéssel fogadta a világ. Napilapok taglalták a sejtés történetét. A szakértők elragadtatott hangnemben méltatták Wiles csodálatos bizonyítását. Ugyanakkor megkezdődött a részletek alapos elemzése, a bizonyítás helyességének gondos ellenőrzése. Pár hónap elteltével Nicholas M. Katz amerikai matematikus komoly hiányosságot talált Wiles érvelésében. A nehéz hét esztendő első felében elsősorban Viktor A. Kolyvagin orosz, Karl Rubin amerikai és Matthias Flach német kutatók munkája nyomán nagyerejű módszer körvonalazódott, amellyel igen szép új összefüggéseket találtak egész együtthatós egyenletek (racionális számok köréből való) megoldhatósága és kongruenciaként való megoldhatósága között. Wiles érvelése használta a Kolyvagin–Rubin–Flach-módszert, és a hiányosság éppen ennek az alkalmazásában volt.

A hiba ellenére a szakértők optimisták maradtak. Egyfelől azt hangsúlyozták, hogy a bizonyítás „ép” részeiből még mindig végtelen sok görbére következik a TSW-sejtés. Másfelől a hiányosságot pontosan kifejező állításról úgy vélték, hogy igaz, és hogy csak idő kérdése a bizonyítás. Így nyilatkozott például Wiles tanára, John Coates is az 1994 tavaszán Budapesten tartott előadásán.

Wiles egy idő után más utat választott. A Kolyvagin–Rubin–Flach-módszer helyett visszatért egy korábbi elgondolásához, amellyel 1991-ben próbálkozott – akkor még eredménytelenül. Ezúttal egykori diákjával, Richard Taylorral együttműködve sikerült életet lehelniük a módszerbe. 1994. október 24-ére elkészült az Utolsó Tétel bizonyítását tartalmazó két kézirat. Az egyik, a hosszabb tartalmazza a hétéves munka eredményeit. A másiknak Taylor a társszerzője. Ebben kapott helyet a korábbi hibás láncszemet pótló algebrai tétel.

A szakértő bírálók ezúttal már nem találtak kivetnivalót a csudálatos bizonyításban. A két dolgozat 1995-ben megjelent az *Annals of Mathematics*-ban; valójában együtt alkotják a folyóirat 130 oldal terjedelmű májusi számát. Azóta szerte a világban több egyetemi szemináriumon dolgozták fel a bizonyítást. Találtak néhány egyszerűsítést, és a módszerek hatókörét sikerült kiterjeszteni. Többen is úgy vélik, hogy a teljes TSW-sejtés bizonyítása elérhető közelségbe került.

Az Utolsó Tétel 350 év után végre bizonyítást nyert. Ragyogó teljesítményéért Wiles méltán kapta meg 1997-ben a Wolfskehl Díjat. És mi lesz ezután? Maradt-e még a Fermat-sejtéshez fogható nagy kihívás a matematikában? A válasz feltétlenül *igen*. Ott van például a *Laglands-program*, hogy csak egyet említsünk. A Robert P. Laglands amerikai matematikus által körvonalazott, grandiózus sejtésekből álló elmélet, aminek csak kicsinyke része a TSW-sejtés és immár a Wiles-tétel. Amire még inkább igaz, hogy káprázatos, mély kapcsolatokat jósol az analízis, a geometria és a számelmélet szemre távoleső vidékei között, és ami komoly rokonságban van a fizikusok álmával, az alapvető kölcsönhatásokat összefogó Nagy Egyesített Elmélettel. Megannyi csudálatos bizonyítás várja felfedezőjét.

A BARKOCHBA JÁTÉK ÉS AZ INFORMÁCIÓELMÉLET

RÉNYI ALFRÉD

1. 1. BEVEZETÉS

Elterjedt előítélet, hogy a matematikát a képletek teszik nehézzé. Ennek éppen az ellenkezője igaz: képletek nélkül a matematika sokkal nehezebb volna. A képleteknek éppen az a célja, hogy megkönnyítsék a bonyolult matematikai gondolatmenetek elvégzését. A képlet-nyelv tömörebb, szabatosabb, csökkenti a hibalehetőségeket és könnyebben ellenőrizhető, mint a mindennapi nyelv. A képlet ugyanúgy segédeszköze a matematikusnak, mint a számológép.

Mi hát a „nehéz” a matematikában? Abban ugyanis – a matematikusokon kívül – szinte mindenki egyetért, hogy a matematika nehéz; egy ennyire elterjedt közhitnek kell, hogy legyen valami reális alapja. Az igazi nehézséget kétségtelenül a matematikai fogalmak és gondolkodásmód alkotják. Ha azonban valaki ezen a nehézségen túljutott, az alapfogalmakat teljesen megértette, és a szabatos matematikai gondolkodásmódban bizonyos gyakorlatra tett szert, a nehézség mintegy szertefoszlik.

A matematika lényegét a matematikai fogalmak és a matematikai gondolkodásmód alkotják.

A matematika tanításában és a matematikai tudományos ismeretterjesztésben ezért a hangsúlyt az alapvető fogalmak magyarázatára, a matematikai gondolkodásmód megismertetésére kell helyezni. Vonatkozik ez arra is, amikor a matematika egy új ágát, irányát igyekszik valaki ismertetni. Ugyanis bár beszélhetünk a matematikai gondolkodásmódról, mint olyanról, ami a matematika minden fejezetében érvényesül, emellett azonban a matematika minden egyes ágának, fejezetének megvan a maga sajátos gondolkodásmódja. A valószínűségszámítást például csak az értheti meg igazán, aki hozzászokott ahhoz, hogy véletlenszerűen változó mennyiségekben, valószínűségi változókban gondolkodjék.

Az információelméletnek, a matematika és közelebbről a valószínűségszámítás nagy jelentőségű új ágának is megvan a maga sajátos gondolkodásmódja, ha tetszik, „filozófiája”. Jelen ismertetésben a hangsúlyt az információelméleti gondolkodásmód ismertetésére igyekszem fektetni, az információelmélet szemléletmódjába próbálom az olvasót bevezetni. Meggyőződésem, hogy aki e szemléletmódot megértette és elsajátította, az sokkal könnyebben meg fogja érteni az információelmélet konkrét eredményeit.

Az információelmélet alapfogalmait és sajátos szemléletmódját a közismert Barkochba játék példáján keresztül igyekszem ismertetni. Az információelmélet konkrét eredményeinek ismertetése nem célja e cikknek. Igyekeztem a képletek használatát a minimumra szorítani. A valószínűségszámítás elemeinek ismeretét azonban fel kell tételeznem.[30]

2. 2. A BARKOCHBA JÁTÉK ÉS AZ INFORMÁCIÓ MÉRÉSE

A Barkochba játéknál az egyik játékos – nevezzük a „felelő”-nek – gondol valamire (vagy valakire), a másodiknak – őt nevezzük „kérdező”-nek – ki kell találnia, hogy a „felelő” mire gondolt; ecélból tetszőleges,

[30] A cikk megértéséhez szükséges elemi valószínűségszámítási ismereteket az olvasó megtalálhatja pl. [1] első három fejezetében vagy a [2] könyvben.

igennel vagy nemmel megválaszolható kérdéseket tehet fel, amelyekre a „felelő”-nek legjobb tudása szerint, az igazságnak megfelelően kell válaszolnia. E válaszokból kell a „kérdező”-nek kitalálnia, hogy mire gondolt a „felelő”. Akik gyakran játsszák e játékot, általában megállapodnak abban, hogy mire lehet gondolni. Akinek van gyakorlata, tudja, hogy viszonylag könnyen kitalálható élő vagy történelmi személyek neve (feltéve, hogy a „kérdező” tud arról a személyről, akire a „felelő” gondolt). Pl. ha Pascalra gondolt az első játékos, a kérdezés mehet a következőképpen:

Személy?	Igen
Él?	Nem
Művész?	Nem
Tudós?	Igen
1900 után született?	Nem
1700 után született?	Nem
1600 után született?	Igen
Angol vagy olasz?	Nem
Német vagy Holland?	Nem
Francia?	Igen
Matematikus?	Igen
Filozófus is volt?	Igen
Janzenista volt?	Igen

Ezek után a második játékos már közölheti, hogy kitalálta, hogy Pascalról van szó. Látszik a kérdésekből, hogy amikor a kérdező odáig eljutott, hogy egy francia matematikusról van szó a XVII. században, megfordult fejében Pascal, Descartes, Desargues, Fermat neve; a két utóbbit kizárta azzal a kérdéssel, hogy filozófus is volt-e, míg Descartes-ot az utolsó kérdéssel zárta ki, úgyhogy ezek után szinte biztosra vehette, hogy Pascalról van szó.

Mármost vizsgáljuk meg azt, hogy általában hány kérdéssel lehet valamit a Barkochba játékban kitalálni. Az, hogy hány kérdéssel sikerül kitalálni a gondolt dolgot, persze függ a kérdező gyakorlatától, műveltségétől, fantáziájától, pszichológiai érzékétől, sőt bizonyos mértékben a véletlentől is, de döntő mértékben mégis attól függ, hogy milyen tág azoknak a dolgoknak a köre, amelyekre a játékosok megállapodása szerint gondolni lehet. Ha például szabad olyasmire gondolni, mint „az a jelző, amivel a New York Times tudósítója jellemezné Cleopatra ruháját, ha Cleopatra feltámadna és mint egyiptomi államfő megjelenne az ENSZ-ben”, a kérdezőnek nyilván jóval több kérdésre van szüksége, és nemcsak a kérdezés, de a kérdések helyes megválaszolása sem egyszerű. Ahhoz, hogy választ kapjunk arra, hogy hány kérdéssel lehet valamit kitalálni, próbáljuk egyszerűsíteni a problémát azáltal, hogy kikötjük, hogy a „felelő” csak egy hatjegyű számra gondolhat. (6-nál kevesebb jegyű számokat is megengedünk, mert 0-k eléírásával ezek 6 jegyűvé tehetők; pl.

313=000 313
.)

Ebben az esetben azoknak a dolgoknak a száma, amelyekre gondolhat a „felelő”,

1 000 000
. Könnyen belátható, hogy első kérdésként célszerű pl. azt a kérdést feltenni: a gondolt hatjegyű szám kisebb-e, mint

500 000

? Ez esetben ugyanis akár igen a válasz, akár nem, az első kérdésre kapott válasz után a még számba jövő lehetőségek száma

500 000
-re csökken. (Ha a válasz „igen”, a szám

000 000
és

499 999
között kell, hogy legyen, míg, ha a válasz „nem”, a szám

500 000
és

999 999
között kell, hogy legyen.) Persze azt is kérdezhetjük, hogy a szám legalább

500 000
-e; ez tulajdonképpen ugyanaz a kérdés másképp megfogalmazva. Hasonlóképpen kérdezhetjük, hogy a szám páros-e, ill. bármilyen más olyan kérdést feltehetünk, amelyre

500 000
hatjegyű szám esetében „igen” és a többi

500 000
hatjegyű szám esetében „nem” a válasz. Könnyű belátni azonban, hogy csak ezek a célszerű kérdések, mert ha olyan kérdést teszünk fel, amelyre

500 000
-nél több szám esetében „igen” a válasz, úgy, ha igenlő választ kapunk, kedvezőtlenebb helyzetben vagyunk, mint pl. ha azt kérdeztük volna, hogy páros-e a gondolt szám. Ha viszont olyan kérdést teszünk fel, amelyre

500 000
-nél kevesebb szám esetében „igen” a válasz, nemleges válasz esetében kerülünk kedvezőtlenebb helyzetbe. Hasonlóképpen a második kérdéssel a számításba jövő számok számát megint meg tudjuk felezni, azaz le tudjuk szállítani

250 000
-re a harmadik kérdéssel

125 000
-re, a negyedikkel

62 500
-ra, az ötödikkel

31 250
-re, a hatodikkal

15 625
-re. Mivel

15 625
páratlan szám, ennyi lehetőséget nem lehet két egyenlő részre bontani, hanem csak úgy, hogy az egyik válasz esetén 7813, a másiknál 7812 lehetőség maradjon nyitva. Így jól választott hetedik kérdés után a fennmaradó lehetőségek száma legfeljebb 7813, a nyolcadik után hasonlóképpen 3907, a kilencedik után 1954, a tizedik után

977, a tizenegyedik után 489, a tizenkettedik után 245, a tizenharmadik után 123, a tizennegyedik után 62, a tizenötödik után 31, a tizenhatodik után 16, a tizenhetedik után 8, a tizennyolcadik után legfeljebb 4, a tizenkilencedik után legfeljebb 2. Így tehát – ha előbb nem – a huszadik kérdésre biztosan kitalálhatjuk a gondolt számot. Ha meggondoljuk, hogy miért éppen 20 kérdés adódott, könnyű észrevenni, hogy ez azon múlott, hogy

219=524 288<1 000 000<1 048 576=220.

Ebből az is látszik, hogy 20 kérdéssel nemcsak

1 000 000
közül lehet kitalálni a gondolt dolgot, hanem még

1 048 576
közül is, de már ha

1 048 577
dologra szabad gondolni, akkor 20 kérdés nem biztos, hogy elég a kitalálásra.[31] Ugyanis minden egyes kérdéssel felére tudjuk csökkenteni a még számításba jövő lehetőségek számát; így tehát ha eredetileg

220
dolog jött tekintetbe,

k
kérdésre kapott válasz után a még megmaradó lehetőségek száma már csak

220–k
lesz és így 20 kérdés után – mivel

20=1
– már csak egyetlen lehetőség marad. Könnyű belátni, hogy ha a lehetőségek száma eredetileg pontosan

220
volt, és az ismertetett kérdezésmódot alkalmazzuk, akkor a gondolt dolgot mindig pontosan a 20-ik kérdésre fogjuk kitalálni, és sohasem előbb. Persze ez a kérdezésmód teljesen gépies, és ezáltal a játék érdektelenné válik; az említett kérdezési módszerrel egy számítógép is tud Barkochbát játszani.

Az elmondottakból az is következik, hogy ha

N
-féle dologra szabad gondolni, a gondolt dolgot mindig ki lehet találni

K
kérdéssel, ahol

K
a legkisebb egész szám, amelyre

2K≥N
, tehát

K
a legkisebb egész szám, amelyre

K≥log2N

[31] Amerikában egyébként a Barkochba játék neve „húsz kérdés”, és ott úgy játsszák, hogy legfeljebb 20 kérdéssel kell kitalálni a gondolt dolgot.

.32

Az eddig elmondottakból úgy tűnik, mintha minden egyes kérdés megfogalmazása az előző kérdésekre kapott választól kellene, hogy függjön. Ez azonban nincs így. Ha nem számít az, hogy a kérdés minél egyszerűbben legyen megfogalmazható, meg lehet a kérdéseket előre adni úgy, hogy a válaszokat a kérdező csak az összes kérdés feltevése után kapja meg. Például, ha csak a 0, 1, 2, 3, 4, 5, 6, 7 számok valamelyikére szabad gondolni, egy lehetséges optimális kérdésrendszer (amely persze 3 kérdésből áll, mivel

8=23
) a következő:

1. kérdés: A gondolt szám az

1 2 3 4
számok egyike?

2. kérdés: A gondolt szám az

1 2 5 6
számok egyike?

3. kérdés: A gondolt szám az

1 3 5 7
számok egyike?

E 3 kérdésre kapott válaszból a gondolt szám egyértelműen kitalálható a következőképpen:

Ha a válaszok a 3 kérdésre			
a fenti sorrendben			akkor a gondolt szám
igen	igen	igen	1
igen	igen	nem	2
igen	nem	igen	3
igen	nem	nem	4
nem	igen	igen	5
nem	igen	nem	6
nem	nem	igen	7
nem	nem	nem	0

E kérdezési rendszer lényegét még jobban megvilágítja a következő megjegyzés. Tegyük fel, hogy az első

2*N*
nem negatív egész számra lehet gondolni. E számokat állítsuk elő a kettes számrendszerben; könnyen beláthatjuk, hogy a 0, 1, …,

2*N*–1

32

log2*N*
itt és a következőkben az

N
szám 2 alapú logaritmusát jelenti.

számok mindegyike a 2-es számrendszerben egy pontosan

N

jegyből álló számként állítható elő, ha megengedjük, hogy egy szám előállítása zérussal, ill. zérusokkal is kezdődhet. Pl. ha

$N=3$

, az első nyolc szám előállítása:

0 = 000 4 = 100 1 = 001 5 = 101 2 = 010 6 = 110 3 = 011 7 = 111

Mármost, ha a „felelő” csak a 0, 1, 2, …, 7 számok valamelyikére gondolhat, a gondolt számot egyszerűen kitalálhatjuk a következő 3 kérdéssel:

1. A gondolt számot a 2-es számrendszerben 3 jegyű számként felírva az első jegy 0-e?
2. A gondolt számot a 2-es számrendszerben 3 jegyű számként felírva a második jegy 0-e?
3. A gondolt számot a 2-es számrendszerben 3 jegyű számként felírva a harmadik jegy 0-e?

Ha pl. a 3 kérdésre a válaszok: igen–nem–nem, akkor a gondolt szám a 2-es számrendszerben felírva 011, tehát a szám 3.

Mármost állapodjunk meg abban, hogy a gondolt szám, ill. dolog kitalálásához szükséges információ mennyiségén a kitaláláshoz szükséges kérdések számát értjük, feltéve, hogy a lehető legjobb kérdezési rendszert használjuk. Ez azt jelenti, hogy az információ egységének az egyetlen egy igen–nem válaszban foglalt információt választjuk. Az információ egységét *bit*-nek nevezzük.[33]

Így tehát pl. ha valaki elkészíti a keresztnevek egy listáját, amely 512 keresztnevet tartalmaz,[34] és a játékosok megállapodnak abban, hogy csak e listában szereplő keresztnévre szabad gondolni, úgy, mivel

$512=29$

, a gondolt keresztnév kitalálásához 9 bit információra van szükség, és – mivel egy kérdésre kapott válasz legfeljebb egy bitet tartalmazhat – tehát legalább 9 kérdés kell a gondolt keresztnév kitalálásához. A mondottakhoz azok megvilágítására néhány megjegyzést fűzünk.

1. *megjegyzés.* Az előbb azt mondottuk, hogy egy igen–nem válasz *legfeljebb* egy bit információt tartalmaz. Valóban, ha pl. az előző példában az első kérdést úgy tesszük fel, hogy nem 256, hanem kevesebb esetben igenlő a válasz, akkor nemleges válasz esetén még 256-nál több lehetőség maradt; ez esetben a válasz 1 bitnél kevesebb információt nyújtott.
2. *megjegyzés.* Ahhoz, hogy valóban 9 kérdéssel ki tudjuk találni, hogy az 512 keresztnév közül a felelő melyikre gondolt, szükséges, de nem elégséges, hogy minden egyes kérdés 1 bit információt nyújtson, vagyis, hogy mindegyik kérdés olyan legyen, amelyre az 512 keresztnév közül pontosan 256 esetben igen és 256 esetben nem a válasz. Lehetséges ugyanis, hogy két kérdés külön-külön jó, de nem illenek össze: külön-külön ugyan mindegyikre a válasz egy bit információt tartalmaz, de a kérdésekre kapott válaszban foglalt információk részben vagy egészben fedik egymást, vagyis legalább részben ugyanazt az információt kapjuk meg kétszer. Kirívó példa erre, ha pl. az első kérdést másodszorra megismételjük; ez esetben a második kérdésre kapott válasz ugyanazt az 1 bit információt nyújtja, mint az első kérdésre adott válasz és így a két válasz – bár külön-külön mindegyik 1 bit információt tartalmaz – együtt nem 2 bit, hanem csak 1 bit információt nyújt. Hasonlóképpen pl. az előbb kétféle módon is megadtunk 3-3 kérdést, amelyekkel ki

33A „bit” szó az angol nyelvből származik és a *bi*nary dig*it* (= kettes számrendszerbeli számjegy) rövidítése.

34A naptárban csak 437 keresztnév van felsorolva, azonban számos olyan keresztnév van, amelyhez nem tartozik névnap, pl. Györgyi, Romeo, Héraklész, Trisztán, Izolda stb.; némi fáradsággal a listát ki lehet egészíteni 512-re.

lehet biztosan találni, hogy a 0, 1, 2, 3, 4, 5, 6, 7 számok közül melyikre gondolt a „felelő”. Ha azonban az első kérdéssorozat első kérdése után (A gondolt szám az 1, 2, 3, 4 számok egyike-e?) a második kérdéssorozat első kérdését tesszük fel (A számot a kettes számrendszerben felírva az első jegy 0-e?), a két kérdésre kapott válasz együtt nem ad 2 bit információt. Ugyanis az első válaszból azt tudjuk meg, hogy a szám az 1, 2, 3, 4 számok között (vagy a 0, 5, 6, 7 számok között) van-e, a második válaszból viszont azt tudjuk meg, hogy a szám a 0, 1, 2, 3 számok között (vagy a 4, 5, 6, 7 számok között) van-e. Így tehát aszerint, hogy a két válasz hogyan hangzik, a még tekintetbe jövő számokat az alábbi táblázat tünteti fel:

	A gondolt szám az
Ha a válaszok	alábbi számok egyike
igen, igen	1, 2, 3
igen, nem	4
nem, igen	0
nem, nem	5, 6, 7

Így tehát például abban az esetben, ha mindkét kérdésre igen a válasz, a még fennmaradó lehetőségek száma 3, és így további 2 kérdésre van szükség. Más szóval a második kérdés ugyan 1 bit információt tartalmazott, de ennek egy része[35] nem volt számunkra új információ, hanem olyan, amivel már rendelkeztünk. Ha viszont az első kérdéssorozat első kérdése után (A gondolt szám az 1, 2, 3, 4 számok között van-e?) a második kérdéssorozat második kérdését tesszük fel (tehát a második válaszból azt tudjuk meg, hogy a gondolt szám a 0, 1, 4, 5 számok között van-e?), akkor a még tekintetbe jövő számokat az alábbi táblázat mutatja:

Ha a válaszok	A még tekintetbe jövő számok
igen, igen	1, 4
igen, nem	2, 3
nem, igen	0, 5
nem, nem	6, 7

Ebben az esetben tehát akármilyen válaszokat kaptunk is a két kérdésre, már csak 2 lehetőség marad nyitva és így a harmadik kérdésre biztosan kitalálhatjuk a gondolt számot. Tehát nemcsak az igaz, hogy a két válasz külön-külön egy-egy bit információt tartalmaz, hanem az is, hogy együtt 2 bit információt adnak, tehát a két válaszban foglalt információ még részben sem fedi egymást, a második kérdésre kapott válaszban foglalt információ teljes egészében új.

3. 3. INFORMÁCIÓ ÉS BIZONYTALANSÁG

A Barkochba játéknál a kérdező, mielőtt az első kérdést feltette volna, bizonytalanságban van arra vonatkozóan, hogy a másik játékos mire gondolt. A bizonytalansága minden egyes kérdésre kapott válasz után csökken, míg végül, amikor kitalálta a másik gondolatát, teljesen megszűnik. A bizonytalanság tehát tulajdonképpen nem más, mint információhiány, azaz negatív információ, míg az információ nem más, mint a bizonytalanság csökkenése

[35] Belátható, hogy a második kérdésre kapott válasz

(1/4)log2(256/27)□0,8
bit új információt és

(1/4)log2(27/16)□0,2
bit olyan információt ad, ami már az előző válaszban benne volt.

(negatív bizonytalanság). Bizonytalanság és információ tehát tulajdonképpen ugyanazt jelenti, más-más oldalról nézve, csak előjelben különböznek egymástól. Így tehát célszerű a bizonytalanság nagyságát azzal mérni, hogy mennyi információ szükséges annak teljes megszüntetéséhez. Mondhatjuk tehát, hogy ha a Barkochba játéknál megállapodunk, hogy 512 keresztnév valamelyikére lehet csak gondolni, akkor a kérdezés megkezdése előtt a kérdező bizonytalansága a gondolt keresztnevet illetőleg 9 bit; ez a bizonytalanság a lehető legjobb (vagyis a fennmaradó lehetőségek halmazát mindig két egyenlő részre felező) kérdezés mellett minden egyes válasz után 1 bittel csökken, és így 9 válasz után szűnik meg teljesen. A játék folyamán bármely időpontban az addig kapott információ és a még fennálló bizonytalanság összege (vagyis a már megszerzett és a még hiányzó információ összege) 9-cel egyenlő. A bizonytalanság mennyiségét az információelméletben entrópiának is nevezik.[36]

Vizsgáljuk most meg azt a kérdést, hogy ha a Barkochba játéknál a játékosok abban állapodnak meg, hogy

N
dolog valamelyikére lehet gondolni, és

N
nem egyenlő 2-nek egy hatványával, akkor mekkora a kérdező bizonytalansága a játék kezdetén a gondolt dologra vonatkozólag?

Tegyük fel, hogy

$2K<N<2K+1$
. Ez esetben nyilván

$K+1$
kérdéssel lehet csak kitalálni a gondolt dolgot. Viszont nyilván azért a bizonytalanság kisebb, mint ha tényleg

$2K+1$
(tehát több) lehetőség állna fenn, de nagyobb, mint ha csak

$2K$
(tehát kevesebb) lehetőség állna fenn. Mivel

$K=[\log2N]$
(itt és a következőkben

$[x]$
az

x
valós szám egészrészét jelöli), tehát, ha a kérdező bizonytalanságát (a helyzet entrópiáját)

$H(N)$
-nel jelöljük:

$[\log2N]<H(N)<[\log2N]+1.$

Meg fogjuk mutatni az eddigi meggondolásainkkal összhangban, hogy ilyen estekben ésszerű az

(1)	$H(N)=\log2N$

[36] Az információelméleti entrópia-fogalomnak a statisztikus fizika entrópia-fogalmával való kapcsolatával itt nem foglalkozhatunk részletesen, csak azt jegyezzük meg, hogy a két fogalom szorosan összefügg. A statisztikus fizikában egy fizikai rendszer entrópiáján az illető rendszer rendezetlenségének mértékszámát értjük; ez a szám felfogható úgy is, mint annak mértékszáma, hogy mennyire bizonytalan az adott rendszer részecskéinek állapota.

ún. *Hartley-képlettel* értelmezni a fennálló bizonytalanságot. Az egyszerűség kedvéért ezt azon a példán keresztül mutatjuk be, amikor

$N=3$
; tetszőleges

N
-re lényegében ugyanígy végezhető el a meggondolás.

Tegyük fel tehát, hogy a játékosok abban állapodtak meg, hogy csak 3 dolog valamelyikére lehet gondolni, pl. az

a

,

b

,

c
betűk egyikére. Mivel

$2^1=2<3<4=2^2$
, tehát a játék kezdetén fennálló bizonytalanság nyilván 1 bitnél nagyobb, de 2 bitnél kisebb. Azt, hogy a bizonytalanság ez esetben

$\log_2 3=1,584\,96\ldots,$
a következőképpen láthatjuk be. A kérdező pl. elsőszörre kérdezheti, hogy a gondolt betű az

a
betű-e; ha a válasz igen, már ki is találta a gondolt betűt; ha azonban a válasz nemleges, fel kell tenni a kérdést, hogy a gondolt betű a

b
betű-e. Akármi is erre a kérdésre a válasz, ebből már a kérdező megtudja, hogy mi volt a gondolt betű. Ha sokszor megismétlik ezt a játékot és a felelő mind a 3 betűre ugyanolyan sokszor gondol, akkor az esetek kb.

$(1/3)$
-ában 1 kérdés,

$(2/3)$
-ában 2 kérdés, átlagban tehát

$1\cdot(1/3)+2\cdot(2/3)=(5/3)=1,6666\ldots$
kérdésre van szükség. Így tehát mondhatjuk, hogy a bizonytalanság ennél a játéknál legfeljebb

$(5/3)$
; valójában azonban a bizonytalanság még kisebb, ugyanis az említett „stratégia” (a játék sokszori ismétlése estében) nem a lehető legjobb. Ha a két játékos a játékot nem egyszer, hanem sokszor játssza, megadható egy jobb stratégia. A kérdező nem külön-külön kérdez a betűkre, hanem megkéri a felelőt, hogy egyszerre gondoljon az

a

,

b

,

c betűkből álló 5 tagú betűsorozatra, és megpróbálja e betűsorozatot egyszerre kitalálni. Mivel az *a* , *b* , *c* betűkből 35=243 öttagú betűsorozat képezhető és 243<256=28 , tehát a szóban forgó 5 betűből álló betűsorozatot a kérdező 8 kérdéssel biztosan ki tudja találni, és így egy betű kitalálására átlagban (8/5)=1,6 kérdés jut, ami valamivel kevesebb, mint (5/3) .

Még jobb eredményt kapunk, ha a kérdező egyszerre kérdez az *a* , *b* , *c* betűkből álló 17 tagú betűsorozatra. Az ilyen sorozatok („szövegek”) száma 317=129 140 163<227=134 217 728, tehát a 17 betűt 27 kérdéssel ki lehet találni és így egy betűre átlagban (27/17)=1,588 23… kérdés esik. Általában, ha egy *n* tagú, az *a* , *b* , *c* betűkből álló betűsorozatra egyszerre kérdezünk, és *k*(*n*) jelöli 2-nek a legkisebb olyan hatványát, amely nagyobb, mint

3^n
, tehát

(2)

$$2^{k(n)-1}<3^n<2^{k(n)},$$

akkor

$k(n)$
kérdéssel az egész

n
betűből álló betűsorozatot kitalálhatjuk és így egy betűre

$(k(n)/n)$
kérdés jut. A (2) egyenlőtlenségből (2 alapú logaritmust véve) azonban következik, hogy

$k(n)-1<n\log_2 3<k(n)$
és így

$(k(n)/n)-(1/n)<\log_2 3<(k(n)/n)$,
vagyis

(3)

$$\log_2 3<(k(n)/n)<\log_2 3+(1/n).$$

Így tehát

$(k(n)/n)$
sohasem lesz kisebb, mint

$\log_2 3$
, azonban azt tetszőleges pontossággal meg fogja közelíteni felülről; ha ugyanis pl. azt akarjuk, hogy az egy betűre jutó kérdések száma

$\log_2 3+(1/10\,000)$
-nél kisebb legyen, ehhez csak

n
-et kell

10 000
-nél nagyobbra választani. Így tehát indokolt azt mondani, hogy ha az

a
,

b
,

c
betűk közül valamelyikre lehet gondolni, úgy a kérdező bizonytalansága a gondolt számra vonatkozólag

$\log_2 3=1{,}584\,96\ldots$

Ez azonban nem úgy értendő, hogy (átlagban)

$\log_2 3$
kérdéssel ki lehet találni a gondolt számot, hanem úgy, hogy ennél kevesebb kérdés átlagban nem lehet elég és tetszőleges, ennél nagyobb számnál átlagban kevesebb kérdéssel ki lehet találni a számot alkalmas játékrendszer mellett. (Érdekes megjegyezni, hogy a fent kapott

$(27/17)$
felső becslés a kérdések minimális számára kevesebb, mint

$(4/1000)$
-del haladja csak meg

$\log_2 3$
-at, tehát sokkal jobban közelít, mint amire a (3) egyenlőtlenség alapján biztosan számítani lehetett.) Így tehát bebizonyítottuk, hogy (a bizonytalanság fogalma értelmezésének kézenfekvő kiterjesztése mellett) ha a másik játékos az

a
,

b
,

c
betűk egyikére gondolt, úgy a gondolt betűre vonatkozó bizonytalanságunk

$\log_2 3=1,584\,96\ldots$
bit. Pontosan ugyanúgy látható be, hogy ha

N
dologra lehet gondolni a Barkochba játéknál, úgy a kérdező bizonytalansága a kérdezés megkezdése előtt

$\log_2 N$
bit. Ha

N
éppen 2-nek egy hatványa,

$N=2^k$
, akkor persze

$\log_2 N=k$
, és így a kapott eredmény speciális esetként tartalmazza azt a legelőször tárgyalt esetet, amikor

2^k
dologra szabad gondolni.

4. 4. A SHANNON-FÉLE FORMULA

Térjünk vissza ahhoz a példához, amikor a Barkochbát játszó játékosok abban állapodnak meg, hogy csak 512 keresztnév egyikére lehet gondolni. Ebben az esetben, mint láttuk, a kérdező bizonytalansága a kérdezés megkezdése előtt 9 bit. Közelebbről megvizsgálva a kérdést, könnyű észrevenni, hogy amikor ezt beláttuk, tulajdonképpen hallgatólagosan feltettük, hogy az 512 keresztnév mindegyikére ugyanolyan valószínűséggel gondolt a felelő (pl. az 512 keresztnevet egy-egy cédulára felírja, e cédulákat egy kalapba teszi, megkeveri és találomra kihúz egy cédulát, és az azon álló keresztnévre gondol). Ha ez nem így van, akkor állításunk nem érvényes és módosításra szorul. Ha pl. a felelő annak a személynek a keresztnevére gondol, akivel legutoljára beszélt telefonon, nem jogosult az a feltevés, hogy mind az 512 keresztnév egyformán valószínű, hiszen az egyes keresztnevek gyakoriságai erősen eltérnek. Sokkal valószínűbb, hogy azt, akivel beszélt, Évának, Jánosnak, Andrásnak vagy Máriának hívják, mint hogy Anasztáziának, Szerafinnak, Kleofásnak vagy Upornak. Vizsgáljuk tehát meg a következő, az eddig tárgyaltnál általánosabb kérdést: Jelöljék

x_1
,

x_2
, …,

x_N
azokat a különböző „dolgokat”, amelyekre a Barkochba játéknál gondolni lehet. (

x_1
,

x_2
, …,

x_N
lehetnek nevek vagy általában szavak, számok, tárgyak stb.) Tegyük fel, hogy a kérdező tudja, hogy ezekre a dolgokra a másik játékos milyen valószínűséggel szokott gondolni; jelölje

P_k
annak a valószínűségét, hogy a másik játékos

x_k
-ra gondolt. (Ez esetben természetesen a

P_k
számok pozitívak, és

$P_1+P_2+\ldots+P_N=1$
; ez azt fejezi ki, hogy másra, mint

x_1
,

x_2
, …,

x_N
valamelyikére, nem gondolhat a „felelő” a játékosok megállapodása szerint.)

Kérdés, mekkora ebben az esetben a kérdező bizonytalansága a kérdezés megkezdése előtt? E szám nyilván csak a

P_1
,

P_2
, …,

P_N

valószínűségektől fog függeni, és x_1, x_2, …, x_N-től nem (hiszen pl. teljesen mellékes, hogy magára a gondolt keresztnévre gondolok-e, vagy a keresztnév sorszámára a keresztnevek egy listájában). Így tehát x_k-t mindig helyettesíthetjük k-val, azaz x_k sorszámával, más szóval feltehetjük, hogy az x_1, x_2, …, x_N sorozat az 1, 2, …, N számok sorozata. Jelölje a szóban forgó bizonytalanságot $H(P_1, P_2, \ldots, P_N)$. A $H(P_1, P_2, \ldots, P_N)$ számot a $P_1, P_2, \ldots, P_N$ *valószínűségeloszlás entrópiájának nevezzük*. *C. Shannon* (és tőle függetlenül *N. Wiener*) megmutatta, hogy

$$H(P_1, P_2, \ldots, P_N) = P_1 \log_2(1/P_1) + P_2 \log_2(1/P_2) + \ldots + P_N \log_2(1/P_N). \quad (4)$$

A (4) képletet *Shannon-féle formulának* nevezik.[37] A Shannon-formula helyességét először egy példán keresztül mutatjuk be. Vizsgáljuk a következő példát! Tegyük fel, hogy az első játékos a játék előtt két pénzdarabot feldob úgy, hogy ő látja, de a kérdező nem látja a dobás eredményét. A kérdezőnek azt kell kitalálni, hogy a dobások eredménye 1. mindkét érmével fej, 2. mindkét érmével írás, vagy 3. az egyik érmével fej, a másikkal írás volt-e. A harmadik lehetőségen belül nem teszünk különbséget aközött, hogy melyik érmével dobott a

[37] A Shannon-formula speciális esetként tartalmazza a Hartley-formulát; ugyanis (4)-ből $H((1/N),(1/N),\ldots,(1/N)) = \log_2 N$.

játékos fejet és melyikkel írást (a két érmét ugyanis egyformának tételezzük fel, amikor is e két eset valójában nem is különböztethető meg).

Így tehát az 1, 2 és 3 lehetőségek valószínűségei

$P_1=(1/4)$

,

$P_2=(1/4)$

,

$P_3=(1/2)$
. Nyilván a kérdező a következőképpen járhat el. Először megkérdezi, hogy a harmadik eset következett-e be; ha a válasz igenlő, már ki is találta, hogy mi történt, csak nemleges válasz esetében kell feltennie a második kérdést, ti. hogy a második lehetőség következett-e be. Így tehát az eseteknek kb. a felében 1 kérdéssel, a másik felében 2 kérdéssel ér célhoz, vagyis átlagban

$(1+2/2)=1{,}5$
kérdéssel ki tudja találni, hogy mi történt.

Míg tehát, ha 3 lehetőség valószínűsége egyenlő, mint láttuk, átlagban

$\log_2 3=1{,}584\,96\ldots$
kérdésre van szüksége, ha e lehetőségek valószínűségei

$(1/4),(1/4),(1/2)$
, akkor átlagban 1,5 kérdés is elegendő. A Shannon-formula erre az esetre nézve ugyanazt az eredményt adja, mivel

$(1/4)\log_2 4+(1/4)\log_2 4+(1/2)\log_2 2=(1/4)\cdot 2+(1/4)\cdot 2+(1/2)\cdot 1=(3/2)=1{,}5$.
A tárgyalt példa alapján már kitalálhatjuk, hogy ha a számításba jövő lehetőségek nem egyformán valószínűek, akkor nem arra kell törekednünk, hogy kérdéseinkkel a lehetőségek *számát* csökkentsük a felére, hanem arra, hogy a még fennmaradó lehetőségeket úgy csoportosítsuk két osztályba, hogy a két osztályba sorolt lehetőségek valószínűségeinek összege legyen egyenlő (vagy ha ez nem lehetséges, közel egyenlő). Az előbbi példában ez pontosan keresztülvihető volt; az általános esetben ez nem mindig érhető el; tetszőlegesen megközelíthető azonban az ilyen kérdezési módszer akkor, ha megint arra kérjük a másik játékost, hogy ne egy dologra gondoljon, hanem egyszerre sokra, és ezeket együtt igyekszünk kitalálni. Ennek szabatos bizonyítása hosszadalmas, ezért a bizonyításnak csak a gondolatmenetét vázoljuk, a részletek kidolgozása nélkül.

Ha az első játékos

n
-szer egymás után gondol az

x_1

,

x_2

, …,

x_N
dolgok egyikére, és minden alkalommal az

x_k
-t

$(k=1, 2, \ldots, N)$

P_k valószínűséggel választja (az előző választásoktól függetlenül), akkor a nagy számok törvénye szerint körülbelül nP_1-szer fogja x_1-et, nP_2-ször x_2-t, …, nP_N-szer x_N-et választani, és a gondolt dolgok minden egyes valószínű sorozatának valószínűsége ennek megfelelően körülbelül

$$P_1^{nP_1} \cdot P_2^{nP_2} \cdots P_N^{nP_N}$$

lesz.

Jelölje S a valószínű sorozatok számát, akkor tehát (mivel a valószínűtlen sorozatok együttes valószínűsége igen kicsiny és így a valószínű sorozatok együttes valószínűsége közel 1 lesz),

$$S \cdot P_1^{nP_1} \cdot P_2^{nP_2} \cdots P_N^{nP_N} \approx 1,$$

vagyis

$$S \approx \left[\left(\frac{1}{P_1}\right)^{P_1}\left(\frac{1}{P_2}\right)^{P_2}\cdots\left(\frac{1}{P_N}\right)^{P_N}\right]^n. \qquad (5)$$

E közül a S, közel egyenlő valószínűségű sorozat közül egyet nyilván körülbelül $\log_2 S$ kérdéssel tudunk kitalálni, és így átlagban egy gondolt dologra $(1/n)\log_2 S$ kérdés jut.

Viszont (5)-ből

$$(1/n)\log_2 S \approx P_1\log_2(1/P_1)+P_2\log_2(1/P_2)+\ldots+P_N\log_2(1/P_N).$$

Tehát a Shannon-formula valóban megadja, hogy átlagban körülbelül hány kérdéssel lehet *majdnem biztosan* kitalálni, hogy a másik játékos x_1

,

$x2$
, …,

xN
közül melyikre gondol.

Figyeljük meg, hogy ebben az esetben már nem állíthatjuk, hogy

$H(P1,P2,…,PN)$
(ill. annál tetszőleges kevéssel több) kérdés *biztosan* elegendő átlagban a gondolt dolog kitalálásához, csak azt, hogy ennyi kérdéssel 1-hez tetszőlegesen közeli valószínűséggel („majdnem biztosan”) ki tudjuk azt találni. Ez azonban inkább csak elvileg jelent különbséget.

5. 5. A BARKOCHBA JÁTÉK ÉS A KÓDOLÁS

Befejezésül az információmennyiség fogalmának gyakorlati jelentőségét kíséreljük meg érzékeltetni oly módon, hogy megmutatjuk, hogyan lehet a Barkochba játékból kiindulva az információ kódolásának problémájához eljutni.

Tegyük fel, hogy a kérdező feljegyzéseket készít magának a kérdéseire kapott válaszokról: valahányszor a válasz „igen”, egyest ír fel, míg ha a válasz „nem”, zérust ír fel egészen addig, míg a gondolt dolgot ki nem találta. Ily módon a kapott válaszok sorozatának megfelel egy, a válaszok (kérdések) számával egyenlő sok, 0 vagy 1 jelből álló jelsorozat. Ez a jelsorozat egyértelműen jellemzi a gondolt dolgot, hiszen azt a kérdező éppen a kapott válaszok alapján találta ki. Tetszőleges adatoknak előírt jelekből álló jelsorozatok megfeleltetését ezen adatok *kódolásának* nevezzük. Így tehát arra az eredményre jutottunk, hogy tetszőleges adat 0 és 1 jelekkel való kódolásánál pontosan annyi jegyből álló jelsorozat szükséges, ahány kérdéssel az illető adat (a szóba jövő lehetőségek közül) a Barkochba játékban kitalálható. Így, amikor megállapítottuk, hogy hány kérdéssel lehet valamit kitalálni a Barkochba játékban, ezzel egyben azt a kérdést is megoldottuk, hogy hány 0 vagy 1 értékű jellel lehet azt „kódolni”, vagyis az információ kódolásának egy alapproblémáját is megoldottuk. Nem térhetünk itt ki a kódolás elméletének bonyolultabb problémáira, ti. a jelek zajos csatornán át való továbbításának kérdésére. Csak jelezzük, hogy e problémák a Barkochba játék olyan módosításának felelnek meg, ahol a „felelő” nem mindig válaszol az igazságnak megfelelően, hanem időnként hazudik. A „hazudós Barkochba” mint játék persze sokkal nehezebb, mint a játék szokásos „igazmondó” formája, ugyanúgy ahogy a zajos csatornán át való információtovábbítás kérdése is sokkal bonyolultabb, mint az információ egyszerű kódolásának problémája.

Az információelmélet iránt érdeklődő olvasó figyelmébe ajánljuk az irodalomjegyzékben [3]–[8] alatt felsorolt munkákat.

IRODALOMJEGYZÉK

[1] [1] Rényi A.: *Valószínűségszámítás*, Tankönyvkiadó, Bp., 1966.

[2] [2] Gnyegyenko, B. V.–Hincsin, A. Ja.: *Bevezetés a valószínűségszámításba*, Művelt Nép Kiadó, Bp., 1954.

[3] [3] Jaglom, A. M.–Jaglom, I. M.–Hincsin, A. Ja.: *Az információelmélet matematikai alapjai*, Műszaki Kiadó, Bp., 1959.

[4] [4] Rényi, A.: *Wahrscheinlichkeitstheorie mit einem Anhang über Informationstheorie*, Deutscher Verlag der Wissenschaften, 2. kiadás, Berlin, 1966. című könyv függeléke.

[5] [5] Feinstein, A.: *Foundations of information theory*, Mc.Graw-Hill, New York, 1958.

[6] [6] Fano, R. M.: *Transmission of information. A statistical theory of communication*, Wiley, New York, 1961.

[7] [7] Woodward, P. M.: *Probability and information theory with applications to radar*, Pergamon Press, London, 1955.

[8] [8] Reza, F. M.: *Bevezetés az információelméletbe*, Műszaki Kiadó, Bp., 1966.

PRÍMKÉPLETEK

LACZKOVICH MIKLÓS

A 31, 331, 3331, 33 331, 333 331, 3 333 331, 33 333 331 számok mind prímek. Ezt úgy is kifejezhetjük, hogy a $(10n–7)/3$ képlet prímszámot szolgáltat az $n=2$, 3, …, 8 értékek mindegyikére. A $333\,333\,331=(109–7)/3$ szám azonban már összetett, mert osztható 17-tel. Ezt az osztás elvégzése nélkül is beláthatjuk a kongruencia jelölésének felhasználásával. Azt mondjuk, hogy a kongruens b-vel modulo m (képletben $a\equiv b \pmod{m}$), ha a és b azonos maradékot adnak m-mel osztva; azaz, ha $a–b$ osztható m-mel. Könnyű ellenőrizni, hogy azonos modulushoz tartozó kongruenciákat szabad összeadni és szorozni. Azaz, ha

$a \equiv b \pmod m$
és

$c \equiv d \pmod m$
, akkor

$a+c \equiv b+d \pmod m$
és

$a \cdot c \equiv b \cdot d \pmod m$
.

Mivel

$10^2 = 17 \cdot 6$
, így

$10^9 = 100^4 \cdot 10 \equiv (-2)^4 \cdot 10 \equiv -10 \equiv 7 \pmod{17}$
, tehát

$10^9 - 7$
osztható 17-tel. Meg lehet mutatni egyébként, hogy

$33\ldots331$
sohasem osztható a 2, 3, 5, 7, 11, 13, 37 prímek egyikével sem. Nem ismeretes, hogy van-e végtelen sok

$33\ldots331 = (10^n - 7)/3$
alakú prímszám. Ez a kérdés minden bizonnyal épp olyan nehéz, mint hogy van-e végtelen sok

$2^n - 1$
alakú, ún. Mersenne-prím. Az utóbbi probléma pedig több száz éve megoldatlan.

Az alábbi táblázatban a bekeretezett (nem áthúzott) számok sokáig prímeknek bizonyulnak:

$\boxed{19}$								
~~21~~	$\boxed{23}$							
~~25~~	~~27~~	$\boxed{29}$						
~~31~~	~~33~~	~~35~~	$\boxed{37}$					
~~39~~	~~41~~	~~43~~	~~45~~	$\boxed{47}$				
~~49~~	~~51~~	~~53~~	~~55~~	~~57~~	$\boxed{59}$			
~~61~~	~~63~~	~~65~~	~~67~~	~~69~~	~~71~~	$\boxed{73}$		
~~75~~	~~77~~	~~79~~	~~81~~	~~83~~	~~85~~	~~87~~	$\boxed{89}$	
~~91~~	~~93~~	~~95~~	~~97~~	~~99~~	~~101~~	~~103~~	~~105~~	$\boxed{107}$
~~109~~	~~111~~	~~113~~	....					

De mégsem lesz mindegyikük prím. A

k
-adik bekeretezett szám ugyanis

$17 + (1+2+\ldots+k) \cdot 2 = k^2 + k + 17$

, és ez biztosan összetett

$k=16$
-ra, hiszen ekkor

$k^2+k+17=k(k+1)+17=17^2$
. Meglepő módon a

k^2+k+17
képlet a

$k=0$
, 1, …, 15 helyettesítési értékek mindegyikére prímet ad.

Ha meg akarjuk keresni azokat a

p
pozitív egészeket, amelyekre

k^2+k+p
prímszám a

$k=0$
, 1, …,

$p-2$
értékek mindegyikére, akkor a következőképpen okoskodhatunk. Először is (

$k=0$
-t véve)

p
prímszám kell, hogy legyen. Ha

$p>3$
, akkor az

1^2+1+p
és

2^2+2+p
számok nem lehetnek sem 3-mal, sem 5-tel oszthatók, így

p
3-mal osztva 2-t, 5-tel osztva pedig 1-et vagy 2-t ad maradékul. Mivel

p
páratlan is, ezért a 30-cal való osztási maradéka csak 11 vagy 17 lehet. Egy további feltétel, hogy

$4p-1$
-nek is prímnek kell lennie. Tegyük fel ugyanis, hogy

$d\mid 4p-1$
és

$1<d<4p-1$
. Ekkor

$d \leq (4p-1)/3$ és $d=2x+1$, ahol tehát $x<(d/2) \leq (4p-1/6) \leq p-2$, ha $p>5$. Így $x2+x+p$-nek prímnek kell lennie. Azonban d osztója $4(x2+x+p)=(2x+1)2+4p-1=d2+(4p-1)$-nek, tehát $x2+x+p$-nek is, hiszen páratlan. Ez csak úgy lehetséges, ha $d=2x+1=x2+x+p>x2+x+2$, $x2-x+1<0$, ami lehetetlen.

Azt kaptuk tehát, hogy vagy $p=2$, 3, 5, vagy pedig p olyan $30k+11$ vagy $30k+17$ alakú prímszám, amelyre $4p-1$ is prím. A 100-nál kisebb számok közül ezeket a feltételeket csak a 2, 3, 5, 11, 17, 41, valamint a 71 elégítik ki. Ezek közül a $p=2$, 3, 5, 11, 17, 41 számokra valóban teljesül, hogy $k2+k+p$ prímszám a $k=0$, 1, …,

$p-2$ értékek mindegyikére. A 71-re ez már nem igaz, mert $2^2+2+71=77$ összetett. Sokáig megoldatlan volt, hogy a 41 után van-e még ilyen tulajdonságú szám. Csak az 1960-as években sikerült bebizonyítani, hogy ilyen szám nem létezik.

□

A fentiek alapján természetesen vetődik fel az a kérdés, hogy van-e olyan képlet, amelynek a pozitív egészekben felvett értékei mind prímek. Először megmutatjuk, hogy egy egész együtthatós polinom – hacsak nem konstans – nem szolgáltathat ilyen képletet. Ennek bizonyításához két segédtételre lesz szükségünk. Először is belátjuk, hogy ha $p(x)=c_kx^k+c_{k-1}x^{k-1}+\ldots+c_0$ egész együtthatós polinom, akkor $a\equiv b \pmod m$ esetén $p(a)\equiv p(b) \pmod m$. Valóban,

$$p(b)-p(a)=c_k(b^k-a^k)+c_{k-1}(b^{k-1}-a^{k-1})+\ldots+c_1(b-a). \tag{1}$$

Itt b^i-a^i-ből $(b-a)$ kiemelhető: $b^i-a^i=(b-a)(b^{i-1}+b^{i-2}a+\ldots+ba^{i-2}+a^{i-1})$. Ha (1) jobb oldalának minden tagjából kiemeljük $b-a$-t, akkor azt kapjuk, hogy $p(b)-p(a)=(b-a)\cdot A$, ahol A egész szám, hiszen a, b, c_0, …,

c_k mindannyian egészek. Így $p(b)-p(a)$ osztható $b-a$-val. Ha $a\equiv b \pmod{m}$, akkor m osztója $b-a$-nak, tehát $p(b)-p(a)$-nak is, azaz $p(a)\equiv p(b) \pmod{m}$.

A másik állítás, amelyre szükségünk van, az, hogy ha a $p(x)=c_kx^k+c_{k-1}x^{k-1}+\ldots+c_0$ polinom nem konstans és a főegyütthatója pozitív, akkor $p(x)$ minden előre megadott számnál nagyobb lesz, ha x elég nagy. Ezt így láthatjuk be: Jelöljük a $|c_{k-1}|+\ldots+|c_1|+|c_0|$ számot c-vel. Ha $x>1$, akkor $1<x<x^2<\ldots<x^{k-1}$, tehát

$$p(x)\geq c_kx^k-|c_{k-1}|x^{k-1}-\ldots-|c_1|x-|c_0|\geq \geq c_kx^k-|c_{k-1}|x^{k-1}-\ldots-|c_1|x^{k-1}-|c_0|x^{k-1}= =c_kx^k-cx^{k-1}=x^{k-1}(c_kx-c).$$

Ha $x>c/c_k$, akkor $c_kx-c>0$, tehát

$x>1$ alapján $p(x)>ckx-c$. Ha tehát azt akarjuk, hogy $p(x)$ nagyobb legyen egy előre megadott $K>0$ számnál, akkor x-et elég úgy választani, hogy nagyobb legyen 1 és $(c+K)/ck$ mindegyikénél. Ekkor ugyanis a fentiek szerint $p(x)>ckx-c>K$ teljesülni fog.

Világos, hogy ha p nem konstans, és a főegyütthatója negatív, akkor $p(x)$ minden előre magadott számnál kisebb lesz, ha x elég nagy.

Most már könnyen beláthatjuk, hogy ha az egész együtthatós $p(x)$ polinom nem konstans, akkor a $p(n)$ értékek nem lehetnek mind prímek. Válasszunk ugyanis egy olyan n_0 pozitív egészet, amelyre $|p(n_0)|=m>1$. Ha $n=n_0+i\cdot m$ ($i=1, 2, \ldots$), akkor $n\equiv n_0 \pmod{m}$, tehát $p(n)\equiv p(n_0)=\pm m\equiv 0 \pmod{m}$, azaz

$p(n)$
osztható

m
-mel. De tudjuk, hogy

$|p(x)|>m$
, ha

x
elég nagy; mondjuk, ha

$x>x_0$
. Ekkor minden elég nagy

i
-re

$n=n_0+i\cdot m>x_0$
, így

$|p(n)|>m$
és

$p(n)$
osztható

m
-mel, tehát

$p(n)$
összetett.

A polinomoknál jóval komplikáltabb képletekről is beláthatjuk, hogy nem szolgáltathatnak csupa prímszámot. Ennek tárgyalásához szükségünk lesz az ún. „kis Fermat-tételre”, ami azt állítja, hogy ha

p
prím és

n
nem osztható

p
-vel, akkor

$n^{p-1}\equiv 1$

$(\mathrm{mod}\ p)$
. Ezt így láthatjuk be:

Az

n
,

$2n$
, …,

$(p-1)n$ számok csupa különböző maradékot adnak p-vel osztva. Valóban, ha $in \equiv jn \pmod{p}$, akkor $p \mid in-jn=(i-j)n$. Mivel p prím és n nem osztható p-vel, ebből következik, hogy $p \mid i-j$. Ha tehát $1 \leq i$, $j \leq p-1$, akkor szükségképpen $i=j$.

Mivel az n, $2n$, …, $(p-1)n$ számok egyike sem osztható p-vel, ebből következik, hogy a p-vel vett osztási maradékaik az 1, 2, …, $p-1$ számok (esetleg más sorrendben). Így $n \cdot 2n \cdots (p-1)n \equiv 1 \cdot 2 \cdots (p-1) \pmod{p}$, azaz $n^{p-1}(p-1)! \equiv (p-1)! \pmod{p}$

. Ezért

$p\square(np–1–1)(p–1)!$
, tehát

$p\square(np–1–1)$
, hiszen

p
nem osztója

$(p–1)!$
-nak.

Ennek felhasználásával lássuk be például, hogy az

$f(n)=22n+18n+1$
képlet nem szolgáltathat csupa prímet. Az

$f(1)=41$
érték mindenesetre prím, ezért

$2240\equiv1 (\mathrm{mod}\ 41)$
és

$1840\equiv1$

$(\mathrm{mod}\ 41)$
. Ebből következik, hogy

$2240k\equiv1(\mathrm{mod}\ 41)$
és

$1840k\equiv1(\mathrm{mod}\ 41)$
minden

k
pozitív egészre. Ha tehát

$n=40k+1$
, akkor

$22n+18n+1=2240k+1+1840k+1+1= =22\square2240k+18\square1840k+1\equiv22+18+1=41\equiv0\ (\mathrm{mod}\ 41),$
és így

$f(40k+1)$
nem lehet prím, ha

$k>0$
.

Most bizonyítsuk be, hogy az

$f(n)=100(n2+n)n3+n2+2n+1$
képlet sem szolgáltathat csupa prímet! (Ez kicsit nehezebb lesz.) Először is válasszunk egy olyan

$n0$
pozitív egészet, amelyre

$p=f(n_0)>n_0^2+n_0$. Ilyen például $n_0=1$, amikor is $p=1601>2$. Ha p összetett szám, akkor máris beláttuk, hogy f értékei a pozitív egészekben nem mind prímek. Ha p prím (mint az adott esetben is), akkor tekintsük az $n=n_0+ip(p-1)$ alakú számokat, ahol $i=1$, 2, …. Belátjuk, hogy ezekre az n-ekre $f(n)$ mindig osztható p-vel. Valóban, rögzítsünk egy ilyen n-et, és tekintsük a $q(x)=100xn^3+n^2+2n+1$ polinomot! Tudjuk, hogy ha $a\equiv b \pmod{m}$, akkor $q(a)\equiv q(b) \pmod{m}$. Mivel $n\equiv n_0 \pmod{p}$, ezért $q(n)\equiv q(n_0) \pmod{p}$, azaz

$$f(n)=100(n^2+n)n^3+n^2+2n+ \\ +1\equiv 100(n_0^2+n_0)n^3+n^2+2n+1 \pmod{p}. \tag{2}$$

Ugyanígy, felhasználva, hogy

$n \equiv n_0 \pmod{p-1}$, azt kapjuk, hogy

$n^3+n^2+2n \equiv n_0^3+n_0^2+2n_0 \pmod{p-1}$, tehát

$n^3+n^2+2n = n_0^3+n_0^2+2n_0+c(p-1)$, ahol

c pozitív egész. Ezt (2)-be helyettesítve azt kapjuk, hogy

(3)	$f(n) \equiv 100(n_0^2+n_0)^{n_0^3+n_0^2+2n_0+c(p-1)}+1 \equiv$ $\equiv 100(n_0^2+n_0)^{n_0^3+n_0^2+2n_0} \cdot (n_0^2+n_0)^{c(p-1)}+1 \pmod{p}$.

Mivel

$1 \le n_0^2+n_0 < p$, ezért a kis Fermat-tétel szerint

$(n_0^2+n_0)^{(p-1)} \equiv 1 \pmod{p}$, és így, tekintve, hogy

c pozitív egész, azt kapjuk, hogy

$(n_0^2+n_0)^{c(p-1)} \equiv 1 \pmod{p}$. Ezt (3)-ba helyettesítve:

$f(n) \equiv 100(n_0^2+n_0)^{n_0^3+n_0^2+2n_0}+1 = p \equiv 0 \pmod{p}$, tehát

$f(n)$ valóban osztható

p-vel. Mivel

f szigorúan monoton növő, ezért

$i>0$-ra

$n = n_0+ip(p-1) > n_0$-ból

$f(n) > f(n_0)$ következik, Ezért

$f(n)$ összetett szám lesz. Így például

f

értéke az

$1+1601\cdot1600=2\ 561\ 601$
helyen osztható 1601-gyel, és ezért összetett.

A következő általános tétel hasonló módszerrel bizonyítható. *Legyen*

$f(n)$
olyan képlet, amely

$p(n)^{q(n)}$
alakú kifejezések szorzatainak összege, ahol

$p(n)$
és

$q(n)$
egész együtthatós polinomok, melyek közül a

$q(n)$
-ek főegyütthatói pozitívak. Ha az

$f(n)$
(

$n=1$
, 2, ...) értékek között van végtelen sok különböző, akkor ezek között az értékek között van végtelen sok összetett szám.

Itt a végtelen sok különböző értékre vonatkozó feltétel lényeges, mert pl. az

$f(n)=18+(-1)^n$
képlet a fenti alakú és

f
minden értéke prím.

□

A fenti tétel azt állítja, hogy ha

$f(n)$
csupa prímszámot ad minden

n
pozitív egészre, akkor az

f
-et definiáló képlet a fent leírtaknál bonyolultabb kell, hogy legyen. Ha például olyan prímképletet keresünk, amelyben csak az összeadás, kivonás, szorzás és hatványozás műveleteit használhatjuk, akkor szerepelnie kell benne olyan „emeletes hatványnak” (azaz kitevőben szereplő hatványnak), amelynek a kitevője tartalmazza az

n
változót. A legegyszerűbb ilyen képlet:

2^{2^n}
. Ez persze

$n>0$
-ra nem ad prímeket, mert mindig páros. De mi a helyzet a

$22n+1$
képlettel? Az

$n=0$
, 1, 2, 3, 4 számokat behelyettesítve a 3, 5, 17, 257,

65 537
értékeket kapjuk; mindnyájan prímek. Ennek alapján Pierre de Fermat azt sejtette (1640 körül), hogy

$Fn=22n+1$
minden

n
-re prímszám. Ez a sejtés csaknem 100 éven át megoldatlan maradt, mert a következő szám:

$F5=225+1=4\ 294\ 967\ 297$
már olyan nagy, hogy annak eldöntése, hogy prím-e vagy sem, reménytelennek látszott. Valóban, gondoljunk csak bele: ha sem kalkulátor, sem számítógép nem állna rendelkezésünkre, el tudnánk-e dönteni, hogy

$F5$
prím-e? Volna-e jobb ötletünk, mint elosztani

$F5$
-öt minden nála kisebb számmal? Persze elég volna csupán a

$\sqrt{}(F5)$
-nél nem nagyobb számokkal osztani, hiszen ha egy

n
szám összetett, akkor van olyan

d
osztója, amelyre

$1<d\leq\sqrt{}n$
. Az adott esetben tehát csak a

65 536
-nál nem nagyobb számokkal kellene osztani; sőt, ezek közül elég lenne a prímeket venni, melyek száma kb. 6500. Ha minden osztás két percet vesz igénybe, még ez is 200 óra munka volna; nem csoda, ha olyan sokáig senki nem vállalkozott rá.

Végül is Leonhard Euler volt az, aki 1732-ben a kérdést eldöntötte. Euler felismerte, hogy

Fn
prímosztói csak speciális alakúak lehetnek, nevezetesen

Fn
minden prímosztója

k□$2n+1+1$
alakú, ahol

k pozitív egész. Ezt a következőképpen láthatjuk be. Legyen p egy prímosztója F_n-nek. Mivel p páratlan, ezért a kis Fermat-tétel szerint $2^{p-1}\equiv 1 \pmod{p}$. Legyen d a legkisebb pozitív egész, amelyre $2^d\equiv 1 \pmod{p}$. Ekkor a 2-hatványok p-vel való osztási maradékai d szerint periodikus sorozatot alkotnak. Valóban, $2^d\equiv 1=2^0$ alapján $2^{d+1}\equiv 2=2^1$, $2^{d+2}\equiv 2\cdot 2^1=2^2$, $2^{d+3}\equiv 2\cdot 2^2=2^3$, és hasonlóan, $2^{d+i}\equiv 2^i$ minden i-re (a kongruenciákat (mod p) értve). Ebből következik, hogy $2^{s\cdot d}\equiv 1$ minden $s=1$, 2, …-re. Másrészt d választása folytán

$2i$□1 ha $0<i<d$, ezért a periodicitásból következik, hogy $2m$ akkor és csak akkor ad 1 maradékot p-vel osztva, ha m osztható d-vel.

Mármost $22n+1\equiv1(\mathrm{mod}\ p)$, ugyanis $22n+1-1=(22n+1)$□$(22n-1)$, tehát $22n+1-1$ osztható Fn-nel, tehát p-vel is. A fentiek szerint ebből következik, hogy d □ $2n+1$. Megmutatjuk, hogy szükségképpen $d=2n+1$. Valóban, $d<2n+1$ esetén d osztója lenne $2n$-nek is. Ebből viszont $22n\equiv1$ következne, holott $22n=Fn-1\equiv-1(\mathrm{mod}\ p)$, hiszen p □ Fn

. Ezzel beláttuk, hogy

d=2n+1
. Mivel pedig

2p–1≡1(mod p)
, ezért

2n+1=d □ p–1,
tehát

p–1=k□2n+1,
azaz

p=k□2n+1+1,
ahol

k
pozitív egész.

Valójában Euler azt is belátta, hogy ha

n≥2
és a

p
prím osztója

Fn
-nek, akkor

p=k□2n+2+1,
ahol

k
pozitív egész. Meg lehet mutatni ugyanis, hogy ha a

p
prím

8k+1
alakú, akkor

2(p–1)⁄2≡1(mod p)
. A fenti okoskodásban ezért

2n+1=d □ (p–1)⁄2,
tehát

(p–1)⁄2=k□2n+1,
azaz

p=k□2n+2+1,
ahol

k
pozitív egész.

Ezt

$n=5$
-re alkalmazva azt kapjuk, hogy

$F5$
minden prímosztója

$27\square k+1=128k+1$
alakú. Ez jelentősen leszűkíti a lehetőségeket. A 2-nél nagyobb prímek ugyanis 128-cal osztva 64-féle maradékot adhatnak (nevezetesen az 1, 3, …, 127 számokat), és azt várhatjuk, hogy a

65 536
-nál kisebb prímeknek durván

1⁄64
-ed része lesz

$128k+1$
alakú. Ez kb.

$6500⁄64<102$
prímet jelent. Ezért annak eldöntéséhez, hogy

$F5$
prímszám-e, a 6500 osztás helyett várhatóan a legrosszabb esetben is elég száz-egynéhány osztást elvégezni. Euler elszánta magát ennek az elvégzésére. Óriási szerencséje volt; az első

$128k+1$
alakú prím, 257, ugyan nem osztója

$F5$
-nek, de a második, 641, már igen:

$F5=4\ 294\ 967\ 297=641\square 6\ 700\ 417$
. Ezzel Fermat sejtése megdőlt: az

$Fn=22n+1$
képlet nem állít elő minden

n
-re prímszámot.

Egyébként a következő Fermat-szám,

$F6$
prímtényezőkre bontását csak 1880-ban sikerült megtalálni:

$F6=274\ 177\square 67\ 280\ 421\ 310\ 721$
. Az ezt követő Fermat-szám,

$F7$
faktorizációját 1970-ben találták meg; eszerint

$F7$
egy 17-jegyű és egy 22-jegyű prím szorzata. Azóta kiderült, hogy

F_n
összetett szám minden

$5 \le n \le 21$
-re (bár nem mindegyiküknek sikerült meghatározni a prímtényezős alakját). Sok, nagyobb indexű Fermat-számot is megvizsgáltak, de még egy prímet sem találtak közöttük.

A fentiekből az a tanulság, hogy hacsak egy képletről nem tudjuk eleve, hogy valamilyen oknál fogva minden értéke prímszám lesz, akkor nem várhatjuk el, hogy ezt „szívességből” megtegye, még akkor sem, ha az első néhány értéke prím (mint pl. az

n^2+n+17

,

n^2+n+41
vagy a

$2^{2^n}+1$
képletek esetében).

Eddig senki nem talált olyan képletet, amely egész számokból és az

n
változóból az összeadás, kivonás, szorzás és hatványozás műveleteinek segítségével épül fel, és amely minden pozitív egész

n
-re prímszámot szolgáltat (azt is feltéve persze, hogy a képlet végtelen sok különböző értéket ad). A következőkben megmutatjuk, hogy a feltételek enyhítésével már találhatunk ilyen képleteket.

Először olyan képleteket tekintünk, amelyekben nemcsak egészek, hanem tetszőleges valós konstansok is szerepelhetnek. Ekkor persze az

x
szám egészrészét megadó

$[x]$
függvény használatát is meg kell engednünk, mert különben nem tudjuk biztosítani, hogy a képlet egész számokat szolgáltasson. (E függvény definíciója a következő:

$[x]$
az az egyértelműen meghatározott egész szám, amelyre

$[x] \le x < [x]+1$
. Így pl.

$[3/2]=1$

,

$[\pi]=3$

,

$[7]=7$

,

$[-4]=-4$

,

$[-3/2]=-2$

,

$[-\pi]=-4$

.) Ezek felhasználásával már viszonylag egyszerű képletet adhatunk az

n

-edik prímszámra, amelyet a következőkben

p_n

-nel fogunk jelölni. Megmutatjuk, hogy van olyan

α

valós szám, amelyre

(4)

$$p_n=[10^{n^2}\alpha]-10^{2n-1}[10^{(n-1)^2}\alpha] \quad (n=1,2,\ldots).$$

Képezzünk ugyanis a prímszámok

p_n

sorozatából egy végtelen tizedestörtet a következőképpen! A tizedestört egészrésze 0 lesz. A tizedesvessző után írjuk le egymás után a prímszámokat, megfelelő számú 0-val elválasztva őket! Az elválasztó 0-k számát úgy választjuk meg, hogy

p_n

utolsó számjegye éppen a tizedesvessző utáni

n^2

-edik helyre kerüljön. Tehát az 1., 4., 9., 16. jegy a 2-es, 3-as, 5-ös, 7-es lesz, a 24. és a 25. 1-es, a 35. ismét 1-es, a 36. 3-as stb. Legyen az így definiált végtelen tizedestört

α

, tehát legyen

$\alpha=0{,}200\ 300\ 005\ 000\ 000\ 70\ldots0110\ldots0130\ldots.$

Ekkor

$A_n=[10^{n^2}\alpha]$

az az egész szám, amelynek a jegyeit az

α

-nak a tizedesvessző utáni első

n^2

jegye adja. A konstrukcióból adódik, hogy

A_n

olyan

n^2

-jegyű szám, amely éppen

p_n
jegyeivel végződik. Nyilvánvaló, hogy az

A_n
első

$(n-1)^2$
jegyéből képzett szám

$A_{n-1}=[10^{(n-1)^2}\alpha]$
lesz. Ha tehát

A_{n-1}
-et kiegészítjük

$n^2-(n-1)^2=2n-1$
darab 0-val, azaz megszorozzuk

10^{2n-1}
-gyel, és az így kapott számot kivonjuk

A_n
-ből, akkor éppen

p_n
-et kapjuk. Ezzel (4)-et beláttuk.

A fenti konstrukcióban hallgatólagosan felhasználtuk, hogy

p_n
jegyeinek száma legfeljebb

$n^2-(n-1)^2=2n-1$
, mert különben

p_n
„nem férne el”. Megmutatjuk, hogy

p_n
jegyeinek száma legfeljebb

n
.

Nevezzünk egy

m
számot négyzetmentesnek, ha nem osztható 1-nél nagyobb négyzetszámmal. Ez azzal ekvivalens, hogy az

m
prímtényezős felbontásában minden prím első hatványon szerepel; azaz, hogy

m
különböző prímek szorzata. Bármely

k

egész szám előáll mint egy négyzetszám és egy négyzetmentes szám szorzata. Ha ugyanis k-t elosztjuk azzal a legnagyobb osztójával, amely négyzetszám, akkor a hányados nyilván négyzetmentes lesz. (Ha k négyzetmentes, akkor 1-gyel osztunk.)

Írjuk fel a p_n-nél nem nagyobb számokat b^2c alakban, ahol c négyzetmentes. Itt $b^2 \le p_n$, tehát $b \le \sqrt{p_n}$. A c szám különböző prímek szorzata, melyek mindegyike legfeljebb p_n, hiszen $c \le p_n$. Ezért c-t úgy kapjuk, hogy a p_1, p_2, …, p_n prímek közül néhányat összeszorzunk. Ezt 2^n-féleképpen tehetjük meg (annyiféleképpen, ahány részhalmaza van a $\{p_1, p_2, \ldots, p_n\}$ halmaznak). A b számot tehát legfeljebb $\sqrt{p_n}$

-féleképpen, a c számot pedig legfeljebb 2^n-féleképpen választhatjuk meg. Mivel b^2c alakban minden, p_n-nél nem nagyobb szám előáll, ebből következik, hogy $p_n \le \sqrt{p_n} \cdot 2^n$, $\sqrt{p_n} \le 2^n$, és így $p_n \le 4^n$. Mivel $4^n < 10^n$, ezért p_n legfeljebb n-jegyű.

Ez a becslés lényegesen javítható: valójában p_n jegyeinek száma alig nagyobb n jegyeinek számánál. Így pl. $p_{100} = 541$, $p_{1000} = 7919$, $p_{10\,000} = 104\,729$, $p_{10^{10}}$ 11-jegyű, $p_{10^{100}}$ pedig 102-jegyű. Az utóbbi két állítás abból következik, hogy $n(\log n + \log\log n - 1{,}5) < p_n < n(\log n + \log\log n + 8)$, valamint $n \log n < p_n$ minden

$n\geq 2$
-re. E képletekben

$\log n$
az ún. természetes vagy

e
alapú logaritmus, amelynek alapja

$e=\lim_{n\to\infty}(1+(1/n))^n=2{,}718\,28\ldots$.

A (4) képletre visszatérve megállapíthatjuk, hogy a puszta érdekességén kívül más haszna nemigen van; például nem használhatjuk fel prímszámok gyártására. Ahhoz ugyanis, hogy a képlet segítségével

p_n
értékét meghatározhassuk, tudnunk kellene

α
értékét. Ehhez viszont már ismernünk kellene az összes prímszámot. Egyébként vannak még egyszerűbb prímképletek is: létezik például egy

$c>1$
valós szám úgy, hogy

$[c^{3^n}]$
prímszám lesz

n
minden pozitív egész értékére. Ennek a képletnek ugyanaz a baja, mint (4)-nek:

c
meghatározásához végtelen sok prímszámot kell felhasználni.

□

Most rátérünk azokra a prímképletekre, amelyekben a konstansok csak egész számok lehetnek, de a felhasználható műveleteket nem korlátozzuk. Megmutatjuk, hogy ha az alapműveleteken kívül felhasználhatjuk az

$[x]$
(egészrész),

$\{x\}=x-[x]$
(törtrész) függvényeket, valamint a

$\sum_{i=k}^{n} a_i=a_k+a_{k+1}+\ldots+a_n$ és $\prod_{i=k}^{n} a_i=a_k \cdot a_{k+1} \cdot \ldots \cdot a_n$
jelöléseket, akkor szintén kaphatunk képleteket

p_n
-re. Jelöljük

$\pi(x)$
-szel az

x
számnál nem nagyobb prímek számát. Először

$\pi(x)$
-re adunk képletet.

Ha az

$n>2$
szám prím, akkor az

$(n/2)$
,

$(n/3)$
, …,

$(n/n-1)$
számok egyike sem egész, tehát az

$\{(n/2)\},\{(n/3)\},\ldots,\{(n/n-1)\}$
törtrészek egyike sem 0. Ekkor tehát a

$\prod i=2n-1[-\{(n/i)\}]$
szorzat minden tényezője

-1
(mert egy

-1
és 0 közé eső szám egészrésze). A szorzat értéke tehát

-1
, hiszen a tényezők száma

$n-2$
, ami páratlan. Ha viszont

n
összetett, akkor

n/i
egész szám lesz legalább egy

$2\leq i\leq n-1$
-re, és ekkor

$[-\{(n/i)\}]=[-0]=0$
alapján a fenti szorzat értéke 0. Így

$x\geq 3$
esetén a

$-\sum n=3x\prod i=2n-1[-\{(n/i)\}]$
képlet értéke

$\pi(x)-1$
, hiszen a szumma

n

-edik tagja -1, ha n prím, és 0, ha n összetett. ($\pi(x)$-ből azért kell 1-et levonni, mert a 2-t kizártuk a szummából.) Azt kaptuk tehát, hogy

$$\pi(x)=1-\sum_{n=3}^{x}\prod_{i=2}^{n-1}\left[-\left\{\left(n/i\right)\right\}\right] \quad (x=3,4,\ldots). \tag{5}$$

Valamivel egyszerűbb képletet is kaphatunk $\pi(x)$-re az ún. Wilson-tétel felhasználásával. Ez azt állítja, hogy ha p prím, akkor $(p-1)!\equiv-1 \pmod p$.

Ezt így láthatjuk be: Az állítás $p=2$, 3-ra nyilvánvaló, tehát feltehetjük, hogy $p\geq5$. A kis Fermat-tétel bizonyításakor megmutattuk, hogy ha n nem osztható p-vel, akkor az n, $2n$, …, $(p-1)n$ számok p-vel vett osztási maradékai az 1, 2, …, $p-1$ számok (esetleg más sorrendben). Ebből következik, hogy az 1, 2, …,

$p-1$ számok között pontosan egy olyan i szám van, amelyre $n \cdot i \equiv 1 \pmod p$. Nevezzük ezt a számot n *reciprokának* (mod p). Ha $n \not\equiv m \pmod p$, akkor n és m reciprokai különbözők. Valóban, $n \cdot i \equiv m \cdot i \equiv 1$-ből következik, hogy $p \mid ni - mi = (n-m)i$, tehát $p \mid n-m$, hiszen $1 \le i \le p-1$. Ha egy $1 \le n \le p-1$ szám reciproka önmaga, akkor $n^2 \equiv 1$, $p \mid n^2-1=(n-1)(n+1)$, tehát $n=1$ vagy $n=p-1$.

A fentiekből következik, hogy a 2, 3, …, $p-2$ számok mindegyikét a reciprokával párosítva diszjunkt párokat kapunk. Mivel az egy párhoz tartozó számok szorzata

p-vel osztva 1-et ad maradékul, ezért $2\cdot 3\cdots(p-2)\equiv 1 \pmod p$, tehát $(p-1)!\equiv -1 \pmod p$.

A Wilson-tétel megfordítható: ha $n>1$ osztója $(n-1)!+1$-nek, akkor n prím. Valóban, ha $1<d<n$, akkor $d \mid (n-1)!$. Ha d osztója volna n-nek, akkor $n \mid (n-1)!+1$ alapján d is osztója volna $(n-1)!+1$-nek, ami lehetetlen. Így n nem osztható a 2, …, $n-1$ számok egyikével sem, tehát prím. Összefoglalva: az $((n-1)!+1)/n$ tört akkor és csak akkor egész, ha $n=1$ vagy n prím. Ebből következik, hogy

$$\sum_{n=1}^{x}\left[-\left\{\frac{(n-1)!+1}{n}\right\}\right]=-(x-\pi(x)-1),$$

hiszen összetett n-re a szumma n-edik tagja -1, míg prím n-re és $n=1$-re 0. Ebből azt kapjuk, hogy

$$\pi(x)=x-1+\sum_{n=1}^{x}\left[-\left\{\frac{(n-1)!+1}{n}\right\}\right] \quad (x=1,2,\dots). \tag{6}$$

Az (5) és (6) képletek mindegyikét felhasználhatjuk p_n kifejezésére. Jelöljük φ-vel azt a függvényt, amelyre

$$\varphi(x)=\begin{cases}1, & \text{ha } x=0,1,2,\dots\\ 0, & \text{ha } x=-1,-2,\dots.\end{cases}$$

(A φ függvényt csak az egész számokban értelmezzük.) A φ függvényre könnyen adhatunk képletet, pl. $\varphi(x)=1+\left[\frac{1}{3x+2}\right]$ minden x egész számra. Mármost $k\le p_n$ esetén $\pi(k)\le n$, tehát $\varphi(n-\pi(k))=1$, míg $k>p_n$ esetén

$\pi(k)>n$
, tehát

$\varphi(n-\pi(k))=0$
. Ebből következik, hogy

$p_n=\sum_{k=1}^{4n}\varphi(n-\pi(k))$.
Valóban, a szumma tagjainak értéke

$k=1$
, …,

p_n
esetén 1, egyébként pedig 0. Mivel

$p_n\leq 4n$
, ezért a szumma értéke

p_n
. Ha ebben a formulában

$\varphi(x)$
helyébe

$1+[1/(3x+2)]$
-t írunk,

$\pi(x)$
-et pedig (6)-tal helyettesítjük, akkor a következő képletet kapjuk

p_n
-re:

$p_n=4n+\sum_{k=1}^{4n}[(1/3n-3k+5-3\sum_{i=1}^{k}[-\{((i-1)!+1/i)\}])]$ $(n=1,2,\ldots)$.
(Itt (6) helyett (5)-öt is alkalmazhatjuk, de ekkor apró módosítás szükséges, tekintve, hogy (5) csak

$x\geq 3$
-ra érvényes.)

Természetesen jó volna olyan prímképletet találni, amelyben a felhasznált műveletek száma minimális. Már említettük, hogy nem ismeretes olyan prímképlet, amelyben csak összeadás, szorzás és hatványozás szerepel. A következőkben a célunk annak bizonyítása, hogy *van olyan prímképlet, amelyben csak összeadás, kivonás, szorzás,*

*[*x*]*
,

$\sqrt{x}$
és maximum szerepel. Ennek ismertetését messziről kell kezdenünk.

□

Diophantosz görög matematikus volt, aki a harmadik században élt Alexandriában. „Aritmetika” című művében egész együtthatós egyenletek egész, illetve racionális megoldásait vizsgálta, és az ilyen egyenleteket azóta diofantoszi vagy diofantikus egyenleteknek nevezik. Ilyenek például

(7) $x2–2y2=1$, $x2–60y2=1$, $x2–61y2=1$,

vagy

(8) $x3+y3=u3+v3$, $x4+y4=u4+v4$, $x5+y5=u5+v5$.

A (7) alatt felsorolt három egyenlet mindegyikének megoldása

$x=1$

,

$y=0$;
ezeket triviális megoldásnak hívjuk. Nemtriviális megoldások is léteznek: az első egyenlet legkisebb pozitív egész megoldása

$x=3$

,

$y=2$
, a másodiké pedig

$x=31$

,

$y=4$
. A harmadiké viszont

$x=1\ 766\ 319\ 049$

,

$y=226\ 153\ 980$

.

A (8) alatt felsorolt egyenletek triviális megoldásai azok, amelyekre

$x=u$

,

$y=v$
vagy

$x=v$

,

$y=u$
. Az első két egyenletnek vannak nem-triviális megoldásai is, pl.

93+103=13+123
, illetve

1334+1344=594+1584
. A harmadiknak viszont nem ismerjük egyetlen nem-triviális megoldását sem, de az sincs bizonyítva, hogy

ilyen megoldás nem létezik.

Ezek a jelenségek tipikusak: előfordul, hogy egész egyszerűnek látszó diofantikus egyenletek megoldása a legnagyobb nehézségbe ütközik, sőt gyakori eset, hogy azt sem tudjuk eldönteni, hogy a kérdéses egyenletnek van-e egyáltalán megoldása. Ezek a tapasztalatok vezették David Hilbertet 1900-ban az alábbi kérdéshez: van-e olyan algoritmus (valamely mechanikus, automatikus eljárás), amely bármely megadott diofantikus egyenletről el tudja dönteni, hogy van-e megoldása? Ha Hilbert korában már léteztek volna számítógépek, nyilván úgy tette volna fel a kérdést, hogy van-e ilyen számítógépes program. A probléma megoldására 70 évig kellett várni. Ahhoz, hogy a megoldást megérthessük, be kell vezetnünk három fogalmat.

Azt mondjuk, hogy természetes számoknak egy végtelen sorozata *rekurzív*, ha van olyan algoritmus (azaz számítógépes program), amely bármely számról el tudja dönteni, hogy tagja-e a sorozatnak vagy sem. Például a 2-hatványok sorozata rekurzív, mert bármely

n

pozitív egészről el tudjuk dönteni, hogy 2-hatvány-e vagy sem: addig osztjuk 2-vel, amíg páratlan számot nem kapunk. Ha ez a szám 1, akkor

n

2-hatvány; egyébként pedig nem. A prímszámok sorozata is rekurzív, hiszen bármely számról el tudjuk dönteni, hogy prím-e: csak azt kell megvizsgálni, hogy osztható-e valamely, nála kisebb, de 1-nél nagyobb számmal. (Azzal most nem foglalkozunk, hogy ennek eldöntése mennyi ideig tart; a lényeg az, hogy az algoritmus véges sok lépésben elvégezhető legyen.)

Egy sorozat akkor és csak akkor rekurzív, ha van olyan számítógépes program, amely a sorozat tagjait növekvő sorrendben kinyomtatja. Ha ugyanis a sorozat rekurzív, akkor írhatunk olyan programot, amely egymás után megvizsgálja a természetes számokat és eldönti, hogy tagjai-e a sorozatnak, és ha egy

n

számról úgy találja, hogy igen, akkor

n

-et kinyomtatja. Megfordítva, ha van egy nyomtatóprogram, amely a sorozat tagjait növekvő sorrendben kinyomtatja, akkor a következő algoritmust készíthetjük annak eldöntésére, hogy egy

n

szám tagja-e a sorozatnak. Indítsuk el a nyomtatóprogramot, és várjunk addig, amíg egy

n

-nél nagyobb szám megjelenik. Ha az eddig kinyomtatott számok között

n

szerepel, akkor tagja a sorozatnak, egyébként pedig nem (hiszen később már nem kerülhet sorra).

Más a helyzet, ha van ugyan olyan program, amely a sorozat elemeit kinyomtatja, de nem feltétlenül növekvő sorrendben. (Az ilyen sorozatokat *rekurzíve felsorolható* sorozatoknak nevezzük.) Egy ilyen sorozat esetében, ha egy

n

szám megjelenik a kinyomtatott számok között, akkor tagja a sorozatnak. Ha viszont

n

nem tagja a sorozatnak, akkor ezt esetleg soha nem tudjuk meg, hiszen soha nem lehetünk biztosak abban, hogy

n

nem lesz-e később kinyomtatva. Tehát egy rekurzíve felsorolható sorozat nem feltétlenül rekurzív. A matematikai logika egyik nevezetes felfedezése, hogy rekurzíve felsorolható, de nem rekurzív sorozatok valóban léteznek, sőt konkrétan meg is adhatók. (Ezt úgy kell érteni, hogy konkrétan megadható olyan program, amely kinyomtatja a sorozat elemeit. A sorozatot magát nem tudjuk megadni abban az értelemben, hogy a tagjait felsoroljuk növekvő sorrendben, hiszen akkor a sorozat rekurzív volna.)

A harmadik fogalom, amelyre szükségünk lesz, a diofantikus sorozat fogalma. Egy diofantikus egyenlet általános alakja (az egyenlet jobb oldalának bal oldalra vitele után)

$P(x_1,x_2,\ldots,x_k)=0$
, ahol

$P(x_1,x_2,\ldots,x_k)$
egész együtthatós polinom; azaz olyan kifejezés, amely az

x_1
, …,

x_k
változókból és egész számokból képződik az összeadás, kivonás és szorzás műveleteinek felhasználásával. Egy sorozatot akkor nevezünk *diofantikusnak*, ha létezik egy egész együtthatós

$P(x_1,\ldots,x_k,x_{k+1})$
polinom a következő tulajdonsággal: egy

y
természetes szám akkor és csak akkor tagja a sorozatnak, ha a

$P(x_1,\ldots,x_k,y)=0$
diofantikus egyenletnek van megoldása a természetes számok körében, azaz ha vannak

$x_1,\ldots,x_k$
természetes számok úgy, hogy

$P(x_1,\ldots,x_k,y)=0$
. Ekkor azt mondjuk, hogy a sorozatot a

$P(x_1,\ldots,x_k,x_{k+1})$
polinom generálja. A négyzetszámok sorozata például diofantikus, mert egy

y
szám akkor és csak akkor négyzetszám, ha az

$x^2-y=0$
(egyváltozós) egyenlet megoldható a természetes számok körében. Tehát a négyzetszámok sorozatát az

$x_1^2-x_2$
polinom generálja.

Nem nehéz belátni, hogy minden diofantikus sorozat rekurzíve felsorolható. Tegyük fel ugyanis, hogy a sorozatot a

$P(x_1,\ldots,x_k,x_{k+1})$
polinom generálja. A rövidség kedvéért nevezzük *vektornak* a természetes számokból álló

$k+1$
-hosszúságú számsorozatokat. Ekkor az összes vektort felsorolhatjuk egyetlen végtelen sorozatban. Ezt úgy tehetjük meg, hogy elsőként felírjuk az

$(0,\dots,0)$ vektort, majd azokat, amelyekben $x_1+\dots+x_{k+1}=1$ (ezekből $k+1$ van), majd azok jönnek, amelyekben $x_1+\dots+x_{k+1}=2$, és így tovább. Minden egyes $(x_1,\dots,x_{k+1})$ vektorra számítsuk ki a $P(x_1,\dots,x_k,x_{k+1})$ értéket! Ha ez nulla, akkor nyomtassuk ki x_{k+1}-et! Ha nem nulla, akkor ugorjunk át a következő vektorra! Nyilvánvaló, hogy ilyen módon éppen azokat az y számokat nyomtattuk ki, amelyekre a $P(x_1,\dots,x_k,y)=0$ diofantikus egyenlet megoldható, és ezzel megmutattuk, hogy a kérdéses sorozat rekurzíve felsorolható.

Mármost Hilbert problémájának megoldásában a kulcslépés annak bizonyítása volt, hogy minden rekurzíve felsorolható sorozat szükségképpen diofantikus. A három fogalom logikai kapcsolata tehát a következő:

rekurzív $\Rightarrow$ rekurzíve felsorolható $\Leftrightarrow$ diofantikus.

Ebből pedig már következik, hogy *nincs olyan algoritmus, amely bármely megadott diofantikus egyenletről el tudná dönteni, hogy van-e megoldása*. Vegyünk ugyanis egy olyan sorozatot, amely rekurzíve felsorolható, de nem rekurzív. Mivel ez a sorozat szükségképpen diofantikus, létezik egy $P(x_1,\dots,x_k,x_{k+1})$ polinom, amely generálja. Ha létezne olyan algoritmus, amely bármely diofantikus egyenletről el tudja dönteni, hogy van-e megoldása, akkor minden y-ról eldönthetnénk, hogy tagja-e a sorozatnak vagy sem, hiszen a $P(x_1,\dots,x_k,y)=0$ diofantikus egyenletet az állítólagos algoritmussal megvizsgálva megállapíthatnánk, hogy megoldható-e vagy sem. Ez azonban lehetetlen, mert akkor a sorozatunk rekurzív lenne, holott olyan sorozatból indultunk ki, ami nem az.

A Hilbert-problémának ez a negatív megoldása nem jelenti azt, mintha találtunk volna egy olyan diofantikus egyenletet, amelyről sohasem dönthetjük el, hogy van-e gyöke vagy sem. Elvileg elképzelhető, hogy előbb-utóbb mindegyik diofantikus egyenletről kideríthetjük, hogy megoldható-e. Ebben az esetben azonban a módszernek egyenletről egyenletre változnia kell; általános, minden egyenletre egyaránt alkalmazható algoritmus nincs.

□

Most térjünk vissza a prímszámokhoz! Mivel a prímszámok sorozata rekurzív, ezért rekurzíve felsorolható, tehát diofantikus. Így létezik egy

$P(x_1,\ldots,x_k,x_{k+1})$

polinom, amely a prímszámok sorozatát generálja. Képezzük a

$Q(x_1,\ldots,x_k,x_{k+1})=x_{k+1}(1-2P^2(x_1,\ldots,x_k,x_{k+1}))$

polinomot! Ha Q-ban az x_1, …, x_k, x_{k+1} változók helyére természetes számokat helyettesítünk, akkor két eset lehetséges.

(i) $P(x_1,\ldots,x_k,x_{k+1})=0$. Ekkor x_{k+1} prím (mert P a prímeket generálja), és $Q(x_1,\ldots,x_k,x_{k+1})=x_{k+1}$.

(ii) $P(x_1,\ldots,x_k,x_{k+1})\neq 0$. Ekkor $1-2P^2(x_1,\ldots,x_k,x_{k+1})\leq 1-2<0$ és így

$Q(x_1,\ldots,x_k,x_{k+1})\leq 0.$

Azt kaptuk tehát, hogy Q-ba természetes számokat helyettesítve vagy prímet, vagy nempozitív számot kapunk. Másrészt így minden prímet megkapunk, mert ha p prím, akkor $P(x_1,\ldots,x_k,p)=0$ alkalmas

$x_1, \ldots, x_k$-ra, tehát $Q(x_1,\ldots,x_k,p)=p$. A fentieket összefoglalva megállapíthatjuk, hogy a

$$\max(Q(x_1,\ldots,x_k,x_{k+1}),2) \tag{9}$$

kifejezés a változók nemnegatív egész értékeire mindig prímet ad, és minden prímet megad. Ilyen tulajdonságú Q polinomok explicite is megadhatók; sajnos mindegyikük komplikált. Van közöttük 10-változós, ennek a foka azonban nagyobb 10^{45}-nél. Van 5-ödfokú ilyen polinom is; ez azonban 42 változót tartalmaz. A jelenleg ismert legegyszerűbb egy 26-változós, 25-ödfokú polinom, amely kinyomtatva 9 sort foglal el a [3] könyv 115-116. oldalán.

A bonyolultságtól eltekintve (9) majdnem ideális prímképletnek tekinthető; csak a többváltozós jellege zavaró egy kicsit. Felmerül a kérdés, nem lehetne-e hasonló, de egyváltozós képletet nyerni. Esetleg lemondhatnánk arról, hogy a képlet *minden* prímet előállítson, megelégednénk végtelen sok prím előállításával is. Egy egyszerű ötlettel (9) azonnal egyváltozóssá tehető: írjunk az x_i változók helyébe egy-egy $f_i(n)$ függvényt. Az így kapott

$$\max(Q(f_1(n),\ldots,f_{k+1}(n)),2) \tag{10}$$

képlet egyváltozós, és ha az $f_i(n)$ érték nemnegatív egész minden $i=1, \ldots, k+1$-re és $n=0, 1, 2, \ldots$-re, akkor (10) minden n-re prímet fog előállítani.

Látszólag készen vagyunk. Azonban a dolog mégsem ilyen egyszerű: ha pl.

f_1 , …, f_{k+1} gyanánt egész együtthatós polinomokat választunk, akkor a (10) képlet csak véges sok különböző értéket fog előállítani. Ugyanis ebben az esetben

$Q(f_1(n),\dots,f_{k+1}(n))=q(n)$ is egész együtthatós polinom lesz. Ha q konstans, akkor $\max(q(n),2)$ is konstans. Ha q nem konstans és a főegyütthatója negatív, akkor minden elég nagy n -re $q(n)<0$, tehát $\max(q(n),2)=2$. Ha viszont q nem konstans és a főegyütthatója pozitív, akkor minden elég nagy n -re $q(n)>2$, tehát $\max(q(n),2)=q(n)$. Ez azt jelentené, hogy $q(n)$ minden elég nagy n -re prímszám, amiről már beláttuk, hogy lehetetlen (egy nem konstans egész együtthatós polinom az n végtelen sok értékére összetett számot ad). Ez az eset tehát nem fordulhat elő!

Ez a jelenség első pillantásra hihetetlennek tűnik: a $Q(x_1,\dots,x_k,x_{k+1})$ polinomnak végtelen sok pozitív értéke van (hiszen minden prímet felvesz), de akárhogy helyettesítünk egész együtthatós polinomokat a változók helyébe, a kapott $Q(f_1(n),\dots,f_{k+1}(n))=q(n)$

polinom vagy konstans, vagy pedig negatív minden elég nagy n-re. Ez a különös jelenség azonban már egész egyszerű polinomok körében is fellép. Meg lehet mutatni, hogy a $Q(x,y)=(x^2+1)(1-2(x^2-2y^2-1)^2)$ polinom is rendelkezik ezzel a tulajdonsággal.

Ha el akarjuk érni, hogy a (10) képlet végtelen sok prímet szolgáltasson, a legegyszerűbb az $f_1,\dots,f_{k+1}$ függvényeket úgy megválasztani, hogy minden $(a_1,\dots,a_{k+1})$ vektor előálljon $(f_1(n),\dots,f_{k+1}(n))$ alakban. Ekkor (10) minden prímet elő fog állítani. Ha egy $(a_1,\dots,a_{k+1})$ vektorhoz létezik egy n természetes szám, melyre $f_1(n)=a_1,\dots,f_{k+1}(n)=a_{k+1}$, akkor azt fogjuk mondani, hogy az $f_1,\dots,f_{k+1}$ függvények *kódolják* az $(a_1,\dots,a_{k+1})$ vektort. Olyan függvényeket keresünk tehát, amelyek minden, nemnegatív egészekből álló vektort kódolnak. Mint láttuk, ezt polinomokkal nem érhetjük el, ezért fel kell használnunk más függvényeket is. Ekkor azonban egy újabb technikai nehézség lép fel. Ha szeretnénk minél egyszerűbb képleteket használni, akkor az f_i függvények nemcsak a pozitív egészekből, hanem a tetszőleges egészekből álló vektorokat is kódolni fogják. A Q polinom a negatív egészekben felvehet pozitív összetett számot is, és akkor (10) értéke nem lesz minden n-re prím. Ezen a következő módszerrel segíthetünk. Legyen $m=4(k+1)$, és tekintsük a

$$Q_1(y_1,\dots,y_m)= =Q(y_1^2+y_2^2+y_3^2+y_4^2,y_5^2+y_6^2+y_7^2+y_8^2,\dots \dots,y_{m-3}^2+y_{m-2}^2+y_{m-1}^2+y_m^2)$$

polinomot. Ezt tehát úgy kapjuk Q-ból, hogy mindegyik x_i változó helyére négy új változó négyzetösszegét írjuk. Most felhasználjuk azt a nevezetes tételt, amely szerint minden természetes szám előáll négy négyzetszám összegeként (lásd [1], 237. oldal). Nyilvánvaló, hogy

Q_1
-ben a változók helyére tetszőleges egészeket helyettesítve ugyanazokat az értékeket kapjuk, mint amikor

Q
-ban a változók helyére tetszőleges természetes számokat helyettesítünk. Így a

$\max(Q_1(y_1,\dots,y_m),2)$
képlet prímszámot szolgáltat, valahányszor

$y_1,\dots,y_m$
egészek, továbbá minden prímszámot megkapunk már akkor is, ha az

y_i
-k helyére természetes számokat helyettesítünk.

Olyan függvényeket fogunk gyártani, amelyek minden

m
-hosszúságú, természetes számokból álló vektort kódolnak. Ezeket az

y_i
változók helyére írva egyváltozós prímképletet kapunk. A konstrukciót csak

$m=2$
-re és

$m=3$
-ra adjuk meg, de világos lesz, hogy minden

m
-re elvégezhető.

Lássuk be először, hogy az

$([\sqrt{n}],n-[\sqrt{n}]^2)$
függvénypár minden olyan

(b_1,b_2)
számpárt kódol, amelyre

$b_1\geq b_2\geq 0$
. Legyen ugyanis

$n=b_1^2+b_2$
. Ekkor

$b_1^2\leq n<b_1^2+2b_1+1$
(hiszen

$b_2\leq b_1$
), tehát

$b_1\leq\sqrt{n}<b_1+1$
. Így

$[\sqrt{n}]=b_1$
és

$n-[\sqrt{n}]^2=(b_1^2+b_2)-b_1^2=b_2$.

Ha most a_1, a_2 tetszőleges természetes számok, akkor $a_1+a_2\geq a_2$, tehát van olyan n, amelyre $[\sqrt{n}]=a_1+a_2$ és $n-[\sqrt{n}]^2=a_2$. Ekkor $[\sqrt{n}]-(n-[\sqrt{n}]^2)=a_1$ és $n-[\sqrt{n}]^2=a_2$, tehát az $([\sqrt{n}]-n+[\sqrt{n}]^2,n-[\sqrt{n}]^2)$ függvénypár minden természetes számokból álló számpárt kódol.

Most tekintsük az $m=3$ esetet! Belátjuk először, hogy a

$g_1(n)=[\sqrt[4]{n}]$, $g_2(n)=[\sqrt{n}]-g_1(n)^2$,

$g_3(n)=n-(g_1(n)^2+g_2(n))^2$ függvények minden olyan (b_1,b_2,b_3) vektort kódolnak, melyekre $b_1\geq b_2\geq b_3\geq 0$. Valóban, legyen $n=(b_1^2+b_2)^2+b_3$. Ekkor $(b_1^2+b_2)^2\leq n<(b_1^2+b_2)^2+2(b_1^2+b_2)+1$ (hiszen $b_3\leq b_2$), tehát

$b_1^2+b_2\leq\sqrt{n}<b_1^2+b_2+1$

és $b_1^2 \le \sqrt{n} < b_1^2+2b_1+1$ (hiszen $b_2 \le b_1$). Így $[\sqrt{n}]=b_1^2+b_2$, $[\sqrt[4]{n}]=b_1$, amiből $g_1(n)=b_1$, $g_2(n)=b_2$ és $g_3(n)=b_3$.

Ha most a_1, a_2, a_3 tetszőleges természetes számok, akkor $a_1+a_2+a_3 \ge a_2+a_3 \ge a_3$, tehát van olyan n, amelyre $g_1(n)=a_1+a_2+a_3$, $g_2(n)=a_2+a_3$ és $g_3(n)=a_3$. Ebből következik, hogy a g_1-g_2, g_2-g_3 és g_3 függvények minden természetes számokból álló számhármast kódolnak.

Ezt a konstrukciót minden m-re elvégezhetjük. Ha az így kapott függvényeket $Q1$-be helyettesítjük, akkor végül is a következő tételt kapjuk.

Létezik olyan $f(n)$ *kifejezés, amely az* n, $[\sqrt{n}]$, $[\sqrt[4]{n}]$, …, $[\sqrt[2^r]{n}]$ *függvények egész együtthatós polinomja (azaz a fenti függvényekből és egész számokból kapható az összeadás, kivonás és szorzás műveleteinek segítségével), és amelyre a* $\max(f(n),2)$ *(*$n=0$*, 1, …) számok halmaza pontosan a prímszámok halmazával egyenlő.*

Ilyen alakban nemcsak a prímszámok állíthatók elő. Ha ismét áttekintjük a tételhez vezető gondolatmenetet, láthatjuk, hogy bármely rekurzíve felsorolható sorozatnak van ilyen előállítása. Mindazonáltal a prímszámok előállításával kapcsolatban felmerül néhány érdekes kérdés.

1. A prímszámokat előállító képletekben mennyi az r minimális értéke? (A fenti gondolatmenet $r=39$-et ad, hiszen k-ra az ismert legkisebb érték 9, és $r=m-1=4(k+1)-1$.) Lehet-e pl. olyan prímképletet megadni, amely csak az n és $[\sqrt{n}]$ függvényeket használja?

2. Megadható-e olyan, fenti alakú

f
függvény, amelyre

$f(n)=pn$
minden

n
-re?

IRODALOM

[1] [1] Erdős Pál és Surányi János: *Válogatott fejezetek a számelméletből*, Polygon, Szeged, 1996.

[2] [2] G. H. Hardy and E. M. Wright: *An Introduction to the Theory of Numbers*, Oxford University Press, Oxford, 1975.

[3] [3] P. Ribenboim: *The Little Book of Big Primes*, Springer, 1991.

[4] [4] D. Shanks: *Solved and Unsolved Problems in Number Theory*, Chelsea, New York, 1985.