Sudoku és a véges geometriák

Diplomamunka

Írta: Kiss-Tóth Christian
Matematikus szak

Témavezető:

Kiss György, egyetemi docens
Geometriai Tanszék

Eötvös Loránd Tudományegyetem
Természettudományi Kar
2011

Tartalomjegyzék

Bevezetés

Diplomamunkámban a manapság rejtvény formájában igen elterjedt és népszerű sudoku táblázattal és annak matematikai vonatkozásaival foglalkozom.

A táblázat legközismertebb változatát *Harold Garns* vezette be 1979-ben, amely 9×9 mezőből áll, 9 darab 3×3-as blokkra van osztva és 1-től 9-ig tartalmazza a számokat úgy, hogy minden sorban, oszlopban és blokkban minden szám pontosan egyszer szerepel. Ennek a konstrukciónak az egy feltétellel gyengébb változata, amikor elfelejtjük a táblázat blokkokra osztását és csak a sorokat, illetve oszlopokat tekintjük a *latin négyzet*. Definiálhatunk azonban a blokkokon kívül más partíciókat is, amelyre ugyanazt követeljük meg, mint a sorokra és oszlopokra: ennek az általánosításnak a neve *design*. A dolgozatban vizsgált szimmetrikus sudoku egy ilyen speciális design.

Munkánk másik központi témája a latin négyzetek ortogonalitásához kapcsolódik. Két ($n \times n$-es) latin négyzetet – és így speciálisan sudoku táblázatot és designt – akkor nevezünk merőlegesnek, ha az általuk meghatározott n^2 rendezett számpár mindegyike különböző. Az ilyen ortogonális rendszerek elemszámára a latin négyzetek esetében az $n - 1$ egy triviális és közismert felső becslés, ezt fogjuk vizsgálni sudoku táblázatok esetén.

Dolgozatomat három nagy fejezetre tagoltam, az első fejezetben összefoglalom azokat az affin és főként projektív geometriákra vonatkozó állításokat, tételeket, illetve konstrukciókat, amelyek a továbbiakban szükségesek lesznek. Az itt közölt tételekhez többségében nem írunk bizonyítást, ezek megtalálhatóak a [4]-ben.
A második fejezetben összefoglalom a klasszikus sudoku táblázatra vonatkozó állításokat. Ezek az eredmények megtalálhatóak a munkám kiindulási alapját képező [1] cikkben, illetve magyar nyelven [2]-[3]-ban, a konkrét számolások illetve konstrukciók azonban saját munka eredményei.

A cikkben szereplő alapgondolatokat és módszereket sikerült általánosítanom a nagyobb, 5×5-ös méretű táblázatok esetére, ezeket az eredményeimet ismertetem a harmadik fejezetben. Mivel a sorok, oszlopok, illetve blokkok partíciója itt már nem ad szükségszerűen szép és jól leírható kitöltéseket, megvizsgáltam hogyan érdemes még partíciókat definiálni, illetve az újonnan meghatározott partíciókat is figyelembe véve mi mondható el a kitöltések és a merőleges rendszerek (elem)számáról.

Köszönetnyilvánítás

Szeretném megköszönni témavezetőmnek, Kiss Györgynek az érdekes, ugyanakkor alapjait tekintve mindenki számára könnyen megérthető témát, valamint a felmerülő kérdéseimre adott részletes válaszait és a dolgozatom alapos átnézésére szakított időt.

1. Alapfogalmak, tételek

Munkánk során látni fogjuk, hogy egy sudoku tábla mezőit el tudjuk látni koordinátákkal, és ezáltal egy táblázatra úgy tudunk majd tekinteni mint egy vektortérre. Számos esetben használni fogjuk a (vektortérből származtatott) projektív, illetve affin geometriákat, így ebben a fejezetben ezeknek az alapvető és kevésbé közismert tulajdonságait foglaljuk össze.

1.1. Affin geometriák

1.1 definíció: Affin síknak egy olyan $(\mathcal{P}, \mathcal{E}, \mathcal{I})$ hármast nevezünk, ahol $\mathcal{P}$ a sík pontjainak, $\mathcal{E}$ a sík egyeneseinek a halmaza, $\mathcal{I}$ pedig egy $\mathcal{P} \times \mathcal{E}$ reláció a következő tulajdonságokkal:

A1) Bármely két különböző ponthoz egyértelműen van egy olyan egyenes, amely mindkettejükkel relációban áll.

A2) Ha egy P pont nem áll relációban egy e egyenessel, akkor pontosan egy olyan egyenes van $\mathcal{E}$-ben, amelyik relációban áll P-vel, de nem áll relációban egyetlen olyan ponttal sem, amely relációban áll e-vel.

A3) $\mathcal{E}$ minden eleme legalább kettő $\mathcal{P}$-beli elemmel áll relációban.

A4) $\mathcal{P}$ minden eleme legalább három $\mathcal{E}$-beli elemmel áll relációban.

A szokásos szóhasználattal élve azt mondjuk, hogy ha egy pont és egy egyenes relációban állnak, akkor a pont rajta van az egyenesen, és ha két egyeneshez van olyan pont, amely mindkettőn rajta van, akkor azok metszik egymást, ellenkező esetben párhuzamosak. Az affin sík tehát a klasszikus euklidészi síknak egy általánosítása, az utolsó két axióma csak arról gondoskodik, hogy az elfajuló eseteket kizárja.

1.2 állítás: Ha egy $\mathcal{A}$ affin síknak van olyan egyenese, amelyre n pont illeszkedik, akkor:

- $\mathcal{A}$ minden egyenesén n pont van,
- $\mathcal{A}$ minden pontján $n+1$ egyenes megy,
- $\mathcal{A}$ összesen n^2 pontot és n^2+n egyenest tartalmaz.

Ezt az n számot az **affin sík rendjé**nek nevezzük.

1.3 definíció: Az **affin tér**en egy olyan vektorteret értünk, amelyben eltávolítottuk az origó kitüntett szerepét. **Affin altér**nek a lineáris alterek eltoltjait nevezzük, az affin altér dimenziója a neki megfelelő lineáris altér dimenziója. Egy altér különböző eltoltjai diszjunktak, és egy **párhuzamossági osztály**t alkotnak.

Látható, hogy ha kétdimenziós vektortérből indulunk ki, akkor éppen affin síkot kapunk, ha a nulla dimenziós altereket tekintjük a pontoknak, az egy dimenziósakat pedig egyeneseknek, az illeszkedés a tartalmazás. A q elemű véges test feletti n dimenziós vektortérből képezett affin teret $AG(n,q)$-val jelöljük.

1.2. Projektív geometriák

Az affin alterekről lényegében nem is fogunk többet használni, most rátérünk a projektív geometriák bevezetésére és főbb tulajdonságainak ismertetésére, bizonyítására.

1.2.1. Projektív geometriák fogalma

1.4 definíció: Projektív síknak egy olyan $\Pi = (\mathcal{P}, \mathcal{E}, \mathcal{I})$ hármast nevezünk, ahol $\mathcal{P}$ a sík pontjainak, $\mathcal{E}$ a sík egyeneseinek a halmaza, $\mathcal{I}$ pedig egy $\mathcal{P} \times \mathcal{E}$ reláció a következő tulajdonságokkal:

P1) Bármely két különböző ponthoz egyértelműen van egy olyan egyenes, amely mindkettejükkel relációban áll.

P2) Bármely két különböző egyeneshez egyértelműen van egy olyan pont, amely mindkettejükkel relációban áll.

P3) $\mathcal{E}$ minden eleme legalább három $\mathcal{P}$-beli elemmel áll relációban.

P4) $\mathcal{P}$ minden eleme legalább három $\mathcal{E}$-beli elemmel áll relációban.

A klasszikus euklidész síkra **P2** nem teljesül, de kis módosítással ez a probléma orvosolható. Vegyünk hozzá a síkhoz új pontokat úgy, hogy minden párhuzamossági osztálynak feleljen meg egy új pont, ezeket hívjuk ideális pontoknak (az euklidészi sík pontjait pedig nevezzük közönséges pontoknak). Bővítsük ki az $\mathcal{I}$-t úgy, hogy minden egyenes álljon relációban azzal az ideális ponttal amelyik párhuzamossági osztálynak ő megfelel, továbbá vegyünk be a geometriánkba egy új (ideális) egyenest is, amely az ideális pontokat tartalmazza. Meggondolható, hogy ez kielégíti a négy projektív axiómát, így kapjuk a **klasszikus projektív síkot**.

Az **1.2**-es állításhoz hasonló projektív esetben is elmondható:

1.5 állítás: Ha egy Π projektív síknak van olyan egyenese, amelyre $n+1$ pont illeszkedik, akkor:

- Π minden egyenesén $n+1$ pont van,
- Π minden pontján $n+1$ egyenes megy,
- Π összesen n^2+n+1 pontot és ugyanennyi egyenest tartalmaz.

Ezt az n számot a **projektív sík rendjé**nek nevezzük. Ha tehát egy projektív sík valamely egyenesén véges sok pont van, akkor minden egyenesén ugyanennyi pont, illetve minden ponton át ugyanennyi egyenes halad.

Projektív síkot úgy is kaphatunk, hogy veszünk egy háromdimenziós V vektorteret, az egydimenziós altereket pontoknak, a kétdimenziós altereket pedig egyeseknek nevezzük, az illeszkedés pedig a halmazelméleti tartalmazás. Ekkor bármely két ponthoz egyértelműen vagy olyen egyenes amely mindkét pontot tartalmazza (az unió által generált altér) és a dimenzió-tétel miatt bármely két különböző kétdimenziós altér metszete egydimenziós, vagyis a második feltétel is teljesül. Ennek mintájára most definiáljuk az n-dimenziós projektív teret.

1.6 definíció: Legyen V_{n+1} a $GF(q)$ véges test feletti $n+1$ dimenziós vektortér, a projektív értelemben k-dimenziós alterek legyenek a vektortér $k+1$ dimenziós lineáris alterei, az illeszkedés pedig a tartalmazás. Az így kapott teret nevezzük az n-dimenziós projektív térnek, és $PG(n,q)$-val jelöljük.

Projektív tereket nem csak véges, hanem tetszőleges $\mathbb{K}$ test felett is definiálhatunk, nekünk azonban erre a dolgozat folyamán nem lesz szükségünk, csak a $PG(n,q)$ térről fogunk beszélni. Ebben az esetben is használhatjuk a pont, egyenes, sík, hipersík szóhasználatot a $0,1,2$ illetve $n-1$ dimenziós alterek megnevezésére. Meggondolható, hogy:

1.7 állítás: $PG(n,q)$ összesen $\frac{q^{n+1}-1}{q-1}$ pontot tartalmaz, és minden egyenesen $q+1$ pont van.

1.2.2. Projektív terek analitikusan

Ebben a pontban a $PG(n,q)$ projektív tér pontjait, illetve alakzatait fogjuk leírni analitikusan. V_{n+1} elemeinek megfeleltethetünk egy $n+1$ hosszú vektorokat, és ennek megfelelően a pontok (az 1 dimenziós alterek) jellemezhetőek az altér generátorvektorával. Mivel egy egydimenziós altérnek $q-1$ generátora van, amelyek egymás nem nulla skalárszorosai így kikötjük, hogy ha λ egy nem nulla testelem, akkor $(x_1, x_2, \ldots, x_{n+1}) \sim (\lambda x_1, \lambda x_2, \ldots, \lambda x_{n+1})$, vagyis a két vektor ugyanazt a pontot határozza meg. Ezt nevezzük a pont **homogén koordinátá**jának (amely tehát skalárszorzó erejéig egyértelmű).

A hipersíkokat is megadhatjuk koordinátákkal, hiszen egy vektortérben minden alteret meghatározza az ortogonális kiegészítője, és mivel a hipersík esetében ez egy egydimenziós altér, így a síkot meghatározza ennek a vektornak a homogén koordinátája. A koordinátákat hipersík esetén szögletes zárójelbe tesszük. Egy $(x_1, \ldots, x_{n+1})$ pont pedig akkor és csak akkor van rajta az $[u_1, \ldots, u_{n+1}]$ hipersíkon, ha a két vektor skalárszorzata nulla.

Felmerül a kérdés, hogy az így kapott koordinátákat használva hogyan tudunk alakzatokat leírni? Ha H egy $n+1$ változós polinom, akkor milyen feltételeket kell H-nak teljesítenie ahhoz, hogy a $H = 0$ egyenlet egy alakzatot határozzon meg? Azt szeretnénk, hogy minden olyan $\underline{x}$ vektorra melyre $H(\underline{x}) = 0$, $H(\lambda \underline{x}) = 0$ is teljesüljön. Ez pedig pontosan akkor igaz, ha H egy homogén (k-fokú) polinom. Az előző bekezdésből látszik például, hogy ha $k = 1$, akkor H egy hipersíkot határoz meg. Alacsonyabb dimenziójú altereket pedig hipersíkok metszeteként (így analitikusan lineáris egyenletrendszerekkel) adhatunk meg.

Vizsgáljuk meg érdekességképpen, hogy mi lehetne például az $Y = X^2$ parabola egyenlete, ha átmegyünk a projektív síkra. A klasszikus sík egy (x, y) pontját megfeleltethetjük a projektív sík $(x, y, 1)$ pontjának, ami azt jelenti, hogy egy közönséges $(x_1, x_2, x_3) \sim \left(\frac{x_1}{x_3}, \frac{x_2}{x_3}, 1\right)$ pont pontosan akkor van rajta a parabolán, ha $\frac{x_2}{x_3} = \left(\frac{x_1}{x_3}\right)^2$, azaz $x_1^2 = x_2 x_3$. Közönséges pontoknál ez ekvivalens az euklideszi síkon ismert képlettel, az ideális pontokat tekintve pedig a projektív síkon a $(0, x_2, 0) \sim (0, 1, 0)$ pont is rajta van a görbén.

1.8 definíció: Legyen adott két projektív tér, $\mathcal{S}$ és $\mathcal{S}'$. Kollineációnak egy $\mathcal{S} \to \mathcal{S}'$ illeszkedéstartó leképezést nevezünk. Ha $\mathcal{S} = \mathcal{S}'$, akkor az $\mathcal{S}$ tér egy kollineációjáról beszélünk.

Ha σ $GF(q)$ egy automorfizmusa, akkor azt a vektortéren koordinátánként alkalmazva kollineációt kapunk. Kollineációt határoz meg továbbá minden $(n+1) \times (n+1)$-es invertálható mátrix is. A projektív geometria alaptétele kimondja, hogy lényegében más kollineáció nincs is:

1.9 tétel: (a projektív geometria alaptétele) A $PG(n, q)$ tér minden kollineációja előáll úgy, mint egy testautomorfizmus és egy lineáris leképzés kompozíciója.

Mi azonban csak lineáris transzformációkkal fogunk számolni.

1.2.3. Másodrendű felületek, és alapvető tulajdonságaik

A továbbiakban vizsgálni fogjuk a projektív terek diszjunkt egyenesekkel történő befedhetőségét, ahol nagy szerepet kapnak a másodrendű felületek, így ebben a pontban definiálni fogjuk ezeket, és ismertetjük a főbb tulajdonságaikat is. Az egyszerűség kedvéért feltesszük, hogy q nem páros, vagyis a testünk karakterisztikája nem 2, mert ebben az esetben az elkövetkezendő tételeket egy kicsit módosítani kellene, hogy igazak legyenek, viszont a mi esetünkben, amikor használni fogjuk ezeket az eszközöket q mindig páratlan lesz.

1.10 definíció: Másodrendű felületen olyan ponthalmazt értünk, melynek elemeire teljesül, hogy $\underline{x}^\top A \underline{x} = 0$, ahol A szimmetrikus mátrix.

1.11 megjegyzés: Ha A az identitás, akkor a pontok koordinátáinak a $\sum x_i^2 = 0$ összefüggést kell kielégítenie, míg például $n = 2$ esetén az $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}$ a fentebb ismertetett parabolát határozza meg. Könnyen látható az is, hogy ilyen alakban pontosan a homogén másodfokú polinomok állnak elő. Az A mátrixhoz tartozó kvadratikus alakot $Q_n(\underline{x})$-el fogjuk jelölni.

Vegyük észre, hogy ha alkalmazunk a téren egy M lineáris transzformációt (megváltoztatjuk a koordinátarendszert), akkor másodrendű felület képe továbbra is másodrendű felület lesz, hiszen ha x képe Mx, akkor átalakítva a feltételt:

$$0 = \underline{x}^\top A \underline{x} = \underline{x}^\top M^\top \left(\left(M^{-1}\right)^\top A M^{-1} \right) M \underline{x},$$

vagyis az $\left(M^{-1}\right)^\top A M^{-1}$ mátrixhoz tartozó felület éppen az eredeti felület M szerinti képe lesz.

Ettől a ponttól kezdve feltesszük, hogy $n = 3$, vagyis a három dimenziós projektív tér másodrendű felületeire vonatkozó állításokat fogjuk ismertetni. Ezeket többnyire bizonyítás nélkül közöljük, azok elolvashatók *Kiss György - Szőnyi Tamás: Véges geometriák* című könyvének negyedik fejezetében.

1.12 definíció: Egy másodrendű felületet szingulárisnak nevezünk, ha a koordinátarendszer változtatásával elérhető, hogy a hozzá tartozó kvadratikus alak kevesebb mint $n + 1$ változót tartalmazzon. Ez ekvivalens azzal, hogy $\det A = 0$.

1.13 tétel: Nem szinguláris másodrendű felülethez tartozó kvadratikus alak a koordinátarendszer megváltoztatásával a következő két alak valamelyikére hozható:
1. $Q_3(\underline{x}) = x_1x_2 + x_3x_4$, vagy
2. $Q_3(\underline{x}) = f(x_1, x_2) + x_3x_4$, ahol f irreducibilis homogén másodfokú polinom.

Ez azt jelenti tehát, hogy a háromdimenziós térben a (projektív szempontból nézve) különböző másodrendű felületek száma kettő. Azokat a másodrendű felületeket, amelyek az 1.-es alakra hozhatóak **hiperbolikus**, amelyek pedig a 2.-es alakra hozhatók **elliptikus** másodrendű felületnek nevezzük.

1.14 tétel: A $PG(3, q)$ térben minden
a) elliptikus másodrendű felületre $q^2 + 1$,
b) hiperbolikus másodrendű felületre $(q + 1)^2$ pont illeszkedik.

Vegyük észre, hogy egy L lineáris transzformációt alkalmazunk a térre, akkor a másodrendű felület mátrixának a determinánsa az inverzmátrix determinánsának a négyzetével szorzódik. Ez azt jelenti, hogy az a tulajdonság, hogy a determináns négyzetelem-e vagy sem a q elemű testben bázistranszformáció során nem változik. Ez azt jelenti tehát, hogy a determináns négyzetelem voltával is lehet karakterizálni a másodrendű felületet, a karakterizáció azonban a test elemszámától függ. Ha például $q = 3$, akkor az elliptikus felületek mátrixának a determinánsa nem, míg a hiperbolikusaké pedig négyzetelem. $q = 5$ esetén ez pont fordítva van.

Ha egy projektív teret szeretnénk egyenesekkel (r dimenziós alterekkel) diszjunkt módon lefedni, akkor szükséges feltétel, hogy az altér pontjainak a száma osztója legyen a tér pontjainak a számával. Mivel egy r dimenziós (al)térben összesen $\Theta(r) = \frac{q^{r+1}-1}{q-1}$ pont van, így $\Theta(r)\big|\Theta(n)$ azzal ekvivalens, hogy $r + 1\big|n + 1$. Igaz azonban a tétel megfordítása is:

1.15 tétel: A $PG(n, q)$ térnek akkor és csak akkor létezik r-dimenziós alterekből álló diszjunkt befedése, ha $r + 1\big|n + 1$.

A mi esetünkben, ha $r = 1$ és $n = 3$ akkor az oszthatóság teljesül, vagyis a háromdimenziós tér befedhető egyenesekkel. Mivel minden egyenesen $q+1$ pont van, és a tér összes pontjainak a száma $q^3 + q^2 + q + 1$, összesen $q^2 + 1$ egyenesre lesz szükségünk.

A tér egy fedését lényegében két lépésben fogjuk előállítani. Először megvizsgáljuk, hogy egy másodrendű felületnek milyen feltételek mellett létezik egyenesekkel történő befedése, majd ezután a teret ilyen felületekkel fogjuk diszjunkt módon lényegében teljesen lefedni (a kimaradó pontok két egyenest fognak alkotni). Lássuk először a következő egyszerű állítást:

1.16 állítás: Ha egy másodrendű felület tartalmazza egy egyenes három pontját, akkor tartalmazza a teljes egyenest is.

BIZONYÍTÁS: Ha az $\underline{u}$, $\underline{v}$ valamint a $\mu\underline{u}+\lambda\underline{v}$ pontok is rajta vannak az A mátrix által meghatározott felületen, az azt jelenti, hogy $\underline{u}^\top A\underline{u} = \underline{v}^\top A\underline{v} = \underline{u}^\top A\underline{v} = 0$, amiből következik, hogy minden $\alpha\underline{u} + \beta\underline{v}$ alakú pont kielégíti a szükséges feltételt, hogy rajta legyen a felületen. ♣

Meghatározható, hogy mekkora az a legnagyobb dimenziós altér, amely teljes egészében rajta van egy másodrendű felületen. Az ezt az alteret meghatározó vektortérbeli altérnek a dimenziója a felület **projektív indexe**, ezzel kapcsolatban igaz a következő tétel:

1.17 tétel: A $PG(3,q)$ térben minden elliptikus felület indexe 1, minden hiperbolikus felület indexe 2. A hiperbolikus felületekre teljesül továbbá, hogy minden pontján át pontosan két olyan egyenes halad, amely teljes egészében a felületen halad, ez összesen $\frac{2(q+1)^2}{q+1} = 2(q+1)$ egyenest jelent, amelyek a felület két diszjunkt fedésévé állnak össze.

A konstrukciónk lényege az lesz, hogy megadunk a térben $q-1$ darab diszjunkt hiperbolikus másodrendű felületet, illetve két további, ezektől diszjunkt egyenest. Ha ez sikerülne, akkor a térnek egy befedését kaptuk, hiszen a felületek különkülön fedhetőek egyenesekkel, és kimaradó két egyenessel megkaptuk a tér összes $(q-1)(q+1)^2 + 2(q+1) = (q^2+1)(q+1)$ pontját. Az alábbiakban mutatunk egy konstrukciót, ami egy ilyen hiperboloidsereget szolgáltat.

Legyen f egy irreducibilis homogén másodfokú polinom, és tekintsük a következő $q-1$ felületet:

$$\mathcal{F}_t : f(x_1,x_2) + t \cdot f(x_3,x_4) = 0,$$

ahol t végigfut az összes nem nulla testelemen. Megmutatjuk, hogy ezek a felületek diszjunktak és hiperbolikusak. Mivel f-et irreducibilisnek választottuk, így ha u és

v nem mindketten nullák, akkor $f(u,v) \neq 0$, mert ha 0 volna, akkor a homogenitás miatt (u,v) bármely skalárszorosára is nulla lenne f értéke, ami azt jelentené, hogy f-ből kiemelhető lenne egy elsőfokú tényező. Ez pedig azt jelenti, hogy minden olyan olyan $\underline{x}$ vektor, amelyre teljesül hogy az első és az utolsó két koordinátája közül is legalább az egyik nem nulla, pontosan egy felületen lesz rajta, méghozzá azon, amelyre $t = -\frac{f(x_1,x_2)}{f(x_3,x_4)}$ (a megjegyzésünk miatt értelmes és nem nulla).

Ebből pedig már egyrészről következik az, hogy a felületek diszjunktak, másrészt az $x_1 = x_2 = 0$, $x_3 = x_4 = 0$ diszjunkt egyenespárt leszámítva minden pont rajta van valamelyik felületen, ami csak úgy lehetséges, ha mind a $q-1$ felület $(q+1)^2$ pontot tartalmaz, vagyis hiperbolikus.

A fenti gondolat szemléletes tartalmát úgy foghatjuk meg a legjobban, ha tekintjük a euklidészi térből képezett projektív teret, és legyen $f(x,y) = x^2 + y^2$. Ekkor a projektív egyenletek a klasszikus térben az $x^2 + y^2 = tz^2$ hiperboloidsereget adják meg, és a két kimaradó egyenes egyike a z tengely, a másik pedig egy ideális egyenes.

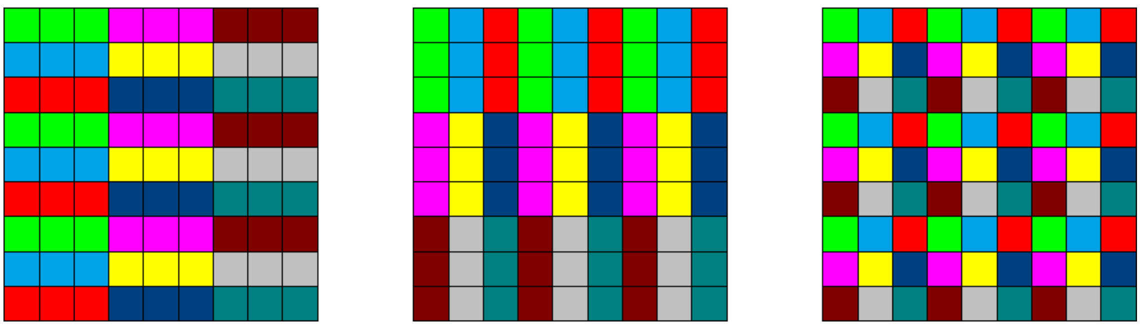
2. A 3×3-as Sudoku

Klasszikus sudokunak azt a mindenki által ismert 9×9-es táblazatot nevezzük, amelybe úgy írjuk be a számokat 1-től 9-ig, hogy minden sorban, oszlopban, illetve 3×3-as blokkban minden szám pontosan egyszer szerepel. Mivel mi ennél többet fogunk megkövetelni a kitöltésről, így a precízség kedvéért először definiáljuk a latin négyzetet:

2.1 definíció: $n \times n$-es latin négyzeten egy olyan táblázatot értünk, amely az $1, 2, \ldots n$ számokat tartalmazza úgy, hogy minden sorban és oszlopban minden szám pontosan egyszer szerepel.

A latin négyzetek általánosítása az úgynevezett *design*. Legyen $S_1, S_2, \ldots S_n$ a táblázat egy partíciója úgy, hogy minden partíció n mezőt tartalmaz, és ebben az esetben azt követeljük meg, hogy a partíció minden osztálya minden számot pontosan egyszer tartalmazzon. Természetesen több partíciót is tekinthetünk, és megkövetelhetjük, hogy a kitöltésünk mindegyik partícióra nézve egyszerre legyen design. Ebben az értelmezésben tehát a latin négyzetek a sorok és az oszlopok partíciójára, míg a sudoku a blokkok partíciójára nézve is design.

Megvizsgálható, hány különböző módon tölthető ki egy ilyen sudoku, viszont egyrészt az ilyen kitöltések száma túl sok, másrészt nem is mondhatóak el olyan szép tulajdonságok, mint amilyenek a továbbiakban ki fognak jönni, ha egy kicsit többet követelünk meg. Tekintsük az alábbi három partíciót:



Az első partíciót *törött sor*-nak, a másodikat *törött oszlop*-nak, a harmadikat pedig *pozíció*-nak fogjuk nevezni.

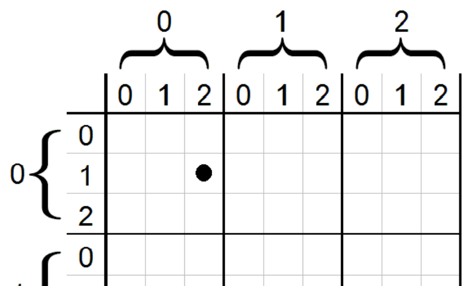
2.2 definíció: Az olyan 9×9-es táblázatot amely a sorok, oszlopok, blokkok, törött sorok, törött oszlopok, és pozíciók partíciójára nézve is design, *szimmetrikus sudoku*-nak nevezzük.

A táblázat szimmetrikus kitöltései már nagyságrendekkel kevesebben vannak, mint a klasszikus kitöltések, és sokkal szebb algebrai tulajdonságokkal bírnak.

2.1. Szimmetrikus kitöltések száma és tulajdonságai

Ebben a fejezetben tehát arra a kérdésre keressük a választ, hogy egy szimmetrikus sudokunak milyen kellemes algebrai illetve geometriai tulajdonságai vannak, és ezen tulajdonságok felhasználásával meghatározzuk azt is, hogy hány különböző szimmetrikus sudoku létezik. A számok permutálásával egymásba vihető megoldásokat természetesen nem tekintjük különbözőnek.

Mielőtt azonban erre rátérnénk, hogy eszköz legyen a kezünkben, koordinátákkal fogjuk ellátni a táblázat mezőit. Három egymás mellett lévő blokk unióját *nagy sor*-nak, három egymás alatt lévő blokk unióját pedig *nagy oszlop*-nak nevezzük. Ennek megfelelően a táblázat három nagy sor/oszlop uniója, amelyek mindegyike három sort/oszlopot tartalmaz. Mivel minden mezőt egyértelműen meghatároz, hogy melyik nagy sorban/oszlopban, illetve azon belül melyik sorban/oszlopban van. Innen már adódik a táblázat mezői és a $V = GF(3)^4$ vektortér vektorai közötti megfeleltetés, ha ellátjuk a nagy (és a bennük lévő kis) sorokat és oszlopokat is a $GF(3)$ test elemeivel mint koordinátákkal:



Minden mezőnek négy koordinátája lesz (x_1, x_2, x_3, x_4), ahol x_1 a nagy sort, x_2 a sort, x_3 a nagy oszlopot, x_4 pedig az oszlopot határozza meg, az ábrán pöttyel jelölt mező koordinátái például $x_1 = 0$, $x_2 = 1$, $x_3 = 0$ és $x_4 = 2$.

Következő lépésben vegyük észre, hogy ezen koordináták segítségével egyszerűen leírhatóak a fenti partíciók, amelyek egyébként alteret alkotnak V-ben.

- sorok: $x_1 = \alpha$, $x_2 = \beta$,
- oszlopok: $x_3 = \alpha$, $x_4 = \beta$,
- blokkok: $x_1 = \alpha$, $x_3 = \beta$,
- törött sorok: $x_2 = \alpha$, $x_3 = \beta$,
- törött oszlopok: $x_1 = \alpha$, $x_4 = \beta$,
- pozíciók: $x_2 = \alpha$, $x_4 = \beta$.

Az olyan halmazok, amelyeket azon vektorok alkotnak, ahová ugyanazt a számot írtuk kilenc (diszjunkt, azonos elemszámú) halmaz uniójára bontják a teret, és ezekre a halmazokra teljesül, hogy az egyes partíciók osztályait pontosan egy pontban metszik. Az ilyen tulajdonságú halmazokat a továbbiakban az egyszerűség kedvéért *jó*-nak fogjuk nevezni. Jelöljön $\mathcal{H}$ egy ilyen jó halmazt.

2.3 definíció: Az x, y vektorok *Hamming*-távolságán azt a $d(x, y)$ számot értjük, ahány koordinátában a két vektor különbözik egymástól.

A vektorok Hamming-távolságának bevezetésével leírhatóak a jó halmazok, vizsgálatunk első kulcsfontosságú tétele a következő:

2.4 tétel: Egy $\mathcal{H}$ halmaz akkor és csak akkor jó, ha minden $x, y \in \mathcal{H}$ esetén $d(x, y) = 3$.

BIZONYÍTÁS: Legyen $\mathcal{H}$ egy jó halmaz. A partíciók választásából azonnal adódik, hogy $d(x, y) \geq 3$, hiszen ellenkező esetben lenne két vektor, amely legalább két koordinátában azonos, vagyis $x_i = y_i$ és $x_j = y_j$ valamely $i \neq j$ indexek esetén. Ez azonban azt jelentené, hogy valamelyik partíció egyik osztályában az egyik szám kétszer szerepel: ha például $i = 1$, $j = 2$, akkor az egyik sorban nem lenne minden szám különböző, ez pedig ellentmondás.

A másik irány igazolásához adjuk össze $d(x, y)$-t az összes $x \neq y$ párra. 72 ilyen rendezett pár van. Legyen ez a mennyiség S. S-et úgy is kiszámolhatjuk, hogy megnézzük az egyes koordináták mennyivel járulnak hozzá ehhez a távolságösszeghez. Vizsgáljuk meg a kilenc darab első koordinátát. Ezek között pontosan annyi 0-ás, 1-es és 2-es van, ahány $\mathcal{H}$-beli elem van az egyes nagy sorokban. Viszont minden nagy sor három sor (blokk, törött oszlop) uniója amelyek mindegyikében pontosan egy $\mathcal{H}$-beli elem található, tehát $GF(3)$ minden eleme háromszor szerepel első koordinátaként. Ez pedig azt jelenti, hogy az első koordináta a $9 \cdot 6 = 54$-el járul hozzá a távolságösszeghez, hiszen minden vektor hat másik vektortól tér el az első koordinátájában. Mivel a négy koordináta minden tekintetben szimmetrikus, így ez az állítás elmondható a másik három koordináta esetében is, vagyis kaptuk, hogy a Hamming-távolságok összege pontosan $4 \cdot 54 = 216$.

A két eredményt összevetve pedig következik az egyik irány, hiszen a távolságösszeg egyrészt legalább $72 \cdot 3 = 216$, másrészt pontosan 216, így az első bekezdésbeli egyenlőtlenségben minden x, y esetén egyenlőség áll.

A másik irány igazolása hasonlóan működik. Legyen $\mathcal{H}$ egy olyan kilenc elemű halmaz, amelyben bármely két vektor távolsága három, meg kell mutatni, hogy egy tetszőleges osztályban pontosan egy darab $\mathcal{H}$-beli elem található. Egy osztály úgy írható le koordinátákkal, hogy $x_i = \alpha$, $x_j = \beta$. A skatulya elv miatt, az i-dik koordinátában minden testelem pontosan háromszor fordul elő. Ha ugyanis lenne olyan, ami négyszer is szerepel, akkor ez a négy vektor közül lenne kettő amelyik egy másik koordinátában is megegyezik, és ez ellentmond annak, hogy a távolságuk három. Így tehát pontosan három olyan vektor van, amelyiknek az i-dik koordinátája α, és ennek a három vektornak a j-dik koordinátája különböző, tehát pontosan egy olyan van, ahol ez β. Ezzel az állítást igazoltuk. $\clubsuit$

2.5 megjegyzés: Ha a bizonyításban nem használjuk ki, hogy egy nagy sor három sor uniója, és így az első koordinátában szükségszerűen mindhárom testelem háromszor fordul elő, akkor az előfordulásokat p, q, r-el jelölve az első koordináta hozzájárulására azt kapjuk, hogy $p\cdot(q+r)+q\cdot(p+r)+r\cdot(p+q) = 81-(p^2+q^2+r^2) \leq 54$ a számtani és négyzetes közép közötti összefüggésből. Ez pedig azt jelenti, hogy a távolságösszeg legfeljebb 216, és a maximumot az említett esetben éri el. Ebből mi csak annyit fogunk használni, hogy a jó halmazokat most már úgy is jellemezhetjük, hogy bármely két elemének távolsága legalább három.

A bizonyítás folytatásaképpen vegyük észre, hogy egy jó halmaz minden eltoltja is jó, hiszen a partíciók valójában nem mások mint egy altér, és annak a nyolc különböző eltoltja, ami azt jelenti, hogy egy halmaz eltolása nem változtat a „jóságán”. Mivel minden halmaznak van olyan eltoltja, amely tartalmazza az origót, így feltehetjük, hogy $\underline{0} \in \mathcal{H}$. Ekkor teljesül a következő tétel:

2.6 tétel: Minden origót tartalmazó jó halmaz alteret alkot V-ben.

Bizonyítás: A tétel bizonyításához egy lemmát fogunk belátni, melyből már szinte azonnal következni fog a tétel:

2.7 lemma: Bármely $x, y \in \mathcal{H}$ esetén $2x + 2y \in \mathcal{H}$.

Bizonyítás: Legyen $x, y \in \mathcal{H}$ és $z = 2x + 2y$. Ha $x = y$ akkor az állítás semmitmondó, ellenkező esetben pedig a lemmából tudjuk, hogy $d(x, y) = 3$, és az általánosság megszorítása nélkül feltehetjük, hogy a két vektor az első koordinátájában egyezik meg, azaz $x_1 = y_1 = c$. Ekkor z első koordinátája is c, mivel $z_1 = 2x_1 + 2y_1 = 4c = c$. $\mathcal{H}$-ban azonban x-en és y-on kívül pontosan egy olyan vektor van, melynek az első koordinátája c, és ezt a vektort x és y egyértelműen meghatározza. Ez azért igaz, mert ha $i > 1$, akkor z_i különbözik x_i-től és y_i-től, ugyanakkor $x_i \neq y_i$, vagyis z_i-re már csak egy lehetőség adódik. Ez a lehetőség pedig éppen $2x_i + 2y_i$ lesz, hiszen ha $\{x_i, y_i\} = \{0, 1\}$ akkor $z_i = 2 \cdot 0 + 2 \cdot 1 = 2$, ha $\{x_i, y_i\} = \{0, 2\}$ akkor $z_i = 2 \cdot 0 + 2 \cdot 2 = 1$ végül ha $\{x_i, y_i\} = \{1, 2\}$ akkor $z_i = 2 \cdot 1 + 2 \cdot 2 = 0$, ami éppen az amit szerettünk volna. ♣

Egy vektortér részhalmaza pontosan akkor altér, ha tartalmazza bármely elemének a skalárszorosát, illetve bármely két elemének az összegét.
A skalárszoros nyilvánvaló, hiszen ha $a \in \mathcal{H}$, akkor $0 \cdot a$ és $1 \cdot a$ benne van $\mathcal{H}$-ban, $2a \in \mathcal{H}$ pedig az előző állításból következik $x = a$, $y = 0$ választással. Ha $a, b \in \mathcal{H}$, akkor az állításunkat alkalmazva $2a + 2b \in \mathcal{H}$, amiből viszont $2 \cdot (2a + 2b)$ azaz a két vektor összege is benne van $\mathcal{H}$-ban. ♣

2.8 megjegyzés: A lemma bizonyításában nem használtuk fel, hogy $\mathcal{H}$ tartalmazza az origót, így az minden jó halmazra teljesül, ez az állítás egyébként azzal ekvivalens, hogy minden jó halmaz affin altér. Egy halmaz akkor affin altér, ha bármely két pontjának összekötő egyenesét tartalmazza, a háromelemű test felett pedig egy egyenesnek három pontja van, és az x-et és y-t összekötő egyenes harmadik pontja éppen $2x + 2y$.

A következő lépésben meghatározzuk, hogy hány jó altér van $AG(4, 3)$-ban. Egy altér megadható a két bázisvektorával, legyen ez u és v, koordinátákkal leírva: $u = (u_1, u_2, u_3, u_4)$ illetve $v = (v_1, v_2, v_3, v_4)$. Úgy kell megválasztanunk u-t és v-t, hogy a generált altér bármely két különböző x, y vektorára teljesüljön, hogy $d(x, y) \geq 3$. Ez azzal ekvivalens, hogy $d(x - y, 0) \geq 3$, vagyis minden nemnulla vektornak legfeljebb egy nulla koordinátája lehet.

Felírva egy generált vektor általános alakját az a tény, hogy van két nulla eleme azt jelentené, hogy az

$$\alpha u_i + \beta v_i = 0,$$
$$\alpha u_j + \beta v_j = 0,$$

egyenletrendszernek van nemtriviális megoldása, vagyis a determinánsa nulla. Ez azzal ekvivalens, hogy $\frac{u_i}{v_i} = \frac{u_j}{v_j}$. A feltételek teljesülése tehát azt jelenti, hogy olyan bázisvektorokat kell keresnünk, amelyre az $\frac{u_i}{v_i}$ arányok különbözőek, így köztük mind a négy lehetséges érték $0, 1, 2, \infty$ előfordul. Ebből a következtetésből már felírhatjuk a szóbajövő jó altereket:

2.9 tétel: Összesen nyolc darab jó altér található $AG(4, 3)$-ban.

BIZONYÍTÁS: Az általánosság megszorítása nélkül feltehetjük, hogy $u = (0, 1, u_3, u_4)$, illetve $v = (1, 0, v_3, v_4)$ ahol a hiányzó négy koordináta egyike sem nulla, hiszen ezek linárisan függetlenek, és kell hogy legyenek ilyen alakú vektorok az altérben. Mivel a négyféle arány közül ekkor a 0 és a ∞ már megvan, a fenti következtetés miatt a másik két koordináta közül az egyikben két azonos, a másikban pedig két különböző érték áll a két vektorban (az előbbi esetben 1, az utóbbiban 2 adódik arányként). Ez pedig azt jelenti, hogy valóban nyolc altér van, hiszen u_3, u_4 és v_3 értéke szabadon választható az $\{1, 2\}$ halmazból, v_4 pedig meghatározott azáltal, hogy u_3 és v_3 különböznek vagy egyformák. ♣

Az egyszerűbb hivatkozás érdekében először írjuk fel ezt a nyolc lehetséges alteret:

$$\begin{aligned}
&1 : \langle(0, 1, 1, 1), (1, 0, 1, 2)\rangle, && 2 : \langle(0, 1, 1, 1), (1, 0, 2, 1)\rangle, \\
&3 : \langle(0, 1, 1, 2), (1, 0, 1, 1)\rangle, && 4 : \langle(0, 1, 1, 2), (1, 0, 2, 2)\rangle, \\
&5 : \langle(0, 1, 2, 1), (1, 0, 1, 1)\rangle, && 6 : \langle(0, 1, 2, 1), (1, 0, 2, 2)\rangle, \\
&7 : \langle(0, 1, 2, 2), (1, 0, 1, 2)\rangle, && 8 : \langle(0, 1, 2, 2), (1, 0, 2, 1)\rangle.
\end{aligned}$$

A kilenc szám meghatároz kilenc jó halmazt, mint láttuk ezek mindegyike szükségszerűen affin altér, melyekhez tartozó lineáris alterek legyenek $\mathcal{H}_1, \mathcal{H}_2, \ldots, \mathcal{H}_9$. Nyilvánvalóan az egy lehetőség, hogy mindegyik $\mathcal{H}_i$ altér azonos, és a teret valójában egy lineáris altér eltoltjaival fedjük le. Mivel összesen nyolc lehetséges alterünk van, ez **nyolc lehetséges kitöltése a sudokunak**. Minden altérhez egy kitöltés tartozik, hiszen a számok permutációjával egymásba vihető kitöltéseket nem tekintjük különbözőnek.

Nézzük most azt az esetet, amikor különböző alterek eltoltjait is használjuk a lefedéshez, vagyis van olyan i és j, hogy $\mathcal{H}_i \neq \mathcal{H}_j$. Ekkor $\mathcal{H}_i \cap \mathcal{H}_j$ vagy csak az origó, vagy egy egyenes. Ha csak az origó lenne a metszéspont, az azt jelentené, hogy a két altér direkt összege kiadja az egész teret. Ekkor viszont bármely két eltolt (legyenek ezek $\mathcal{H}_i + v_i$ és $\mathcal{H}_j + v_j$) is metszi egymást, hiszen a feltétel szerint van olyan $u_i \in \mathcal{H}_i$ és $u_j \in \mathcal{H}_j$, melyre $u_i - u_j = v_j - v_i$, ekkor viszont $u_i + v_i = u_j + v_j$, és ez a vektor benne van mindkét eltoltban. Mi viszont olyan altereket keresünk, amelynek vannak diszjunkt eltoltjai, így ez az eset nem állhat fent, tehát ha használunk is különböző altereket, akkor azok páronként egy egyenesben metszik egymást.

Végignézve a szóbajövő altérpárokat (28 pár van) az alábbiak metszik egyenesben a másikat:

$$1-2,\ 1-3,\ 1-6,\ 1-7,\ 2-4,\ 2-5,\ 2-8,\ 3-4,$$
$$3-5,\ 3-8,\ 4-6,\ 4-7,\ 5-6,\ 5-7,\ 6-8,\ 7-8.$$

Ha egy nyolccsúcsú gráfként akarjuk ábrázolni a metszéseket (két csúcs össze van kötve, ha a megfelelő alterek metszete egyenes) akkor egy olyan teljes páros gráfot kapunk, amelyben az egyik osztály a $\{1, 4, 5, 7\}$ a másik pedig a $\{2, 3, 6, 8\}$. Ez egyrészt azt jelenti, hogy minden altér négy másikat metsz egyenesben, másrészt mivel páros gráfban nincsen háromszög, három altér sem választható ki úgy, hogy bármelyik kettő egyenesben messe egymást. Így csak az olyan sudoku kitöltéseket kell megvizsgálni, ahol a jó halmazok két különböző lineáris altér eltoltjai.

Az alterekből összesen $\frac{8 \cdot 4}{2} = 16$ megfelelő pár képezhető, már csak azt kell meghatároznunk, hogy egy adott $\mathcal{H}_1, \mathcal{H}_2$ pár eltoltjaival hányféle különböző módon lehet lefedni a teret. Legyen a $\mathcal{H}_1$-et és $\mathcal{H}_2$-t tartalmazó hipersík $\mathcal{J}_0$, (ilyen van, mivel a két altér uniója egy háromdimenziós alteret feszít) két eltoltja pedig $\mathcal{J}_1$ és $\mathcal{J}_2$, ahol a három (affin) hipersík lefedi a teret. Tudjuk, hogy $\mathcal{H}_1$ és $\mathcal{H}_2$ eltoltjai mindig teljes egészében benne vannak valamelyik $\mathcal{J}_i$-ben, ami azt jelenti, hogy egy $\mathcal{J}_i$ hipersíkot kizárólag az egyik altér eltoltjaival fedhetünk le, mert két különböző altér eltoltja (ha ugyanabban az affin hipersíkban van) szintén egyenesben metszi egymást.

Ott tartunk tehát, hogy mind a három hipersík esetében el kell döntenünk, hogy melyik altér eltoltjaival fedjük le. Ez elvileg nyolc lehetőség, de az a két eset amikor mindhárom hipersík esetén ugyanazt az alteret választottuk nem számít, mert azt már leszámoltuk a fenti nyolc kitöltés valamelyikében. Ez pedig azt jelenti, hogy összesen $16 \cdot 6 = 96$ olyan kitöltés van, amikor a kilenc jó halmaz két különböző altér eltoltjaiként áll elő, vagyis összesen:

$8 + 96 = 104$ különböző szimmetrikus kitöltése van a klasszikus sudoku táblázatnak.

2.2. Páronként merőleges Sudoku táblázatok

Ebben a pontban egy új fogalmat tárgyalunk. Ha veszünk két kitöltött sudoku táblázatot, akkor azokat képzeletben egymásra téve a 81 mező 81 rendezett számpárt határoz meg. 1-től 9-ig a számokból ugyanennyi különböző pár képezhető, így felmerül a kérdés, hogy lehetséges-e az, hogy minden lehetséges számpár pontosan egyszer szerepel a két táblázat által meghatározott párok között. A kérdés persze nem csak sudokura, hanem már latin négyzetek esetén is értelmes és érdekes, és mivel a válasz igen, így az ilyen kitöltésekre külön elnevezést vezetünk be.

2.10 definíció: Két $n \times n$-es latin négyzetet (legyenek ezek L_1 és L_2) egymásra merőlegesnek nevezünk, ha minden (i, j) pár esetén egyértelműen létezik olyan mező, melyen L_1-ben i, L_2-ben pedig j áll.

Milyen n-re létezik két egymásra merőleges $n \times n$-es latin négyzet? Létezik-e esetlegesen több, páronként merőleges latin négyzetek rendszere is? Ha igen, akkor legfeljebb hány ilyen négyzet van? Ezeket a kérdéseket fogjuk most megválaszolni a sudokunak mind a sima, mind pedig a szimmetrikus kitöltései esetén, először azonban lássunk pár egyszerűbb állítást, illetve felső becslést az ilyen ortogonális rendszerekről illetve a maximális elemszámukról.

2.11 állítás: Ha két latin négyzet egymásra merőleges, akkor ha valamelyikükben (akár mind a kettőben) megpermutáljuk a szimbólumokat, továbbra is merőslegesek maradnak.

2.12 következmény: Páronként merőleges $n \times n$-es latin négyzetek rendszerének maximális elemszáma $n - 1$.

Bizonyítás: Az előző állításból következik, ha van egy ilyen rendszer, akkor feltehető, hogy mindegyik latin négyzet felső sorában sorban vannak a számok 1-től n-ig. Ez azt jelenti, hogy az összes azonos (x, x) alakú párt már „kipipáltuk" és így a második sor első mezőjén a különböző négyzetekben csak különböző számok állhatnak, vagyis a négyzetek száma legfeljebb n. Mivel azonban ezen a mezőn nem állhat 1-es a bal felső elem miatt, így meg is kaptuk az $n - 1$-es felső becslést. ♣

Ezt a gondolatot már konkrétan alkalmazhatjuk merőleges Sudoku rendszerek maximális elemszámának a becslésére:

2.13 tétel: A páronként merőleges sudoku táblázatok maximális száma:
a) klasszikus esetben legfeljebb hat.
b) szimmetrikus esetben legfeljebb négy.

Bizonyítás: A bizonyítás lényegében ugyanaz mint az előző következményé. Ha feltesszük, hogy a legfelső sorban sorban mennek a számok, akkor a második sor első elemei különböznek, továbbá a klasszikus esetben $1, 2, 3$, szimmetrikus esetben pedig 4 és 7 sem állhat ezen a mezőn. Ebből következik, hogy ha van egy merőleges rendszerünk, akkor az az első esetben legfeljebb hat, a másodikban pedig legfeljebb négy táblázatból állhat. ♣

A továbbiakban azt fogjuk igazolni, hogy mind a két becslés éles. Ezt persze megtehetnénk úgy, hogy mutatunk mind a két esetben hat illetve négy egymásra merőleges táblázatot és készen vagyunk, mi azonban most a sudoku táblázatunknak a három dimenziós projekív térrel való kapcsolatát, és az első fejezetben ismertett tételeket szeretnénk használni a konstrukcióink létrehozására.

Tekintsük először a klasszikus esetet. Hat olyan táblázatot fogunk mutatni, ahol az egyes szimbólumok által meghatározott halmazok affin alteret alkotnak, és mindegyik altér ugyanannak a lineáris altérnek az eltoltja (az egyszerűség kedvéért az ilyen táblázatokat lineárisnak fogjuk nevezni). Milyen feltételt kell kielégítenie ennek az altérnek, hogy az általa meghatározott lineáris kitöltés jó legyen? Tudjuk, hogy mind a három partíció egy-egy altérnek az eltoltjai, így az kell nekünk, hogy a kitöltést és a három partíciót meghatározó altér csak az origóban messe egymást. Ekkor ugyanis bármely két eltolt is egy pontban metszi egymást, és így minden szimbólum minden osztályban pontosan egyszer szerepel.

Vegyük észre továbbá, hogy ha két ilyen táblázatunk van, akkor a merőlegesség feltétele is egyszerűen megfogalmazható, hiszen az azzal ekvivalens, hogy a két altér csak az origóban metszi egymást. Ekkor mint láttuk bármely két eltolt ismételten egy pontban metszi egymást, és mivel minden rendezett párhoz tartozik egy affin altér pár, így egy egyértelmű mező is, ahol ők metszik egymást, tehát minden rendezett pár pontosan egyszer fog szerepelni.

Legyen tehát a sorokat, az oszlopokat és a blokkokat meghatározó három altér $\mathcal{P}_1, \mathcal{P}_2$ és $\mathcal{P}_3$. Ekkor a fenti gondolatmenet miatt hat olyan alteret keresünk, amelyek egymást és a $\mathcal{P}_i$ altereket is csak az origóban metszik. Az egész eddigi vizsgálatunk során úgy tekintettünk a sudokura mint egy affin geometriára, most viszont a továbbiakban a vektortérből származatott projektív geometriát fogjuk vizsgálni, és az ezzel kapcsolatban ismertetett tételeket használjuk annak érdekében hogy megmutassuk: a hat éles becslés.

A vektortér két dimenziós alterei a projektív geometriában egyeneseknek, az egy dimenziósak pedig pontoknak felelnek meg. Ha két altér csak az origóban metszi egymást, akkor azok projektív értelemben diszjunkt egyenesek. Legyenek a $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3$ altereknek megfelelő egyenesek L_1, L_2 és L_3. Ezek közül L_1 és L_2 diszjunktak, míg L_3 egy-egy pontban metszi L_1-et és L_2-t. Az tehát a feladatunk, hogy mutassunk a projektív térben hat, egymástól illetve az L_i egyenesektől diszjunkt egyeneshalmazt.

Ha találunk a projektív térnek egy olyan egyenesekkel történő befedését, amelynek eleme L_1 és L_2 akkor készen lennénk, hiszen egy ilyen befedés egyrészről 10 egyenesből áll, másrészről mivel diszjunktak így L_3 pontosan négyet metsz közülük (amelyek közül az egyik L_1 a másik pedig L_2) vagyis a maradék hat egyenes pont megfelel a követelményeknek.

Az első fejezetben ismertett konstrukciót fogjuk alkalmazni és konkrétan végigszámolni a táblázatok megkonstruálása érdekében. A fedés két hiperbolikus felületből és a kimaradó két egyenesből fog állni úgy, hogy az egyik felület teljes egészében tartalmazza az L_i egyeneseket. Ekkor a keresett hat egyenes a másik felület egyik fedéséből és a két kimaradó egyenesből fog állni.

Első lépésben határozzuk meg azoknak a másodrendű felületeknek a mátrixát (egyenletét) amelyek tartalmazzák a három kitüntetett egyenesünket. Vegyük észre, hogy a konstrukciónkban a végső hat egyenes nem csak a három L_i, hanem a hiperboloidon fekvő mind a nyolc egyenestől diszjunkt lesz, így a rendszerünk az azok által meghatározott partícióra is design lesz. Mivel azonban három dimenziós projektív térben egy másodrendű felületet kilenc pont határoz meg (egy 4×4-es szimmetrikus mátrixban 10 elem választható meg szabadon, de mivel a skalárszorosok ugyanazt a felületet adják így valójában csak kilenc szabadsági fokunk van) és ez a három egyes hét pont előírásával is rajta lenne a felületen, így a kapott egyenletben lesznek szabadon választható paraméterek. Ezeket a számolás végén arra fogjuk használni, hogy a hiperboloidunk lévő másik öt egyenes között is legyenek „szépek”.

Egy másodrendű felület általános egyenlete:

$$ax_1^2 + bx_2^2 + cx_3^2 + dx_4^2 + ex_1x_2 + fx_1x_3 + gx_1x_4 + hx_2x_3 + ix_2x_4 + jx_3x_4 = 0.$$

Mivel a felület tartalmazza az L_1 egyenest, így a $(0,0,0,1), (0,0,1,0), (0,0,1,1)$ pontok kielégítik az egyenletet, amiből $c = d = j = 0$. Hasonlóképpen L_2 miatt $a = b = e = 0$, L_3-ból pedig $i = 0$ következik, vagyis a keresett felületek

$$fx_1x_3 + gx_1x_4 + hx_2x_3$$

alakúak. Ha ezen belül még f-et is nullának választjuk, akkor a partíciókat meghatározó $x_2 = x_4 = 0$ egyenes is rajta lesz a felületen. Ha nem elfajuló másodrendű felületet szeretnénk kapni, akkor innen már csak két választási lehetőségünk van, vagy $g = h = 1$, vagy $g = 1$, $h = 2$. Ha mi a második lehetőséget választjuk, akkor a felület tartalmazni fogja még az alábbi négy egyenest:

$$\begin{aligned} x_1 - x_3 = x_2 - x_4 = 0, &\qquad x_1 + x_3 = x_2 + x_4 = 0, \\ x_1 - x_2 = x_3 - x_4 = 0, &\qquad x_1 + x_2 = x_3 + x_4 = 0. \end{aligned}$$

amelyből igazándiból az első kettő az érdekes, mert az éppen a táblázat főátlóját illetve mellékátlóját határozza meg. Jelöljük ezt a felületet $\mathcal{G}_2$-vel, a hozzá tartozó mátrix legyen G_2.

Szeretnénk meghatározni a fedés többi egyenesét. Ezt úgy fogjuk megtenni, hogy vesszük az első fejezetben megkonstruált fedését a térnek, az egyszerűség kedvéért legyen $f(x,y) = x^2 + y^2$ (ez meghatározza az $\mathcal{F}_1$ és $\mathcal{F}_2$ felületeket, mátrixuk legyen F_1 és F_2) és keresünk egy olyan M lineáris transzformációt, amely $\mathcal{F}_2$-t átviszi G_2-be. Ekkor ez a transzformáció $\mathcal{F}_1$-et is átviszi egy $\mathcal{G}_2$-től diszjunkt felületbe (jelölje ezt $\mathcal{G}_1$), és az egyenesek képe is egyenes lesz. Az egyszerűség kedvéért az inverztranszformációt határozzuk meg, abból ugyanis az első fejezetben ismertett képlettel megkapjuk $\mathcal{G}_1$ mátrixát. Egy ilyen megfelelő mátrix például:

$$M^{-1} = \begin{pmatrix} 1 & 2 & 2 & 0 \\ 1 & 1 & 0 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 1 & 0 & 2 \end{pmatrix}, \qquad M = \begin{pmatrix} 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \end{pmatrix},$$

Ellenőrizhető, hogy $\left(M^{-1}\right)^{\top} F_2 M^{-1} \sim G_2$. Innentől kezdve pedig a konkrét példa hat merőleges sudoku táblázat előállítására már csak számolás, az elméletből tudjuk, hogy a becslés éles. A $\mathcal{G}_1$ felület mátrixa:

$$G_1 = \left(M^{-1}\right)^\top F_1 M^{-1} = \left(M^{-1}\right)^\top M^{-1} = \begin{pmatrix} 1&0&1&0\\ 0&1&0&1\\ 1&0&2&0\\ 0&1&0&2 \end{pmatrix},$$

vagyis az egyenlete:

$$x_1^2 + x_2^2 - x_3^2 - x_4^2 - x_1x_3 - x_2x_4 = 0.$$

Az $\mathcal{F}_1$ felületen lévő egyik egyenessereg:

$$\begin{aligned}
F_{11}\colon&\quad x_2 + x_3 + x_4 = 0, & x_2 &= x_1 + x_3,\\
F_{12}\colon&\quad x_1 = x_3 + x_4, & x_3 &= x_2 + x_4,\\
F_{13}\colon&\quad x_1 = x_2 + x_3, & x_2 &= x_3 + x_4,\\
F_{14}\colon&\quad x_3 = x_1 + x_2, & x_4 &= x_2 + x_3.
\end{aligned}$$

Ha ezekre az egyenesekre alkalmazzuk az M transzformációt, akkor a $\mathcal{G}_1$ felület egyik egyenesseregét kapjuk. Egyenes képét úgy tudjuk a legegyszerűbben megkapni, hogy azon síkoknak a normálvektorainak a képét határozzuk meg, amelyek metszeteként az egyenes előáll. Az F_{11} egyeneshez tartozó két normálvektor a $(0, 1, 1, 1)$ és az $(1, 2, 1, 0)$, ezeknek a képei pedig az $(1, 0, 0, 2)$ és a $(0, 2, 1, 1)$ vektorok. Ennek alapján a keresett egyenessereg:

$$\begin{aligned}
G_{11}\colon&\quad x_1 = x_4, & x_2 &= x_3 + x_4,\\
G_{12}\colon&\quad x_4 = x_1 + x_2 + x_3, & x_2 &= x_3,\\
G_{13}\colon&\quad x_2 = x_1 + x_3 + x_4, & x_1 + x_2 + x_4 &= 0,\\
G_{14}\colon&\quad x_1 + x_2 + x_3 = 0, & x_1 &= x_2 + x_3 + x_4.
\end{aligned}$$

Az eredeti fedésben kimaradó két egyenes ($x_1 = x_2 = 0$ és $x_3 + x_4 = 0$) képeinek egyenlete pedig:

$$\begin{aligned}
E_1\colon&\quad x_2 + x_3 + x_4 = 0, & x_4 &= x_1 + x_3,\\
E_2\colon&\quad x_3 = x_2 + x_4, & x_1 + x_3 + x_4 &= 0.
\end{aligned}$$

Meghatároztunk tehát hat alteret amelyekből képezett lineáris kitöltések merőleges rendszert alkotnak. Annak érdekében, hogy ez a konstrukció kézzelfogható legyen, írjuk is fel ezt a hat táblázatot:

1	2	3	4	5	6	7	8	9
7	8	9	1	2	3	4	5	6
4	5	6	7	8	9	1	2	3
6	4	5	9	7	8	3	1	2
3	1	2	6	4	5	9	7	8
9	7	8	3	1	2	6	4	5
8	9	7	2	3	1	5	6	4
5	6	4	8	9	7	2	3	1
2	3	1	5	6	4	8	9	7

G_{11}

1	2	3	4	5	6	7	8	9
8	9	7	2	3	1	5	6	4
6	4	5	8	7	9	3	1	2
3	1	2	6	4	5	8	7	9
7	8	9	1	2	3	4	5	6
5	6	4	8	9	7	2	3	1
2	3	1	5	6	4	8	9	7
9	7	8	3	1	2	6	4	5
4	5	6	7	8	9	1	2	3

G_{12}

1	2	3	4	5	6	7	8	9
5	6	4	8	9	7	2	3	1
9	7	8	3	1	2	6	4	5
2	3	1	5	6	4	8	9	7
6	4	5	9	7	8	3	1	2
7	8	9	1	2	3	4	5	6
3	1	2	6	4	5	9	7	8
4	5	6	7	8	9	1	2	3
8	9	7	2	3	1	5	6	4

G_{13}

1	2	3	4	5	6	7	8	9
4	5	6	7	8	9	1	2	3
7	8	9	1	2	3	4	5	6
5	6	4	8	9	7	2	3	1
8	9	7	2	3	1	5	6	4
2	3	1	5	6	4	8	9	7
9	7	8	3	1	2	6	4	5
3	1	2	6	4	5	9	7	8
6	4	5	9	7	8	3	1	2

G_{14}

1	2	3	4	5	6	7	8	9
9	7	8	3	1	2	6	4	5
5	6	4	8	9	7	2	3	1
8	9	7	2	3	1	5	6	4
4	5	6	7	8	9	1	2	3
3	1	2	6	4	5	9	7	8
6	4	5	9	7	8	3	1	2
2	3	1	5	6	4	8	9	7
7	8	9	1	2	3	4	5	6

E_1

1	2	3	4	5	6	7	8	9
6	4	5	9	7	8	3	1	2
8	9	7	2	3	1	5	6	4
9	7	8	3	1	2	6	4	5
2	3	1	5	6	4	8	9	7
4	5	6	7	8	9	1	2	3
5	6	4	8	9	7	2	3	1
7	8	9	1	2	3	4	5	6
3	1	2	6	4	5	9	7	8

E_2

Ezzel tehát megkonstruáltunk hat páronként merőleges klasszikus sudoku táblázatot, amelyre még az is teljesül, hogy minden táblázatban a pozíciók által alkotott partíciókban, és a fő-, illetve mellékátlóban is minden szám pontosan egyszer szerepel. Természetesen több merőleges rendszer létezik, egy másik megoldást kapunk például ha a $\mathcal{G}_1$ felületről a másik egyenessereget választjuk.

Térjünk most rá a szimmetrikus esetre. A konstrukció itt is lineáris kitöltésekből fog állni, viszont a korábbi eredményeinket felhasználva sokkal könnyebb dolgunk van az élesség igazolására. Négy olyan alteret keresünk, amelyek páronként csak az origóban metszik egymást, viszont azt láttuk, hogy a nyolc lehetséges altér (amelyek lineáris megoldást határoznak meg) éppen két olyan négyes csoportot alkot, amelyekre a feltétel teljesül, így rögtön két rendszert is kaptunk. Végignézve a kitöltéseket kapjuk, hogy erre a rendszerre is teljesül, hogy minden fő-, és mellékátlóban minden szám pontosan egyszer szerepel.

Ebben az esetben érdekességképpen belátható az is, hogy más megoldás nem lehetséges. Tegyük fel ugyanis, hogy a rendszerben van egy A kitöltés, amelyik két alteret használ, eredményeink szerint a két négyes csoportból egyet-egyet: $\mathcal{A}_1$-et és $\mathcal{A}_2$-t. Ha veszünk egy tetszőleges B kitöltést, és a hozzá felhasznált $\mathcal{B}$ alteret (ha kettőt használtunk fel, akkor az egyiket) akkor teljesül, hogy $\mathcal{B}$ valamelyik $\mathcal{A}_i$-t egyenesben metszi. Ekkor azonban a kitöltésben szereplő eltoltak is vagy diszjunktak, vagy egyenesben metszik egymást, de egyik esetben sem szerepel a hozzájuk tartozó számpár pontosan egyszer a táblázatok egymásra rakásakor, ami ellentmondás. Ha pedig a rendszer összes táblázata lineáris akkor csak a már ismertetett két megoldás lehetséges.

Ezzel egyébként azt is igazoltuk, hogy ha egy szimmetrikus táblázat nem lineáris, hanem két alteret használ, akkor ő semmelyik másik kitöltésre sem lehet merőleges, vagyis ha egy merőleges rendszer legalább kételemű, akkor az valamelyik négyelemű rendszer részhalmaza.

3. Az 5×5-ös Sudoku

3.1. Szimmetrikus megoldások 5×5-ös esetben

A 3×3 esetben ismertett eredményeket követően azt vizsgáljuk meg, hogyan menthető át az alapgondolat nagyobb méretű sudoku táblázatok „szép kitöltéseinek” tárgyalására. Vizsgálatunk tárgya az 5×5-ös sudoku lesz, ahová 1-től 25-ig kell beírnunk a számokat. Először azt kell definiálnunk, hogy mik lesznek azok a partíciók az ábrában, amelyekre megköveteljük, hogy minden szám minden osztályban pontosan egyszer forduljon elő. A korábbi hat követelményt (sor, oszlop, blokk, pozíció, törött sor, törött oszlop) értelemszerűen fenntartjuk, viszont a táblázat méretének növekedése miatt ez még nem tartja kordában a kitöltések számát, és nem is ad szép leírhatóságot az origót tartalmazó jó halmazokra. Természetes módon teljesül, hogy bármely két vektor Hamming-távolsága legalább három (és így bármely nem nulla vektornak legfeljebb egy nulla koordinátája van) de az nem igaz, hogy bármely két vektorpárnak a távolsága három lenne, illetve hogy ezen igények mellett minden jó halmaz alteret alkotna.

Ha veszünk egy blokkot, illetve a benne lévő fő/mellékátlót, akkor az egy nagy oszlopban (illetve sorban) lévő öt ilyen átló is egy szép szimmetrikus halmazt alkot. Az átlókat a blokkokon belül eltoltva egy újabb partíciót kapunk (nevezhetjük őket *törött átlóknak*). Ezeket könnyedén leírhatjuk a bevezetett koordinátákra felírt egyenletekkel: $x_1 = \alpha$, $x_2 \pm x_4 = \beta$, illetve $x_3 = \alpha$, $x_2 \pm x_4 = \beta$.
A változók szimmetriája érdekében megköveteljük még, hogy az $x_2 = \gamma$, $x_1 \pm x_3 = \delta$, illetve $x_4 = \gamma$, $x_1 \pm x_3 = \delta$ egyenletekkel leírható partíciókra is teljesüljenek a feltételek. Az ilyen típusú halmazokat nevezhetjük *eltolt törött soroknak/oszlopoknak*. Ilyen megszorításokkal igaz a korábbi eredményünk:

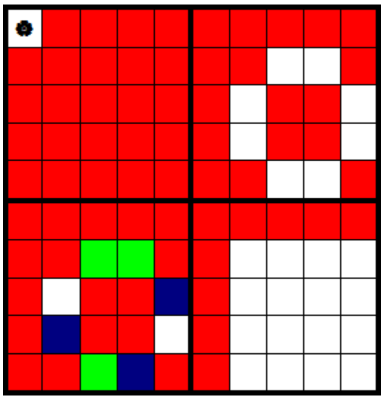
3.1 tétel: Minden origót tartalmazó jó halmaz szükségszerűen lineáris alteret alkot.

BIZONYÍTÁS: A koordináták segítségével tudunk úgy beszélni halmazokról, hogy a 01-es blokk, illetve a 02-es oszlop, ahol értelemszerűen az első az első nagy sor második blokkját, a második az első nagy oszlop harmadik oszlopát jelöli (a $(0, y, 1, w)$ és az $(x, y, 0, 2)$ alakú elemek halmaza).

A bizonyítás indirekt, feltesszük, hogy van olyan $\mathcal{S}$ jó halmaz, amely nem alkot alteret. Vegyük észre, hogy ha a halmaz összes vektorára alkalmazzuk az

$$\mathcal{A}_\alpha : (x, y, z, w) \to (x, \alpha y, z, \alpha w)$$

transzformációt, akkor $\alpha \neq 0$ esetén továbbra is egy jó halmazt kapunk. Ez azért igaz, mert ez a transzformáció $GF(5)^4$-nek egy automorfizmusa amely a partíciókat fixen hagyja (az osztályokat egymás között mozgatja), továbbá ha az eredeti jó halmazunk nem volt altér, akkor a transzformáltja sem lesz altér. Mivel minden blokk a jó halmaznak pontosan egy elemét tartalmazza, így a halmazunkban van $(0, y, 1, w)$ alakú vektor (ahol $y \neq 0$), így a transzformáció segítségével feltehető, hogy van olyan w, melyre $(0, 1, 1, w) \in \mathcal{S}$.



Tekintsük most a fenti ábrát. Ez a sudoku bal felső 2×2 blokkját ábrázolja, ahol a $(0, 0, 0, 0)$-ás mezőn lévő pötty azt mutatja, hogy az szükségszerűen benne van a halmazunkban. A piros hátterű mezők azok, amelyek a követelmények miatt nem lehetnek benne egy origót tartalmazó jó halmazban. Alkalmazva az előző gondolatot feltehető, hogy a $(0, 1, 1, 2)$ illetve a $(0, 1, 1, 3)$ mezők valamelyike eleme $\mathcal{S}$-nek. Nézzük meg, hogy az 10 blokk melyik mezője lehet benne a halmazunkban az egyik, illetve másik esetben.

Ha $(0, 1, 1, 2) \in \mathcal{S}$ akkor a zöld színű mezők alapból nem jönnek számításba. Vegyük azonban észre, hogy a kék színű mező sem lehet a halmazban, mert ha valamelyik kék színű mező benne lenne, akkor az egész 02-es oszlop kiesik, ami ellentmondás.

Ezt könnyen meggondolhatjuk, hiszen az origó miatt ebből az oszlopból csak az $(x, 1, 0, 2)$ és $(x, 4, 0, 2)$ alakú 8 mező jönne számításba, amelyből $(0, 1, 1, 2)$ kizárja az első, a kék mezők pedig kizárják a második négyet. Ha tehát $(0, 1, 1, 2) \in \mathcal{S}$ akkor azt kaptuk, hogy csak az $(1, 2, 0, 1)$-et vagy $(1, 3, 0, 4)$-et választhatjuk az 10 blokkból. A másik esetben ugyanezzel a gondolatmenettel kapjuk, hogy $(0, 1, 1, 3)$ párja $(1, 3, 0, 1)$ illetve $(1, 2, 0, 4)$ lehet.

Ezen négy esetben pedig teljesül, hogy a feltételeket figyelembe véve már csak egyféle olyan $\mathcal{S}$ halmaz van, amely tartalmazza a két megfelelő vektort, méghozzá a két vektor által kifeszített altér. Egy lépésben kihúzzuk azokat a mezőket, amelyek már nem lehetnek benne $\mathcal{S}$-ban, és megnézzük, hogy van-e olyan blokk, ahonnan még nem választottunk mezőt, de erre már csak egy lehetőség van. Ha vannak ilyenek, akkor ezeket a mezőket mind beválasztjuk a halmazba, és megismételjük a lépést. Ez az algoritmus mind a négy esetben tökéletesen egyformán működik (abban az értelemben, hogy a blokkokat tekintve ugyanabban a sorrendben tölthető ki a táblázat) mi az első esetben ismertetjük ezeket a lépéseket.
Tegyük fel tehát, hogy $(0, 0, 0, 0), (0, 1, 1, 2), (1, 2, 0, 1) \in \mathcal{S}$.

1) $04 \to (0, 4, 4, 3)$, $40 \to (4, 3, 0, 4)$, $11 \to (1, 3, 1, 3)$.

2) $02 \to (0, 2, 2, 4)$, $20 \to (2, 4, 0, 2)$, $12 \to (1, 4, 2, 0)$, $13 \to (1, 0, 3, 2)$, $14 \to (1, 1, 4, 4)$, $21 \to (2, 0, 1, 4)$, $31 \to (3, 2, 1, 0)$, $41 \to (4, 4, 1, 1)$, $24 \to (2, 3, 4, 0)$, $42 \to (4, 0, 2, 3)$, $33 \to (3, 4, 3, 4)$, $34 \to (3, 0, 4, 1)$, $43 \to (4, 1, 3, 0)$ $44 \to (4, 2, 4, 2)$.

3) $03 \to (0, 3, 3, 1)$, $30 \to (3, 1, 0, 3)$, $22 \to (2, 1, 2, 1)$, $23 \to (2, 2, 3, 3)$, $32 \to (3, 3, 2, 2)$.

Ami valóban éppen a kifeszített alteret adja. Ezzel az állításunkat beláttuk. ♣

A tételünk fényében már tudjuk, hogy minden jó $\mathcal{S}$ halmaz szükségszerűen altér, meg kell határoznunk, hogy hány olyan altér van, amely kielégíti a követelményeket. Legyen a két generátorvektor $(0, 1, a_1, a_2)$ és $(1, 0, b_1, b_2)$. Ekkor a generált altér a $(\beta, \alpha, \alpha a_1 + \beta b_1, \alpha a_2 + \beta b_2)$ alakú vektorok halmaza, ahol $\alpha, \beta \in GF(5)$. Az a tény, hogy $\mathcal{S}$ minden kiválasztott halmazt pontosan egy pontban metsz azt jelenti, hogy az alábbi egyenletrendszereknek egyértelműen létezik megoldása:

$$\beta = c,\ \alpha = d$$
$$\beta = c,\ \alpha a_1 + \beta b_1 = d$$
$$\beta = c,\ \alpha a_2 + \beta b_2 = d$$
$$\alpha = c,\ \alpha a_1 + \beta b_1 = d$$
$$\alpha = c,\ \alpha a_2 + \beta b_2 = d$$
$$\alpha a_1 + \beta b_1 = c,\ \alpha a_2 + \beta b_2 = d$$
$$\beta = c,\ \alpha \pm (\alpha a_2 + \beta b_2) = d$$
$$\alpha a_1 + \beta b_1 = c,\ \alpha \pm (\alpha a_2 + \beta b_2) = d$$
$$\alpha = c,\ \beta \pm (\alpha a_1 + \beta b_1) = d.$$
$$\alpha a_2 + \beta b_2 = c,\ \beta \pm (\alpha a_1 + \beta b_1) = d.$$

Az első feltétel automatikusan teljesül, a következő négy azt jelenti, hogy a_1, a_2, b_1, b_2 koordináták egyike sem nulla. Az utolsó öt egyenletnek ha egyértelműen létezik megoldása, az azzal ekvivalens, hogy az egyenletrendszer együtthatómátrixának determinánsa nem nulla. Ezeket felírva az alábbi feltételek adódnak:

$$a_1 b_2 \neq b_1 a_2$$
$$a_2 \neq 1, 4$$
$$\pm a_1 b_2 \neq b_1(1 \pm a_2)$$
$$b_1 \neq 1, 4$$
$$\pm a_1 b_2 \neq a_2(1 \pm b_1).$$

(a_1, a_2)-re a második feltétel miatt nyolc lehetőség van, hasonlóképpen a negyedik feltétel miatt (b_1, b_2)-re is nyolc lehetőség adódik. Végignézve ezeket az eseteket azt kapjuk, hogy minden (a_1, a_2) párhoz pontosan kettő jó (b_1, b_2) pár tartozik. Ez azt jelenti tehát, hogy összesen $8 \times 2 = 16$ altér alkot jó halmazt. Ezek az alterek:

$$\begin{array}{lll}
1 : \langle(0,1,1,2),(1,0,2,3)\rangle, & 2 : \langle(0,1,1,2),(1,0,3,2)\rangle, & 3 : \langle(0,1,1,3),(1,0,2,2)\rangle \\
4 : \langle(0,1,1,3),(1,0,3,3)\rangle, & 5 : \langle(0,1,2,2),(1,0,2,4)\rangle, & 6 : \langle(0,1,2,2),(1,0,3,1)\rangle \\
7 : \langle(0,1,2,3),(1,0,2,1)\rangle, & 8 : \langle(0,1,2,3),(1,0,3,4)\rangle, & 9 : \langle(0,1,3,2),(1,0,2,1)\rangle \\
10 : \langle(0,1,3,2),(1,0,3,4)\rangle, & 11 : \langle(0,1,3,3),(1,0,2,4)\rangle, & 12 : \langle(0,1,3,3),(1,0,3,1)\rangle \\
13 : \langle(0,1,4,2),(1,0,2,2)\rangle, & 14 : \langle(0,1,4,2),(1,0,3,3)\rangle, & 15 : \langle(0,1,4,3),(1,0,2,3)\rangle
\end{array}$$
$$16 : \langle(0,1,4,3),(1,0,3,2)\rangle.$$

Ettől a ponttól kezdve pedig a megoldás váza változtatás nélkül átmenthető a 3×3-as esetből. Be kell tehát osztanunk a 625 mezőt 25 darab páronként diszjunkt jó halmazba, melyeket jelöljünk $\mathcal{S}_1, \mathcal{S}_2, \ldots \mathcal{S}_{25}$-el. Ezen halmazoknak az origót tartalmazó eltoltjai legyenek $\mathcal{S}'_1, \mathcal{S}'_2, \ldots \mathcal{S}'_{25}$. Egy $\mathcal{S}'_i$ és egy $\mathcal{S}'_j$ halmaz metszete nem lehet csak az origó, mert akkor az eltoltjaik se lennének diszjunktak, tehát a két altér vagy egybeesik, vagy egy egyenesben metszi egymást. Végignézve a jó altereket azt kapjuk, hogy az alábbi párok metszik egymást egyenesben:

$1-2$,	$1-3$,	$1-8$,	$1-12$,	$1-14$,	$1-15$,	$2-4$,	$2-7$,
$2-11$,	$2-13$,	$2-16$,	$3-4$,	$3-6$,	$3-10$,	$3-13$,	$3-16$,
$4-5$,	$4-9$,	$4-14$,	$4-15$,	$5-6$,	$5-7$,	$5-10$,	$5-11$,
$5-16$,	$6-8$,	$6-9$,	$6-12$,	$6-15$,	$7-8$,	$7-9$,	$7-12$,
$7-14$,	$8-10$,	$8-11$,	$8-13$,	$9-10$,	$9-11$,	$9-16$,	$10-12$,
$10-15$,	$11-12$,	$11-14$,	$12-13$,	$13-14$,	$13-15$,	$14-16$,	$15-16$.

Látható, hogy minden altérhez pontosan hat másik olyan altér található, amely őt egy egyenesben metszi, de nincs három olyan, amelyik közül bármely kettő metszete egy egyenes lenne. Ez azt jelenti tehát, hogy a jó halmazaink vagy ugyanannak, vagy két különböző altérnek az eltoltjai.

Az első esetben 16 megfelelő kitöltése van a sudokunak, hiszen mind a 16 jó altérhez pontosan egy kitöltés tartozik. A második esetben a két (egymást egyenesben) metsző alteret $\frac{16 \cdot 6}{2} = 48$ féleképpen választhatjuk ki. A két altérhez egyértelműen van egy őket tartalmazó $\mathcal{H}_0$ hipersík, amely az eltoltjaival $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ és $\mathcal{H}_4$ lefedi az egész teret. Akárcsak a 3×3-as esetnél most is fennáll, hogy minden hipersíkot vagy teljes egészében vagy az egyik, vagy a másik altér eltoltjai fednek le, hiszen a két különböző altér az eltolt hipersíkon is egy egyenesben metszené egymást. Így mind a 48 altérpár esetében eldönthetjük, hogy $1-4$, $2-3$, $3-2$ vagy $4-1$ arányban fedik le az eltoltakat (a $0-5$ és $5-0$ esetekben visszakapnánk az első esetet). Az egyes alesetekben pedig rendre $5, 10, 10$ és 5-féleképpen választhatom ki, hogy melyik altérhez melyik eltoltak tartoznak, amit lefednek. Ez azt jelenti tehát, hogy minden $\mathcal{S}'_i, \mathcal{S}'_j$ esetében $5 + 10 + 10 + 5 = 30$ kitöltés van, vagyis **az összes jó kitöltések száma $\mathbf{16 + 48 \cdot 30 = 1456}$**.

3.2. Páronként merőleges 5×5-ös és nagyobb méretű táblázatok

Ebben a fejezetben először ismertetjük az 5×5-ös esetre vonatkozó konstrukciókat és becsléseket, majd megemlítünk egy nagyobb méretű sudoku rendszerekre vonatkozó tételt.

3.2 tétel: A páronként merőleges 5×5-ös sudoku táblázatok maximális száma:
a) klasszikus esetben legfeljebb 20.
b) szimmetrikus esetben legfeljebb 8.

BIZONYÍTÁS: A bizonyítás ugyanúgy működik, mint a 3×3-as esetben, feltehetjük, hogy az egyes táblázatokban a legfelső sorban sorba mennek a számok. Ekkor a második sor első mezőjére a klasszikus esetben 20, (az $1, 2, 3, 4, 5$ esik ki) míg a szimmetrikus esetben csak nyolc lehetőségünk van: $8, 9, 13, 14, 18, 19, 23, 24$. Az első sor miatt azonban különböző táblázatokban ezen a mezőkön különböző számoknak kell szerepelnie, így a 20 és a 8 felső becslés a merőleges rendszerek elemszámára.
A második részt úgy is igazolhatjuk, ha eszükbe jut, hogy a 01-es blokkban nyolc helyre írhattuk az origón lévő számot. Ha ismételten feltesszük, hogy a felső sorban sorrendben vannak a számok, akkor a 01-es blokkban az 1-est a különböző táblázatokban különböző mezőre kell írjuk, és a 8 ismét adódik felső becslésként. ♣

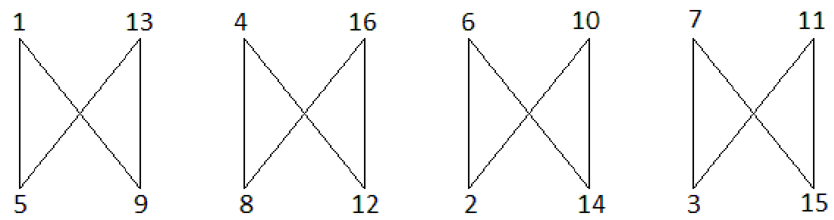
Szerencsénkre ebben az esetben is mindkét becslés éles, és ráadásul a tanusító rendszerek választhatók olyannak, hogy minden kitöltés lineáris. A klasszikus esetet nem fogjuk végigszámolni, csak megmutatjuk, hogy a 2.2-es fejezetben ismertetett elgondolás itt is helytálló.

A sorok, az oszlopok és a blokkok továbbra is három egyenest határoznak meg, (L_1, L_2, L_3) melyek közül a L_1 és L_2 diszjunktak, és mindketten metszik L_3-at. Keresünk tehát 20 páronként és az L_i egyenesektől is diszjunkt egyenest. Azok a felületek amelyek mind a három egyenest tartalmazzák ugyanúgy néznek ki, mint a 3×3-as esetben, és ha az $x_1x_4 - x_2x_3 = 0$ egyenletű felületet választjuk, akkor arra a szükséges tulajdonságok mellett teljesül az is, hogy a pozíciókat illetve az átlókat meghatározó egyenesek is rajta vannak a felületen.

Mivel ez a felület hiperbolikus (tartalmaz egyenest) így ha vesszük a térnek a már ismertetett fedését amely $q = 5$ esetén négy felületből és két egyenesből áll, akkor van olyan leképzés amely a négy felület egyikét (szabadon megválaszthatjuk, hogy melyiket) átviszi a mi felületünkbe. Ez a transzformáció a teljes fedést átviszi egy olyan fedésbe, amely szintén négy felületet és két egyest fog tartalmazni, és az egyik felületen rajta lesznek az L_i egyenesek. Ha veszünk a másik három felületen egy-egy fedő egyenessereget (egy felület $q + 1 = 6$ egyenessel fedhető le) és a két kimaradó egyenest az összesen 20 számunkra megfelelő egyenes lesz. Ez a 20 egyenes pedig ebben az esetben is meghatároz ez klasszikus sudoku táblázatokból álló ortogonális rendszert, amelyre még az is teljesül, hogy minden táblázatban minden pozícióban és a két átlóban is minden szimbólum pontosan egyszer szerepel.

A szimmetrikus esetben ha felrajzoljuk azt a 16 pontú gráfot, amely az altereket reprezentálja és az egymást egyenesben metsző halmazok között fut él, akkor ismét egy (ugyan nem teljes) páros gráfot kapunk, amelynek mind a két osztálya nyolc elemű (és minden pont foka hat). Ennek a két osztálynak megfelelő nyolc-nyolc kitöltés egy-egy a feltételeknek megfelelő merőleges rendszert alkot, amely igazolja, hogy a becslés ebben az esetben is éles.

Érdekességképpen megjegyezzük, hogy mivel a páros gráf nem teljes, így ebben az esetben nem igaz, hogy egy két alteret használó kitöltés nem lehet benne egy merőleges rendszerben. Annak érdekében, hogy kicsit körbejárjuk ezt a kérdést, hogy hogyan nézhet ki egy szimmetrikus kitöltésekből álló merőleges rendszer, vizsgáljuk meg az említett gráfot. Ha az osztályok között vesszük a gráf komplementerét akkor a következő gráfot kapjuk:



Az ábra négy kör uniójára esik szét, és az alterek közül az összekötöttek, illetve az egy osztályban lévők metszik egymás az origóban. A könnyebbség kedvéért nevezzük a lineáris kitöltéseket A, a nem lineárisakat B típusú táblázatnak.

Ha van a rendszerben egy B típusú táblázat, akkor a hozzá felhasznált két altér különböző négyesből való, és mindkettőhöz két olyan altér van a másik osztályból, melyet az origóban metsz. Ez azt jelenti, hogy a többi táblázatot csak ebből a négy altérből konstruálhatjuk, amire már nincs is túlságosan sok lehetőségünk.
A szimmetria miatt feltehető, hogy a választott pár a $4-9$, a többi táblázathoz felhasználható alterek az $1, 8, 12, 13$. Ezekből négy A és négy B típusú táblázatot készíthetünk, amelyek mindegyike merőleges lesz a már meglévő kitöltésünkre. Ha beválasztunk még egy B típusú táblázatot, akkor ehhez a rendszerhez már nem választható újabb kitöltés, így az legfeljebb két elemű. Ha csak A típusúakat válasz-tunk, akkor abból legfeljebb kettőt tudunk: vagy a 4-es, vagy a 9-es altér két szomszédja által meghatározottakat.

Ha a rendszerben nincs B típusú táblázat, akkor minden kitöltés A típusú, innen is két lehetőség adódik. Ha használtunk különböző osztálybeli alteret, akkor azok ugyanabból a négy hosszú körből valók, és a rendszerhez csak a kör másik két altere által meghatározott kitöltés vehető hozzá, így legfeljebb négyelemű.
Ha nem használtunk ilyet, akkor a rendszer valamelyik nyolcelemű rendszer részeként áll elő, így kijött az is, hogy a maximális elemszám (nyolc) csak a két teljes osztály kiválasztásával érhető el. Ezzel a lehetséges merőleges rendszereket leírtuk.

Megjegyezzük még, hogy a klasszikus eset tetszőleges q prímhatvány esetére általánosítható. Ebben az esetben a rendszer elemszámára $q(q-1)$ a felső becslés, és ez éles is, hiszen ha tekintjük azt (a már bizonyítottan létező) $q-1$ felületből és két egyenesből álló fedést, amelyre teljesül, hogy az egyik felület tartalmazza mind a három L_i egyenest, akkor a maradék $q-2$ felületen lévő $q+1$ egyenest tartalmazó sereghez hozzávéve a kimaradó két egyenest éppen $(q-2)(q+1)+2 = q^2-q$ számunkra megfelelő egyenest kaptunk vagyis a konstrukció ebben az esetben is működik, és bír a már korábban látott extra tulajdonságokkal is.

Hivatkozások

[1] Rosemary A. Bailey, Peter J. Cameron and Robert Connelly: *Sudoku, gerechte designs, resolutions, affine space, spreads, reguli, and Hamming codes*, The American Mathematical Monthly **115** 2008, 383 – 404

[2] Kiss György: *Latin négyzetek, szép sudoku megoldások és véges geometriák I.*, Középiskolai Matematikai és Fizikai Lapok, 2010. március

[3] Kiss György: *Latin négyzetek, szép sudoku megoldások és véges geometriák II.*, Középiskolai Matematikai és Fizikai Lapok, 2010. április

[4] Kiss György - Szőnyi Tamás: *Véges Geometriák*, Polygon Kiadó, Szeged, 2001.