Wiener Gábor

Gráfok hosszú körei és útjai

Habilitációs tézisek

Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
2015.

Tartalomjegyzék

Bevezetés

Jelen értekezés a szerzőnek a Budapesti Műszaki és Gazdaságtudományi Egyetem Villamosmérnöki és Informatikai Karán indított habilitációs eljárásához készült. Célja, hogy a szerző (részben társszerzőkkel közös), a PhD fokozat megszerzését követő tudományos eredményeinek egy részét egységes keretben mutassa be. Az eredményeket 11 tézispontban rendszerezzük. Ezt követi a feldolgozott témakör áttekintése és az egyes eredmények bővebb kifejtése angol nyelven. Szinte minden eredményhez közöljük a részletes bizonyításokat, kivételt csak a nagyon technikai, illetve más bizonyításokhoz rendkívül hasonlító esetekben teszünk. Természetesen ezek a bizonyítások is megtalálhatók a szerző idevágó publikációiban. A tézisek a szerző [4, 42, 60, 61, 62, 63, 64, 65] publikációira épülnek, melyek közül [4] és [65] társszerzője Makoto Araya, [42] társzerzője Salamon Gábor. A [4, 61, 62, 63, 64, 65] publikációk az elmúlt 5 évben jelentek meg.

A gráfelméletben központi szerepet játszik a Hamilton-kör és a Hamilton-út probléma, vagyis annak eldöntése, hogy egy adott gráfnak van-e Hamilton-köre, illetve -útja. Egyikükre sem ismert jól használható szükséges és elégséges feltétel, sőt mindkét probléma NP-teljes. Hasonlóan nehezek a gráfok egyéb hosszú köreivel és útjaival, illetve speciális feszítőfáival kapcsolatos problémák is; ezek egy része speciális esetként tartalmazza a Hamilton-kör, illetve -út problémát. A kapcsolódó kutatások ennek, és a téma fontosságának köszönhetően meglehetősen szerteágazók. Jelen disszertációban három különböző aspektusból vizsgáljuk a kérdést.

Az első fejezetben olyan gráfokat vizsgálunk, melyek maguk nem rendelkeznek Hamilton-körrel (-úttal), de bármely csúcsukat elhagyva már olyan gráfot kapunk, melynek van Hamilton-köre (-útja). Ezek az úgynevezett hypohamiltonian (hypotraceable) gráfok. (Magyar nyelvű terminológia hiányában az angol elnevezéseket használjuk.) A legkisebb hypohamiltonian gráf a jól ismert Petersen-gráf. A téma vizsgálata Sousselier 1963-as cikkével [46] kezdődött, melyben a Petersen-gráf egy általánosítása segítségével végtelen sok hypohamiltonian gráfot talált. 1964-ben Herz, Gaudin és Rossi [23] belátta, hogy a Petersen-gráfnál kisebb hypohamiltonian gráf nem létezik. 1997-re sikerült meghatározni, hogy pontosan mely csúcsszámokra létezik hypohamiltonian gráf (elsősorban Chvátal [11] és Thomassen [49] munkájának köszönhetően, az i-re a pontot Aldred, McKay és Wormald [2] tette fel). Grötschel 1977-ben megmutatta, hogy a hypohamiltonian gráfok használhatók az utazóügynök probléma egészértékű programozási megoldásához (a Gomory-féle cutting-plane módszert használva), így alkalmazásaik rendkívül szerteágazók, a hálózatok és chipek tervezésétől a DNS-szekvenálásig. Hatékony megoldást elsősorban kis méretű hypohamiltonian gráfok esetén kaphatunk. Bár számos cikk foglalkozik hypohamiltonian gráfokkal (kiváló, bár nem kimondottan friss összefoglaló Holton és Sheehan cikke [26]), valójában elég keveset tudunk róluk. Nem ismert például, hogy létezik-e négyszeresen összefüggő hypohamiltonian gráf, sőt az sem, hogy létezik-e olyan, amelynek nincs 3 fokú csúcsa (nyilvánvaló ugyanakkor, hogy minden hypohamiltonian gráf háromszorosan összefüggő). A hypotraceable gráfokról még ennél is jóval kevesebbet tudunk. Sokáig azt sejtették, hogy ilyenek nem is léteznek [30], sőt egy ideig az is kérdéses volt, hogy létezik-e olyan gráf, melyben egyik csúcson sem megy át az összes leghosszabb út (a kérdést 1966-ban vetette fel Gallai [17], 1969-ben válaszolta meg – igenlően – Walther [58]). Az első, 40 csúcsú hypotraceable gráfot Horton találta 1976-ban (ld. [67, 51]), a legkisebb ismert hypotraceable gráfnak 34 csúcsa van, ez Thomassen nevéhez fűződik [49]. Nem ismert érdemi alsó becslés a legkisebb hypotraceble gráf méretére. Ennek az az egyik magyarázata, hogy az összes ismert hypotraceable gráf hypohamiltonian gráfok segítségével készült, lényegében Thomassen két módszerét használva [49, 51]. (Az ugyanakkor ismert, hogy ha $n \geq 42$, akkor létezik n csúcsú hypotraceable gráf [49].)

Az 1976-ig ismert hypohamiltonian gráfok jórészt a Petersen-gráf általánosításaként, illetve Chvátal úgynevezett flip-flopjainak segítségével [11] álltak elő és egyikük sem volt síkbarajzolható. Ez motiválta Chvátalt, amikor felvetette, hogy egyáltalán léteznek-e síkbarajzolható hypohamiltonian gráfok (és ha igen, léteznek-e ilyenek, amelyek még 3-regulárisak is). Az első síkbarajzolható hypohamiltonian gráfot 1976-ban találta Thomassen [51], ennek 105 csúcsa volt, 1979-ben pedig Hatzel [22] talált egy 57 csúcsú hypohamiltonian síkgráfot. 1993-ban Holton és Sheehan [26] tette fel a kérdést, hogy vajon létezik-e ennél kisebb hypohamiltonian síkgráf. C. Zamfirescu és T. Zamfirescu [68] 2007-ben talált egy 48 csúcsú ilyet, a szerző pedig (Makoto Arayaval közösen) 2011-ben egy 42 csúcsút [65]. A legkisebb ismert hypohamiltonian síkgráf mérete 40, ezt Jooyandeh, McKay, Östergård, Pettersson és C. Zamfirescu [29] találta 2014-ben.

A síkbarajzolható esetben még kevesebbet tudunk a hypohamiltonian és hypotraceable gráfokról. 2011-ig még az sem volt ismert, hogy kellően nagy n-re létezik-e n csúcsú hypohamiltonian, illetve hypotraceable síkgráf (Holton és Sheehan meg is említi az előbbit a terület megoldatlan problémái között [26]). 2011-ben Makoto Arayaval közösen sikerült megválaszolnunk e kérdéseket: megmutattuk, hogy minden $n \geq 76$ esetén létezik n csúcsú síkbarajzolható hypohamiltonian gráf, illetve minden $n \geq 180$ esetén létezik n csúcsú síkbarajzolható hypotraceable gráf [65]. A becsléseket 2014-ben 42-re, illetve 156-ra javították Jooyandeh és szerzőtársai [29].

A síkbarajzolható 3-reguláris gráfok Hamilton-köreinek problémája több, mint fél évszázadon át a gráfelmélet egyik központi kérdése volt, hiszen Tait sejtéséből, miszerint minden háromszorosan összefüggő, 3-reguláris síkgráfnak van Hamilton-köre, következett volna a híres négyszín-sejtés [48]. Bár Tait sejtését 1946-ban megcáfolta Tutte [55], 1968-ig, a Grinberg-tétel [19] felfedezéséig nagyon nehéz volt további ellenpéldákat találni. Chvátal 1973-as, 3-reguláris hypohamiltonian síkgráfokra vonatkozó kérdése ennek megfelelően cseppet sem tűnt könnyűnek. Az első ilyen gráfot Thomassen találta 1981-ben, ennek 94 csúcsa van. 2011-ig nem is sikerült ennél kisebb példát találni és az sem volt ismert, hogy kellően nagy páros n esetén létezik-e n csúcsú 3-reguláris hypohamiltonian síkgráf (mindkét kérdés szerepel Holton és Sheehan cikkében [26] a megoldatlan problémák között.) Aldred, Bau, Holton és McKay 2000-es cikkéből [1] ugyanakkor következett, hogy nincs 42 vagy kevesebb csúcsú ilyen gráf. Makoto Arayaval közösen 2011-ben sikerült mindkét kérdést megválaszolnunk: mutattunk egy 70 csúcsú 3-reguláris hypohamiltonian síkgráfot (melynél kisebb ma sem ismert) és bebizonyítottuk, hogy minden $n \geq 86$ esetén létezik n csúcsú 3-reguláris hypohamiltonian síkgráf [4]. A 86-os korlátot 2015-ben 74-re javították [69].

A második fejezetben egy feszítőfa-optimalizálási problémára adunk közelítő algoritmusokat. A feszítőfa-optimalizálási problémák tipikusan gyakorlatban felmerülő feladatokkal állnak szoros kapcsolatban (pl. hálózatok tervezése, routing) [41, 18, 39, 45, 6]. A cél egy összefüggő gráf valamilyen célfüggvény szerint optimális feszítőfájának megtalálása; nagyon gyakori, hogy a gráf egy Hamilton-útja (ha létezik) optimális feszítőfa, ilyenkor a feladat persze

NP-nehéz, ezért a pontos (de lassú) megoldások helyett a közelítő algoritmusok kerülnek előtérbe. Az általunk vizsgált MINLST (Minimum Leaf Spanning Tree) probléma is ide tartozik: a cél olyan feszítőfa megtalálása, melynek a lehető legkevesebb levele (vagyis 1 fokú csúcsa) van. Lu és Ravi 1996-ban megmutatta [38], hogy erre az optikai hálózatok, vízgazdálkodási rendszerek tervezésekor is hasznos problémára még közelítő algoritmust sem lehet adni (hacsak $P = NP$ nem teljesül). Optimalizálási szempontból a MINLST feladat nyilván ekvivalens azzal a problémával, amikor olyan feszítőfát keresünk, melynek a lehető legtöbb belső csúcsa (azaz nem levele) van. Ez a probléma (Maximum Internal node Spanning Tree – MAXIST) azonban már approximálható: 2008-ban Salamon Gáborral közösen lineáris idejű 2-approximációt sikerült megadnunk, melynek finomításával $\frac{3}{2}$-approximációt kaptunk karom-mentes gráfokra és lineáris futásidejű $\frac{6}{5}$-approximációt 3-reguláris gráfokra [42]. A cikk közlése óta az approximációs faktort számos alkalommal javították, a legjobb ismert faktor általános gráfokra $\frac{3}{2}$ [35], 1 fokú csúcs nélküli gráfokra pedig $\frac{4}{3}$ [36].

A harmadik fejezetben az első két fejezet megközelítéseit egyesítve a hypohamiltonian és hypotraceable tulajdonságokat kiterjesztjük az említett feszítőfa-optimalizálási problémára (és egy útfedéssel kapcsolatos problémára is). Az egyesített megközelítés hatékonyságát mutatja, hogy a segítségével sikerült megválaszolni Gargano, Hammar, Hell, Stacho és Vaccaro [18] egy nyitott kérdését. Egy összefüggő gráf minimális levélszámát a feszítőfái levélszámának minimumaként definiáljuk, azzal a kiegészítéssel, hogy ha a gráfnak van Hamilton-köre, akkor a kérdéses szám nem 2, hanem 1. Egy gráfot l-levél-kritikusnak nevezünk, ha a minimális levélszáma l és bármely csúcsát elhagyva a minimális levélszám $l-1$. Könnyen látható, hogy a 2-levél-kritikus gráfok épp a hypohamiltonian gráfok, a 3-levél-kritikus gráfok pedig a hypotraceable gráfok. A 3.1 alfejezetben megmutatjuk, hogy nem csak $l = 2,3$, hanem tetszőleges $l \geq 2$ egész esetén léteznek l-levél-kritikus gráfok, sőt elegendően nagy n esetén létezik n csúcsú síkbarajzolható, 3-reguláris l-levél-kritikus gráf is [62, 63]. A hypohamiltonian és hypotraceable gráfok szerkezetéről nagyon keveset lehet tudni, az egyik ilyen eredmény Thomassen hypotraceable 2-töredékeket karakterizáló lemmája [49]. Ennek egy levél-kritikus gráfokra vonatkozó általánosítását bizonyítjuk be a 3.2 alfejezetben [62, 63].

A következő definíciók Garganotól és szerzőtársaitól származnak [18]. Egy fát póknak nevezünk, ha legfeljebb egy olyan csúcsa van, melynek foka nagyobb, mint 2; a pók középpontja a 2-nél nagyobb fokú csúcs (ha van ilyen, egyébként tetszőleges csúcs tekinthető a középpontnak). Egy gráf pókszerű, ha bármely v csúcsához létezik a gráfnak olyan feszítő pókja, melynek középpontja v. Nyilvánvaló, hogy a Hamilton-úttal rendelkező gráfok pókszerűek és könnyen látható, hogy ugyanez igaz a hypotraceable gráfokra is. Garganoék (egyik) kérdése az volt, hogy léteznek-e egyéb pókszerű gráfok is. A 3.3 alfejezetben először megmutatjuk, hogy a korábban talált levél-kritikus gráfok közül bizonyosak út-kritikusak is (vagyis bármely csúcsukat elhagyva a csúcsok fedéséhez szükséges diszjunkt utak száma eggyel csökken – korábban ilyen gráfok csak a 2 úttal fedhető esetben voltak ismertek) [64], majd ezt a tulajdonságot felhasználva Hamilton-út nélküli, nem hypotraceable, pókszerű gráfokat konstruálunk. Sőt, azt is megmutatjuk, hogy tetszőleges H gráf esetén létezik olyan Hamilton-út nélküli, nem hypotraceable, pókszerű gráf, mely feszített részgráfként tartalmazza H-t [64].

A negyedik fejezet olyan, hipergráfok nyomairól szóló tételeket tartalmaz, melyek hálózatok hibatűréséhez, precízebben a hiperkocka bizonyos (hibás) csúcsait elkerülő hosszú útjaihoz és köreihez kapcsolódnak. Hipergráfok nyomait (vagyis az alaphalmaz valamely részhalmazára vett megszorításait) régóta vizsgálják; Vapnik és Chervonenkis [56] klasszikus cikke 1971-ben jelent meg. Ebben a cikkben már szerepel (implicit formában) a többnyire Sauer tételeként [43] ismert állítás (melyet az említetteken kívül bebizonyított Perles és Shelah [44] is, de már Erdős is sejtette). A tétel szerint minden n elemű alaphalmazon adott, legalább $\sum_0^{r-1}\binom{n}{i}+1$ különböző halmazt tartalmazó halmazrendszernek van olyan r elemű R halmazon vett nyoma, amely R minden részhalmazát tartalmazza. Ennek az állításnak és Bondy egy tételének [8] közös általánosítását adjuk a 4.1 alfejezetben [60]. Ebből az általános állításból és Turán tételéből [54] következik a 4.2 alfejezet fő eredménye, mely szerint $m \geq 2n$ esetén minden n elemű alaphalmazon adott, m halmazt tartalmazó halmazrendszernek van olyan $\frac{n^2}{2m-n-2}$ elemű halmazon vett nyoma, melyben minden halmaz multiplicitása legfeljebb $\frac{n^2}{2m-n-2}+1$ [60]. Ezt a tételt használta Fink és Gregor [14] annak bizonyítására, hogy elegendően nagy n esetén az n-dimenziós hiperkockából egy legfeljebb $\frac{n^2}{10}+\frac{n}{2}+1$ elemű X csúcshalmazt törölve, a kapott gráfnak van $2^n-2|X|$ hosszú köre (ennél hosszabb kör tetszőleges X esetén nem várható el, hiszen a hiperkocka páros gráf). Ez volt az első olyan eredmény, amelyben négyzetes nagyságrendű hibás csúcsot engedtek meg (korábban X méretét $(n-1)$-gyel, $(2n-4)$-gyel, majd $(3n-7)$-tel kellett felülről korlátozni). Hasonló, négyzetes nagyságrendű eredményt bizonyított az említett tétel segítségével Dvořák és Koubek [13] körök helyett utakról.

Jelölések. A disszertációban szereplő gráfok mind véges, egyszerű, irányítatlan, összefüggő gráfok. A G gráf csúcshalamzát $V(G)$, élhalmazát $E(G)$ jelöli. Az a és b csúcsok közti élet (a,b)-vel, az $a_1,a_2,\ldots,a_k$ csúcsokon átmenő kört $(a_1,a_2,\ldots,a_k)$-val jelöljük. $G[X]$ jelöli a G gráf Xcsúcshalmaza által feszített részgráfját, $e_G(X)$ a $G[X]$ gráf élszámát, $G-X$ pedig azt a gráfot, amit G-ből az X csúcshalmaz törlésével kapunk, $G-v := G-\{v\}$. Ha H részgráfja G-nek ,akkor $G\setminus H$ az a gráf, melynek csúcshalmaza $V(G)$, élhalmaza $E(G)\setminus E(H)$.

A v csúcs fokát a G gráfban $d_G(v)$ jelöli (ha világos, hogy melyik gráfról van szó, akkor egyszerűen $d(v)$), az X és Y csúcshalmazok közt futó élek számát pedig $d_G(X,Y)$. $d_G(X) := d_G(X,V(G)\setminus X)$, $d_G(X,v) := d_G(X,\{v\})$.

$G\cup H$ a G és H gráfok diszjunkt uniója, de használjuk a jelölést akkor is, ha G és H ugyanazon gráf részgráfjai, ilyenkor $G\cup H$ csúcshalmaza $V(G)\cup V(H)$, élhalmaza $E(G)\cup E(H)$.

Legyen H a G gráf részgráfja, $X\subseteq V(G)$. Ekkor $H+X$ jelöli G-nek azt a részgráfját, melynek csúcsai $V(H)\cup X$, élei pedig H és $G[X]$ élein kívül a $V(H)$ és X közti G-beli élek; $H+v := H+\{v\}$ bármely $v\in V(G)$-re. Legyen a és b a G gráf két csúcsa, ekkor $G+(a,b)$ jelöli azt a gráfot, melyet G-ből az (a,b) él G-hez adásával kapunk.

Tézispontok

1. Minden elegendően nagy n egész esetén létezik n csúcsú síkbarajzolható hypohamiltonian, illetve hypotraceable gráf (sőt, az első esetben $n \geq 76$, a másodikban $n \geq 180$ elég) – a hypohamiltonian eset Holton és Sheehan egy 1993-as problémájának [26] megoldása. (Theorem 1.5 és Theorem 1.13. A 76-os korlátot azóta 42-re, a 180-as korlátot 156-ra javították [28, 29].) (Forrás: [65], közös eredmények Makoto Arayaval.)

2. A legkisebb síkbarajzolható hypohamiltonian gráfnak legfeljebb 42, a legkisebb síkbarajzolható hypotraceable gráfnak legfeljebb 162 csúcsa van. (Theorem 1.1, és Corollary 1.3. A becsléseket azóta 40-re és 154-re javították [28, 29].) (Forrás: [65], közös eredmények Makoto Arayaval.)

3. Minden elegendően nagy páros n egész esetén létezik n csúcsú 3-reguláris síkbarajzolható hypohamiltonian, illetve hypotraceable gráf (sőt, az első esetben $n \geq 86$, a második esetben $n \geq 356$ elég) – a hypohamiltonian eset Holton és Sheehan egy 1993-as problémájának [26] megoldása. (Corollary 1.17 és Corollary 1.18. A 86-os korlátot azóta 74-re javították [69].). (Forrás: [4], közös eredmények Makoto Arayaval.)

4. A legkisebb síkbarajzolható 3-reguláris hypohamiltonian gráfnak legfeljebb 70, a legkisebb síkbarajzolható 3-reguláris hypotraceable gráfnak legfeljebb 340 csúcsa van – a hypohamiltonian eset ugyancsak Holton és Sheehan egy 1993-as problémájának [26] megoldása. (Theorem 1.16 és Corollary 1.18.) (Forrás: [4], közös eredmények Makoto Arayaval.)

5. Lineáris futásidejű 2-approximációs algoritmus a MAXIST problémára (maximális belső csúcsú feszítőfa keresése), $\frac{3}{2}$-approximáció claw-free gráfokra, lineáris futásidejű $\frac{6}{5}$-approximáció 3-reguláris gráfokra. (Algorithm 1, Theorem 2.4, Algorithm 2, Theorems 2.6, 2.8. Azóta az approximációs faktort általános gráfokra előbb $\frac{5}{3}$-ra [32], majd $\frac{3}{2}$-re [35], 1 fokú csúcs nélküli gráfokra $\frac{7}{4}$-re [40], $\frac{5}{3}$-ra [32], $\frac{3}{2}$-re [35], majd $\frac{4}{3}$-ra [36] javították. Forrás: [42], közös eredmények Salamon Gáborral.)

6. Minden $l \geq 2$ egészre léteznek l-levél-kritikus és l-levél-stabil gráfok, sőt minden elegendően nagy n-re létezik n csúcsú l-levél-kritikus és l-levél-stabil gráf. (Theorem 3.9, Theorem 3.10, Remark, 33. oldal. Forrás: [62, 63])

7. l-levélkritikus 2-töredékek karakterizációja (Thomassen hypotraceable 2-töredékeket karakterizáló lemmájának [51] általánosítása). (Theorem 3.17. Forrás: [62, 63]).

8. Minden $\mu \geq 2$ egészre léteznek μ-út-kritikus gráfok, sőt minden elegendően nagy n-re létezik n csúcsú μ-út-kritikus gráf. (Theorem 3.20. Forrás: [64])

9. Léteznek olyan nem hypotraceable arachnoid gráfok, amiknek nincs Hamilton-útja – Gargano, Hammar, Hell, Stacho és Vaccaro 2002-es problémájának [18] megoldása. Sőt,

tetszőleges H gráfhoz létezik olyan Hamilton-út nélküli, nem hypotraceable arachnoid gráf, mely H-t feszített részgráfként tartalmazza. (Theorem 3.22, Remark, 40. oldal. Forrás: [63, 64].)

10. Az $(n,m) \triangleright (r,s)$ reláció karakterizációja a letömörítési technika segítségével – Bondy [8] és Sauer [43] tételeinek egy közös általánosítása. (Theorem 4.4. Forrás: [60].)

11. $m \geq 2n$ és $r = \lceil \frac{n^2}{2m-n-2} \rceil$ esetén $(n,m) \triangleright (r, r+1)$. Sőt, minden $\mathscr{A} \in MSH(n,m)$ esetén létezik olyan $\left\lceil \frac{n^2}{2m-n-2} \right\rceil$ elemű $X \subseteq [n]$ halmaz, melyre $\mathscr{A}$-t az $[n]-X$ halmazra megszorítva, a kapott hipergráfban minden él multiplicitása legfeljebb $\left\lceil \frac{n^2}{2m-n-2} \right\rceil + 1$. (Theorem 4.6, Theorem 4.7. Forrás: [60, 61].)

1. fejezet

Hypohamiltonian and hypotraceable graphs

A graph is called *hypohamiltonian* if it is not hamiltonian but deleting any vertex gives a hamiltonian graph; a well-known example is the Petersen graph. The study of hypohamiltonian graphs started in 1963, with the paper of Sousselier [46], who managed to find an infinite sequence of hypohamiltonian graphs on $6k+10$ vertices for every integer $k \geq 0$ by generalizing the Peopen prob tersen graph. In 1975 Doyen and Van Dienst [12] found another generalization and a sequence of hypohamiltonian graphs on $3k+10$ vertices for every integer $k \geq 0$. In 1973 Chvátal [11] invented the so-called flip-flops (that we will use in Chapter 3) and obtained many new hypohamiltonian graphs.

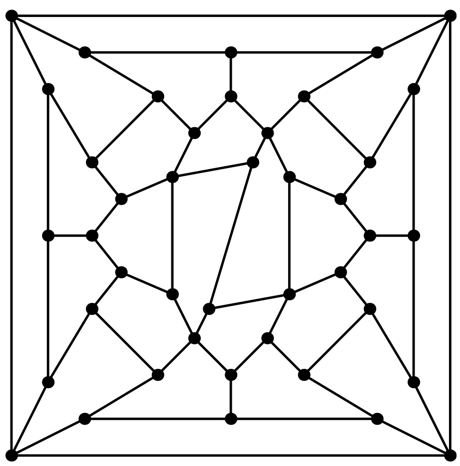
Herz, Gaudin, and Rossi [23] in 1964 proved that the Petersen graph is the smallest hypohamiltonian graph, and Aldred, McKay, and Wormald [2] in 1997, finalizing the efforts of many others (Herz, Duby, and Vigue [24], Chvátal [11], Thomassen [49], Collier and Scmeichel [10]) proved that a hypohamiltonian graph on n vertices exists if and only if $n = 10, 13, 15, 16$ or $n \geq 18$.

A graph is called *hypotraceable* if it is not traceable, but deleting any vertex gives a traceable graph. The existence of such graphs was an open problem till 1975, when Horton found such a graph on 40 vertices (see [67, 51]) disproving the conjecture of Kapoor, Kronk, and Lick [30]. Actually, even the existence of graphs without concurrent longest paths was an open question from 1966 to 1969 (raised by Gallai [17] and settled by Walther [58]). The smallest known hypotraceable graph (having 34 vertices) is due to Thomassen [49], who also proved that hypotraceable graphs on n vertices exist for every $n \geq 42$ [49].

Hypohamiltonian and hypotraceable graphs were extensively studied in the last five decades, see e.g. the papers [50, 51, 52, 53, 22, 68, 28, 29] and the excellent survey by Holton and Sheehan [26]. However, not much is known about their structure, especially in the case of hypotraceable graphs, e.g. all known hypotraceable graphs are constructed using hypohamiltonian graphs. There are still a lot of open questions, even among the very natural ones, like whether there exists a 4-connected hypohamiltonian or hypotraceable graph. Hypohamiltonian graphs are easily seen to be 3-connected, hypotraceable graphs are easily seen to be 2-connected and 3-edge-connected, on the other hand no hypohamiltonian or hypotraceable graph without a vertex of degree 3 is known.

1.1. Planar hypohamiltonian and hypotraceable graphs

All graphs obtained by the flip-flop technique or generalizations of the Petersen graph are non-planar. This fact led Chvátal to ask whether there exist planar hypohamiltonian graphs [11]. This was answered by Thomassen [51], who found such a graph on 105 vertices in 1976. Hatzel [22] found a smaller planar hypohamiltonian graph, having 57 vertices in 1979. Holton and Sheehan [26] asked about the minimum size of planar hypohamiltonian graphs. Hatzel's bound was improved to 48 by Zamfirescu and Zamfirescu [68] in 2007. M. Araya and the author have found a planar hypohamiltonian graph on 42 vertices [65] in 2011 (see also Theorem 1.1) and the currently known smallest such graph has 40 vertices [29, 28].
Using the graph in Theorem 1.1 and a theorem of Thomassen [49], M. Araya and the author constructed a planar hypotraceable graph on 162 vertices (see [65] and Corollary 1.3) improving the (then) best known bound of 186, which was improved further to 154 by Jooyandeh et al. [29, 28] recently.
We have mentioned that not much is known about hypohamiltonian and hypotraceable graphs. This is even more true for the planar case (a nice exception is the theorem of Thomassen [52] stating that every planar hypohamiltonian graph contains a vertex of degree 3); while since 1997 it has been known for which values of n exists a hypohamiltonian graph, Holton and Sheehan [26] mention the open problem whether there exists a planar hypohamiltonian graph on n vertices, provided n is sufficiently large. This problem has been settled in 2011 by M. Araya and the author (see [65] and Theorem 1.5), moreover we proved a similar theorem for hypotraceable graphs (Theorem 1.13). We showed that for every integer $n \geq 76$ there exists a planar hypohamiltonian graph on n vertices and for every integer $n \geq 180$ there exists a planar hypotraceable graph on n vertices. The bounds were improved recently to 42 and 156 by Jooyandeh et al. [29, 28].
Zamfirescu [66] denoted the smallest number of vertices of a planar k-connected graph, in which every j vertices are omitted by some longest cycle (path) by $\overline{C_k^j}$ ($\overline{P_k^j}$). In this section we also improve on the (then) best known bounds concerning the numbers $\overline{C_3^1}$, $\overline{C_3^2}$, $\overline{P_3^1}$, $\overline{P_3^2}$.
Let us consider now the following graph Γ.



Theorem 1.1 (Araya-Wiener, 2011 [65]) Γ *is a planar hypohamiltonian graph.*

Proof. Γ is obviously planar and has 42 vertices, 67 edges, and 27 faces, of which one is a quadrilateral and the others are all pentagons. To prove that Γ is not hamiltonian we use a theorem of Grinberg [19].

Theorem 1.2 (Grinberg, 1968 [19]) *Suppose a plane graph has a hamiltonian cycle, such that there are f_i i-gons in the exterior of the hamiltonian cycle and f'_i i-gons in the interior of the hamiltonian cycle. Then*

$$\sum_i (i-2)(f_i - f'_i) = 0.$$

For the graph Γ the sum in Grinberg's theorem cannot be 0, since there is only one face of Γ whose degree is not congruent to 2 modulo 3, from which the nonhamiltonicity of Γ follows. To see that Γ is hypohamiltonian indeed, we have to show that the deletion of any vertex of Γ gives a hamiltonian graph. Since the drawing of Γ in Figure 1 is centrally symmetric, we only have to check 21 cases. The hamiltonian cycles of all graphs obtained by deleting one vertex of Γ can be found in [65]. □

An easy corollary of the above theorem is the existence of a planar hypotraceable graph on 162 vertices, improving the bound of 186 in [68]. The construction is based on graph Γ and a method of Thomassen [49] for creating hypotraceable graphs using hypohamiltonian graphs.

Corollary 1.3 (Araya-Wiener, 2011 [65]) *There exists a planar hypotraceable graph on* 162 *vertices.*

Proof. Let Γ_4 be the following graph. Let the neighbours of a vertex v of degree 3 in graph Γ be x,y,z. Take 4 vertex-disjoint copies of $\Gamma - v$, and let the copies of x (resp. y,z) be x_1,x_2,x_3,x_4 (resp. y_1,y_2,y_3,y_4, z_1,z_2,z_3,z_4). Now identify the vertex x_1 with x_2 and the vertex x_3 with x_4 and add the edges $(y_1,y_3),(z_1,z_3),(y_2,y_4),(z_2,z_4)$ to the graph.
It is obvious that Γ_4 has 162 vertices and it is also easy to see that it is planar. By a theorem of Thomassen [49] Γ_4 is hypotraceable, since Γ is hypohamiltonian by Theorem 1.1. □

Another corollary concerns some of the numbers $\overline{C_k^j}$ (and $\overline{P_k^j}$), that are defined in [66] as the smallest number of vertices of a planar k-connected graph, in which every j vertices are omitted by some longest cycle (path). In the book by Voss [57] the following bounds can be found for $\overline{C_3^1}, \overline{C_3^2}, \overline{P_3^2}$, and $\overline{P_3^1}$: $\overline{C_3^1} \leq 57$, $\overline{C_3^2} \leq 6758$, $\overline{P_3^1} \leq 224$, $\overline{P_3^2} \leq 26378$. These bounds were improved by Zamfirescu and Zamfirescu [68]: based on their 48 vertex hypohamiltonian planar graph they showed that $\overline{C_3^1} \leq 48$, $\overline{C_3^2} \leq 4277$, $\overline{P_3^1} \leq 188$, and $\overline{P_3^2} \leq 16926$. Now using our graph Γ we can derive even better bounds. The proof of the bounds is based on the technique of Corollary 2 in [68].

Corollary 1.4 (Araya-Wiener, 2011 [65]) $\overline{C_3^1} \leq 42$, $\overline{C_3^2} \leq 3701$, $\overline{P_3^1} \leq 164$, $\overline{P_3^2} \leq 14694$.

Now we prove the main theorem of this section.

Theorem 1.5 (Araya-Wiener, 2011 [65]) *There exists a planar hypohamiltonian graph on n vertices for every integer $n \geq 76$.*

We will use the following definition and lemma several times in the proof of Theorem 1.5.

Definition 1.6 *Let G be a graph with a* 4*-cycle (a,b,c,d). Now* $\mathrm{Th}(G,a,b,c,d)$ *is the graph obtained from G by deleting the edges (a,b) and (c,d) and adding a new* 4*-cycle (a',b',c',d') and the edges (a,a'), (b,b'), (c,c'), (d,d') to G.*

We call the function Th the *Thomassen operation*, since it was introduced by Thomassen [53], who used it to show that there exist infinitely many planar cubic hypohamiltonian graphs. The next two lemmas are a slight modification of a claim of Thomassen [53] and the proof is almost the same; we include it for completeness.

Lemma 1.7 *Let G be a planar nonhamiltonian graph having a* 4*-cycle* (a,b,c,d)*. Then* $\mathrm{Th}(G,a,b,c,d)$ *is also a planar nonhamiltonian graph.*

Proof. We use the shorthand notation $\mathrm{Th}(G)$ for $\mathrm{Th}(G,a,b,c,d)$. It is obvious that $\mathrm{Th}(G)$ is planar. Now suppose to the contrary that $\mathrm{Th}(G)$ contains a hamiltonian cycle C. C clearly contains either all four or exactly two of the edges (a,a'), (b,b'), (c,c'), (d,d'). In the first case there exist two vertex-disjoint paths covering all vertices of G with endvertices a,b,c,d, which together with two of the edges (a,b), (b,c), (c,d), (d,a) gives a hamiltonian cycle of G, a contradiction. In the second case there exists a hamiltonian path P of G with its endvertices among a,b,c,d. We show that the endvertices are neighbours, thus again we have a hamiltonian cycle in G, a contradiction. If the endvertices of the path were (say) a and c, then the deletion of a' and c' from the hamiltonian cycle C would give a graph having three components ($\{b'\}$, $\{d'\}$, and P), which is clearly impossible. □

Lemma 1.8 *Let G be a planar hypohamiltonian graph having a* 4*-cycle* (a,b,c,d) *and suppose that the degrees of the vertices* a,b,c,d *are* 3*. Then* $\mathrm{Th}(G,a,b,c,d)$ *is also a planar hypohamiltonian graph.*

Proof. By Lemma 1.7, $\mathrm{Th}(G)$ is planar and nonhamiltonian. We have to show that the deletion of any vertex of $\mathrm{Th}(G)$ gives a hamiltonian graph.

First let us suppose that we delete one of the new vertices a',b',c',d', let it be (say) a'. Consider now a hamiltonian cycle C_d of the graph $G-d$. Since a has degree 3 in G and d is one of its neighbours, C_d uses the edge (a,b). Now it is easy to see that by deleting this edge from C_d and adding the path (b,b',c',d',d,a), we obtain a hamiltonian cycle of $\mathrm{Th}(G)-a'$.

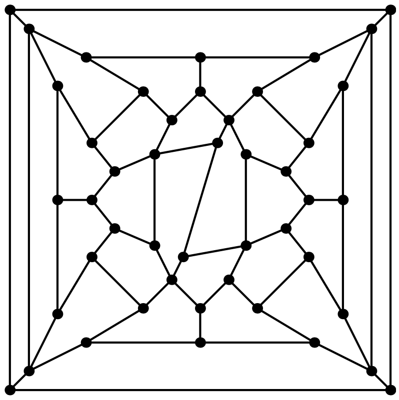
Now suppose we delete a vertex v of G from $\mathrm{Th}(G)$. Without loss of generality we may assume that $v \neq a$. Let us consider a hamiltonian cycle C_v of $G-v$. Since a is in the cycle and has degree 3 in G (and therefore degree at most 3 in $G-v$), C_v contains at least one of the edges (a,d), (a,b).

If C_v contains both (a,b) and (c,d), then replace these edges by the paths (a,a',b',b) and (c,c',d',d); if C_v contains (a,b) and does not contain (c,d), then replace (a,b) by the path (a,a',d',c',b',b); if C_v contains (c,d) and does not contain (a,b), then replace (c,d) by the path (c,c',b',a',d',d); finally if C_v contains none of (a,b) and (c,d), then it contains the edge (a,d) and now replace this edge by the path (a,a',b',c',d',d). In any case we obtain a Hamiltonian cycle of $\mathrm{Th}(G)-v$. □

Now we prove Theorem 1.5 through a sequence of lemmas.

Lemma 1.9 *There exists a planar hypohamiltonian graph on* $42+4m$ *vertices for every integer* $m \geq 0$*.*

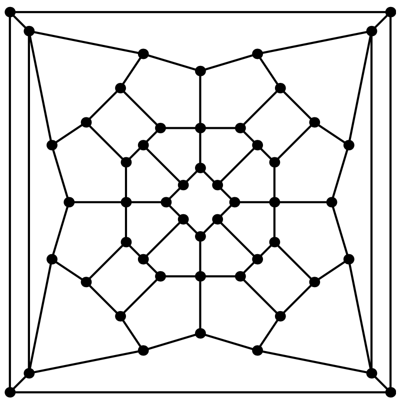
Let a,b,c,d be the vertices of the quadrilateral of graph Γ. Then the graph $\mathrm{Th}(\Gamma,a,b,c,d)$ is the following:



By Lemma 1.7, $\mathrm{Th}(\Gamma, a, b, c, d)$ is planar and nonhamiltonian. To see that it is also hypohamiltonian we have to find hamiltonian cycles of all of its vertex-deleted subgraphs – these can be found in [65]. Since it is obvious that $\mathrm{Th}(G)$ always contains a 4-cycle with vertices of degree 3, applying the Thomassen operation iteratively we obtain planar hypohamiltonian graphs on $42 + 4m$ vertices for every integer $m \geq 0$, by Lemma 1.8. □

Lemma 1.10 *There exists a planar hypohamiltonian graph on* $48 + 4m$ *vertices for every integer* $m \geq 0$*.*

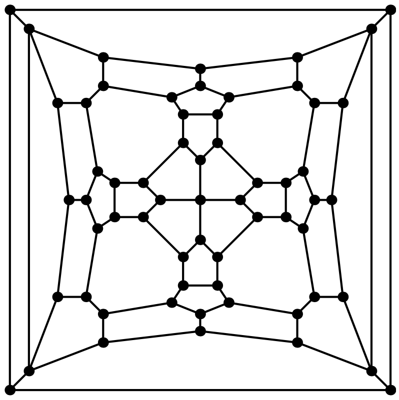
Now we apply the Thomassen operation on the Zamfirescu graph [68]: let a, b, c, d be the vertices of the quadrilateral of the Zamfirescu graph Z. The resulting graph is the following:



By Lemma 1.7, $\mathrm{Th}(Z, a, b, c, d)$ is planar and nonhamiltonian. To see that it is also hypohamiltonian again we have to find H hamiltonian cycles of all of its vertex-deleted subgraphs – these can be found in [65]. Now applying the Thomassen operation iteratively we obtain planar hypohamiltonian graphs on $48 + 4m$ vertices for every integer $m \geq 0$, by Lemma 1.8. □

Lemma 1.11 *There exists a planar hypohamiltonian graph on* $57 + 4m$ *vertices for every integer* $m \geq 0$*.*

Proof. We apply the Thomassen operation on the Hatzel graph [22]: let a, b, c, d be the vertices of the quadrilateral of the Hatzel graph H, the resulting graph can be seen here:



By Lemma 1.7, $\mathrm{Th}(H,a,b,c,d)$ is planar and nonhamiltonian. The hamiltonian cycles of its vertex-deleted subgraphs again can be found in [65]. Now applying the Thomassen operation iteratively we obtain planar hypohamiltonian graphs on $57+4m$ vertices for every integer $m \geq 0$, by Lemma 1.8. □

Lemma 1.12 *There exists a planar hypohamiltonian graph on* $79+4m$ *vertices for every integer* $m \geq 0$.

Proof. Let T be the following graph. Let us take two vertex-disjoint copies of graph Γ and delete a vertex of degree 3 in both copies. Now we identify the neighbours of the deleted vertices (that is, if they are α, β, γ in one of the copies and α', β', γ' in the other, then we identify α with α', β with β', γ with γ'). The graph T has 79 vertices. It is easy to see that T is planar and by Lemma 2.1. of [49], T is hypohamiltonian. To obtain a planar hypohamiltonian graph on $79+4m$ vertices for some $m \geq 1$, we just have to change one of the copies of Γ to a planar hypohamiltonian graph on $42+4m$ vertices (such a graph exists by Lemma 1.9). □

Proof of Theorem 1.5: Now the proof is easy: since $(42,48,57,79)$ is a complete residue system modulo 4, by Lemmas 1.9, 1.10, 1.11, and 1.12, there exists a planar hypohamiltonian graph on n vertices for every integer $n \geq 76$. □

Now we prove a similar theorem concerning hypotraceable graphs.

Theorem 1.13 (Araya-Wiener, 2011 [65]) *There exists a planar hypotraceable graph on n vertices for every integer $n \geq 180$.*

Proof. We use the same method of Thomassen [49] as we used in the proof of Corollary 1.3. Let G_1, G_2, G_3, G_4 be planar hypohamiltonian graphs and let v_i be a vertex of degree 3 in G_i $(i = 1,2,3,4)$. (Such a vertex always exists, see [52].) Let the neighbours of v_i in G_i be x_i, y_i, z_i. Now consider the union of the graphs $G_i - v_i$ $(i = 1,2,3,4)$ and identify the vertices x_1, x_2 and the vertices x_3, x_4 and add to the graph the edges (y_1,y_3), (z_1,z_3), (y_2,y_4), and (z_2,z_4). The resulting graph G is easily seen to be planar and by Lemma 3.1 of [49] also hypotraceable.
We distinguish 4 cases according to the residue of n modulo 4.
Case 1. $n = 4k$ for some $k \geq 42$. Let G_1 and G_2 be the graph Γ, G_3 the Zamfirescu graph Z, and G_4 be a planar hypohamiltonian graph on $4k-126$ vertices ($4k-126 \geq 42$, thus by Lemma 1.9, such a graph exists). Now G has $4k$ vertices.
Case 2. $n = 4k+1$ for some $k \geq 44$. Let G_1 and G_2 be the graph Γ, G_3 the Hatzel graph H, and G_4 be a planar hypohamiltonian graph on $4k-134$ vertices ($4k-134 \geq 42$, thus by Lemma 1.9, such a graph exists). Now G has $4k+1$ vertices.

Case 3. $n = 4k+2$ for some $k \geq 40$. Let G_1, G_2, and G_3 be the graph Γ, and G_4 be a planar hypohamiltonian graph on $4k-118$ vertices ($4k-118 \geq 42$, thus by Lemma 1.9, such a graph exists). Now G has $4k+2$ vertices.
Case 4. $n = 4k+3$ for some $k \geq 45$. Let G_1 be the graph Γ, G_2 the Zamfirescu graph Z, G_3 the Hatzel graph H, and G_4 be a planar hypohamiltonian graph on $4k-138$ vertices ($4k-138 \geq 42$, thus by Lemma 1.9, such a graph exists). Now G has $4k+3$ vertices. □

1.2. Cubic planar hypohamiltonian and hypotraceable graphs

Hamiltonian properties of planar cubic graphs have been investigated extensively since Tait's attempt to prove the four color conjecture based on the proposition that every 3-connected cubic planar graph has a hamiltonian cycle. This proposition was disproved by Tutte [55] in 1946. However, until 1968, when Grinberg [19] proved his famous theorem (Theorem 1.2), such graphs were quite difficult to find. Grinberg's theorem can be easily used to create non-hamiltonian planar cubic graphs, like graph Γ of the previous section. Since 1968, several non-hamiltonian 3-connected planar cubic graphs have been found, the smallest of them is the Barnette-Bosák-Lederberg graph on 38 vertices [9, 33], see also [20]. The graph was discovered by the three scientists independently, about the same time. It is worth mentioning that Lederberg was not a mathematician or a computer scientist, but a molecular biologist (a really succesful one – he won a Nobel Prize in Phisiology or Medicine at the age of 33.) In 1986, Holton and McKay [25] (extending the results of many researchers) showed that there exists no 3-connected cubic planar non-hamiltonian graph on fewer vertices.
Chvátal [11] raised the question in 1973 whether there exists a cubic planar hypohamiltonian graph. This was answered by Thomassen [53], who found a sequence of such graphs on $94+4k$ vertices for every integer $k \geq 0$ in 1981. However, the question whether there exist smaller cubic hypohamiltonian graphs and whether there exists a positive integer N, such that for every integer $n \geq N$ there exists a cubic planar hypohamiltonian graph on n vertices remained open (both questions appear in the survey paper of Holton and Sheehan [26]). From the results of Aldred et al. [1] follows that there is no cubic planar hypohamiltonian graph on 42 or fewer vertices. They showed that every 3-connected, cyclically 4-connected cubic planar non-hamiltonian graph has at least 42 vertices and presented all such graphs on exactly 42 vertices. Since hypohamiltonian graphs are easily seen to be 3-connected and cyclically 4-connected, they must have at least 42 vertices in the cubic case. Moreover, all 42-vertex graphs presented in [1] have exactly one face with a degree not congruent to 2 modulo 3, and it is easy to see that cubic graphs with this property cannot be hypohamiltonian, as was observed by Thomassen [49].
Here we present a cubic planar hypohamiltonian graph on 70 vertices. Using the method of Thomassen for creating an $n+4$ vertex cubic hypohamiltonian graph from an n vertex cubic hypohamiltonian graph [53] this also shows that cubic planar hypohamiltonian graphs on $70+4m$ vertices exist for every even integer $m \geq 0$. Since $70 \equiv 94 \pmod 4$, this is not enough to answer the second open question, however we also give a cubic planar hypohamiltonian graph on 88 vertices, thus proving that cubic planar hypohamiltonian graphs on n vertices exist for every even number $n \geq 86$. Using our graphs on 70 and 88 vertices and another construction method of Thomassen [49], we can also show that a cubic planar hypotraceable graph exists on 340 vertices and on n vertices for every even number $n \geq 356$.
Using the 70-vertex cubic planar hypohamiltonian graph, the bounds on the numbers $\overline{C_3^2}$ and

$\overline{P_3^2}$ we have seen in the previous section are also improved.

We have seen that the size of the smallest cubic planar hypohamiltonian graph is at least 44 and at most 70. The next claims (that are extensions of the observation of Thomassen) may help to obtain a better lower bound. Let us denote the number of edges of a face T by $d(T)$ and for the sake of simplicity let us call a face F an *i-face* ($i = 0, 1, 2$), if $d(F) \equiv i \pmod 3$ and call the 0- and 1-faces together *non-2-faces*.

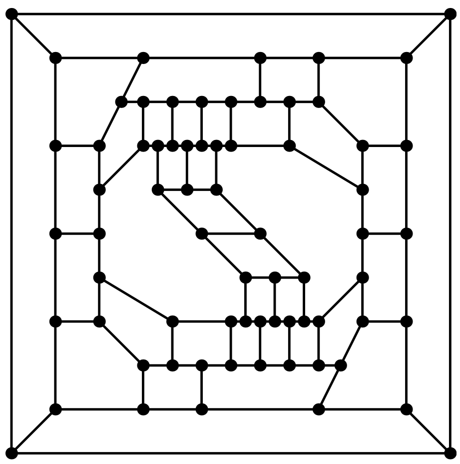
Claim 1.14 *A cubic planar hypohamiltonian graph has at least three non-2-faces.*

Proof. Let D be an arbitrary cubic planar hypohamiltonian graph. If D has only 2-faces, then the deletion of any vertex gives a graph D' with exactly one non-2-face, so D' is not hamiltonian, a contradiction. D cannot have exactly one non-2-face by the observation of Thomassen [49]. So let us assume that D has two non-2-faces A and B. It is easy to see that both A and B should be 0-faces, because the deletion of a vertex that is in one 1-face and two 2-faces gives a graph with exactly one non-2-face. Now the deletion of a vertex not in any of the 0-faces, but adjacent to a vertex that is in exactly one of the 0-faces gives a graph with exactly three 0-faces, of which two have two common edges. These cannot be on the same side of a hamiltonian cycle, therefore the equality in Grinberg's theorem cannot be satisfied, which finishes the proof. □

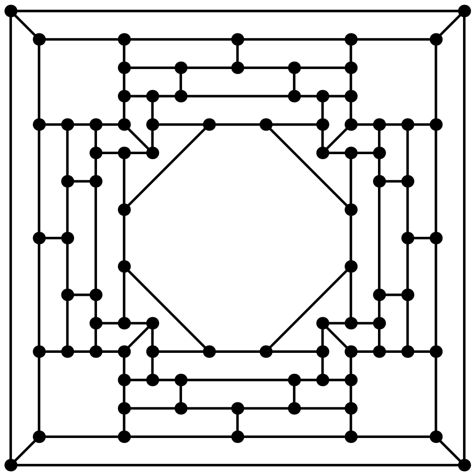
The following claim can be proved similarly.

Claim 1.15 *If a cubic planar hypohamiltonian graph has exactly three non-2-faces, then the three non-2-faces do not have a common vertex, moreover two 1-faces or a 1-face and a 0-face cannot be adjacent.*

Now we construct our (relatively) small cubic planar hypohamiltonian graphs. Let G be the following cubic planar graph on 70 vertices:



and let H be the following cubic planar graph on 88 vertices:



Theorem 1.16 (Araya-Wiener, 2011 [4]) *G and H are cubic planar hypohamiltonian graphs.*

Proof. Both G and H are obviously cubic and planar. Both have one face of degree 4, and four faces of degree 7, such that the face of degree 4 is adjacent to all faces of degree 7 and the degrees of the other faces are congruent to 2 modulo 3. By Proposition 2.1. of [53], G and H are non-hamiltonian (the proof is quite easy using Grinberg's theorem: in order to satisfy the equality in Grinberg's theorem modulo 3, a hamiltonian cycle should separate one of the five faces of degree 4 or 7 from the others, which is impossible in the case of G and H).
Now it remains to show that every vertex-deleted subgraph of G and H is hamiltonian. This can be found in [4]. □

Now we show some corollaries of the previous theorem. The most important corollary is the existence of cubic planar hypohamiltonian graphs on n vertices for every even number $n \geq 86$. This settles an open question in [26].

Corollary 1.17 (Araya-Wiener, 2011 [4]) *There exists a cubic planar hypohamiltonian graph on n vertices for every even number $n \geq 86$.*

Proof. The proof is quite obvious using a method of Thomassen [53]. Let T be a cubic planar hypohamiltonian graph on n vertices having a 4-cycle (a,b,c,d). The graph T' obtained from T by deleting the edges (a,b) and (c,d) and adding a new 4-cycle (a',b',c',d') and the edges (a,a'), (b,b'), (c,c'), (d,d') to T. Now it is easy to see that T' is also a cubic planar hypohamiltonian graph on $n+4$ vertices having a 4-cycle. By applying this operation iteratively on the graphs G and H we obtain cubic planar hypohamiltonian graphs on n vertices for every even number $n \geq 86$. □

Using another construction of Thomassen [51] a similar corollary for hypotraceable graphs can also be proved.

Corollary 1.18 (Araya-Wiener, 2011 [4]) *There exists a cubic planar hypotraceable graph on* 340 *vertices and on n vertices for every even number $n \geq 356$.*

Proof. We use a construction of Thomassen [51]. Let T_1, T_2, T_3, T_4, T_5 be cubic planar hypohamiltonian graphs and let x_i and y_i be adjacent vertices of T_i $(i = 1,2,3,4,5)$. Let furthermore the neighbours of x_i (resp. y_i), other than y_i (resp. x_i) be a_i and b_i (resp. c_i and d_i). Consider the disjoint union of the graphs $T_i - \{x_i, y_i\}$ and add to this graph the edges $(c_1,a_2),(c_2,a_3),(c_3,a_4),(c_4,a_5),(c_5,a_1)$ and the edges (d_1,b_2), (d_2,b_3), (d_3,b_4), (d_4,b_5),

(d_5, b_1). Now the resulting graph T is easily seen to be planar and cubic and by Lemma 3.1. of [51], it is also hypotraceable. If we choose each T_i to be isomorphic with G, then we obtain a cubic planar hypotraceable graph on 340 vertices. To obtain a cubic planar hypotraceable graph on $2k$ vertices for any $k \geq 178$ we just have to change T_1 in this construction to a cubic planar hypohamiltonian graph on $2k-270$ vertices (such a graph exists by Corollary 1.17, since $2k-270 \geq 86$). □

The next corollaries concern planar 3-connected graphs, in which every two vertices or edges are omitted by some longest cycle or path. First we improve a theorem of Schauerte and C. Zamfirescu. In [47] they showed (using a computer) that for any pair of edges e, f there exists a longest cycle in Thomassen's 94-vertex cubic planar hypohamiltonian graph [53] avoiding e and f. Using this observation and a method of T. Zamfirescu [67] they proved that there exists a cubic planar 3-connected graph on 8742 vertices, such that any pair of vertices is missed by a longest cycle.
The same property can also be checked easily for graph G by a computer, i.e. using a software like Mathematica or Maple.

Claim 1.19 *Let e and f be arbitrary edges of G. Then there exists a longest cycle in G that does not contain e and f.*

Corollary 1.20 (Araya-Wiener, 2011 [4]) *There exists a cubic planar 3-connected graph on 4830 vertices, such that any pair of vertices is missed by a longest cycle.*

Proof. We create a graph with the desired properties using a method of T. Zamfirescu [67]. Consider the 70-vertex cubic planar hypohamiltonian graph G, and let $V(G) = \{a_1, a_2, \ldots, a_{70}\}$. Let furthermore G' be the graph obtained from G by the deletion of a_{70} and assume that the neighbours of a_{70} are a_1, a_2, and a_3 in G. Now consider the graph Z consisting of 70 copies of G': $G'_1, G'_2, \ldots, G'_{70}$, such that we draw an edge between two copies G'_i and G'_j if and only if a_i and a_j are adjacent in G. These additional edges are always drawn between two vertices having degree 2 in the copies (that is, copies of a_1, a_2, or a_3). It is easy to see that Z is a cubic planar 3-connected graph on $69 \cdot 70 = 4830$ vertices. By Theorem 1.16, Proposition 1.19, and a theorem of T. Zamfirescu [67], any pair of vertices is missed by a longest cycle in Z. For completeness' sake we reformulate here the proof of Zamfirescu. Since G is hypohamiltonian, it is easy to see that the longest cycle of Z has length $68 \cdot 69 = 4692$ (one copy and one vertex of every other copy must be avoided, otherwise G would be hamiltonian, and a cycle of length 4692 is easy to find using the hypohamiltonicity of G). If the two vertices x and y we would like to avoid by a longest cycle are in the same copy, then simply consider a longest cycle avoiding this copy completely. Thus we may assume that x and y are in different copies. It is easy to see that there is a hamiltonian path between two of the vertices a_1, a_2, a_3 in every vertex-deleted subgraph of G'. Let x' (y') be that copy of a_1, a_2, or a_3 that is not the endvertex of such a hamiltonian path if we delete x (y). Now let us delete x, y, and one vertex from every other copy of G' from Z. Let us delete furthermore the additional edges incident to x' and y'. By Theorem 1.16 and Proposition 1.19 there is a cycle of length 4692 in the remaining graph, which proves the corollary. □

Finally, we improve the bounds of the previous section concerning the numbers $\overline{C_3^2}$ and $\overline{P_3^2}$.

Corollary 1.21 (Araya-Wiener, 2011 [4]) $\overline{C_3^2} \leq 2765$, $\overline{P_3^2} \leq 10902$.

Proof. The method is similar to the one used in Corollary 1.20. Recall that Γ is the planar hypohamiltonian graph on 42 vertices described in the precious section. The graph Γ' is obtained by

deleting any vertex of degree 3 from Γ. Now consider the graph Y consisting of 70 copies of Γ': $\Gamma'_1, \Gamma'_2, \ldots, \Gamma'_{70}$, such that we draw an edge between two copies Γ'_i and Γ'_j if and only if a_i and a_j are adjacent in G. These additional edges are always drawn between two vertices that are copies of the neighbours of the deleted vertex. It is easy to see that Y is a planar 3-connected graph on $41 \cdot 70 = 2870$ vertices. From the hypohamiltonicity of Γ and G, Proposition 1.19, and the mentioned theorem of Zamfirescu [67], any pair of vertices is missed by a longest cycle in Y. None of these properties are lost if we now contract the additional edges of Y (see [67]), obtaining a graph on $41 \cdot 70 - 105 = 2765$ vertices, which proves the first upper bound.
The second bound is proved similarly. First we take four copies of G' and an additional edge between any two copies (these edges are drawn between copies of a_1, a_2, or a_3 again). Denote the graph obtained in this way by X. Now we execute the same procedure as above, but this time we put the copies of Γ' into the graph X and then contract the additional edges to obtain a 3-connected planar graph, where every pair of vertices is missed by a longest path in $69 \cdot 4 \cdot 41 - ((105-3) \cdot 4 + 6) = 10902$ vertices (see [67]). □

2. fejezet

Minimum leaf spanning trees

Spanning tree optimization problems naturally arise in many applications, such as network design and connection routing. Several of these problems have an objective function based on the degrees of nodes of the spanning tree. This model is extremely useful when designing networks where the cost of devices to install depends highly on the needed routing functionality (ending, forwarding, or routing a connection). Typical examples are cost-efficient optical networks [41, 18, 39, 45] and water management systems [6].

In this chapter we are dealing with a problem of this kind. The problem MINLST (Minimum Leaf Spanning Tree) is to find a spanning tree of a given graph having a minimum number of leaves. Since hamiltonian paths (if exist) are the only spanning trees with exactly 2 leaves, MINLST is a generalization of the Hamiltonian path problem and therefore is *NP*-hard. Moreover, it is even hard to approximate: using a result of Karger, Motwani, and Ramkumar [31] concerning the problem of finding the longest path of a graph, Lu and Ravi [38] showed that no constant-factor approximation exists for the problem MINLST, unless $P = NP$.

From an optimization point of view, MINLST is equivalent to the problem of finding a spanning tree with a maximum number of internal nodes (non-leaves). However, we show that this latter problem (called MAXIST – Maximum Internal node Spanning Tree) has much better approximability properties. In Section 2.1 we give a linear time 2-approximation algorithm for the MAXIST problem based on depth first search. In Section 2.2 we show that a refined version of the depth first search algorithm provides a $\frac{3}{2}$-approximation on claw-free graphs (graphs not containing $K_{1,3}$ as an induced subgraph) and a $\frac{6}{5}$-approximation on cubic graphs. It is worth mentioning that for the problem of finding a spanning tree having a maximum number of leaves Lu and Ravi [38] gave a constant factor approximation algorithm, followed by a more efficient, near-linear time approximation [39].

One year after our paper was published, Salamon found the first approximation with a factor of less than 2 [40, 41] for graphs without degree 1 vertices, while the best known approximation has a factor of $\frac{3}{2}$ and is due to Li, Chen, and Wang [35]. For graphs without degree 1 vertices the best known approximation ratio is $\frac{4}{3}$ [36].

2.1. Maximizing the Number of Internal Nodes

In this section, we first give a linear-time algorithm (Algorithm ILST) that finds either a hamiltonian path of a given graph G or a spanning tree of G with independent leaves. Then we prove that such a tree has at least half times as many internal nodes as the optimal one. This shows that Algorithm ILST is a linear-time 2-approximation algorithm for the MAXIST problem.

The number of vertices of graph G is denoted by n, the number of edges by m. $V_i(G)$ ($V_{\geq i}(G)$) denotes the set of nodes having degree exactly i (at least i) in a graph G. $\text{comp}_G(X)$ denotes the number of the connected components of $G[X]$. Finally, given two nodes u and v of a tree T we denote by $P_T(u,v)$ the unique path in T connecting u and v.

Our algorithm is basically a depth-first search. However, it can happen that the leaves of a DFS-tree T are not independent. Thus, a single additional local replacement step might be needed to execute on T.
For a detailed discussion, let us recall that depth first search (DFS) (see for example [34]) is a traversal, that is, it visits the nodes of the graph one by one, such producing a spanning tree (the so-called DFS-tree) T of G *rooted* at some node r. We assign a unique *DFS number* to each node v, which is the rank of v in the order of visiting. Each non-root node v has a unique *parent* u, namely the node succeeding v on the path $P_T(v,r)$. The node v is called a *child* of u, and the nodes of the path $P_T(u,r)$ are the *ancestors* of v. A node having no child is called a *d-leaf*. Note that all d-leaves of T are also leaves of T, and only the root r can be a leaf of T without being a d-leaf. We recall a well-known property of DFS-trees.

Claim 2.1 *Let T be a DFS-tree of the undirected graph G. Then each edge of G connects two nodes of which one is an ancestor of the other in T. This implies that the d-leaves of T form an independent set of G.* □

Though the d-leaves of a DFS-tree T are independent, it may happen that the root of T is a leaf and is adjacent to some d-leaves of T. In this case, an additional replacement step is executed that decreases the number of leaves by one and also makes the leaves independent.

Algorithm 1: Independent Leaves Spanning Tree (ILST)

Input: An undirected graph $G = (V,E)$
Output: A spanning tree T of G with independent leaves
$T \leftarrow$ `DFS(`G`)` ; `// an arbitrary DFS tree of` G
$r \leftarrow$ the root of T;
if *T is not a hamiltonian path and $d_T(r) = 1$ and l is a d-leaf such that $(r,l) \in E(G)$* **then**
  `// ` r ` is a leaf and is adjacent to an other leaf ` l
  $x \leftarrow$ the branching node being closest to l in T;
  $y \leftarrow$ the neighbor of x on the path (l,x);
  Add edge (l,r) to T;
  Delete edge (x,y) from T;
return T;

Algorithm ILST produces a spanning tree, as the replacement step first creates a unique cycle by adding an edge to the DFS-tree and then removes an edge of this cycle. If the replacement step is applied then l and r become internal nodes and y becomes a leaf. Since y is not an ancestor of the other leaves, the spanning tree returned has independent leaves. The DFS-tree can be found in linear time. If we check $(r,l) \in E(G)$ for each d-leaf l during the traversal then the evaluation of the "if" condition needs only constant extra time. Once l is found, finding x and y and executing the replacement need linear time. Thus we have proved

Claim 2.2 *The algorithm ILST gives either a hamiltonian path or a spanning tree whose leaves form an independent set of G in $O(m)$ time.* □

In order to show that ILST is a 2-approximation, first we introduce the *cut-asymmetry* of a graph $G = (V, E)$ as $\mathrm{ca}(G) = \max_{X \subset V, X \neq \emptyset} (\mathrm{comp}_G(X) - \mathrm{comp}_G(V \setminus X))$. Lemma 2.3 shows a connection between cut-asymmetry and the number of leaves of trees.

Lemma 2.3 *Let T be an arbitrary tree on at least 3 vertices. Then* $\mathrm{ca}(T) = |V_1(T)| - 1$.

Proof. First observe that $\mathrm{comp}_T(V_1(T)) - \mathrm{comp}_T(V \setminus V_1(T)) = |V_1(T)| - 1$, since $V_1(T)$ is an independent set and $V \setminus V_1(T)$ spans a subtree. This implies $\mathrm{ca}(T) \geq |V_1(T)| - 1$.
To show that $\mathrm{ca}(T) \leq |V_1(T)| - 1$ let $X \subset V$ be a set of vertices for which $\mathrm{ca}(T) = \mathrm{comp}_T(X) - \mathrm{comp}_T(V \setminus X)$. For the sake of convenience, let $x = \mathrm{comp}_T(X)$ and $\bar{x} = \mathrm{comp}_T(V \setminus X)$. Then $e_T(X) = |X| - x$, and $e_T(V \setminus X) = n - |X| - \bar{x}$, thus

$$\sum_{v \in X} d_T(v) = 2e_T(X) + e_T(X, V \setminus X) = 2e_T(X) + n - 1 - (e_T(X) + e_T(V \setminus X))$$
$$= 2(|X| - x) + x + \bar{x} - 1 = 2|X| - x + \bar{x} - 1. \quad (2.1)$$

Observe that each internal node of X contributes to $\sum_{v \in X} d_T(v)$ by at least 2, yielding

$$|V_1(T) \cap X| \geq 2|X| - \sum_{v \in X} d_T(v) \quad (2.2)$$

Therefore, by (2.1) and (2.2), for the number of leaves of T, we have $|V_1(T)| \geq |V_1(T) \cap X| \geq 2|X| - \sum_{v \in X} d_T(v) \geq x - \bar{x} + 1 = \mathrm{ca}(T) + 1$, finishing the proof of the lemma. □

Now we apply the above lemma to prove the approximation ratio.

Theorem 2.4 (Salamon-Wiener, 2008 [42]) *The algorithm ILST is a 2-approximation for the* MAXIST *problem.*

Proof. We have seen that the algorithm is polynomial (actually, linear), so we only have to prove the approximation factor. Let T^* be a spanning tree with a maximum number of internal nodes, and let T be a spanning tree given by the algorithm. If T is a hamiltonian path, we are done, otherwise we apply Lemma 2.3: $|V_1(T^*)| = \mathrm{ca}(T^*) + 1 \geq \mathrm{comp}_{T^*}(V_1(T)) - \mathrm{comp}_{T^*}(V \setminus V_1(T)) + 1 \geq |V_1(T)| - |V \setminus V_1(T)| + 1 = 2|V_1(T)| - n + 1$, since $V_1(T)$ is an independent set of G (and thus also of T^*) by Claim 2.2. Thus $|V_{\geq 2}(T^*)| = n - |V_1(T^*)| \leq 2(n - |V_1(T)|) = 2|V_{\geq 2}(T)|$, proving the theorem. □

Notice that in DFS – and so in Algorithm ILST – the way of selecting the next node to visit is not fully specified. It says only that an unvisited neighbor of the currently visited node must be chosen. In Section 2.2, we present a refined version of DFS, which applies a node selection rule to (partially) resolve the non-deterministic behaviour of the original algorithm. We can profit of this refinement by obtaining a better approximation ratio for claw-free and cubic graphs.

2.2. Claw-free and Cubic Graphs

In this section, we deal with claw-free graphs (graphs not containing $K_{1,3}$ as an induced subgraph), and cubic graphs (3-regular graphs). First we present a refined version of the original DFS algorithm, called RDFS. Then we prove that RDFS approximates the MAXIST problem within a factor of $\frac{3}{2}$ for claw-free graphs, and within a factor of $\frac{6}{5}$ for cubic graphs.
RDFS is a depth first search in which we specify how to choose the next node of the traversal in the cases when DFS itself would choose arbitrarily from several candidates. The main idea

Algorithm 2: Refined DFS (RDFS)

```
Input: An undirected graph G = (V,E)
Output: An RDFS tree T of G
begin
    T ← (V,∅);
    foreach v ∈ V(G) do
        dfs[v] ← 0 ;                              // the DFS number of v
        actdeg[v] ← d_G(v) ;   // the number of non-visited neighbors
        of v
    k ← 0 ;              // the number of already visited vertices
    r ← a random vertex of G;
    RDFSNode(r);
    return T;

// Traversing from a node v
function RDFSNode(v)
begin
    k ← k+1;
    dfs[v] ← k;
    foreach neighbor w of v do actdeg[w] ← actdeg[w] − 1;
    ;
    while actdeg[v] > 0 do
        // We refine the original DFS by specifying how to
           choose
        // the next node to visit.
△       w ← a neighbor of v that has not been visited yet and that minimizes actdeg[.];
        Add edge (v,w) to T;
        RDFSNode (w);
```

is to select the vertex that has the minimum number of non-visited neighbors. For this purpose, we use the array "actdeg" to maintain the number of non-visited neighbors of each node.
RDFS differs from DFS only at line (△), where this latter one would choose a non-visited neighbor of v arbitrarily. Recall that DFS runs in linear time. At line (△) we make at most Δ steps to find the minimum, and this line is executed at most once for each edge of G. Thus the total running time is $O(\Delta m)$, where Δ is the maximum node-degree of G.
A tree produced by RDFS is called an *RDFS-tree*. We use the following notation. Let T be an RDFS-tree, and let l be a d-leaf of T such that $d_G(l) \geq 2$. $c(l)$ stands for the neighbor of l having the greatest DFS number such that $(l, c(l)) \in E(G) \setminus E(T)$ ($c(l)$ exists, because $d_G(l) \geq 2$). Let $g(l)$ denote the neighbor of $c(l)$ along the path $P_T(c(l), l)$. We also use the shorthand notions v_1, v_2, v_3 for the numbers $|V_1(T)|$, $|V_2(T)|$, and $|V_1(T)|$, respectively.
Now we prove a useful lemma concerning the degree of the node $g(l)$ in T.

Lemma 2.5 *Let T be an RDFS-tree and let l be a d-leaf of T. Then $d_T(g(l)) = 2$.*

Proof. Let us denote the set of vertices having DFS number greater than or equal to the DFS number of a vertex v by Y_v. It is obvious that $l \in Y_{c(l)}$.
Consider now that step of RDFS when we choose $g(l)$ to be the next visited vertex. By Rule (△) of RDFS, $d_{G[(Y_{g(l)}]}(g(l)) \leq d_{G[Y_{g(l)}](l)}$. By the definition of $c(l)$ and $g(l)$, $d_{G[(Y_{g(l)}]} = 1$, thus $d_{G[(Y_{g(l)}]}(g(l)) = 1$ (since $d_{G[(Y_{g(l)}]}(g(l)) \geq 1$ is obvious). Therefore $g(l)$ has only one child (and one parent, namely $c(l)$), so $d_T(g(l)) = 2$ indeed. □

Now we prove the approximation ratio for claw-free graphs.

Theorem 2.6 (Salamon-Wiener, 2008 [42]) *RDFS is a $\frac{3}{2}$-approximation for the* MAXIST *problem for claw-free graphs.*

Proof. We have seen that the algorithm is polynomial, so we have to check the approximation ratio. Let G be an arbitrary connected claw-free graph on n vertices and let T be an RDFS-tree of G. First notice that $d_T(v) \leq 3$ for any $v \in V(T) = V(G)$, otherwise the node v and three of its children would induce a subgraph $K_{1,3}$ in G, because of Claim 2.1. Thus our aim is to show that $v_2 + v_3 \geq \frac{2}{3} i_{opt}$, where i_{opt} is the number of internal nodes of an optimal spanning tree. Since T is a tree, we have $v_1 = v_3 + 2$.
Now we would like to find many nodes of degree 2 in T in order to show that the number of internal nodes is large.
For this we use Lemma 2.5. The problem is that in general, the nodes $g(l)$ (having degree 2 in T) are not necessarily distinct. However, we show that for claw-free graphs this is not the case.

Lemma 2.7 *Let T be an RDFS-tree of G and let l and l' be d-leaves of T, such that $d_G(l) \geq 2$ and $d_G(l') \geq 2$. Then $g(l) \neq g(l')$.*

Proof of Lemma 2.7. Suppose to the contrary that $g(l) = g(l')$. It is obvious that $c(l)$ and $c(l')$ are ancestors of l and l', respectively, thus from $d_T(g(l)) = 2$ (Lemma 2.5) follows that $c(l) = c(l')$. Now consider the induced subgraph $S = G[\{c(l), l, l', g(l)\}]$. The vertices l, l' and $g(l)$ are all adjacent to $c(l)$ in G. On the other hand, l and l' are d-leaves of T, thus cannot be adjacent in G. Moreover, $g(l)$ cannot be adjacent to either l or l' in $G \setminus T$, because $g(l)$ clearly has greater DFS number than $c(l) = c(l')$. Since $(l, g(l))$ and $(l', g(l))$ are not edges of T either (otherwise $g(l)$ could not be a common ancestor of l and l'), the induced subgraph S is isomorphic to $K_{1,3}$, a contradiction. □

Thus we have found as many nodes of degree 2 in T as the number of those d-leaves that has degree at least 2 in G. Let us denote the number of vertices having degree 1 in G by w. These

vertices are clearly leaves of any spanning tree of G, so the optimal spanning tree has at most $\min(n-w, n-2)$ internal nodes.
Now we consider two cases.
Case 1. $w = 0$. Now every d-leaf has degree at least two in G, thus $v_2 \geq v_1 - 1$. Since $v_1 = v_3 + 2$ and $n = v_1 + v_2 + v_3$, we have $v_2 + v_3 \geq \frac{2}{3}(n-2)m \geq \frac{2}{3}i_{opt}$.
Case 2. $w \geq 1$. It suffices to show that $v_2 + v_3 \geq \frac{2}{3}(v_1 + v_2 + v_3 - w)$. Since now the graph G has a vertex of degree one, the root of T has degree one in G, because of Rule (△) of RDFS. Thus now we have $v_2 \geq v_1 - w$, so it is enough to show that $v_2 + v_3 \geq \frac{2}{3}(2v_2 + v_3)$, that is, $v_3 \geq v_2$, which is equivalent to $v_1 - 2 \geq v_2$. If this latter inequality holds then we are done. Otherwise $v_2 \geq v_1 - 1$ holds, from which $v_2 + v_3 \geq \frac{2}{3}(n-2)m \geq \frac{2}{3}i_{opt}$ follows just like in Case 1. □

Theorem 2.8 (Salamon-Wiener, 2008 [42]) *RDFS is a $\frac{6}{5}$-approximation algorithm for the* MAXIST *problem for cubic graphs.*

Proof. We have to check the approximation ratio. Let G be an arbitrary connected cubic graph on n vertices and let T be a spanning tree of G given by RDFS. Obviously $d_T(v) \leq 3$ for any $v \in V(T) = V(G)$, thus $v_1 = v_3 + 2$. We show that $v_2 \geq 4v_1 - 6$, then some elementary computation gives $|V_{\geq 2}(T)| \geq \frac{5}{6}n - \frac{4}{3} \geq \frac{5}{6}(n-2)$, from which the theorem follows.
Let l be a d-leaf of T. Now l has two neighbors in $G \setminus T$, one of them is $c(l)$, call the other one $c'(l)$. It is obvious that $d_T(c(l)) = 2$ and also $d_T(c'(l)) = 2$ if $c'(l)$ is not the root. Furthermore, $d_T(g(l)) = 2$ by Lemma 2.5. Now let $h(l)$ be the only neighbor of $g(l)$ in $G \setminus T$. As a consequence of Rule (△) of RDFS and of Claim 2.1 we obtain that $h(l)$ is an ancestor of $g(l)$. Then either $h(l)$ is the root or $d_T(h(l)) = 2$. It is easy to check that if l and l' are two distinct d-leaves, then the sets $\{c(l), c'(l), g(l), h(l)\}$ and $\{c(l'), c'(l'), g(l'), h(l')\}$ can have only the root as a common element.
In this way to every d-leaf l we have associated 4 nodes (namely $c(l), c'(l), g(l)$, and $h(l)$). Among these nodes only the root can occur more than once and all the other nodes have degree 2. Since the root can occur at most twice (because of the 3-regularity) and the number of d-leaves is at least $v_1 - 1$, we have found $4(v_1 - 1) - 2$ distinct nodes of degree 2, that is, $v_2 \geq 4v_1 - 6$. □

Remark Actually, we have proved $v_1 \leq \frac{n}{6} + \frac{4}{3}$ for any spanning tree T obtained by RDFS for cubic graphs.

3. fejezet

Leaf-critical and leaf-stable graphs

In this chapter we unify the approaches of Chapter 1 and Chapter 2. More precisely, we extend the notions of hypohamiltonicity and hypotraceablity to spanning tree and path cover optimization problems, where a hamiltonian path (if exists) gives the optimum value (like the problems MINLST and MAXIST of Chapter 2). The usefulness of this unified approach is demonstrated by settling an open question of Gargano, Hammar, Hell, Stacho, and Vaccaro [18].
The *leaf number* of a graph G, denoted by $l(G)$ is the number of vertices of degree 1 in G, that is $l(G) = |V_1(G)|$. The *minimum leaf number* of a connected graph G, denoted by $\mathrm{ml}(G)$ is the minimum number of leaves of the spanning trees of G if G is not hamiltonian and 1 if G is hamiltonian.
We study nonhamiltonian graphs, whose vertex-deleted subgraphs have the same minimum leaf number. The deletion of a vertex obviously may increase the minimum leaf number of the graph, but there are vertices (e.g. leaves of an optimum spanning tree) whose deletion does not increase, or even decrease $\mathrm{ml}(G)$. However, it is easy to see that by deleting a vertex of a connected graph G, $\mathrm{ml}(G)$ can be decreased by at most one. Thus if G is nonhamiltonian and $\mathrm{ml}(G-v)$ does not depend on v, then either $\mathrm{ml}(G-v) = \mathrm{ml}(G)$ for every $v \in V(G)$ or $\mathrm{ml}(G-v) = \mathrm{ml}(G) - 1$ for every $v \in V(G)$.

Definition 3.1 *Let $l \geq 2$ be an arbitrary integer. A graph G is called l-leaf-critical, if $ml(G) = l$, but for every vertex v of G, $ml(G-v) = l-1$. A graph G is called l-leaf-stable, if $ml(G) = l$, and for every vertex v of G, $ml(G-v) = l$.*

At first sight it is not obvious whether such graphs exist at all. Actually, 2-leaf-critical and hypohamiltonian graphs are the same. This follows immediately from the definitions and the obvious fact that every hypohamiltonian graph is traceable. 3-leaf-critical graphs are also easy to find: they are the hypotraceable graphs; this also follows immediately from the definitions and the obvious fact that every hypotraceable graph has a spanning tree with 3 leaves.
Having seen that 2-leaf-critical and 3-leaf-critical graphs exist, one might expect that 2-leaf-stable and 3-leaf-stable graphs also exist. Actually, in Section 3.1 we show that l-leaf-stable and l-leaf-critical graphs exist for every integer $l \geq 2$, moreover we show that for n sufficiently large, planar 3-connected l-leaf-stable and l-leaf-critical graphs exist on n vertices and for n even and sufficiently large, cubic planar 3-connected l-leaf-stable and l-leaf-critical graphs exist on n vertices. Our construction is a generalization of a construction of Thomassen [49] that we have already used to obtain cubic hypotraceable graphs from cubic hypohamiltonian graphs (Corollary 1.18). In Section 3.2 we explore some properties of leaf-critical graphs of connectivity 2. Sections 3.3 and 3.4 show some interesting connections between our graphs and some known graph classes. Actually, in Section 3.3 we settle an open problem of

Gargano, Hammar, Hell, Stacho, and Vaccaro [18] concerning the existence of non-traceable, non-hypotraceable arachnoid graphs and in Section 3.4 we show that the graphs constructed in Section 3.1 belong to a family of graphs introduced by Grünbaum [21] in connection with the problem of finding graphs without concurrent longest paths. We conclude the chapter by discussing some open problems.

3.1. Constructions

To construct l-leaf-critical and l-leaf-stable graphs for $l \geq 3$ we use the notion of J-cells [27].

Definition 3.2 *A pair of vertices (a,b) of a graph G is said to be good if there exists a hamiltonian path of G between them. A pair of pairs of vertices of G $((a,b),(c,d))$ is said to be good if there exists a spanning subgraph of G consisting of two vertex-disjoint paths, one between a and b and another one between c and d.*

Definition 3.3 *(Hsu, Lin [27]) The quintuple (H,a,b,c,d) is a J-cell if H is a graph and $a,b,c,d \in V(H)$, such that*

1. *The pairs (a,d), (b,c) are good in H.*
2. *None of the pairs (a,b), (a,c), (b,d), (c,d), $((a,b),(c,d))$, $((a,c),(b,d))$ are good in H.*
3. *For each $v \in V(H)$ there is a good pair in $H-v$ among (a,b), (a,c), (b,d), (c,d), $((a,b),(c,d))$, $((a,c),(b,d))$.*

It is worth mentioning that flip-flops used by Chvátal (see page 11) to obtain many hypohamiltonian graphs [11] are special J-cells: in flip-flops the pair $((a,d),(b,c))$ is also good in H and for each $v \in V(H)$ there is a good pair in $H-v$ among (a,c), (b,d), $((a,b),(c,d))$, $((a,c),(b,d))$. J-cells can be obtained by deleting two adjacent vertices of degree 3 from a hypohamiltonian graph, as was observed by Thomassen, who used them to construct 3-connected hypotraceable graphs [51]. Our graphs are generalizations of this construction. (Actually, Thomassen did not name these graphs and used a somewhat different notation.) It is also easy to see that by adding two vertices u and v and the edges $(u,a),(u,d),(v,c),(v,d),(u,v)$ to a J-cell we obtain a hypohamiltonian graph (this is observed for flip-flops in [11], but the proof also works for J-cells). Thus the smallest J-cell is obtained from the Petersen graph by deleting two adjacent vertices (this J-cell is also a flip-flop).

Let $F_i = (H_i,a_i,b_i,c_i,d_i)$ be J-cells for $i = 1,2,\ldots,k$. Now we define the graphs G_k as follows. G_k consists of vertex-disjoint copies of the graphs $H_1,H_2,\ldots,H_k$, the edges $(b_i,a_{i+1}),(c_i,d_{i+1})$ for all $i = 1,2,\ldots k-1$, and the edges $(b_k,a_1),(c_k,d_1)$. We will consider the graphs H_i as (induced) subgraphs of G_k.

First we prove some useful properties of spanning trees of G_k.

Claim 3.4 *Let $l \geq 2$ and $k \in \{2l-1,2l\}$. Let furthermore A be an arbitrary subset of $B = \{(b_k,a_1),(c_k,d_1)\}$. Then G_k has a spanning tree T with l leaves, such that $E(T) \cap B = A$.*

Proof. Since the pairs (a_i,d_i) and (b_i,c_i) are good in H_i, the graphs $H_i \cup H_{i+1}$ are hamiltonian for each $i = 1,2,\ldots,k$ (where H_{k+1} is considered to be H_1). Let C_i be a hamiltonian cycle of $H_i \cup H_{i+1}$. First let us consider the case $k = 2l$. Let

$$E_1 = \bigcup_{j=0}^{l-1} E(C_{2j+1}) \cup \bigcup_{j=1}^{l-1} (b_{2j},a_{2j+1}), \quad E_2 = \bigcup_{j=0}^{l-1} E(C_{2j+1}) \cup \bigcup_{j=1}^{l-2} (b_{2j},a_{2j+1}) \cup (b_{2l},a_1),$$

$$E_3 = \bigcup_{j=1}^{l} E(C_{2j}) \cup \bigcup_{j=0}^{l-2} (b_{2j+1}, a_{2j+2}), \;\; E_4 = \bigcup_{j=0}^{l-1} E(C_{2j+1}) \cup \bigcup_{j=1}^{l-2} (c_{2j}, d_{2j+1}) \cup (c_{2l}, d_1).$$

Now the graphs $D_i = V(G_{2l}, E_i)$ are connected subgraphs of G_{2l} for $i = 1,2,3,4$ and it is easy to see that for $i = 1,2,3,4$, D_i contains a spanning tree S_i with l leaves, such that $E(S_1) \cap B = \emptyset$, $E(S_2) \cap B = (b_{2l}, a_1)$, $E(S_3) \cap B = B$, $E(S_4) \cap B = (c_{2l}, d_1)$.

For $k = 2l-1$ let P_i be a hamiltonian path of H_i between the vertices b_i and c_i for $i = 1,2$ and let

$$E_5 = E(P_1) \cup \bigcup_{j=1}^{l-1} E(C_{2j}) \cup \bigcup_{j=1}^{l-1} (b_{2j-1}, a_{2j}),$$

$$E_6 = E(P_1) \cup \bigcup_{j=1}^{l-1} E(C_{2j}) \cup \bigcup_{j=1}^{l-2} (b_{2j-1}, a_{2j}) \cup (b_{2l-1}, a_1),$$

$$E_7 = E(P_2) \cup \bigcup_{j=1}^{l-1} E(C_{2j+1}) \cup \bigcup_{j=1}^{l-1} (b_{2j}, a_{2j-1}) \cup (b_1, a_2),$$

$$E_8 = E(P_1) \cup \bigcup_{j=1}^{l-1} E(C_{2j}) \cup \bigcup_{j=1}^{l-2} (c_{2j-1}, d_{2j}) \cup (c_{2l-1}, d_1).$$

Now the graphs $D_i = V(G_{2l-1}, E_i)$ are connected subgraphs of G_{2l-1} for $i = 5,6,7,8$ and it is easy to see that for $i = 5,6,7,8$, D_i contains a spanning tree S_i with l leaves, such that $E(S_5) \cap B = \emptyset$, $E(S_6) \cap B = (b_{2l-1}, a_1)$, $E(S_7) \cap B = B$, $E(S_8) \cap B = (c_{2l-1}, d_1)$. □

Corollary 3.5 *Let $l \geq 2$. Then $ml(G_{2l}) \leq l$ and $ml(G_{2l+1}) \leq l+1$.* □

Claim 3.6 *Let T be a spanning tree of G_k. Then there are at most two indices i, such that all vertices in $V(H_i)$ has degree 2 in T.*

Proof. Suppose that all vertices in (say) $V(H_1)$ has degree 2 in T. Then $d_T(H_1)$ must be even (since $d_T(H_1) = \sum_{v \in V(H_1)} d(v) - 2|E(T[H_1])| = 2|V(H_1)| - 2|E(T[H_1])|$), thus $d_T(H_1)$ is 2 or 4. If $d_T(H_1) = 2$, then $T[H_1]$ is a hamiltonian path of H_1 and by the second property of J-cells the endvertices of the path are either a_1 and d_1 or b_1 and c_1 (w.l.o.g. assume they are a_1 and d_1). Therefore the edges leaving $V(H_1)$ in T are (b_k, a_1) and (c_k, d_1), thus there are no edges between $V(H_1)$ and $V(H_2)$ in T. If $d_T(H_1) = 4$, then $T[H_1]$ is a spanning subgraph of H_1 consisting of two vertex-disjoint paths. By the second property of J-cells, the endvertices of one of the paths are a_1 and d_1 and the endvertices of the other path are b_1 and c_1. Thus in this case there is no path between a_1 and b_1 in $T[H_1]$. It is clear now that if there is an index $i \neq 1,2$, such that all vertices in $V(H_i)$ has degree 2 in T, then T is not connected, a contradiction. □

Claim 3.7 *Let $l \geq 2$. Then $ml(G_{2l+1}) = l+1$ and $ml(G_{2l}) = l$.*

Proof. We have seen (Corollary 3.5) that $\text{ml}(G_{2l+1}) \leq l+1$ and $\text{ml}(G_{2l}) \leq l$, so we have to show that $\text{ml}(G_{2l+1}) \geq l+1$ and $\text{ml}(G_{2l}) \geq l$. Let us start by proving $\text{ml}(G_{2l+1}) \geq l+1$. Assume to the contrary that G_{2l+1} has a spanning tree T with at most l leaves. Then the number of vertices of degree at least 3 in T is at most $l-2$, thus the number of vertices not having degree 2 is at most $2l-2$. This means that there are at least three indices i, such that $V(H_i)$ only contains vertices of degree 2 in T, a contradicition by Claim 3.6. The proof of $\text{ml}(G_{2l}) \geq l$ is the same for $l \geq 3$: if G_{2l} has a spanning tree T with at most $l-1$ leaves, then the number of vertices of degree at least 3 in T is at most $l-3$, thus the number of vertices not having degree 2 is at most $2l-4$. This means that there are at least four indices i, such that $V(H_i)$ only contains vertices of degree 2 in T, once again a contradicition by Claim 3.6. For the case $l = 2$ we have to show that G_4 is not hamiltonian. This also follows easily from Claim 3.6: if C is a hamiltonian cycle of G_4 and e is an edge of $H_1[C]$ (that clearly exists, provided C exists), then $C-e$ is a spanning tree of G_4, such that $V(H_i)$ only contains vertices of degree 2 in $C-e$ for $i = 1,2,3$, a contradiction. □

Lemma 3.8 *Let $k \geq 4$ and v be an arbitrary vertex of G_k. Then $G_k - v$ has a spanning tree with $ml(G_{k-1})$ leaves.*

Proof. Let $l := \text{ml}(G_{k-1})$ and let us suppose w.l.o.g. that $v \in H_k$. Since F_k is a J-cell, at least one of the pairs (a_k,b_k), (c_k,d_k), (a_k,c_k), (b_k,d_k), $((a_k,b_k),(c_k,d_k))$, $((a_k,c_k),(b_k,d_k))$ is good in $H_k - v$. We distinguish six cases, based on which pair is good and construct a spanning tree T' of $G_k - v$ with l leaves in each case.

Case 1: (a_k,b_k) is good in $H_k - v$. Let P be a hamiltonian path between a_k and b_k in $H_k - v$. By Claims 3.4 and 3.7 there exists an l-leaf spanning tree T of G_{k-1}, such that $(b_{k-1},a_1) \in E(T)$ and $(c_{k-1},d_1) \notin E(T)$. Now let $E(T') = E(T) \setminus (b_{k-1},a_1) \cup (b_{k-1},a_k) \cup E(P) \cup (b_k,a_1)$. It is easy to verify that T' is a spanning tree of $G_k - v$ with l leaves.

Case 2: (c_k,d_k) is good in $H_k - v$. The construction is similar to that of the previous case. Let P be a hamiltonian path between c_k and d_k in $H_k - v$. By Claims 3.4 and 3.7 there exists an l-leaf spanning tree T of G_{k-1}, such that $(c_{k-1},d_1) \in E(T)$ and $(b_{k-1},a_1) \notin E(T)$. Let now $E(T') = E(T) \setminus (c_{k-1},d_1) \cup (c_{k-1},d_k) \cup E(P) \cup (c_k,d_1)$. Again, it is easy to verify that T' is a spanning tree of $G_k - v$ with l leaves.

Case 3: (a_k,c_k) is good in $H_k - v$. Let P be a hamiltonian path between a_k and c_k in $H_k - v$. By Claims 3.4 and 3.7 there exists an l-leaf spanning tree T of G_{k-1}, such that $(b_{k-1},a_1) \in E(T)$ and $(c_{k-1},d_1) \notin E(T)$, just like in Case 1. Actually, now we need that T possesses the additional property that $d_T(a_1) = 3$. It is easy to see that T can be chosen this way (consider the edge sets E_2 and E_6 in the proof of Claim 3.4). Let $E(T') = E(T) \setminus (b_{k-1},a_1) \cup (b_{k-1},a_k) \cup E(P) \cup (c_k,d_1)$. It is easy to verify that T' is a spanning tree of $G_k - v$ with l leaves (since a_1 has degree 2 in T').

Case 4: (b_k,d_k) is good in $H_k - v$. Again, the construction is similar to that of the previous case. Let P be a hamiltonian path between b_k and d_k in $H_k - v$. By Claims 3.4 and 3.7 there exists an l-leaf spanning tree T of G_{k-1}, such that $(c_{k-1},d_1) \in E(T)$ and $(b_{k-1},a_1) \notin E(T)$, and we choose T, such that it has the additional property that $d_T(d_1) = 3$. It is easy to see that T can be chosen this way (consider now the edge sets E_4 and E_8 in the proof of Claim 3.4). Let $E(T') = E(T) \setminus (c_{k-1},d_1) \cup (c_{k-1},d_k) \cup E(P) \cup (b_k,a_1)$. It is easy to verify that T' is a spanning tree of $G_k - v$ with l leaves (since d_1 has degree 2 in T').

Case 5: $((a_k,b_k),(c_k,d_k))$ is good in $H_k - v$. Let S be a spanning subgraph of $H_k - v$ consisting of two vertex-disjoint paths, one between a_k and b_k and the other one between c_k and d_k. By Claims 3.4 and 3.7 there exists an l-leaf spanning tree T of G_{k-1}, such that $(b_{k-1},a_1) \in E(T)$ and $(c_{k-1},d_1) \in E(T)$. Let $E(T') = E(T) \setminus (b_{k-1},a_1) \setminus (c_{k-1},d_1) \cup (b_{k-1},a_k) \cup (c_{k-1},d_k) \cup E(S) \cup (b_k,a_1) \cup (c_k,d_1)$. Again, it is easy to verify that T' is a spanning tree of $G_k - v$ with l leaves.

Case 6: $((a_k,c_k),(b_k,d_k))$ is good in $H_k - v$. Let S be a spanning subgraph of $H_k - v$ consisting of two vertex-disjoint paths, one between a_k and c_k and the other one between b_k and d_k. If k

is odd, then $k = 2l+1$ and if k is even, then $k = 2l$ by Claim 3.7. Suppose first that k is odd. Let us consider now the l-leaf spanning tree S_3 of G_{2l}, constructed in the proof of Claim 3.4. S_3 contains the edges (b_{2l}, a_1) and (c_{2l}, d_1), moreover by deleting these two edges from S_3 we obtain the following three components: a hamiltonian path of H_{2l} between b_{2l} and c_{2l}, a path inside H_1 starting at either a_1 or d_1 (say a_1) and a third component that contains d_1 and all other vertices of H_1 not contained in the previous path and all vertices of $H_2, H_3, \dots, H_{2l-1}$. Let now $E(T') = E(S_3) \setminus (b_{2l}, a_1) \setminus (c_{2l}, d_1) \cup (b_{2l}, a_{2l+1}) \cup (c_{2l}, d_{2l+1}) \cup E(S) \cup (b_{2l+1}, a_1) \cup (c_{2l+1}, d_1)$. It can be easily seen that the paths of S connect the components of $S_3 - (b_{2l}, a_1) - (c_{2l}, d_1)$ without creating a cycle, and therefore T' is a spanning tree of $G_{2l+1} - v$ with l leaves. Let us suppose now that k is even and let us consider the l-leaf spanning tree S_7 of G_{2l-1}, constructed in the proof of Claim 3.4. S_7 contains the edges (b_{2l-1}, a_1) and (c_{2l-1}, d_1), moreover by deleting these two edges from S_7 we obtain three components just like in the previous case. Let now $E(T') = E(S_7) \setminus (b_{2l-1}, a_1) \setminus (c_{2l-1}, d_1) \cup (b_{2l-1}, a_{2l}) \cup (c_{2l-1}, d_{2l}) \cup E(S) \cup (b_{2l}, a_1) \cup (c_{2l}, d_1)$. The same argument we have seen in the previous case shows that T' is a spanning tree of $G_{2l} - v$ with l leaves. □

Theorem 3.9 (Wiener, 2015 [62, 63]) *G_{2l+1} is $(l+1)$-leaf-critical for $l \geq 2$.*

Proof. We have seen that $\mathrm{ml}(G_{2l+1}) = l+1$ (Claim 3.7), so we have to show that $\mathrm{ml}(G_{2l+1} - v) = l$ for every vertex $v \in V(G_{2l+1})$. $\mathrm{ml}(G_{2l+1} - v) \leq l$ follows from Lemma 3.8, while $\mathrm{ml}(G_{2l+1} - v) \geq l$ is obvious. □

Theorem 3.10 (Wiener, 2015 [62, 63]) *G_{2l} is l-leaf-stable for $l \geq 2$.*

Proof. We have seen that $\mathrm{ml}(G_{2l}) = l$ (Claim 3.7), so we only have to show that $\mathrm{ml}(G_{2l} - v) = l$ for every vertex $v \in V(G_{2l})$. $\mathrm{ml}(G_{2l} - v) \leq l$ follows from Lemma 3.8. Now we prove that $\mathrm{ml}(G_{2l} - v) \geq l$. Let us suppose w.l.o.g. that $v \in H_{2l}$ and let us assume to the contrary that there exists a spanning tree T of $G_{2l} - v$ with at most $l-1$ leaves. Let a be an arbitrary neighbour of v in H_{2l}. Then $T' := T + (a, v)$ is a spanning tree of G_{2l} with at most l leaves. The number of vertices of degree at least 3 in T' is therefore at most $l-2$, thus the number of vertices not having degree 2 is at most $2l-2$. It is easy to see that two of these vertices, namely a and v are in H_{2l}: v is a leaf of T' and a has degree at least 3 in T', since a cannot be a leaf of T, otherwise T' would also have at most $l-1$ leaves, which is impossible by Claim 3.7. Thus there are at least 3 of the H_i's contain only vertices of degree 2 in T, which is a contradiction by Claim 3.6. □

Remark. By choosing the J-cells appropriétly, we obtain l-leaf-critical and l-leaf-stable graphs possessing some additional properties. It is easy to see that the graphs G_k are 3-connected for $k \geq 4$. A J-cell is said to be cubic, if the vertices a, b, c, d have degree 2 and the other vertices have degree 3. It is straightforward that if all J-cells used in the construction are cubic, then G_k is also cubic, while if all J-cells used are planar, then G_k is also planar. Since J-cells can be obtained from hypohamiltonian graphs by deleting two neighbouring vertices of degree 3 and planar hypohamiltonian graphs containing neighbouring vertices of degree 3 exist on n vertices for every n sufficiently large [65], it is easy to see that for n sufficiently large, planar l-leaf-stable and l-leaf-critical graphs exist on n vertices. Since J-cells obtained from a cubic hypohamiltonian graph are cubic and for n even and sufficiently large, cubic, planar hypohamiltonian graphs exist on n vertices [4], for n even and sufficiently large, cubic, planar l-leaf-stable and l-leaf-critical graphs also exist on n vertices. The smallest l-leaf-critical (l-leaf-stable) graph that can be obtained using our construction has $16l - 8$ ($16l$) vertices for $l \geq 3$ (for $l \geq 2$), using the J-cell obtained from the Petersen graph as F_i for all i.

3.2. Leaf-critical graphs of connectivity 2

We have seen that not much is known about the structure of hypohamiltonian and hypotraceable graphs, and obviously the same holds for leaf-critical graphs as well. (Though it is easy to see that all leaf-critical graphs are 2-connected and 3-edge-connected, but not necessarily 3-connected). In this section we give a characterization of the so-called 2-fragments of leaf-critical graphs generalizing a lemma of Thomassen (Lemma 5.1 of [51]).

Definition 3.11 *Let G be a non-complete graph with connectivity k and $X = \{x_1, x_2, \ldots, x_k\}$ be a cut of G. Let furthermore H be one of the components of $G - X$. Then $H + X$ is called a k-fragment of G, and X is called the vertices of attachment of H.*

Definition 3.12 *Let G be a graph, $a, b \in V(G)$. A subgraph F of G is said to be (a,b)-nice if at least one of the following three properties hold.*

1. *F is a tree and $l(F) \leq ml(G-a) - 1$.*
2. *F is a tree, $l(F) \leq ml(G-a)$ and a or b is a leaf of F.*
3. *F is a forest with two components, such that $l(F) \leq ml(G-a) + 1$, both a and b are leaves of F and they are in different components of F.*

If it does not cause any misunderstanding we just use the shorthand term nice, instead of (a,b)-nice.

Lemma 3.13 *Let G be a leaf-critical graph of connectivity 2 and $\{a,b\}$ a cut of G. Then $G-a-b$ has two components.*

Proof. Let the components of $G-a-b$ be $H_1, \ldots, H_r$ and assume to the contrary that $r \geq 3$. Let $l = \mathrm{ml}(G) - 1$. Since G is leaf-critical, $G-a$ has a spanning tree F_b and $G-b$ has a spanning tree F_a with l leaves. Let $A_i = F_a[H_i + a]$ and $B_i = F_b[H_i + b]$ for $i = 1, 2, \ldots, r$. Let furthermore $l_i(a) = 1$ if a is a leaf of A_i for $i = 1, 2, \ldots, r$ and let $l_i(a) = 0$ otherwise. Similarly, let $l_i(b) = 1$ if b is a leaf of B_i and $l_i(b) = 0$ otherwise. $A_1 \cup A_2 \cup \ldots \cup A_r = F_a$ and $B_1 \cup B_2 \cup \ldots \cup B_r = F_b$, thus $\sum_{i=1}^{r} l(A_i) = l + \sum_{i=1}^{r} l_i(a)$ and $\sum_{i=1}^{r} l(B_i) = l + \sum_{i=1}^{r} l_i(b)$. Let e be an edge between b and H_1 and f be an edge between a and H_1 (such edges clearly exist). Now $A_1 \cup A_2 \cup \ldots \cup A_{r-1} \cup B_r + e$ is a spanning tree of G with $\sum_{i=1}^{r-1} l(A_i) + l(B_r) - \sum_{i=1}^{r-1} l_i(a) - l_r(b)$ leaves and $B_1 \cup B_2 \cup \ldots \cup B_{r-1} \cup A_r + f$ is a spanning tree of G with $\sum_{i=1}^{r-1} l(B_i) + l(A_r) - \sum_{i=1}^{r-1} l_i(b) - l_r(a)$ leaves (since none of a and b is a leaf of any of the two spanning trees, because $r-1 \geq 2$). Thus these two spanning trees of G have $\sum_{i=1}^{r}(l(A_i) + l(B_i)) - \sum_{i=1}^{r}(l_i(a) + l_i(b)) = 2l$ leaves altogether. Therefore (at least) one of them has at most l leaves, a contradiction, since $l = \mathrm{ml}(G) - 1$. □

Lemma 3.14 *Let G_1 be a 2-fragment of the $(l+1)$-leaf-critical graph G with vertices of attachment a and b. Then G_1 has no (a,b)-nice spanning forest, but for any $v \in V(G_1)$, $G_1 - v$ has an (a,b)-nice spanning forest.*

Proof. We start along the same lines as in the previous proof. Let $l_1 = \mathrm{ml}(G_1 - a)$ and let the other 2-fragment of G with vertices of attachment a and b be G_2 (by Lemma 3.13 there are no more 2-fragments with the same vertices of attachment). Since G is $(l+1)$-leaf-critical, $G-a$ has a spanning tree F_b with l leaves and $G-b$ has a spanning tree F_a with l leaves. Let $A_1 = F_a[V(G_1 - b)], A_2 = F_a[V(G_2 - b)], B_1 = F_b[V(G_1 - a)], B_2 = F_b[V(G_2 - a)]$. Let furthermore $l_i(a) = 1$ if a is a leaf of A_i for $i = 1, 2$ and let $l_i(a) = 0$ otherwise. Similarly, let $l_i(b) = 1$ if

b is a leaf of B_i for $i = 1,2$ and let $l_i(b) = 0$ otherwise. $A_1 \cup A_2 = F_a$ and $B_1 \cup B_2 = F_b$, thus $l(A_1) + l(A_2) = l + l_1(a) + l_2(a)$ and $l(B_1) + l(B_2) = l + l_1(b) + l_2(b)$. Now we show that a and b are not adjacent in G. Suppose they are and consider the spanning trees $A_1 \cup B_2 + (a,b)$ and $A_2 \cup B_1 + (a,b)$ of G. These two trees have $l(A_1) + l(A_2) + l(B_1) + l(B_2) - l_1(b) - l_2(b) - l_1(a) - l_2(a) = 2l$ leaves altogether, thus (at least) one of them has at most l leaves, which contradicts the fact that G is $(l+1)$-leaf-critical.

Now let e be an edge between a and B_2 if a is a leaf of A_1 and let e be an edge between A_1 and b otherwise. Similarly, let f be an edge between a and B_1 if a is a leaf of A_2 and let f be an edge between A_2 and b otherwise. (such edges clearly exist, since G is 2-connected). Consider now the spanning trees $A_1 \cup B_2 + e$ and $A_2 \cup B_1 + f$ of G. Since a and b are not adjacent we have $l(A_1 \cup B_2 + e) = l(A_1) + l(B_2) - \max(l_2(b), l_1(a))$ and $l(A_2 \cup B_1 + f) = l(A_2) + l(B_1) - \max(l_2(a), l_1(b))$. Therefore these two spanning trees have $l(A_1) + l(B_2) - \max(l_2(b), l_1(a)) + l(A_2) + l(B_1) - \max(l_2(a), l_1(b)) = 2l + l_1(b) + l_2(b) + l_1(a) + l_2(a) - \max(l_2(b), l_1(a)) - \max(l_2(a), l_1(b)) \leq 2l + 2$ leaves altogether. Since none of these trees can have at most l leaves, both of them have exactly $l+1$ leaves, which implies $l(B_1) + l(A_2) = l(A_1) + l(B_2) = l + 2$ and $l_1(b) = l_2(b) = l_1(a) = l_2(a) = 1$, that is a is a leaf of both A_1 and A_2 and b is a leaf of both B_1 and B_2. We claim that B_1 is a minimum leaf spanning tree of $G_1 - a$. Indeed, if a spanning tree T of $G_1 - a$ with less than $l(B_1)$ leaves exists, then the spanning tree $T \cup A_2 + f$ of G has less than $l(B_1) + l(A_2) - 1 = l + 1$ leaves (since a is a leaf of A_2 and the edge f is incident to a), a contradiction. It can be similarly shown that A_1 is a minimum leaf spanning tree of $G_1 - b$. $A_1 \cup A_2$ is a spanning tree of $G - b$ and $B_1 \cup B_2$ is a spanning tree of $G - a$, therefore $l(A_1 \cup A_2) \geq l$ and $l(B_1 \cup B_2) \geq l$. Since a is a leaf of both A_1 and A_2 and b is a leaf of both B_1 and B_2, we have $l(A_1 \cup A_2) = l(A_1) + l(A_2) - 2$ and $l(B_1 \cup B_2) = l(B_1) + l(B_2) - 2$, that is $l(A_1) + l(A_2) \geq l + 2$ and $l(B_1) + l(B_2) \geq l + 2$. Since we have $l(B_1) + l(A_2) = l(A_1) + l(B_2) = l + 2$, this implies $l(B_1) = l(A_1)$ and $l(B_2) = l(A_2)$. Thus both A_1 and B_1 have $l_1 = l(G_1 - a)$ leaves. It is obvious now (by symmetry) that A_2 and B_2 are minimum leaf spanning trees of $G_2 - b$ and $G_2 - a$, respectively. Let $l_2 = l(A_2) = l(B_2)$, then we have $\mathrm{ml}(G_2 - a) = l_2$ and $l_1 + l_2 = l + 2$.

Now we prove that G_1 has no nice spanning forest. First let us show that $\mathrm{ml}(G_1) \geq l_1$. Assume to the contrary that a spanning tree T of G_1 with less than l_1 leaves exists. Then the spanning tree $T \cup A_2$ of G has at most $l_1 - 1 + l_2 - 1 = l$ leaves (since a is a leaf of A_2), a contradiction. The proof of the fact that G_1 has no spanning tree with at most l_1 leaves where a or b is a leaf is basically the same: if T is such a spanning tree, then $T \cup A_2$ or $T \cup B_2$ is a spanning tree of G with at most $l_1 - 1 + l_2 - 1 = l$ leaves, a contradiction. Finally let us show that G_1 has no spanning forest with at most $l_1 + 1$ leaves consisting of two trees, such that a is a leaf of one of the trees and b is a leaf of the other tree. Suppose F is such a forest and consider the spanning tree $F \cup B_2 + e$ of G. This has at most $(l_1 + 1) + l_2 - 2 - 1 = l$ leaves, a contradiction. It is now obvious by symmetry that G_2 has no nice spanning forest either.

Let us prove now that $G_1 - v$ has a nice spanning forest for each $v \in V(G_1)$. If $v = a$ then B_1 is a nice spanning tree of $G_1 - v$, since it has l_1 leaves and b is a leaf of B_1. If $v = b$ then A_1 is a nice spanning tree of $G_1 - v$, since it has l_1 leaves and a is a leaf of A_1. Suppose now that $v \neq a, b$. Since G is $(l+1)$-leaf-critical, $G - v$ has a spanning tree F with l leaves. Let $F_1 = F[V(G_1 - v)]$ and $F_2 = F[V(G_2)]$. Since $(a,b) \notin E(G)$, exactly one of F_1 and F_2 is connected (the one which contains the unique $a - b$ path in F). It is also obvious that the F_i that is not connected has two components: one containing a and the other one containing b. Now we distinguish two cases based on whether F_1 is connected.

Case 1: F_1 is connected. F_1 is obviously a spanning tree of $G_1 - v$ and F_2 is a spanning forest of G_2 with two components, one containing a and the other one containing b. Thus either F_2 has

at least l_2+2 leaves or (at least) one of a and b is not a leaf of F_2. Now we distinguish three cases and show that in each case F has at most l_1-2 leaves in G_1-v.
Case 1.1: $l(F_2)\geq l_2+2$. Then F has at most $l-((l_2+2)-2)=l_1-2$ leaves in G_1-v, since the leaves of F_2 different from a and b are also leaves of F.
Case 1.2: $l(F_2)<l_2+2$ and exactly one of a and b is a leaf of F_2. W. l. o. g. assume b is a leaf and a is not. Then $l(F_2)\geq l_2+1$, since G_2 has no spanning tree with at most l_2 leaves, where b is a leaf and therefore G_2 obviously has no spanning forest with at most l_2 leaves, where b is a leaf. Thus F has at most $l-(l_2+1)-1)=l_1-2$ leaves in G_1-v, since the leaves of F_2 different from b are also leaves of F.
Case 1.3: $l(F_2)<l_2+2$ and none of a and b is a leaf of F_2. Now $l(F_2)\geq l_2$ (since G_2 has no spanning tree with less than l_2 leaves, therefore G_2 obviously has no spanning forest with less than l_2 leaves). F has at most $l-l_2=l_1-2$ leaves in G_1-v, because now each leaf of F_2 is also a leaf of F.
F has at most l_1-2 leaves in G_1-v, $l(F_1)\leq l_1$ (the leaves of F_1 different from the leaves of F can only be a and b) and $l(F_1)=l_1$ is possible only if both a and b are leaves of F_1, therefore F_1 is nice. (Actually, it is easy to check that $l(F_1)<l_1$ is not possible.)

Case 2: F_1 has two components: one containing a and the other one containing b. Now F_2 is a spanning tree of G_2, thus either $l(F_2)\geq l_2+1$ or none of a and b is a leaf of F_2. Here we distinguish two cases, depending on $l(F_2)$.
Case 2.1 $l(F_2)\geq l_2+1$. Then F has at most $l-((l_2+1)-2)=l_1-1$ leaves in G_1-v, since the leaves of F_2 different from a and b are also leaves of F. Now if both a and b are leaves of F_1, then $l(F_1)\leq l_1+1$ and F_1 is a nice spanning forest. If exactly one of a and b is a leaf, then $l(F_1)\leq l_1$. Now we can add an edge to F_1 to obtain a spanning tree of G_1-v with l_1-1 leaves or with l_1 leaves, such that a or b is a leaf. In both cases the spanning tree obtained is nice. Finally, if none of a and b is a leaf, then $l(F_1)\leq l_1-1$ and therefore a spanning tree T of G_1-v with at most l_1-1 leaves exists; by definition T is nice.
Case 2.2: $l(F_2)<l_2+1$. Then none of a and b is a leaf of F_2 and since $\mathrm{ml}(G_2)\geq l_2$, we have $l(F_2)=l_2$. Then F has at most $l-l_2=l_1-2$ leaves in G_1-v, since the leaves of F_2 are also leaves of F. This means that $l(F_1)\leq l_1$ and $l(F_1)=l_1$ is possible only if both a and b are leaves of F_1. Now by adding an edge between the components of F_1 we obtain a nice spanning tree of G_1-v and the proof is finished. □

The properties of Lemma 3.14 characterize the leaf-critical 2-fragments. In order to prove this, we still have to show that every graph possessing these properties is a 2-fragment of some leaf-critical graph. We prove a somewhat stronger lemma.

Claim 3.15 *Let G be a graph, $a,b\in V(G)$. If G has no nice spanning forest, but both $G-a$ and $G-b$ have a nice spanning forest, then $(a,b)\notin E(G)$, $ml(G-b)=ml(G-a)$, and any nice spanning forest F_b (F_a) of $G-a$ ($G-b$) is a tree, such that b (a) is a leaf of F_b (F_a).*

Proof. Let F_b and F_a be a nice spanning forest of $G-a$ and $G-b$, respectively. Since b (a) cannot be a leaf of F_a (F_b), F_a and F_b are trees, thus by definition $l(F_a)\leq \mathrm{ml}(G-b)$ and $l(F_b)\leq \mathrm{ml}(G-a)$. On the other hand, clearly $l(F_a)\geq \mathrm{ml}(G-b)$ and $l(F_b)\geq \mathrm{ml}(G-a)$ (since F_a and F_b are spanning trees), therefore $l(F_b)=\mathrm{ml}(G-a)$ and $l(F_a)=\mathrm{ml}(G-b)$ and by the definition of nice subgraphs b is a leaf of F_b, a is a leaf of F_a. If $(a,b)\in E(G)$ then $F_b+(a,b)$ is a nice spanning tree of G, a contradiction. $\mathrm{ml}(G-b)=\mathrm{ml}(G-a)$ is also easy to prove: no spanning tree F of $G-b$ can have less than $\mathrm{ml}(G-a)$ leaves, otherwise by adding an edge between F and b to F we would obtain a spanning tree of G with at most $\mathrm{ml}(G-a)$ leaves, where b is a leaf, a contradiction. Thus $\mathrm{ml}(G-b)\geq \mathrm{ml}(G-a)$, and $\mathrm{ml}(G-a)\geq \mathrm{ml}(G-b)$ can be proved similarly. □

Lemma 3.16 **(Wiener, 2015 [62, 63])** *Let G be a graph of connectivity 2 and $\{a,b\}$ a cut in G. Let G_1 and G_2 be 2-fragments of G with vertices of attachment a,b, such that G_i has no (a,b)-nice spanning forest, but for any $v \in V(G_i)$, $G_i - v$ has an (a,b)-nice spanning forest for $i = 1,2$. Then G is l-leaf-critical, where $l = ml(G_1 - a) + ml(G_2 - a) - 1$.*

Proof. By Lemma 3.13, $G = G_1 \cup G_2$. Let us prove first that $\mathrm{ml}(G) \leq l$. We have to find a spanning tree of G with at most l leaves: let F_1 be a nice spanning forest of $G_1 - a$ and F_2 be a nice spanning forest of $G_2 - b$. Then by Claim 3.15, F_1 and F_2 are trees and b is a leaf of F_1 and a is a leaf of F_2. By adding an edge between a and F_1 to $F_1 \cup F_2$ we obtain a spanning tree of G with at most $l(F_1) + l(F_2) - 1$ leaves. By Claim 3.15, $l(F_1) = \mathrm{ml}(G_1 - a)$ and $l(F_2) = \mathrm{ml}(G_2 - a)$ and we are done.
Now we prove $\mathrm{ml}(G) \geq l$. Let T be an arbitrary spanning tree of G. Then $T_1 = T[V(G_1)]$ and $T_2 = T[V(G_2)]$ are spanning forests of G_1 and G_2, respectively, such that one of them (say T_1) is a tree and the other one (T_2 then) consists of two trees, such that a and b are in different components. Since G_1 has no nice spanning forest, $l(T_1) \geq \mathrm{ml}(G_1 - a)$ and if $l(T_1) = \mathrm{ml}(G_1 - a)$, then none of a and b is a leaf. Since G_2 has no nice spanning forest, $l(T_2) \geq \mathrm{ml}(G_2 - a)$, furthermore if $l(T_2) = \mathrm{ml}(G_2 - a)$, then none of a and b is a leaf and if $l(T_2) = \mathrm{ml}(G_2 - a) + 1$, then at most one of a and b is a leaf. Thus T has at least $\mathrm{ml}(G_1 - a) - 1$ leaves in $V(G_1 - a - b)$ and at least $\mathrm{ml}(G_2 - a)$ leaves in $V(G_2 - a - b)$, that is T has at least $l = \mathrm{ml}(G_1 - a) + \mathrm{ml}(G_2 - a) - 1$ leaves altogether, which proves $\mathrm{ml}(G) \geq l$ and we have proved $\mathrm{ml}(G) = l$.
Now let us prove that for an arbitrary $v \in V(G)$ we have $\mathrm{ml}(G - v) = l - 1$. Obviously, it suffices to prove that $\mathrm{ml}(G - v) \leq l - 1$, that is $G - v$ has a spanning tree with $l - 1$ leaves. W.l.o.g. assume $v \in V(G_1)$. Suppose first that $v = a$. $G_1 - a$ and $G_2 - a$ have nice spanning forests $N_1 = F_1$ and N_2. By Claim 3.15, N_1 and N_2 are trees with at most $\mathrm{ml}(G_1 - a)$ and $\mathrm{ml}(G_2 - a)$ leaves, respectively and b is a leaf of both N_1 and N_2. Thus $N_1 \cup N_2$ is a spanning tree of $G - a$ with at most $\mathrm{ml}(G_1 - a) + \mathrm{ml}(G_2 - a) - 2 = l - 1$ leaves. The case $v = b$ is proved similarly. Let us suppose now that $v \in V(G_1 - a - b)$. $G_1 - v$ has a nice spanning forest H. Now we distuingish three cases, based on whether the first, second, or third property of nice subgraphs holds for H.
Case 1: H is a tree and $l(H) \leq \mathrm{ml}(G_1 - a) - 1$. Then $H \cup N_2$ is a spanning tree of $G - v$ with at most $(\mathrm{ml}(G_1 - a) - 1) + \mathrm{ml}(G_2 - a) - 1 = l - 1$ leaves.
Case 2: H is a tree, $l(H) \leq \mathrm{ml}(G_1 - a)$ and a or b is a leaf of H. Then $H \cup F_2$ and $H \cup N_2$ are spanning trees of $G - v$ and one of them has at most $\mathrm{ml}(G_1 - a) + \mathrm{ml}(G_2 - a) - 2 = l - 1$ leaves.
Case 3: H is a forest with two components, such that $l(H) \leq \mathrm{ml}(G_1 - a) + 1$, both a and b are leaves of H and they are in different components of H. Let g be an edge between b and F_2. Then $H \cup F_2 \cup g$ is a spanning tree of $G - v$ with at most $(\mathrm{ml}(G_1 - a) + 1) + \mathrm{ml}(G_2 - a) - 2 - 1 = l - 1$ leaves and the proof is complete. □

Now the following generalization of Lemma 5.1 of [49] is an immediate consequence of Lemmas 3.14 and 3.16.

Theorem 3.17 **(Wiener, 2015 [62, 63])** *Let G be a graph, $a,b \in V(G)$. G is a 2-fragment of a leaf-critical graph with vertices of attachment a and b if and only if G has no (a,b)-nice spanning forest, but for any $v \in V(G)$, $G - v$ has an (a,b)-nice spanning forest.* □

Another interesting corollary is that leaf-critical graphs contain 2-fragments of other leaf-critical graphs with a much smaller minimum leaf number.

Corollary 3.18 *If G is an l-leaf-critical graph of connectivity 2, then it contains an r-leaf-critical 2-fragment, where* $r \leq \left\lfloor \frac{l+3}{2} \right\rfloor$.

Proof. Let $\{a,b\}$ be a cut in G and let G_1 and G_2 be the 2-fragments of G with vertices of attachment a,b (by Lemma 3.13, $G = G_1 \cup G_2$). Suppose w.l.o.g. that $\mathrm{ml}(G_1 - a) \leq \mathrm{ml}(G_2 - a)$ and let G_3 be a 2-fragment of a 3-leaf-critical (that is, hypotraceable) graph with vertices of attachment x,y (Such a 2-fragment exists [49]). Let H be the graph obtained by identifying the vertices a and x and the vertices b and y of the graph $G_1 \cup G_3$. By Lemma 3.14 and Lemma 3.16 H is $(\mathrm{ml}(G_1 - a) + 1)$-leaf-critical. Since $\mathrm{ml}(G_1 - a) \leq \frac{l+1}{2}$ and G_1 is a 2-fragment of H, the proof is finished. □

3.3. Path-critical and arachnoid graphs

The *path-covering number* of G, denoted by $\mu(G)$ is the minimum number of vertex-disjoint paths that cover the vertices of G (a path may consist of just one vertex). The *branch number* of G, denoted by $s(G)$ is the minimum number of branch vertices (vertices of degree at least 3) of the spanning trees of G. Gargano, Hammar, Hell, Stacho, and Vaccaro [18] defined the notion of *spanning spiders*: these are spanning trees with at most one branch. The spider is said to be *centred* at the branch vertex (if there is any, otherwise the spider is centred at any of the vertices). They studied the parameter $s(G)$ and graphs with $s(G) \leq 1$. They also defined *arachnoid* graphs; these are graphs that have a spanning spider centred at any of their vertices. Traceable graphs are obviously arachnoid, and Gargano et al. observed that hypotraceable graphs are also easily seen to be arachnoid [18]. However, they did not find any other arachnoid graphs, and asked the question whether they exist. In this section we answer this question in the affirmative, moreover, we show that for any prescribed graph H, there exists a non-traceable, non-hypotraceable, arachnoid graph that contains H as an induced subgraph. To this end we introduce path-critical graphs.

Definition 3.19 *Let $\mu \geq 2$ be an integer. A graph G is μ-path-critical if $\mu(G) = \mu$ and $\mu(G - v) = \mu - 1$ for each $v \in V(G)$.*

It is easy to see that the 2-path-critical graphs are the hypotraceable graphs, but the existence of μ-path-critical graphs for $\mu \geq 3$ is far from from obvious. The next theorem shows that some of the leaf-critical graphs G_k we have constructed in Section 3.1 are also path-critical.

Theorem 3.20 (Wiener, 2015 [64]) *Let $k \geq 0$ be an integer. Then for any $v \in V(G_{4k+5})$ we have $\mu(G_{4k+5} - v) = \mu(G_{4k+5}) - 1 = k+1$, thus G_{4k+5} is $(k+2)$-path-critical.*

Proof. We need the following lemma.

Lemma 3.21 *G_4 has a hamiltonian path P, such that there is no edge of P between H_1 and H_4 and for any vertex $v \in V(G_5)$ there is a hamiltonian path P of $G_5 - v$, such that there is no edge of P between H_1 and H_5.*

Proof. The first part of the claim is easy to see: there is a hamiltonian path of H_i between b_i and c_i and a hamiltonian path of H_{i+1} between a_{i+1} and d_{i+1}, by the first property of J-cells, thus $H_1 \cup H_2$ and $H_3 \cup H_4$ are hamiltonian, therefore there is a hamiltonian path P_1 of $H_1 \cup H_2$ starting at b_2 and a hamiltonian path P_3 of $H_3 \cup H_4$ starting at a_3. Now $E(P_1) \cup (b_2, a_3) \cup E(P_3)$ is a hamiltonian path of G_4 without edges between H_1 and H_4. Let now $F = (H,a,b,c,d)$ be any of the J-cells used in the construction of G_5 and let us check whether (a,b), (a,c), (b,d), (c,d), $((a,b),(c,d))$, or $((a,c),(b,d))$ is good in $H-v$. Let us number the J-cells used to construct G_5, such that $H_3 = H$ in the first four cases, and $H_2 = H$ in the last two cases. If $(a,b) = (a_3,b_3)$ is good in $H_3 - v$, then let P be a hamiltonian path of $H_3 - v$ between a_3 and b_3. We have seen that $H_i \cup H_{i+1}$ is hamiltonian, therefore $H_i \cup H_{i+1}$ has a hamiltonian path starting at any of its vertices. Let P_1 be a hamiltonian path of $H_1 \cup H_2$ starting at b_2 and let P_4 be a hamiltonian path of $H_4 \cup H_5$ starting at a_4. Then $E(P_1) \cup (b_2,a_3) \cup E(P) \cup (b_3,a_4) \cup E(P_4)$ is the edge set of a hamiltonian path of $G_5 - v$ and does not contain any edges between H_1 and H_5. The cases when (a,c), (b,d), or (c,d) is good is dealt with similarly. If $((a,b),(c,d)) = ((a_2,b_2),(c_2,d_2))$ is good in $H_2 - v$, then let Q be the union of the vertex-disjoint $a-b$ and $c-d$ paths that cover all vertices of $H_2 - v$. Let furthermore Q_1 be a hamiltonian path between b_1 and c_1 in H_1, and Q_3 be a hamiltonian path between d_3 and either b_3 or c_3 (say w.l.o.g. b_3) in $G_3 - a_3$. Q_1 and Q_3 exist since F_1 and F_3 are J-cells. Then $E(Q_1) \cup (b_1,a_2) \cup (c_1,d_2) \cup E(Q) \cup (b_2,a_3) \cup (c_2,d_3) \cup E(Q_3) \cup (b_3,a_4) \cup E(P_4)$ is again the edge set of a hamiltonian path of $G_5 - v$ that does not contain any edges between H_1 and H_5. The case when $((a,c),(b,d))$ is good is dealt with similarly. □

Now let us denote $G_{4k+5}[\cup_{i=n}^{m} V(H_i)]$ by $G(n,m)$ for $1 \le n < m \le 4k+5$. It is obvious that if $n \ne 1$ or $m \ne 4k+5$, then $G(n,m)$ is isomorphic to some graph $G_{m-n+1} - (b_{m-n+1},a_1) - (c_{m-n+1},d_1)$, thus $G(n,m)$ is traceable if $m = n+3$ and $G(n,m) - v$ is traceable for any $v \in G(n,m)$ if $m = n+4$ by Lemma 3.21. Since $G(1,4), G(5,8), \dots, G(4k-3,4k)$ and $G(4k+1,4k+5) - v$ are all traceable, the vertices of $G_{4k+5} - v$ can be covered by $k+1$ vertex-disjoint paths, that is $\mu(G_{4k+5} - v) \le k+1$ for any $v \in V(G)$. On the other hand, we show that $\mu(G_{4k+5}) \ge k+2$. Assume to the contrary that there are at most $k+1$ vertex-disjoint paths that cover the vertices of G_{4k+5}. Since G_{4k+5} is connected, it is possible to add some (at most k, but it is irrelevant) edges to these paths to obtain a spanning tree of G_{4k+5} with at most $2k+2$ leaves. On the other hand, by Lemma 3.7, $\mathrm{ml}(G_{4k+5}) \ge 2k+3$, a contradiction. Since for any graph G, $\mu(G) \le \mu(G-v)+1$ is obvious, we have $k+1 \le \mu(G_{4k+5}) - 1 \le \mu(G_{4k+5} - v) \le k+1$, and the theorem is proved. □

Now we are ready to construct non-traceable, non-hypotraceable, arachnoid graphs. Let G_k^j be the graph obtained from G_k by adding j new vertices $u_1, u_2, \dots, u_j$ and edges between u_i and every vertex of G_k to G_k for $i = 1,2,\dots,j$.

Theorem 3.22 (Wiener, 2015 [63, 64]) *G_{4k+5}^k is an arachnoid graph that is neither traceable, nor hypotraceable for any $k \ge 1$.*

Proof. Let $G = G_{4k+5}^k$. We have to show that for any $w \in V(G)$, G has a spanning spider centred at w. Let v be a neighbour of w, such that $v \in G_{4k+5}$ (such a v clearly exists). Now by Theorem 3.20, the vertices of $G_{4k+5} - v$ can be covered by $k+1$ vertex-disjoint paths, thus using the vertices $u_1, \dots, u_k$ (that are all connected to all vertices of G_{4k+5}) a hamiltonian path of $G - v$ is easy to obtain. Now by adding the edge (v,w) to this path we obtain a spanning spider of G centred at w, therefore G is arachnoid, indeed.
Now we show that G is not traceable. Assume to the contrary that there exists a hamiltonian path P of G and let us delete the vertices $u_1, \dots, u_k$ from P. We obtain at most $k+1$ vertex-disjoint paths, such that they cover the vertices of G_{4k+5}, which is a contradiction, by Theorem 3.20.
Finally, we have to show that G is not hypotraceable. It is easy to see that $G - u_i$ is not traceable, the proof is the same as the proof of the non-traceablity of G (by deleting the u_i's we would obtain at most k paths, instead of at most $k+1$). □

Remark. It is easy to see that adding any edges between the u_i's does not make the graph either traceable or hypotraceable (while the arachnoid property is obviously preserved), therefore we can obtain a non-traceable, non-hypotraceable, arachnoid graph that contains any prescribed graph H as an induced subgraph.

3.4. Longest paths avoiding certain vertices

We have mentioned that Walther [58] settled Gallai's problem whether there exist graphs without concurrent longest paths. Exploring this area further, T. Zamfirescu [66] and Grünbaum [21] defined several numbers and graph families based on properties of longest paths and cycles, of which we have already studied the numbers $\overline{C_3^1}$, $\overline{C_3^2}$, $\overline{P_3^1}$, $\overline{P_3^2}$ in Section 1.1 and the numbers $\overline{C_3^2}$ and $\overline{P_3^2}$ again in Section 1.2. Now we are dealing with the graph family $\Pi(j,m)$ introduced by Grünbaum [21]. This family consists of graphs having m more vertices than their longest paths have, such that for each j vertices there is a longest path missing these j vertices; e.g. $\Pi(1,1)$ is the class of hypotraceable graphs.

Theorem 3.23 (Wiener, 2015 [63]) *$G_{k+4} \in \Pi(1,k)$ for all $k \geq 1$.*

Proof. First we prove that for any $v \in V(G_{k+4})$ there is a path in G_{k+4} that misses exactly k vertices, one of which is v. W.l.o.g. we may suppose that $v \in H_3$. Since F_3 is a J-cell, at least one of the pairs (a_3,b_3), (a_3,c_3), (b_3,d_3), (c_3,d_3), $((a_3,b_3),(c_3,d_3))$, $((a_3,c_3),(b_3,d_3))$ is good in $H_3 - v$. It is easy to see that by symmetry reasons we may suppose that either (a_3,b_3) or $((a_3,b_3),(c_3,d_3))$ is good in $H_3 - v$.
If (a_3,b_3) is good, then let P_3 be a hamiltonian path between a_3 and b_3 in $H_3 - v$ and P_2 be a hamiltonian path of $H_1 \cup H_2$ starting at b_2 (since $H_1 \cup H_2$ is hamiltonian, such a path exists). If $k \geq 2$ then let us consider now the graph $H_4 - d_4$. Since F_4 is a J-cell, this graph contains a hamiltonian path between a_4 and either b_4 or c_4. By symmetry reasons we may suppose that there is a hamiltonian path between a_4 and b_4 in $H_4 - d_4$, let us call it P_4. If $k \geq 3$ then the paths P_i are defined similarly for $i = 5,6,\dots,k+2$. Finally, let P_{k+3} be a hamiltonian path of $H_{k+3} \cup H_{k+4}$ starting at a_{k+4} (since $H_{k+3} \cup H_{k+4}$ is hamiltonian, such a path exists). Now

$$P_2 \cup (b_2,a_3) \cup P_3 \cup (b_3,a_4) \cup \dots \cup (b_{k+2},a_{k+3}) \cup P_{k+3}$$

is a path in G_{k+4} missing exactly k vertices, one of which is v.
If $((a_3,b_3),(c_3,d_3))$ is good in $H_3 - v$, then let Q be a spanning subgraph of $H_3 - v$ consisting of two vertex-disjoint paths, one between a_3 and b_3 and the other one between c_3 and d_3. Let P_2 be a hamiltonian path between b_2 and c_2 in H_2 (since F_2 is a J-cell, such a hamiltonian path exists). Since F_4 is a J-cell, $H_4 - d_4$ contains a hamiltonian path between a_4 and either b_4 or c_4. By symmetry reasons we may suppose that there is a hamiltonian path between a_4 and b_4 in $H_4 - d_4$, let it be P_4. If $k \geq 2$ then the paths P_i are defined similarly for $i = 5,6,\dots,k+3$. Finally, let P_{k+4} be a hamiltonian path of $H_{k+4} \cup H_1$ starting at a_{k+4} (since $H_{k+4} \cup H_1$ is hamiltonian, such a path exists). Now

$$P_2 \cup (b_2,a_3) \cup (c_2,d_3) \cup Q \cup (c_3,d_4) \cup (b_3,a_4) \cup P_4 \cup (b_4,a_5) \cup \dots \cup (b_{k+3},a_{k+4}) \cup P_{k+4}$$

is a path in G_{k+4} missing exactly k vertices, one of which is v.
Now we have to show that there is no path in G_{k+4} that misses only $k-1$ vertices. Assume to the contrary that such a path P exists. If there are at least 5 indices i, such that P contains all vertices of H_i, then there are 3 such i's with the additional property that all vertices in H_i have

degree 2 in P. In this case it is possible to add some edges to P to obtain a spanning tree T of G_{k+4} such that all vertices in $V(H_i)$ has degree 2 in T, which is impossible by Claim 3.6. Therefore there are at most 4 such indices, thust there are at least k indices i, for which P does not contain all vertices of H_i. This means that there are at least k vertices of G_{k+4} missing from P, a contradiction. □

Walther [58] constructed connected graphs belonging to $\Pi(1,m)$ for every $m \geq 4$, and T. Zamfirescu [67] constructed 2-connected planar graphs and 3-connected graphs belonging to $\Pi(1,m)$ for every $m \geq 1$. Actually, using one of Thomassen's 3-connected planar hypotraceable graphs [51] instead of Horton's graph in Zamfirescu's construction one obtains 3-connected planar graphs belonging to $\Pi(1,m)$ for every $m \geq 1$. The graphs G_k seem to give the smallest known 3-connected graphs in $\Pi(1,m)$ (using the 8 vertex flip-flop obtained from the Petersen graph as F_i for $i = 1,2,\ldots,k = m+4$) and also the smallest known 3-connected planar graphs in $\Pi(1,m)$ (using a J-cell obtained from a 40 vertex planar hypohamiltonian graph of Jooyandeh et al. [29] as F_i for $i = 1,2,\ldots,m+4$).

3.5. Open problems

Here we mention some open questions related to leaf-critical and leaf-stable graphs and some other topics covered. We have constructed $(l+1)$-leaf-critical and l-leaf-stable graphs for every $l \geq 2$ and explored some properties of leaf-critical graphs of connectivity 2. However, all leaf-critical and leaf-stable graphs we have constructed have connectivity 3. While the 2-leaf-critical (hypohamiltonian) graphs are all 3-connected, there exist 3-leaf-critical (hypotraceable) graphs of connectivity 2 [49] and it is not so difficult to construct 2-leaf-stable graphs of connectivity 2:

Theorem 3.24 (Wiener, 2015 [63]) *Let G be a hypotraceable graph with a cut $\{a,b\}$. Then $(a,b) \notin E(G)$ and $G+(a,b)$ is 2-leaf-stable.*

Proof. Let us denote $G+(a,b)$ by G'. Let G_1 and G_2 be the 2-fragments of G with vertices of attachment a,b (by Lemma 3.13, $G = G_1 \cup G_2$). First we show that G' is traceable (this also implies that $(a,b) \notin E(G)$). Let P be a hamiltonian path of $G-a$ and Q a hamiltonian path of $G-b$. Then $P[V(G_1-a)] \cup Q[V(G_2-b)] + (a,b)$ is a hamiltonian path of G'. On the other hand, G' is not hamiltonian, otherwise G would be traceable. Now we have to show that for any vertex $v \in V(G')$ $G'-v$ is traceable, but not hamiltonian. The former is obvious, since $G-v$ is a subgraph of $G'-v$ and is traceable. Now assume to the contrary that there exists a hamiltonian cycle of $G'-v$. This cycle must contain the edge (a,b), since $G-v$ cannot have a hamiltonian cycle, otherwise G would be traceable. This means that there is a hamiltonian path between a and b in $G-v$, but in this case $G-a-b$ would be connected, a contradiction. □

It would be interesting to construct $(l+1)$-leaf-critical and l-leaf-stable graphs of connectivity 2 with $l \geq 3$.

A pretty natural question concerns the size of the smallest l-leaf-critical and l-leaf-stable graphs. This is known only for hypohamiltonian graphs. Probably it is even more difficult for planar graphs; the size of the smallest planar hypohamiltonian graph is only known to be between 18 and 40 [2], [29].

The next question is about the structure of leaf-critical and leaf-stable graphs. All such graphs known (except the hypohamiltonian graphs) are constructed using hypohamiltonian graphs as building blocks. Is it possible to construct such graphs without using hypohamiltonian graphs or do these graphs always contain J-cells or other graphs obtained from hypohamiltonian graphs (like vertex-deleted hypohamiltonian graphs)? We have mentioned that planar hypohamiltonian graphs contain a vertex of degree 3 [52]. Using the constructions known this property is inherited for leaf-critical and leaf-stable graphs. Are there planar leaf-critical or leaf-stable graphs without a degree 3 vertex?

One of the classical open problems concerning hypohamiltonicity (that we have already mentioned earlier) is whether there exist hypohamiltonian graphs without a degree 3 vertex or even of connectivity at least 4.

We have settled an open problem of Gargano et al. [18] concerning spanning spiders and arachnoid graphs, but they also proposed the more general problem whether there exist arachnoid graphs containing a vertex v, such that v is the center of only spanning spiders S, for which $d_S(v) \geq 4$. This question is still open. Now that we have seen new arachnoid graphs it is worth asking whether there are arachnoid graphs containing *several* vertices v, such that v is the center of only spanning spiders S, for which $d_S(v) \geq d$ for some fixed $d \geq 4$.

There are many open questions concerning the graph families $\Pi(j,m)$, among which the most interesting one is maybe the conjecture that $\Pi(2,2)$ is empty [21], see also [67].

4. fejezet

Traces of hypergraphs

Traces of hypergraphs have been examined for more than 40 years. The classical paper of Vapnik and Chervonenkis [56] that now plays a central role in computational learning theory, statistics, and discrete geometry appeared in 1971. In an implicit form this paper contains the proposition known now as Sauer's theorem [43] (the theorem was also proved independently by Perles and Shelah [44] and was conjectured by Erdős). Traces also have strong connections with other hypergraph problems (e.g. Turán type problems). However, the reason why this topic is included here is that theorems concerning traces can be efficiently used in fault tolerance problems concerning the hypercube (that is, finding long paths or cycles avoiding some faulty vertices or edges of the hypercube), as it was showed by Fink and Gregor [14]. Given a set of (faulty) vertices X of the n-dimensional hypercube, a cycle is said to be a long fault-tolerant cycle if it contains no vertex from X and has length $2^n - 2|X|$ (this is the maximum length that one can expect, since the hypercube is bipartite). Fink and Gregor proved that if $n \geq 15$, then for any X of size at most $\frac{n^2}{10} + \frac{n}{2} + 1$, there exists a long fault-tolerant cycle [14]. This was the first result with a quadratic number of faulty vertices, which is known to be asimptotically optimal (earlier results were about $n-1$ faulty vertices, which was improved to $2n-4$ and later $3n-7$). The key to this result is Theorem 4.6 of the author to be presented soon. A similar result concerning long paths instead of long cycles was achieved by Dvořák and Koubek [13], they also used Theorem 4.6.

We denote the set of the first n positive integers by $[n]$ and the complement of a set $X \subseteq [n]$ by $\overline{X}$. Throughout this chapter the vertex set of a hypergraph is $[n]$, unless it is stated otherwise. We call a hypergraph *simple* if it does not contain multiple edges. Simple hypergraphs will also be called *set systems*. If it does not cause any misunderstanding we identify hypergraphs by their edge set. The *multiplicity* of a set of vertices X in a hypergraph $\mathscr{H}$ is the number of occurences of X as an edge and is denoted by $m_{\mathscr{H}}(X)$. A hypergraph $\mathscr{H}$ is said to be *hereditary* if $A \in \mathscr{H}$ and $B \subseteq A$ implies $B \in \mathscr{H}$. The *trace* of a hypergraph $\mathscr{H}$ on $R \subseteq [n]$, denoted by $\mathscr{H}|_R$, is the not necessarily simple hypergraph obtained by intersecting the edges of $\mathscr{H}$ with the set R, i.e. $\mathscr{H}|_R$ is the multiset $\{H \cap R : H \in \mathscr{H}\}$. An r-trace of a hypergraph $\mathscr{H}$ is a trace of $\mathscr{H}$ on some $R \subseteq [n]$, where $|R| = r$. The *arrow-relation* $(n,m) \to (r,s)$ means that for every hypergraph $\mathscr{H}$ containing m distinct edges there exists an r-trace that contains at least s distinct edges. Bondy [8] observed that $(n,m) \to (n-1,m)$ holds if $m \leq n$. Bollobás [7] showed that $(n,m) \to (n-1,m-1)$ holds if $m \leq \lceil \frac{3}{2}n \rceil$. Sauer [43] (and independently Vapnik and Chervonenkis [56] and Perles and Shelah [44]) proved that $(n,m) \to (r,2^r)$ holds if $m > \sum_{i=0}^{r-1} \binom{n}{i}$. Frankl [15] and independently Alon [3] gave a common generalization of these results. They showed that $(n,m) \to (r,s)$ holds if and only if for every *hereditary* hypergraph $\mathscr{H}$ containing m distinct edges there exists a subset $R \subseteq [n], |R| = r$ such that $\mathscr{H}|_R$ contains at least s distinct edges. (Actually, Alon proved the theorem in a more general setting.) It is easy to check that the first three theorems follow directly from the latter one, indeed. All of these theorems deal with the number of distinct edges of the trace. About other functions of traces not much is known. In Section 4.1 we show that the maximum multiplicity of edges of trace hypergraphs can be characterized using the number of distinct edges of traces of hereditary hypergraphs and prove that Sauer's theorem is an immediate corollary of this characterization. We also obtain Theorem 4.6 as a corollary of this characterization.

4.1. Maximum multiplicity of edges

Definition 4.1 *Let m,n,r,s be positive integers. The relation $(n,m) \triangleright (r,s)$ holds if for any set system $\mathscr{H} \subseteq 2^{[n]}$, $|\mathscr{H}| = m$ there exists $X \subseteq [n]$, $|X| = r$, such that $\forall S \subseteq \overline{X} : m_{\mathscr{H}|_{\overline{X}}}(S) \leq s$.*

For example $(n,m) \triangleright (1,2)$ holds for every m and n obviously. Moreover, $(n,m) \triangleright (1,1)$ holds for $m \leq n$, this is just Bondy's theorem and it is easy to show that $(n,n+1) \not\triangleright (1,1)$ (consider the system containing all 1-element sets and the empty set). More generally, $(n,m) \triangleright (r,2^r)$ and $(n, \sum_{i=0}^{r} \binom{n}{i}) \not\triangleright (r, 2^r - 1)$ can be checked similarly. Now we present some further properties of the relation $\triangleright$ that can be readily proved.

Claim 4.2 *Let m,n,r,s be positive integers.*

1. $(n,m) \triangleright (r,s) \Rightarrow (n,m) \triangleright (r,s+1)$.
2. $(n,m) \triangleright (r,s) \Rightarrow (n,m-1) \triangleright (r,s)$.
3. $(n,m) \triangleright (n-1,m-1)$. □

In order to give a characterization of the relation $\triangleright$, we need the following lemma.

Lemma 4.3 *The relation $(n,m) \triangleright (r,s)$ holds if and only if for any hereditary set system $\mathscr{H} \subseteq 2^{[n]}$, $|\mathscr{H}| = m$ there exists $X \subseteq [n]$, $|X| = r$, such that $\forall S \subseteq \overline{X} : m_{\mathscr{H}|_{\overline{X}}}(S) \leq s$.*

Proof. The only if direction is trivial, now we prove the if direction. A set system $\mathscr{H} \subseteq 2^{[n]}$ is said to be a counterexample for $(n,m) \triangleright (r,s)$ if it contains m sets but the condition of Definition 4.1 is not fulfilled for $\mathscr{H}$. We show that if a counterexample for $(n,m) \triangleright (r,s)$ exists, then a hereditary counterexample also exists, thus proving the if direction of the lemma.
Let $\mathscr{H} \subseteq 2^{[n]}$ be a counterexample for $(n,m) \triangleright (r,s)$ and consider the following functions $D_i : \mathscr{H} \to 2^{[n]}$ $(i = 1,2,\ldots,n)$.

$$D_i(H) = \begin{cases} H \setminus \{i\}, & \text{if } i \in H \text{ and } H \setminus \{i\} \notin \mathscr{H}, \\ H, & \text{otherwise.} \end{cases}$$

The set system $D_i(\mathscr{H}) = \{D_i(H) : H \in \mathscr{H}\}$ is called the *down-compression* of $\mathscr{H}$ on element i. It is obvious that $\forall i : |D_i(\mathscr{H})| = |\mathscr{H}|$ and that $\forall i : D_i(\mathscr{H}) = \mathscr{H}$ holds if and only if $\mathscr{H}$ is hereditary. Moreover, it is also easy to see that if $\mathscr{H}$ is not hereditary, then there exists an $i \in [n]$, such that $\sum_{H \in \mathscr{H}} |D_i(H)| < \sum_{H \in \mathscr{H}} |H|$. Thus for any set system $\mathscr{H} \subseteq 2^{[n]}$ there is a hereditary system $\mathscr{H}' \subseteq 2^{[n]}$ obtained by a sequence of down-compressions from $\mathscr{H}$, such that $|\mathscr{H}'| = |\mathscr{H}| = m$. Now we show that $\mathscr{H}'$ is a counterexample for $(n,m) \triangleright (r,s)$.
Since $\mathscr{H}$ is a counterexample and $\mathscr{H}'$ is obtained by a sequence of down-compressions from $\mathscr{H}$, it suffices to show that the down-compression of a counterexample is also a counterexample. So let $\mathscr{C} \subseteq 2^{[n]}$ be a counterexample, that is, for any set $X \subseteq [n]$ of size r there

is a set $S \subseteq \overline{X} : m_{\mathscr{C}|_{\overline{X}}}(S) > s$. Now we show that $m_{D_i(\mathscr{C})|_{\overline{X}}}(S \setminus \{i\}) \geq m_{\mathscr{C}|_{\overline{X}}}(S) > s$ for every $i \in [n]$. This proves that any down-compression of $\mathscr{C}$ is a counterexample, as we have seen that $|D_i(\mathscr{C})| = |\mathscr{C}| = m$.
To verify $m_{D_i(\mathscr{C})|_{\overline{X}}}(S \setminus \{i\}) \geq m_{\mathscr{C}|_{\overline{X}}}(S)$, first let us assume that $i \notin S$. In this case

$$m_{D_i(\mathscr{C})|_{\overline{X}}}(S \setminus \{i\}) = m_{D_i(\mathscr{C})|_{\overline{X}}}(S) = \left|\{D \in D_i(\mathscr{C}) \ : \ D \cap \overline{X} = S\}\right| =$$

$$= \left|\{D_i(C) \ : \ C \in \mathscr{C}, D_i(C) \cap \overline{X} = S\}\right| \geq \left|\{C \in \mathscr{C} \ : \ C \cap \overline{X} = S\}\right| = m_{\mathscr{C}|_{\overline{X}}}(S),$$

where the inequality holds because $i \notin S \subseteq \overline{X}$ and $C \cap \overline{X} = S$ implies $D_i(C) \cap \overline{X}$. Now let us assume that $i \in S$. Then

$$m_{D_i(\mathscr{C})|_{\overline{X}}}(S \setminus \{i\}) = \left|\{D \in D_i(\mathscr{C}) \ : \ D \cap \overline{X} = S \setminus \{i\}\}\right| \geq$$

$$\geq \left|\{C \setminus \{i\} \ : \ C \in \mathscr{C}, \ (C \setminus \{i\}) \cap \overline{X} = S \setminus \{i\}\}\right| \geq \left|\{C \in \mathscr{C} \ : \ C \cap \overline{X} = S\}\right| = m_{\mathscr{C}|_X}(S),$$

because $D_i(\mathscr{C})$ contains $C \setminus \{i\}$ for every $C \in \mathscr{C}$ (first inequality) and $C \cap \overline{X} = S$ implies $(C \setminus \{i\}) \cap \overline{X} = S \setminus \{i\}$ and the sets $C \setminus \{i\}$ are all distinct for $C \in \mathscr{C}, C \cap \overline{X} = S$ (second inequality). This completes the proof of the lemma. □

Notice that Bondy's theorem follows directly from Lemma 4.3.

Theorem 4.4 (Wiener, 2007 [60]) *The relation $(n,m) \triangleright (r,s)$ holds if and only if for any hereditary set system $\mathscr{H} \subseteq 2^{[n]}$, $|\mathscr{H}| = m$ there exists $X \subseteq [n]$, $|X| = r$, such that $\mathscr{H}|_X$ contains at most s distinct edges.*

Proof. By Lemma 4.3 we only have to prove that the statements
(A) for any hereditary set system $\mathscr{H} \subseteq 2^{[n]}$, $|\mathscr{H}| = m$ there exists $X \subseteq [n]$, $|X| = r$, such that $\forall S \subseteq \overline{X} : m_{\mathscr{H}|_{\overline{X}}}(S) \leq s$
and
(B) for any hereditary set system $\mathscr{H} \subseteq 2^{[n]}$, $|\mathscr{H}| = m$ there exists $X \subseteq [n]$, $|X| = r$, such that $\mathscr{H}|_X$ contains at most s distinct edges
are equivalent. Because $\mathscr{H}$ is hereditary, $m_{\mathscr{H}|_{\overline{X}}}(S) \leq m_{\mathscr{H}|_{\overline{X}}}(\emptyset)$ for any $S \subseteq \overline{X}$, thus (A) is equivalent to the statement
(A') for any hereditary set system $\mathscr{H} \subseteq 2^{[n]}$, $|\mathscr{H}| = m$ there exists $X \subseteq [n]$, $|X| = r$, such that $m_{\mathscr{H}|_{\overline{X}}}(\emptyset) \leq s$.
The empty set can occur more than s times as an edge of $\mathscr{H}|_{\overline{X}}$ only if $\mathscr{H}|_X$ contains more than s distinct edges, because the sets whose restriction to $\overline{X}$ is the empty set must be different on X. This proves (B) ⇒ (A'). To show (A') ⇒ (B), assume that (A') is true and consider the set $X \subseteq [n]$ for which $m_{\mathscr{H}|_{\overline{X}}}(\emptyset) \leq s$. Since $\mathscr{H}$ is hereditary, distinct edges of $\mathscr{H}|_X$ are also distinct edges of $\mathscr{H}$. The restriction of each of these edges to $\overline{X}$ is the empty set, so their number is at most s, thus (B) is true, indeed. □

4.2. Corollaries

Corollary 4.5 (Wiener, 2007 [60]) $(n, \sum_{i=0}^{r} \binom{n}{i} - 1) \triangleright (r, 2^r - 1)$ *holds for any $r \leq n$ positive integers.*

Proof. By Theorem 4.4 we only have to show that for any hereditary set system $\mathscr{H} \subseteq 2^{[n]}$, $|\mathscr{H}| = \sum_{i=0}^{r} \binom{n}{i} - 1$ there exists $X \subseteq [n]$, $|X| = r$, such that $\mathscr{H}|_X$ contains at most $2^r - 1$ distinct edges. This is quite easy to verify: since $\mathscr{H}$ is hereditary, there exists a set $X \subseteq [n]$ of cardinality r, such that $X \notin \mathscr{H}$ (otherwise $|\mathscr{H}| \geq \sum_{i=0}^{r} \binom{n}{i}$ would follow). Now the trace $\mathscr{H}|_X$ does not contain X, thus $\mathscr{H}|_X$ contains at most $2^r - 1$ distinct edges. □

We have already seen that $(n, \sum_{i=0}^{r} \binom{n}{i}) \not\triangleright (r, 2^r - 1)$, so Corollary 4.5 is sharp. It is also worth mentioning that Corollary 4.5 and Sauer's theorem are equivalent: we just have to consider the complement of a set system $\mathscr{H}$ and notice that a trace $\mathscr{H}|_R$ contains $2^{|R|}$ distinct edges if and only if $\overline{\mathscr{H}}|_R$ contains no edge of multiplicity $2^{n-|R|}$. Another easy corollary of Theorem 4.4 is that the relation $\triangleright$ is transitive.

By Bondy's theorem, $(n,m) \triangleright (1,1)$ for $m \leq n$, but we have seen that $(n, n+1) \not\triangleright (1,1)$. By point 2 of Proposition 4.2, this implies $(n,m) \not\triangleright (1,1)$ for $m > n$. From this relation and $(r,r) \triangleright (1,1)$, by the transitivity of the relation $\triangleright$ follows that $(n,m) \not\triangleright (r,r)$ for $m > n$.
Hence for $m > n$, the smallest s for which $(n,m) \triangleright (r,s)$ can be true for some r, is $s = r+1$. If we are interested in those numbers r for which $(n,m) \triangleright (r, r+1)$ holds (for fixed m and n, $m > n$) we only have to find the maximum r having this property, since by point 3 of Proposition 4.2, all positive integers smaller than r also have this property. The next theorem gives a lower bound on this maximum, which is sharp for infinitely many values of m and n.

Theorem 4.6 (Wiener, 2007 [60]) *Let $m \geq 2n$ be positive integers and $r = \lceil \frac{n^2}{2m-n-2} \rceil$. Then $(n,m) \triangleright (r, r+1)$.*

Proof. We use induction on n. For $n = 1$ we have to check $(1,2) \triangleright (1,2)$, which is obvious. Now let $r' = \lceil \frac{(n-1)^2}{2m-(n-1)-2} \rceil$ (obviously $r' \leq r$) and let us assume that $(n-1, m) \triangleright (r', r'+1)$ holds. We have to show that $(n,m) \triangleright (r, r+1)$.
Because of Theorem 4.4, we only have to prove that for any hereditary set system $\mathscr{H} \subseteq 2^{[n]}$ of m sets there exists an r-element set $X \subseteq [n]$, such that $\mathscr{H}|_X$ contains at most $r+1$ distinct edges. So let $\mathscr{H} \subseteq 2^{[n]}$ be a hereditary system of m sets. Now we consider two cases.
Case 1 For every $i \in [n]$, $\{i\} \in \mathscr{H}$. This means that the number of sets of at least 2 elements in $\mathscr{H}$ is $m - n - 1$ (since $\mathscr{H}$ contains n 1-element sets and also the empty set). Consider now that graph G on the vertex set $[n]$ whose edges are the 2-element sets of $\mathscr{H}$. G has n vertices and at most $m - n - 1$ edges. A corollary of Turán's theorem [54], [5, p. 282.] states that a graph having n vertices and e edges has a stable set of size at least $\frac{n^2}{2e+n}$. Thus the graph G contains a stable set X of size $\lceil \frac{n^2}{2(m-n-1)+n} \rceil = \lceil \frac{n^2}{2m-n-2} \rceil = r$.
If $i, j \in X$ $(i \neq j)$, then $\{i, j\} \notin \mathscr{H}$, since X is stable in G. Furthermore, there is no set in $\mathscr{H}$ that contains both i and j, because $\mathscr{H}$ is hereditary. Thus $\mathscr{H}|_X$ does not contain sets of size greater than 1, so the number of distinct sets in $\mathscr{H}|_X$ is at most $|X| + 1 = r + 1$.
Case 2 There is an $i \in [n]$ such that $\{i\} \notin \mathscr{H}$. Then there is no set in $\mathscr{H}$ that contains the element i, because $\mathscr{H}$ is hereditary, thus we can delete the element i from the underlying set $[n]$ without changing $\mathscr{H}$. Now we use the induction hypothesis: $(n-1, m) \triangleright (r', r'+1)$. This implies that a set $X \subseteq [n] \setminus \{i\}$ of size r' exists, such that $\mathscr{H}|_X$ contains at most $r'+1$ distinct edges.
Now for the set $X' = X \cup \{i\}$ we have $\mathscr{H}|_{X'} = \mathscr{H}|_X$, hence $\mathscr{H}|_{X'}$ also contains at most $r'+1$ distinct edges. Since $r' \leq r$, it only remains to show that either X or X' has r elements.
Because $|X| = r' \leq r$ and $|X'| = |X| + 1$, it is enough to prove that $r' + 1 \geq r$. That is, we have to show that

$$\lceil \frac{(n-1)^2}{2m-(n-1)-2} \rceil + 1 \geq \lceil \frac{n^2}{2m-n-2} \rceil.$$

This holds if

$$\frac{(n-1)^2}{2m-(n-1)-2}+1 \geq \frac{n^2}{2m-n-2}.$$

Eliminating the fractions we obtain

$$(2m-3n)(2m-n-2) \geq n^2,$$

which is true, since $m \geq 2n$ and $n \geq 2$.

Note that the lower bound $\frac{n^2}{2e+n}$ following from Turán's theorem is sharp for the graphs whose components are complete graphs of the same size. Therefore considering the hypergraph containing the empty set, all the 1-element sets, and the edges of such a graph we can see that $(n,m) \not\to (r+1, r+2)$ for $r = \lceil \frac{n^2}{2m-n-2} \rceil$, that is, our bound is sharp in these cases. □

For a somewhat stronger form of the previous theorem we need the following definitions. A hypergraph $\mathscr{H}$ is a *minimal simple* hypergraph if it is simple but for every subset X of the vertices the restriction of $\mathscr{H}$ to X is not simple. The set of all minimal simple hypergraphs on the vertex set $[n]$ having m hyperedges is denoted by $MSH(n,m)$.

Theorem 4.7 (Wiener, 2013 [61]) *Let $\mathscr{A} \in MSH(n,m)$. Then there exists a subset $X \subseteq [n]$ of cardinality $\left\lceil \frac{n^2}{2m-n-2} \right\rceil$, such that by deleting X we obtain a hypergraph where every hyperedge has multiplicity at most $\left\lceil \frac{n^2}{2m-n-2} \right\rceil + 1$.*

The proof of this theorem is pretty similar to the proof of Theorem 4.6 and is therefore omitted.

Irodalomjegyzék

[1] R. E. L. Aldred, S. Bau, D. A. Holton, B. D. McKay, Nonhamiltonian 3-connected cubic planar graphs. SIAM J. Disc. Math. 13 (2000), 25–32.

[2] R. E. L. Aldred, B. D. McKay, N. C. Wormald, Small Hypohamiltonian Graphs, J. Combin. Math. Combin. Comput. 23 (1997) 143–152.

[3] N. Alon, On the density of sets of vectors, Discrete Mathematics 46 (1983), 199–202.

[4] M. Araya, G. Wiener, On Cubic Planar Hypohamiltonian and Hypotraceable Graphs, Electronic Journal of Combinatorics 18, P85. (2011)

[5] C. Berge, Graphs and Hypergraphs, North-Holland, Amsterdam, 1973.

[6] D. Binkele-Raible, H. Fernau, S. Gaspers, M. Liedloff, Exact and Parameterized Algorithms for Max Internal Spanning Tree, Algorithmica 65 (2009), 95–128.

[7] B. Bollobás, unpublished, see in [37] Problem 13.10.

[8] J. A. Bondy, Induced subsets, Journal of Combinatorial Theory Ser. B 12 (1972), 201–202.

[9] J. Bosák, Hamiltonian lines in cubic graphs. Theory of Graphs (Internat. Sympos., Rome, 1966) Gordon and Breach, New York; Dunod, Paris (1967), 35–46.

[10] J. B. Collier, E. F. Schmeichel, Systematic searches for hypohamiltonian graphs, Networks 8 (1978), 193–200.

[11] V. Chvátal, Flip-flops in hypohamiltonian graphs, Can. Math. Bull. 16 (1973), 33–41.

[12] J. Doyen, V. Van Dienst, New families of hypohamiltonian graphs, Discrete Mathematics 13 (1975), 225–236.

[13] T. Dvořák, V. Koubek, Long paths in hypercubes with a quadratic number of faults, Information Sciences 179 (2009), 3763–3771.

[14] J. Fink, P. Gregor, Long paths and cycles in hypercubes with faulty vertices, Information Sciences 179 (2009), 3634–3644.

[15] P. Frankl, On the trace of finite sets, Journal of Combinatorial Theory Ser. A 34 (1983), 41–45.

[16] Z. Füredi, J. Pach, Traces of finite sets: extremal problems and geometric applications, in: „Extremal Problems for Finite Sets," Bolyai Society Mathematical Studies 3 (1991), 251–282.

[17] T. Gallai, On directed paths and circuits, Theory of Graphs, P. Erdős and G. Katona (Editors), Academic Press, New York (1968), pp. 115–118.

[18] L. Gargano, M. Hammar, P.Hell, L. Stacho, U. Vaccaro, Spanning spiders and light-splitting switches, Discrete Mathematics 285 (2004), 83–95. (Earlier versions: L. Gargano, P. Hell, L. Stacho, U. Vaccaro, Spanning trees with bounded number of branch vertices, ICALP02, Lecture Notes in Computer Science 2380 (2002), 355–365. and L. Gargano, M. Hammar, There are spanning spiders in dense graphs (and we know how to find them), ICALP03, Lecture Notes in Computer Science 2719 (2003), 802–816.)

[19] E. J. Grinberg, Plane homogeneous graphs of degree three without Hamiltonian circuits, Latvian Math. Yearbook, Izdat. Zinatne, Riga 4 (1968), 51–58. (In Russian)

[20] B. Grünbaum, Convex Polytopes, John Wiley and Sons, New York, 1967.

[21] B. Grünbaum, Vertices missed by longest paths or circuits, Journal of Combinatorial Theory Ser. A 17 (1974), 31–38.

[22] W. Hatzel, Ein planarer hypohamiltonscher Graph mit 57 Knoten, Math. Ann. 243 (1979), 213–216.

[23] J. C. Herz, T. Gaudin, P. Rossi, Solution du probléme No. 29, Revue Francaise de Recherces Opérationelle 8 (1964), pp. 214-218.

[24] J. C. Herz, J. J. Duby, F. Vigué, Recherche systématique des graphes hypohamiltoniens, in: Théorie des Graphes, Dunod, Paris (1967), 153–160.

[25] D. A. Holton, B. D. McKay, The smallest non-hamiltonian 3-connected cubic planar graphs have 38 vertices. Journal of Combinatorial Theory Ser. B 45 (1988), 315–319.

[26] D. A. Holton and J. Sheehan, Hypohamiltonian graphs, The Petersen Graph, Cambridge University Press, New York, 1993.

[27] L.-H. Hsu and C.-K. Lin, Graph Theory and Interconnection Networks, CRC Press, Boca Raton, 2008.

[28] M. Jooyandeh, Recursive Algorithms for Generation of Planar Graphs, PhD thesis, Australian National University (2014)

[29] M. Jooyandeh, B. D. McKay, P. R. J. Östergård, V. H. Pettersson, C. T. Zamfirescu, Planar Hypohamiltonian Graphs on 40 Vertices, preprint, arXiv:1302.2698, (2013)

[30] S. F. Kapoor, H. V. Kronk, and D. R. Lick, On detours in graphs, Canad. Math. Bull. 11 (1968), 195–201.

[31] D. Karger, R. Motwani, G. D. S. Ramkumar, On Approximating the Longest Path in a Graph, Algorithmica 18 (1997), 82–98.

[32] M. Knauer, J. Spoerhase, Better Approximation Algorithms for the Maximum Internal Spanning Tree Problem, Algorithmica 71 (2013), 797–811.

[33] J. Lederberg, Systematics of organic molecules, graph theory and Hamiltonian circuits, Instrumentation Research Laboratory Report, no. 1040, Stanford University, Stanford, Calif., 1144 (1966)

[34] J. van Leeuwen (ed.), Handbook of Theoretical Computer Science A: Algorithms and Complexity, Elsevier, 1990.

[35] W. Li, J. Chen, J. Wang, Deeper Local Search for Better Approximation on Maximum Internal Spanning Trees, Lecture Notes in Computer Science 8737 (2014), 642–653.

[36] X. Li, D. Zhu, Approximating the Maximum Internal Spanning Tree Problem via a Maximum Path-Cycle Cover, Lecture Notes in Computer Science 8889 (2014), 467–478.

[37] L. Lovász, Combinatorial Problems and Exercises, North-Holland, Amsterdam, 1979.

[38] H.-I. Lu, R. Ravi, The Power of Local Optimization: Approximation Algorithms for Maximum-leaf Spanning Tree (DRAFT), CS-96-05, Department of Computer Science, Brown University, Providence, Rhode Island, 1996.

[39] H.-I. Lu, R. Ravi, Approximation for maximum leaf spanning trees in almost linear time, J. Algorithms 29 (1998), 132–141.

[40] G. Salamon, Approximating the Maximum Internal Spanning Tree problem, Theoretical Computer Science 410 (2009), 5273–5284.

[41] G. Salamon, Degree-Based Spanning Tree Optimization, PhD Thesis, `http://doktori.math.bme.hu/Ertekezesek/salamon_dissertation.pdf` (2010)

[42] G. Salamon, G. Wiener, On Spanning Trees with Few Leaves, Information Processing Letters 105 (2008), 164–169.

[43] N. Sauer, On the density of families of sets, Journal of Combinatorial Theory Ser. A 13 (1972), 145–147.

[44] S. Shelah, A combinatorial problem: stability and order for models and theories in infinitary languages, Pacific J. Math. 41 (1972), 247–261.

[45] R. Solis-Oba, 2-Approximation algorithm for finding a spanning tree with maximum number of leaves, Lecture Notes in Computer Science 1461 (Proc. of 6th ESA Symposium) (1998), 441–452.

[46] R. Sousselier, Probleme No. 29: Le Cercle des Irascibles, Revue Française de Recherches Operationelles 7 (1963), 405–406.

[47] B. Schauerte, C. T. Zamfirescu, Regular graphs in which every pair of points is missed by some longest cycle. An. Univ. Craiova, Ser. Mat. Inf. 33 (2006), 154–173.

[48] P. G. Tait, Note on a theorem in geometry of position, Trans. Roy. Soc. Edinburgh 29 (1880), 657-660.

[49] C. Thomassen, Hypohamiltonian and hypotraceable graphs, Discrete Mathematics 9 (1974), 91–96.

[50] C. Thomassen, On hypohamiltonian graphs, Discrete Mathematics 10 (1974), 383–390.

[51] C. Thomassen, Planar and infinite hypohamiltonian and hypotraceable graphs, Discrete Mathematics 14 (1976), 377–389.

[52] C. Thomassen, Hypohamiltonian graphs and digraphs, Theory and Applications of Graphs, Lecture Notes in Mathematics No. 642, Springer, Berlin (1978), pp. 557–571.

[53] C. Thomassen, Planar cubic hypohamiltonian and hypotraceable graphs, J. Comb. Theory B 30 (1981), 36–44.

[54] P. Turán, On an extremal problem in graph theory, Mat. Fiz. Lapok 48 (1941), 436–452. (in Hungarian)

[55] W. T. Tutte, On hamiltonian circuits. J. London Math. Soc. 21 (1946), 98–101.

[56] V. N. Vapnik, A. Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, Th. Prob. Appl. 16 (1971), 264–280.

[57] H. J. Voss, Cycles and bridges in graphs, Kluwer Academic Publishers, Dordrecht, 1990.

[58] H. Walther, Über die Nichtexistenz eines Knotenpunktes, durch den alle längsten Wege eines Graphen gehen, J. Comb. Theory 6 (1969), 1–6.

[59] M. Watanabe, Arrow relations on families of finite sets, Discrete Mathematics 94 (1991), 53–64.

[60] G. Wiener, Edge Multiplicity and Other Trace Functions, Electronic Notes in Discrete Mathematics 29 (2007), 491–495.

[61] G. Wiener, Rounds in combinatorial search, Algorithmica 67 (2013), 315–323.

[62] G. Wiener, Leaf-critic graphs, Proc. of the 3rd Bordeaux Graph Workshop (2014), 101–102.

[63] G. Wiener, Leaf-critical and leaf-stable graphs, Proc. of the 9th Japanese-Hungarian Symposium on Discrete Mathematics and its Applications, Fukuoka (2015), 267–278.

[64] G. Wiener, On non-traceable, non-hypotraceable, arachnoid graphs, Electronic Notes in Discrete Mathematics 49 (2015), to appear

[65] G. Wiener, M. Araya: On planar hypohamiltonian graphs, Journal of Graph Theory 67 (2011), 55–68.

[66] T. Zamfirescu, A two-connected planar graph without concurrent longest paths, Journal of Combinatorial Theory Ser. B 13 (1972), 116–121.

[67] T. Zamfirescu, On longest paths and circuits in graphs, Math. Scand. 38 (1976), 211–239.

[68] C. Zamfirescu and T. Zamfirescu, A planar hypohamiltonian graph with 48 vertices, Journal of Graph Theory 55 (2007), 338–342.

[69] C. T. Zamfirescu, On Hypohamiltonian and Almost Hypohamiltonian Graphs, Journal of Graph Theory 79 (2015), 63–81.