A Barabási-Albert féle véletlen fa szélessége

Szakdolgozat

Katona Zsolt

Témavezető: Móri Tamás

Eötvös Lóránd Tudományegyetem

Természettudományi Kar

2003

1. Bevezetés

Dolgozatomban véletlen gráfokkal azon belül is véletlen fákkal foglalkozom. A véletlen gráfok elméletét Erdős és Rényi alapozta meg az 1960-as évek elején, miután Erdősnek sikerült véletlen módszerekkel megoldani extremális gráfelméleti problémákat. Ezzel a módszerrel ugyanis be lehet bizonyítani bizonyos konstrukciók létezését anélkül, hogy konkrétan megadnánk őket, sőt sok esetben nem is ismerjük őket.

Az első ilyen megoldás Szele [17] nevéhez fűződik, aki 1943-ban belátta, hogy van olyan n csúcsú turnament (teljes gráf, minden éle valamelyik irányban irányítva), amelyben legalább $n!/2^{n-1}$ Hamilton-kör van, hiszen ha az éleket egymástól függetlenül irányítjuk azonos valószínűséggel a két irányba, akkor $n!/2^{n-1}$ a Hamilton körök számának várható értéke. 1947-ben pedig Erdős látta be [8] hasonló módszerrel, hogy az $R(k)$ Ramsey-szám legalább $2^{k/2}$.

A véletlen gráf modellek egymástól függetlenül több szerzőnél jelentek meg, először az Erdős-Rényi féle modell, $\mathcal{G}(n,m)$. Eszerint az összes n csúcsú m élű gráf közül választunk azonos valószínűséggel. Nagyon a hasonló az először Gilbert [10] által bevezetett gráf $\mathcal{G}(n,p)$, ahol n csúcs között húzzuk be a lehetséges éleket egymástól függetlenül azonos p valószínűséggel. (Ha például $p = 1/2$, akkor az összes n csúcsú gráf közül választottunk azonos valószínűséggel.) Az elmélet felépítése, azonban egyértelműen Erdős és Rényi érdeme a két alapvető modell tekintetében. Ezeket a modelleket természetesen sok irányban lehet általánosítani.

Az informatika fejlődésével párhuzamosan nőtt az érdeklődés a véletlen gráfok iránt és olyan kifejezetten számítógépes alkalmazású modellek születtek mint például a bináris keresőfa, ami alkalmas valamilyen módon rendezett adat tárolására. Ez egy épített fa, tehát egy gyökér csúcsból indulunk, amit kezdetben nevezzünk külső csúcsnak. A következő lépésben kössünk hozzá két új csúcsot egy-egy éllel, most az új csúcsok lesznek külsők a gyökér belsővé válik. Minden további lépésben a külső pontok közül választunk egyet egyforma valószínűséggel és belsővé tesszük úgy, hogy két új külső pontot kötünk hozzá.

Az internet világméretűvé válása nyomán felmerült a kérdés, hogy milyen a struktúrája. Sokan (fizikusok, informatikusok) kezdték vizsgálni a World Wide Web (WWW), azaz interneten elérhető oldalak irányított gráfját melynek éleit az egymásra való hivatkozások határozzák meg. Az eddigi véletlen gráf modellek nem támasztották alá elméletileg a mérési eredményeket, ennek nyomán alkotott Barabási Albert-László és Albert Réka egy új véletlen gráf modellt [1].

2. A Barabási-Albert modell

A WWW mellett számos területen találkozhatunk olyan gráfokkal melyek komplex struktúráját nehéz leírni [1]. Ilyen például az a genetikus hálózat, melynek csúcsai fehérjék, élei pedig a köztük létrejövő kémiai reakciók. Hasonlóan hatalmas hálózatot alkot az idegrendszer, melynek csúcsai az idegsejtek, élei pedig a köztük kapcsolatot létesítő axonok, illetve bizonyos társadalomtudományokban felmerülő szervezeti kapcsolatok hálózatai.

A hálózatok nagy mérete miatt teljes leírásuk reménytelen, az interneten becslések szerint legalább $8 \cdot 10^8$ oldal van. Így merül fel, hogy valamilyen véletlen gráffal lehetne modellezni ezeket a hálózatokat. Első megközelítésként a fokszámeloszlást lehet vizsgálni, azaz a $P(k) := P(d_i = k)$ érétékek sorozatát $k = 0, 1, 2, ...$ esetén, ahol d_i az i. csúcs fokszáma. Találtak olyan gráfokat, ahol fel tudták jegyezni a fokszámok tapasztalati eloszlását, ilyen például a filmszínészek gráfja, mely azt írja le, hogy ki kivel szerepelt együtt. Azt találták [1], hogy az eloszlás farka $P(k) \sim k^{-\gamma}$ alakú, ahol $\gamma = 2,3 \pm 0.1$. Másik ilyen példa az USA nyugati elektromos hálózata 4941 csúccsal (erőművek, transzformátorok, alállomások), itt szintén $P(k) \sim k^{-\gamma'}$, ahol a kitevő -4 körül van. A WWW teljes fokszámeloszlását ugyan nem lehet leírni, de bizonyos programok segítségével egy területét igen. Az nd.edu vizsgálata [3] során arra jutottak, hogy az irányított gráf (egy hivatkozásnak iránya van) fokszámeloszlása mind a be-, mind a kifokszámok tekintetében ismét $P(k) \sim k^{-\gamma}$, különböző, $2,1$ és $2,45$-ös γ értékekkel, ki- illetve befokszám esetén. Érdekes példa még a tudományos publikációk hivatkozási gráfja is, itt γ értéke 3 körüli.

Tehát ezekben a példákban mind hasonló fokszámeloszlást figyeltünk meg, vajon megfelel-e ez az Erdős-Rényi féle véletlen gráf modellnek? Ha n csúcs között a lehetséges éleket egymástól függetlenül, p valószínűséggel húzzuk be, akkor a fokszámeloszlás n növelésével aszimptotikusan Poisson np paraméterrel[4]:

$$P(k) \sim e^{-np}\frac{(np)^k}{k!}.$$

A $\mathcal{G}(n,m)$ modellben a fokszámeloszlás aszimptotikusan szintén Poisson, $\frac{2m}{n}$ várható értékkel.

Látható, hogy k-ban mindkét esetben exponenciálisan csökken a nagy fokú pontok aránya, ráadásul a paraméter függ a csúcsok számától. Így ez a modell nem felel meg a mérési eredménynek. Másik probléma a hagyományos modellel, hogy a míg a vizsgált hálózatok folyamatosan nőnek, és egy új csúcsot nem teljesen véletlenszerűen kötünk az eddigiekhez (hanem valamilyen tulajdonsága alapján), addig az Erdős-Rényi modellben rögzített a csúcsok száma.

Egy új modell a Watts és Strogatz által bevezetett [18] "kicsi világ" hálózat. A gráf N csúcsa egy egyenesen helyezkedik el sorban és minden szomszéd és másodszomszéd között vezet egy él. Ezek után minden élet p valószínűséggel újra kötünk, azaz az egyik végét átkötjük egy véletlenül választott csúcshoz (egyenletes eloszlás szerint), azzal a megkötéssel, hogy nem keletkezzen többszörös és hurokél. $p = 0$ esetén a gráf változatlan marad és így az átlagos távolság lineáris függvénye N-nek, míg $p = 1$ esetén minden élet újra kötünk, ekkor az átlagos távolság logaritmikus függvénye N-nek. A fokszámeloszlás az előbbi esetben nyilván konstans 4, míg ha $p \to 1$, akkor tart ahhoz az Erdős-Rényi féle modelléhez, ahol $p = 4/N$ valószínűséggel húzzuk be az éleket.

A Barabási Albert-László és Albert Réka által bevezetett [1] új modell alapja, hogy új csúcsokat veszünk a gráfhoz és már meglévő csúcsokhoz azok fokszámával arányos valószínűséggel kötjük. A nem teljesen precízen leírt eljárás szerint kezdetben tekintsünk m_0 csúcsot és minden lépésben vegyünk egy új csúcsot a gráfhoz $m \leq m_0$ éllel. Ezeket az éleket kössük már meglévő csúcsokhoz a csúcsokon

$$\Pi(i) = \frac{d_i}{\sum_j d_j}$$

valószínűség-eloszlás szerint. Így t lépés után $m_0 + t$ csúcs és mt éle lesz a gráfnak. A szerzők, akik fizikusok számítógéppel mérték a modell szerint létrehozott gráf tulajdonságait és azt találták, hogy $P(k) \sim k^{-\gamma}$ alakú a fokszámeloszlás $\gamma = 2,9 \pm 0,1$ paraméterrel.

Ezen felül [2]-ban a következőképpen érvelnek $\gamma = 3$ mellett. Legyen d_i a v_i csúcs fokszáma, a t időpontban, amikor a gráfnak $t + m_0$ csúcsa és mt éle van. Ha egy új csúcsot adunk hozzá, akkor $md_i/2mt$ a valószínűsége, hogy hozzákötjük v_i-hez, tehát a fokszámot az idő folytonos függvényeként értelmezve,

$$\frac{\partial d_i}{\partial t} = \frac{d_i}{2t}.$$

Azzal a kiinduló feltétellel, hogy $d_i = m$ a $t = i$ időpontban, azt kapjuk, hogy t függvényeként

$$d_i = m(t/i)^{\frac{1}{2}}.$$

Ebből pedig egyszerű számolással adódik $\gamma = 3$.

3. Matematikai eredmények

Dolgozatomban a fenti modellt vizsgálom, azzal az egyszerűsítéssel, hogy $m_0 = m = 1$, azaz egy pontból indulva egyesével veszünk új éleket, illetve pontokat, tehát egy fa keletkezik.

A véletlen fa modellben egy pontból indulunk ki, ez a gyökér. Az első lépésben hozzáveszünk egy új pontot és összekötjük a gyökérrel. Minden további lépésben hozzáveszünk egy új pontot, amelyet egy régi ponttal kötünk össze, minden régi pontot a fokszámával arányos, azaz

$$\Pi(i) = \frac{d_i}{\sum_j d_j}$$

valószínűséggel választva. Ugyanezt adja, ha a második lépéstől kezdve mindig egy élt választunk (egyenlő valószínűséggel) a meglévők közül és ennek egyik végpontjához (ismét egyenlő valószínűséggel választva) hozzákötünk egy új pontot egy új éllel. Nyilvánvaló a két eljárás megegyezősége, hiszen az utóbbi esetben egy x csúcsot $\frac{1}{2} d_x \frac{1}{e} = \frac{d_x}{2e}$ valószínűséggel választottunk, ahol e az élek száma, míg d_x a pont foka.

Móri Tamás általánosította [12] az eljárást a következő módon. Válasszunk egy tetszőleges monoton növő, pozitív egészeken értelmezett, pozitív értékű

φ fügvényt, jelölje S_n a $\sum_{j=1}^{n}\varphi(d_j)$. Az eddigi fokszámarányos valószínűség helyett most

$$\Pi(i)=\frac{\varphi(d_i)}{S_n}$$

valószínűséggekel kössük az új csúcsot az i., már csúcshoz. Különösen érdekes a $\varphi(k)=k+\beta$ eset ($\beta>-1$), hiszen ekkor $S_n=2n+\beta(n+1)$, tehát nem véletlen.

Móri Tamás megmutatta, hogy a k fokú csúcsok aránya 1 valószínűséggel konvergál egy c_k határértékhez, amely, mint k függvénye $k^{(-3+\beta)}$ alakban csökken.

Ezek kívül belátta, hogy maximális fokszám elosztva $n^{\frac{1}{2+\beta}}$-val egy pozitív valószínűségi változóhoz konvergál, amint $n\to\infty$.

Mivel az eredeti modell $\beta=0$-nak felel meg, ez alátámasztja a mérési eredményeket, hiszen ott k^{-3} alakban csökkent a k fokú csúcsok aránya. Az általánosítás révén pedig magyarázatot ad azokra a fent leírt esetekre, amikor a mérés szerint $k^{-\gamma}$ volt az eredmény $2,1$ és 4 közötti γ-kal.

A fenti eredménnyel párhuzamosan Bollobás, Riordan, Spencer és Tusnády vizsgálta ilyen gráfok fokszámeloszlását [6]. Ehhez bevezettek egy precíz modellt [5] a következők miatt. Először is nem világos, hogy miből indul az eredeti algoritmus, milyen élek vannak az eredeti m_0 csúcs között. Másrészt azon csúcsok S halmazának eloszlása, amelyeket egy új csúccsal összekötünk nincs meghatározva azon valószínűségek által, hogy egyes csúcsok S-ben vannak-e. Például, ha egy 4 hosszú körből indulunk és $m=2$, akkor az ötödik csúcsot bármilyen $0\leq\alpha\leq 1/4$ mellett α valószínűséggel köthetjük két szomszédos és $1/2-2\alpha$ valószínűséggel két nem szomszédos csúcshoz. Ugyanis ekkor minden csúcs egyforma valószínűséggel kerül be a két kiválasztott közé és az eredeti modell ezt követeli meg.

Az ilyen problémák elkerülése végett megengedik többszörös- és hurokélek létezését is a gráfban. Először az $m = 1$ esetben induljunk ki egy csúcsból, amelyhez egy hurok él kapcsolódik. Vegyünk egy új csúcsot, amelyet $2/3$ valószínűséggel kössünk hozzá, $1/3$ valószínűséggel pedig egy hurokéllel önmagához. Minden további lépésben vegyünk egy új csúcsot és egy róla lelógó élet. A másik végét pedig kössük valamelyik csúcshoz fokszámarányos valószínűséggel, megengedve a hurokélet, hiszen az új csúcsnak már egy a foka a lelógó él miatt. Irányított gráfot úgy kaphatunk, ha az új csúcsból a régi felé mutatnak az irányított élek.

$m > 1$ esetén készítsünk egy G' gráfot $v'_1, v'_2, ..,$ csúcsokkal az előbbi eljárással, majd vonjuk össze az $v'_1, ..., v'_m$-t egy új v_1 csúccsá, $v'_{m+1}, ..., v'_{2m}$-t v_2-vé, és így tovább. Az összes élet megtartva így persze többszörös- és új hurokélek is keletkezhetnek. Látható, hogy az $m > 1$ esetet olyan egyszerűen kapjuk az $m = 1$ esetből, hogy minden kérdést ez utóbbira érdemes visszaveztni.

Nyilvánvaló, hogy ez a modell nem különbözik lényegesen a fejezet elején leírtaktól, a különbség, hogy itt fa helyett ugyanolyan (csak különböző méretű) fákból álló erdő keletkezik, mindegyiknek a gyökerén egy hurokéllel. A komponensek számának eloszlása könnyen meghatározható, hiszen az n-edik lépésben $1/(2n + 1)$ valószínűséggel keletkezik új komponens, egyébként pedig valamelyik már meglévő fa bővül.

A szerzők az $m = 1$ eset vizsgálatához [6]-ben a következő ekvivalens leírást adják meg. Vegyünk fel egy egyenesen $2n$ pontot és párosítsuk őket úgy, hogy minden párosítás azonos valószínűséggel forduljon elő. Létrehozunk egy irányított gráfot. Induljunk el az egyenesen az egyik irányból (balról) és vonjuk össze a pontokat egészen addig, míg egy pár jobb oldali pontjához érünk (ezt is vegyük bele), így kapjuk a v_1 csúcsot. Most egyesítsük a következő jobb oldali pontig a következőket, így kapjuk a v_2 csúcsot, és így tovább v_n-ig. Az éleket úgy vegyük fel, hogy minden pár jobb oldali pontjának megfelelő csúcsból vezessen a baloldalinak megfelelőbe egy irányított él. Indukcióval könnyen bizonyítható, hogy így ugyanolyan eloszlása lesz a keletkező gráfoknak, mint az eredeti eljárással. Így sikerült egy nemrekurzív ekvivalens jellemzését adni a gráfnak, amelynek segítségével belátták a következő eredményt:

Legyen $\#_m^n(d)$ a d befokú pontok száma az n csúcsú m paraméterű gráfban és

$$\alpha_{m,d} = \frac{2m(m+1)}{(d+m)(d+m+1)(d+m+2)}.$$

Ha $m \geq 1$ és $\varepsilon > 0$ rögzített, akkor $n \to \infty$ mellett 1-hez tartó valószínűséggel

$$(1-\varepsilon)\alpha_{m,d} \leq \frac{\#_m^n(d)}{n} \leq (1+\varepsilon)\alpha_{m,d},$$

minden $0 \leq d \leq n^{1/15}$ mellett.

Bollobás és Riordan [5]-ban a fent leírt, többszörös és hurokéleket is megengedő modellben vizsgálta a gráf átmérőjét. $m = 1$ esetén megmutatták, hogy ha γ a $\gamma e^{1+\gamma} = 1$ egyenlet megoldása és $\varepsilon > 0$ rögzített, akkor 1 valószínűséggel a legnagyobb komponens átmérő $(1/\gamma - \varepsilon)\log n$ és $(1/\gamma + \varepsilon)\log n$ közé esik.

$m \geq 2$-re viszont rögzített $\varepsilon > 0$ mellett 1 valószínűséggel a gráf összefüggő és az átmérőjére

$$(1-\varepsilon)\log n/\log\log n \leq \mathrm{diam}(G_m^n) \leq (1+\varepsilon)\log n/\log\log n$$

teljesül.

4. A fa szélessége és magassága

A fát a gyökértől elindulva bontsuk szintekre (a gyökér a 0. szint). Jelölje $X[n,k]$ a k-adik szint méretét az n-edik lépés után (az első lépés után van egy él). Ezek a valószínűségi változók meghatározzák a fa alakját, $W_n := \max\{X[n,k] : 1 \leq k\}$ a *szélessége*, míg $H_n := \max\{k \geq 1 : X[n,k] \neq 0\}$ a *magassága*.

Mivel a magasság szoros kapcsolatban van az átmérővel, [5] eredményeiből levezethető, hogy a H_n magasság az általunk vizsgált fa modellben $\log n$ nagyságrendű. Móri Tamás megmutatta ezt jobb konstansokkal és általánosítva a $\varphi(k) = k + \beta$ esetre. Ezek szerint 1 valószínűséggel

$$\frac{1+\beta}{2+\beta} \leq \liminf_{n\to\infty} \frac{H_n}{\log n},$$

$$\limsup_{n\to\infty} \frac{H_n}{\log n} \leq \min_{z>0} \frac{1+\frac{1+\beta}{2+\beta}z}{\log(1+z)}.$$

A felső becslés értéke $\beta = 0$ esetén $1,79556$, β-nak monoton növő függvénye és határértéke e. A $z = e-1$ helyettesítéssel $\limsup(H_n/\log n) \leq 1 + \frac{1+\beta}{2+\beta}e$.

Érdemes ezt az eredményt összehasonlítani más fák magasságával. Rényi és Szekeres belátták [15], hogy ha egyeneletes eloszlás szerint választjuk a fát, akkor

$$\lim_{n\to\infty} P(\frac{H_n}{\sqrt{2n}} < x) = H(x),$$

ahol $H(x) = 4x^{-3}\pi^{5/2}\sum_{p=1}^{\infty} p^2 \exp(-\pi^2 p^2/x^2)$.

Bináris keresőfákra ismert, hogy $H_n \sim c\log n$, ahol $c = 4,31107..$ a $c\log(2e/c) = 1$ egyenlet 2-nél nagyobb megoldása.

Célunk a szélesség meghatározása, melynek alapja Chauvin, Drmota és Jabbour-Hattab módszere [7], akik bináris kereső fákra bebizonyították, hogy $W_n \sim \frac{n}{\sqrt{\pi \log n}}$.

1. Tétel. *1 valószínűséggel*

$$\frac{X[n,k]}{2n/\sqrt{\pi \log n}} = \exp\left(-\frac{(k-\frac{1}{2}\log n)^2}{\log n}\right) + \mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right),$$

ha $n \to \infty$ egyenletes hibataggal $k \geq 0$-ra.

2. Következmény. *Ha $n \to \infty$, akkor 1 valószínűséggel*

$$\frac{W_n}{2n/\sqrt{\pi \log n}} = 1 + \mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right),$$

Látható, hogy az eredmény nem sokban különbözik a bináris keresőfáknál tapasztaltaktól, míg jelentősen eltér az egyenletes eloszlás esetén kialakuló $\sqrt{n}$ nagyságrendtől.

A Barabási féle fa alakja, tehát nagyon hasonlít a bináris keresőfáéhoz, a magasság és szélesség is 2 körüli szorzóval tér el.

5. Tétel bizonyítása

A bizonyítás alapja a következő, komplex generátorfüggvény vizsgálata. Legyen

$$G_n(z) = \sum_{k=0}^{\infty} X[n,k+1]z^k.$$

Az összegzés valójában legfeljebb n-ig megy, hiszen $k \geq n$-re $X[n,k] = 0$. Jelölje $\mathcal{F}_n$ az első n lépés által generált σ-algebrát. A generátorfüggvényt várható értékével elosztva martingált kapunk $\mathcal{F}_n$-re nézve.

3. Lemma.

$$M_n(z) := \frac{G_n(z)}{\mathbf{E}G_n(z)}$$

martingál $\mathcal{F}_n$-re nézve.

BIZONYÍTÁS: $X[n+1,k]$ feltételes eloszlása $\mathcal{F}_n$-re nézve a következő:

$$P(X[n+1,k]=X[n,k]+1|\mathcal{F}_n)=\begin{cases}\frac{X[n,k]+X[n,k-1]}{2n}, & \text{ha } k>1,\\ \frac{X[n,1]}{2n}, & \text{ha } k=1,\end{cases}$$

és nyilván

$$P(X[n+1,k]=X[n,k]|\mathcal{F}_n)=1-P(X[n+1,k]=X[n,k]+1|\mathcal{F}_n).$$

Ekkor

$$\mathbf{E}(X[n+1,1]|\mathcal{F}_n)=(X[n,1]+1)\frac{X[n,1]}{2n}+X[n,1]\left(1-\frac{X[n,1]}{2n}\right)=$$
$$=\frac{X[n,1]}{2n}+X[n,1]=\frac{2n+1}{2n}X[n,1]$$

ha $k=1$ és

$$\mathbf{E}(X[n+1,k]|\mathcal{F}_n)=(X[n,k]+1)\frac{X[n,k]+X[n,k-1]}{2n}+$$
$$+X[n,k]\left(1-\frac{X[n,k]+X[n,k-1]}{2n}\right)=\frac{2n+1}{2n}X[n,k]+\frac{1}{2n}X[n,k-1]$$

ha $k>1$. Tehát azt kapjuk, hogy

$$\mathbf{E}(G_{n+1}(z)|\mathcal{F}_n)=\frac{2n+1}{2n}G_n(z)+\frac{z}{2n}G_n(z)=\frac{2n+1+z}{2n}G_n(z).$$

Ebből a teljes várható érték tételével rögtön látható, hogy

$$\mathbf{E}G_n(z)=\prod_{j=1}^{n-1}\frac{2j+1+z}{2j},$$

hiszen $G_1(z)=1$ azonosan. Tehát $M_n(z)$ valóban martingál.

$\square$

Határozzuk meg a várható érték nagyságrendjét.

4. Lemma. *Minden kompakt C komplex halmazon*

$$\mathbf{E}G_n(z)=\frac{n^{(1+z)/2}}{\Gamma(\frac{3+z}{2})}+\mathcal{O}\left(n^{(\Re z-1)/2}\right)$$

egyenletesen $z\in C$-re, ha $n\to\infty$.

BIZONYÍTÁS: Az előbbi lemmából

$$\mathbf{E}G_n(z) = \prod_{j=1}^{n-1} \frac{2j+1+z}{2j} = \prod_{j=1}^{n-1} \frac{j+\frac{1+z}{2}}{j} = \frac{2n}{1+z} \prod_{j=0}^{n-1} \frac{j+\frac{1+z}{2}}{j+1}.$$

Vezessük be az $\alpha = \frac{1+z}{2}$ jelölést. Ekkor az utóbbi produktum a $(-1)^n\binom{-\alpha}{n}$ binomiális együttható. Ennek nagyságrendjét [9], [7] szerint lehet meghatározni az

$$\frac{1}{2\pi i} \oint (1-x)^{-\alpha} x^{-n-1} dx$$

integrál segítségével. Az adódik, hogy

$$(-1)^n \binom{-\alpha}{n} = \frac{n^{\alpha-1}}{\Gamma(\alpha)} + \mathcal{O}(n^{\alpha-2}),$$

egyenletesen, minden kompakt halmazon. Tehát

$$\mathbf{E}G_n(z) = \frac{n}{\alpha}\left(\frac{n^{\alpha-1}}{\Gamma(\alpha)} + \mathcal{O}(n^{\alpha-2})\right) =$$

$$= \frac{n^{\alpha}}{\Gamma(\alpha+1)} + \mathcal{O}(n^{\alpha-1}) = \frac{n^{(1+z)/2}}{\Gamma(\frac{3+z}{2})} + \mathcal{O}\left(n^{(\Re z-1)/2}\right),$$

egyenletesen minden kompakt halmazon, amint $n \to \infty$. □

A következőkben az $M_n(z)$ martingál konvergenciáját fogjuk vizsgálni. Ehhez vizsgáljuk az L^2-beli viselkedését, pontosabban $G_n(z)$ kovarianciafüggvényét.

5. Lemma. *Minden $z_1, z_2 \in \mathbb{C}$ esetén*

$$C_{n+1}^G(z_1,z_2) := \mathbf{E}(G_{n+1}(z_1)G_{n+1}(z_2)) =$$

$$= \sum_{j=1}^{n}\left(b_j(z_1,z_2)\prod_{k=j+1}^{n} a_k(z_1,z_2)\right) + \prod_{j=1}^{n} a_j(z_1,z_2),$$

ahol

$$a_k(z_1,z_2) = 1 + \frac{1+(z_1+z_2)/2}{k} \quad \textit{és} \quad b_k(z_1,z_2) = \frac{1+z_1z_2}{2k}\mathbf{E}G_k(z_1z_2).$$

BIZONYÍTÁS: Megadunk egy lineáris rekurziót $C_n^G(z_1,z_2)$-re. Legyen k_n+1 az a valószínűségi változó, mely megadja, hogy hányadik szinten lesz az $n+1$-edik lépésben vett új csúcs. Ekkor $G_{n+1}(z)-G_n(z)=z^{k_n}$, hiszen eggyel nő a k_n+1-edik szint mérete, a többi nem változik. Ekkor

$$C_{n+1}^G(z_1,z_2)=\mathbf{E}\left[\mathbf{E}\left((G_n(z_1)+z_1^{k_n})(G_n(z_2)+z_2^{k_n})|\mathcal{F}_n\right)\right]=$$
$$=C_n^G(z_1,z_2)+\mathbf{E}\left[\mathbf{E}\left(G_n(z_1)z_2^{k_n}+z_1^{k_n}G_n(z_2)+z_1^{k_n}z_2^{k_n}|\mathcal{F}_n\right)\right].$$

Mivel ismerjük a $(k_n|\mathcal{F}_n)$ feltételes eloszlást, így

$$P(k_n=k)=\begin{cases}\frac{X[n,k]+X[n,k+1]}{2n}, \text{ ha } k>0,\\ \frac{X[n,1]}{2n}, \text{ ha } k=0\end{cases}$$

behelyettesítésével

$$\mathbf{E}(G_n(z_1)z_2^{k_n}|\mathcal{F}_n)=\frac{X[n,1]}{2n}G_n(z_1)+\sum_{k=1}^{\infty}\frac{X[n,k]+X[n,k+1]}{2n}G_n(z_1)z_2^k=$$
$$=\frac{1+z_2}{2n}G_n(z_1)G_n(z_2),$$

hasonlóan

$$\mathbf{E}(G_n(z_2)z_1^{k_n}|\mathcal{F}_n)=\frac{1+z_1}{2n}G_n(z_1)G_n(z_2),\quad \mathbf{E}(z_1^{k_n}z_2^{k_n}|\mathcal{F}_n)=\frac{1+z_1z_2}{2n}G_n(z_1z_2),$$

tehát

$$C_{n+1}^G(z_1,z_2)=\left(1+\frac{1+(z_1+z_2)/2}{n}\right)C_n^G(z_1,z_2)+\frac{1+z_1z_2}{2n}\mathbf{E}G_n(z_1z_2).$$

Ez a lineáris rekurzió pedig pontosan a felírt képletet adja, tekintve, hogy $C_1^G(z_1,z_2)=1$ bármely z_1,z_2-re. □

Mit mondhatunk ezután $M_n(z)$ L^2-beli viselkedéséről?

6. Következmény. *$\{M_n(z) : n\in\mathbb{N}\}$ korlátos L^2-ben, ha $|z-1|<\sqrt{2}$. Emiatt létezik $M(z)\in L^2$ valószínűségi változó, hogy $M_n(z)\to M(z)$ 1 valószínűséggel és L^2-ben, ha $n\to\infty$ és $z\in\mathcal{H}:=\{w\in\mathbb{C}:|w-1|<\sqrt{2}\}$.*

Bizonyítás: Az előző, 5. lemma jelölését használva

$$\prod_{k=j+1}^{n} a_k(z_1, z_2) = \left(\frac{n}{j}\right)^{1+(z_1+z_2)/2} \left(1 + \mathcal{O}\left(\frac{1}{j}\right)\right).$$

Vezessük be az $A \ll B$ jelölést, ha létezik $c > 0$ konstans úgy, hogy $A \leq cB$. Így a 4. lemma alapján

$$C_n^G(z_1, z_2) = \frac{1}{2}(1 + z_1 z_2) \sum_{j=1}^{n} \frac{\mathbf{E}G_j(z_1, z_2)}{j} \prod_{j=k+1}^{n} a_k(z_1, z_2) + \prod_{j=1}^{n} a_k(z_1, z_2) \ll$$

$$\ll \sum_{j=1}^{n} j^{\Re(1+z_1 z_2)/2-1} \left(\frac{n}{j}\right)^{\Re(z_1+z_2)/2+1} + n^{\Re(z_1+z_2)/2+1} \ll$$

$$\ll n^{\Re(z_1+z_2)/2+1} \sum_{j=1}^{n} j^{-\Re(3+z_1+z_2-z_1 z_2)/2}.$$

Ebből

$$C_n^M(z_1, z_2) := \mathbf{E}(M_n(z_1) M_n(z_2)) =$$

$$= \frac{\mathbf{E}(G_n(z_1) G_n(z_2))}{\mathbf{E}G_n(z_1)\mathbf{E}G_n(z_2)} \ll \sum_{j=1}^{n} j^{-\Re(3+z_1+z_2-z_1 z_2)/2}.$$

Tehát ha

$$\Re \frac{3 + z + \overline{z} - z\overline{z}}{2} = \frac{3}{2} + \Re z - \frac{|z|^2}{2} > 1,$$

akkor az összeg n-nek korlátos függvénye. A fenti egyenlőtlenség pedig pont $\mathcal{H}$-n teljesül, így $M_n(z)$ korlátos L^2-ben, ha $z \in \mathcal{H}$. □

Sőt, mint látni fogjuk, ha $z_1, z_2 \in \mathcal{H}$, akkor teljesül $\Re(3+z_1+z_2-z_1 z_2)/2 > 1$ is, így $C_n^M(z_1, z_2) \to C^M(z_1, z_2)$ egyenletesen $\mathcal{H}^2$ kompakt részhalmazain és $C^M(z)$ holomorf $\mathcal{H}^2$-en.

Az egyenletes konvergenciát úgy bizonyítjuk, mint [7]. Először szükségünk lesz a következő állításra.

7. Állítás. *Legyen* $I = (1-\sqrt{2}, 1+\sqrt{2})$. *Ekkor* $(M(t))_{t\in I}$*-nek van* $\tilde{M}$ *folytonos változata, amelyre minden kompakt* $C \subseteq I$ *esetén*

$$\mathbf{E}\left(\sup_{t\in C} |\tilde{M}|^2\right) < \infty.$$

Általánosabban, ha $\gamma : \mathbb{R} \to \mathcal{H}$ *folytonosan differenciálható görbe, akkor van* $(M_n(\gamma(t)))_{t\in R}$*-nek folytonos* $\tilde{M}_\gamma$ *változata, hogy minden kompakt* $C \subseteq \mathbb{R}$ *esetén*

$$\mathbf{E}\left(\sup_{t\in C} |\tilde{M}_\gamma(t)|^2\right) < \infty.$$

Bizonyítás: Nyilvánvalóan $\overline{M_n(z)} = M_n(\overline{z})$, így minden $z_1, z_2 \in \mathcal{H}$-ra

$$\mathbf{E}(|M(z_1) - M(z_2)|^2) = C^M(z_1, \overline{z_1}) + C^M(z_2, \overline{z_2}) - 2\Re(C^M(z_1, \overline{z_2}). \tag{1}$$

Ha $C' \subseteq \mathcal{H}$ kompakt, akkor C^M regularitása miatt, sorba fejtve a második tagig

$$C^M(z_1, \overline{z_1}) + C^M(z_2, \overline{z_2}) - 2\Re(C^M(z_1, \overline{z_2}) \leq K|z_1 - z_2|^2 \tag{2}$$

valamilyen $K > 0$ konstanssal minden $z_1, z_2 \in C'$-re. Tehát (1) és (2) alapján

$$\mathbf{E}(|M(z_1) - M(z_2)|^2) \leq K|z_1 - z_2|^2. \tag{3}$$

Így a Kolmogorov kritérium miatt ([16], 25. oldal), létezik egy folytonos $\tilde{M}$ változat és

$$\mathbf{E}\left(\sup_{s,t\in C'} \frac{|\tilde{M}(t) - \tilde{M}(s)|}{|t-s|^\alpha}\right)^2 < \infty$$

minden $\alpha \in (0, 1/2)$-re. Tehát minden kompakt $C \subseteq I$ esetén

$$\mathbf{E}\left(\sup_{t\in C} |\tilde{M}(t)|^2\right) < \infty.$$

Legyen ezek után $\gamma : \mathbb{R} \to \mathcal{H}$ folytonosan differenciálható. Tekintsük az $M_{n,\gamma} = M_n(\gamma(t))$ martingálokat ($t \in \mathbb{R}$). Ekkor (3) a következő alakot ölti:

$$\mathbf{E}(|M_\gamma(t_1) - M_\gamma(t_2)|^2) \leq K|\gamma(t_1) - \gamma(t_2)|^2 \leq K'|t_1 - t_2|^2$$

valamilyen $K' > 0$-val. Tehát $(M_\gamma(t))_{t\in\mathbb{R}}$-nek van folytonos $\tilde{M}_\gamma$ változata, amelyre

$$\mathbf{E}\left(\sup_{t\in C}|\tilde{M}_\gamma(t)|^2\right) < \infty$$

minden kompakt $C \subseteq \mathbb{R}$ esetén. □

(M_n) egyenletes konvergenciája a következő állításból következik, amely eredetileg [11]-ból való.

8. Állítás. *Minden kompakt $C \subseteq I$ esetén C-n 1 valószínűséggel egyenletesen*

$$M_n \to M,$$

$$\mathbf{E}\left(\sup_{t\in C}|M_n(t) - M(t)|^2\right) \to 0,$$

ha $n \to \infty$. Általánosabban, ha $\gamma : \mathbb{R} \to \mathcal{H}$ folytonosan differenciálható görbe, akkor legyen $M_{n,\gamma}(t) = M_n(\gamma(t))$ és $M_\gamma(t) = M(\gamma(t))$. Ekkor ugyanez igaz $(M_{n,\gamma})$-ra.

Bizonyítás: Legyen $[a,b] \subseteq I$. Az előző állítással kapott $\tilde{M}$ egy olyan valószínűségi változó, melynek értékei $E = \mathcal{C}([a,b],\mathbb{C})$ elemei, azaz folytonos komplex értékű függvények. Jelöljük $\mathcal{E}$-vel az E Borel halmazait, és $\mathcal{F}_\infty := \sigma(\mathcal{F}_n, n \geq 1)$, ekkor $\tilde{M}$ nyilván $\mathcal{F}_\infty$-mérhető.

Legyen $\varphi_t : E \to \mathbb{C}$ az a folytonos lineáris funkcionál, amely $X \in E$-hez $X(t)$-t rendeli. Ekkor $\mathbf{E}(\varphi_t(\tilde{M})|\mathcal{F}_n) = \varphi_t(\mathbf{E}(\tilde{M}|\mathcal{F}_n))$ m.m. Mivel $\tilde{M}$ M-nek folytonos változata, így minden $t \in [a,b]$-re 1 valószínűséggel $\varphi_t(\tilde{M}) = M(t)$. Tehát $M_n(t) = \mathbf{E}(\tilde{M}|\mathcal{F}_n)(t)$ m.m., azaz $M_n = \mathbf{E}(\tilde{M}|\mathcal{F}_n)$ m.m.

Most alkalmazzuk egy tételt vektorértékű martingálokról ([14], Proposition V-2-6, 104.o.). Eszerint

$$M_n = \mathbf{E}(\tilde{M}|\mathcal{F}_n) \to \mathbf{E}(\tilde{M}|\mathcal{F}_\infty) \text{ m.m. és } L^2_E\text{-ben is.}$$

Mivel $\tilde{M}$ $\mathcal{F}_\infty$-mérhető, így

$$\sup_{t\in[a,b]} |M_n(t) - \tilde{M}(t)| \to 0 \text{ m.m., ha } n \to \infty, \tag{4}$$

$$\mathbf{E}\left(\sup_{t\in[a,b]} |M_n(t) - \tilde{M}(t)|^2\right) \to_{n\to\infty} 0.$$

Tehát az állítás első része (4)-ből következik, hiszen $M(t) = \tilde{M}(t)$ m.m. minden $t \in [a, b]$-re.

A második rész ugyanígy bizonyítható, ha az $(M_{n,\gamma})$ és $\tilde{M}_\gamma$ martingálokra alkalmazzuk az előző lépéseket.

□

9. Következmény. *$M_n(z)$ és a deriváltjai egyenletesen konvergensek $\mathcal{H}$ kompakt részhalmazain.*

Bizonyítás: Az 8. állítás alapján M_n egyenletesen konvergens a $\gamma(t) = 1 + \rho e^{it}$ körön minden $0 < \rho < \sqrt{2}$ esetén, így mivel $|s - 1|\rho$ esetén

$$M_n(s) = \frac{1}{2\pi i}\oint_\gamma \frac{M_n(z)}{z-s}dz,$$

a Cauchy formula szerint, így M_n és a deriváltjai egyenletesen konvergensek $\mathcal{H}$ kompakt részhalmazain. □

Az 1. tétel bizonyításához szükség lesz $G_n(z)$ további becslésére két lemmában. Először $\mathbf{E}|G_n(z)|^2$-et becsüljük.

10. Lemma. *Minden $\delta > 0$-val, $|z - 1| \leq \sqrt{2} - \delta$ esetén*

$$\mathbf{E}|G_n(z)|^2 = \mathcal{O}(n^{\Re z+1}),$$

$\sqrt{2}-\delta \le |z-1| \le \sqrt{2}$ esetén

$$\mathbf{E}|G_n(z)|^2 = \mathcal{O}(n^{\Re z+1}\log n),$$

egyenletes hibatagokkal, ha $n \to \infty$. Ezen kívül, ha $C \subseteq \mathbb{C}-\mathcal{H}$ kompakt (azaz $z \in C$-re $|z-1| \ge \sqrt{2}$), akkor

$$\mathbf{E}|G_n(z)|^2 = \mathcal{O}(n^{(|z|^2+1)/2}\log n)$$

egyenletesen $z \in C$-re.

Bizonyítás: Idézzük fel a 6. következmény bizonyításában kapott becslést. Ez alapján

$$\mathbf{E}|G_n(z)|^2 \ll n^{\Re z+1}\sum_{j=1}^{n} j^{-(3-|z|^2)/2-\Re z}$$

Mivel pont az 1 körüli $\sqrt{2}$ sugarú kör az, ahol j kitevője kisebb, mint egy, ezért $|z-1| \le \sqrt{2}-\delta$ sugarú körön a kitevő legfeljebb $-1-\delta' < -1$, tehát

$$\mathbf{E}|G_n(z)|^2 \ll n^{\Re z+1}\sum_{j=1}^{n} j^{-1-\delta'} \ll n^{\Re z+1}.$$

Másrészt, ha $\sqrt{2}-\delta \le |z-1| \le \sqrt{2}$, akkor pedig

$$\mathbf{E}|G_n(z)|^2 \ll n^{\Re z+1}\sum_{j=1}^{n} j^{-1} \ll n^{\Re z+1}\log n.$$

Harmadrészt $|z-1| > \sqrt{2}$ esetén a szummát felülről becsülve

$$\mathbf{E}|G_n(z)|^2 \ll n^{\Re z+1}\sum_{j=1}^{n} j^{-(3-|z|^2)/2-\Re z} \ll n^{\Re z+1}\frac{n^{-(1-|z|^2)/2-\Re z}}{-(1-|z|^2)/2-\Re z}.$$

Az egyenletességhez azonban ez nem elég, előfordulhat, hogy a nevező tart 0-hoz, viszont tovább felülről becsülve

$$\mathbf{E}|G_n(z)|^2 \ll n^{\Re z+1}\frac{(n+1)^{-(1-|z|^2)/2-\Re z}-1}{-(1-|z|^2)/2-\Re z} \ll$$

$$\ll n^{(|z|^2+1)/2}\frac{1-e^{-((1-|z|^2)/2-\Re z)\log(n+1)}}{-(1-|z|^2)/2-\Re z}.$$

Ha most a nevező tart 0-hoz, akkor az exponenciális függvényt sorba fejtve látható, hogy

$$\mathbf{E}|G_n(z)|^2 \ll n^{(|z|^2+1)/2} \log n.$$

Ezzel az egyenletességet biztosítottuk, a lemmát bebizonyítottuk. $\square$

Most pedig $G_n(z)$ deriváltját becsüljük.

11. Lemma. *Minden $z \neq 0$ mellett, 1 valószínűséggel*

$$|G_n'(z)| \ll |z|^{-1} n^{(|z|+1)/2} \log n.$$

Bizonyítás: Nyilván $|G_n'(z)| \leq G_n'(|z|)$. Ismert, hogy a fa magassága $H_n \ll c \log n$. Tehát van n_0, hogy $n \geq n_0$-ra $X[n,k] = 0$ 1 valószínűséggel, ha $k > (c+1)\log n$. Így elég nagy n-re 1 valószínűséggel

$$G_n'(|z|) = \sum_{k=0}^{\infty} kX[n,k+1]|z|^{k-1} \leq (c+1)\log n \sum_{k=0}^{\infty} X[n,k+1]|z|^{k-1} =$$
$$= (c+1)\log n \frac{G_n(|z|)}{|z|},$$

ami a bizonyítandó becslés. $\square$

A tétel bizonyításában $G_n(z)$-t becsüljük majd a $|z| = 1$ körvonalon. A $\mathcal{H}$ körlapon kívüli részen a következő lemmával:

12. Lemma. *Minden $K > 0$-ra van $\delta > 0$, hogy 1 valószínűséggel*

$$\sup_{|z|=1, |z-1| \geq \sqrt{2}-\delta} |G_n(z)| = \mathcal{O}\left(\frac{n}{(\log n)^K}\right),$$

ha $n \to \infty$.

Bizonyítás: A Csebisev-egyenlőtlenség szerint

$$P\left(|G_n(z)| \geq \frac{n}{(\log n)^K}\right) \leq \frac{\mathbf{E}|G_n(z)|^2}{(n/(\log n)^K)^2}. \tag{5}$$

A számlálót a 10. lemma alapján becsülhetjük. Ha ugyanis $|z-1| \geq \sqrt{2}$, akkor $|z| = 1$ miatt

$$\mathbf{E}|G_n(z)|^2 << n \log n, \tag{6}$$

ha pedig $\sqrt{2} > |z-1| \geq \sqrt{2} - \delta$ valamilyen kis $\delta > 0$-val, akkor

$$\mathbf{E}|G_n(z)|^2 << n^{\Re z+1} \log n << n^{1+\sqrt{2}\delta-\delta^2/2} \log n, \tag{7}$$

mert az origó körüli 1 sugarú és az 1 körüli $\sqrt{2}$ sugarú kör $\pm i$-ben metszi egymást és ebből pontosan kiszámolhatjuk, hogy mennyi $\Re z$, ha $|z-1| = \sqrt{2} - \delta$.

Tehát (5), (6) és (7) alapján minden $\varepsilon > 0$-hoz van $\delta > 0$ ($\varepsilon = \sqrt{2}\delta - \delta^2/2$-ből visszaszámolva), hogy $|z| = 1$ és $|z-1| \geq \sqrt{2} - \delta$ mellett

$$P\left(|G_n(z)| \geq \frac{n}{(\log n)^K}\right) \leq \frac{\mathbf{E}|G_n(z)|^2}{(n/(\log n)^K)^2} << \frac{(\log n)^{2K+1}}{n^{1-\varepsilon}}. \tag{8}$$

Legyen $\varphi = \arg z_\delta$, ahol z_δ a $|z| = 1$, $|z-1| = \sqrt{2} - \delta$ egyenleteknek eleget tevő komplex számok közül a pozitív képzetes részű, tehát a két kör felső metszéspontja. Legyen $z(n,j) = e^{t(n,j)}$, ahol

$$t(n,j) = \varphi + (2\pi - 2\varphi)\frac{j}{(\log n)^{K+1}},$$

ahol $j = 0, 1, ..., (\log n)^{K+1}$. Ezek a pontok tehát a $|z| = 1$ kör $|z-1| \geq \sqrt{2} - \delta$ ívét osztják fel egyenlő részekre. Az únió valószínűségét a valószínűségek összegével becsülve tehát (8) miatt

$$P\left(|G_n(z(n,j))| \geq \frac{n}{(\log n)^K} \text{ valamely } j\text{-re}\right) \leq \frac{(\log n)^{3K+2}}{n^{1-\varepsilon}}.$$

Vegyük észre, hogy elég kis ε-ra

$$\sum_{n \geq 0} \frac{(\log(n^2))^{3K+2}}{n^{2(1-\varepsilon)}} < \infty,$$

így a Borel-Cantelli lemma miatt 1 valószínűséggel legfeljebb véges sok n kivételével

$$\sup_j |G_{n^2}(z(n^2,j))| \le \frac{n^2}{(\log(n^2))^K}.$$

Az 10. lemma segítségével becsülhetünk az osztópontok között. Tegyük fel, hogy $|z|=1$, $t(n^2,j) < \arg z < t(n^2,j+1)$. Ekkor egyenletesen

$$G_{n^2}(z) = G_{n^2}(z(n^2,j)) + \mathcal{O}(G'_{n^2}(1)(\log(n^2))^{-K-1}) \le$$
$$\le \frac{n^2}{(\log(n^2))^K} + \mathcal{O}(n^2(\log(n^2))^{1-K-1}) << \frac{n^2}{(\log(n^2))^K}.$$

Tehát 1 valószínűséggel legfeljebb véges sok kivétellel

$$\sup_{|z|=1, |z-1|\ge\sqrt{2}-\delta} |G_{n^2}(z)| \le \frac{n^2}{(\log(n^2))^K}.$$

Végül idézzük fel a $G_{n+1}(z) - G_n(z) = z^{k_n}$ egyenlőséget. Ez alapján $1 \le k < 2n+1$-re egyenletesen

$$G_{n^2+k}(z) = G_{n^2}(z) + \mathcal{O}(k) << \frac{n^2}{(\log(n^2))^K} + \mathcal{O}(n) << \frac{n^2+k}{(\log(n^2+k))^K}.$$

Tehát 1 valószínűséggel legfeljebb véges sok kivétellel

$$\sup_{|z|=1, |z-1|\ge\sqrt{2}-\delta} |G_n(z)| \le \frac{n}{(\log(n))^K}.$$

□

Ezek után rátérhetünk közvetlenül a tétel bizonyítására. Az $X[n,k]$ együtthatót a Cauchy-formula segítségével nyerjük ki a $G_n(z)$ generátorfüggvényből:

$$X[n,k+1] = \frac{1}{2\pi i}\int_{|z|=1} \frac{G_n(\xi)}{\xi^{k+1}} d\xi = \frac{1}{2\pi}\int_{-\pi}^{\pi} G_n(e^{it})e^{-kit}dt.$$

Az integrált bontsuk két részre:

$$I_1 := \frac{1}{2\pi}\int_{|t|\le\pi/2-\delta} G_n(e^{it})e^{-kit}dt,$$

$$I_2 := \frac{1}{2\pi}\int_{\pi\geq|t|\geq\pi/2-\delta} G_n(e^{it})e^{-kit}dt,$$

olyan δ-val, amely eleget tesz a 12. lemma feltételének. A lemmát alkalmazva tetszőleges $K>0$-ra becsülhetjük a második integrált a következőképpen:

$$|I_2| \leq \frac{1}{2\pi}\int_{\pi\geq|t|\geq\pi/2-\delta} |G_n(e^{it})|dt << \frac{n}{(\log n)^K}. \tag{9}$$

A $|t|\leq\pi/2-\delta$ esetben használjuk ki, hogy a 9. következmény miatt

$$M_n(e^{it}) = \frac{G_n(e^{it})}{\mathbf{E}G_n(e^{it})}$$

1 valószínűséggel egyenletesen korlátos. A nevező nagyságrendjét a 4. lemma alapján ismerjük és az egyenletes korlátosság miatt a számláló abszolút értéke ennek legfeljebb konstansszorosa. Tehát

$$|G_n(e^{it})| << n^{(1+\Re e^{it})/2} = n\cdot n^{(\Re e^{it}-1)/2} = n\cdot n^{(\cos t-1)/2} =$$
$$= ne^{(\cos t-1)(\log n)/2} << ne^{-c't^2(\log n)}$$

valamilyen $c'>0$ konstanssal. Rögzítsünk egy elég kis pozitív ϑ-t. Ekkor

$$\frac{1}{2\pi}\int_{(\log n)^{-(1-\vartheta)/2}\leq|t|\leq\pi/2-\delta} |G_n(e^{it})|dt << n\int_{(\log n)^{-(1-\vartheta)/2}}^{\infty} e^{-c't^2\log n}dt. \tag{10}$$

Ha Y normális eloszlású valószínűségi változó 0 várható értékkel és σ^2 szórásnégyzettel, akkor

$$\int_\eta^\infty e^{-\frac{t^2}{2\sigma^2}}dt = \sqrt{2\pi}\sigma P(Y\geq\eta) = \sqrt{2\pi}\sigma P\left(\frac{Y}{\sigma}\geq\frac{\eta}{\sigma}\right),$$

ahol $X_1 := Y/\sigma$ már standard normális, legyen X_2 szintén, de független az elsőtől. Ekkor

$$P\left(X_1\geq\frac{\eta}{\sigma}\right) \leq P\left(X_1^2\geq\frac{\eta^2}{\sigma^2}\right) \leq P\left(X_1^2+X_2^2\geq\frac{\eta^2}{\sigma^2}\right) = \exp\left(-\frac{1}{2}\frac{\eta^2}{\sigma^2}\right).$$

Így (10) tovább folytatható, ha $\sigma^2 = 1/(c' \log n)$-et és $\eta = (\log n)^{-(1-\vartheta)/2}$-t helyettesítünk a számolásba.

$$n \int_{(\log n)^{-(1-\vartheta)/2}}^{\infty} e^{-c't^2 \log n} dt << ne^{-c'(logn)^{\vartheta}}.$$

Látható, hogy ez bármilyen pozitív ϑ-ra aszimptotikusan kisebb, mint az $|I_2|$-re (9)-ben kapott becslés.

Tehát a teljes integrálból marad

$$I_0 := \frac{1}{2\pi} \int_{|t| \leq (\log n)^{-(1-\vartheta)/2}} G_n(e^{it}) e^{-kit} dt.$$

Definíció alapján

$$G_n(z) = \mathbf{E} G_n(z) M_n(z) \tag{11}$$

A 4. lemma a következő alakba is írható

$$\mathbf{E} G_n(z) = \frac{n^{(1+z)/2}}{\Gamma(\frac{3+z}{2})} + \mathcal{O}\left(n^{(\Re z - 1)/2}\right) = n \cdot n^{(z-1)/2} \left(\frac{1}{\Gamma(\frac{3+z}{2})} + \mathcal{O}\left(\frac{1}{n}\right)\right)$$

egyenletesen. Ebből, ha $t \to 0$, ráadásul úgy, hogy $|t| \leq (\log n)^{-(1-\vartheta)/2}$, akkor

$$\begin{aligned} \mathbf{E} G_n(e^{it}) &= ne^{\frac{1}{2}(e^{it}-1)(\log n)} \left(\frac{1}{\Gamma(\frac{3+e^{it}}{2})} + \mathcal{O}\left(\frac{1}{n}\right)\right) = \\ &= ne^{-(t^2/4)\log n + (it/2)\log n} \left(1 - \frac{it}{2}\Gamma'(2) - \frac{t^3}{12} i \log n + \mathcal{O}(t^4 \log n)\right). \end{aligned} \tag{12}$$

Másrészt, mivel $M_n(1) = 1$, így

$$M_n(e^{it}) = 1 + itM_n'(1) + \mathcal{O}(t^2). \tag{13}$$

Ekkor (11)-be behelyettesítve (12)-t és (13)-at, kapjuk, hogy k-ban egyenletesen 1 valószínűséggel

$$\begin{aligned} G_n(e^{it}) e^{-kit} &= ne^{it((\log n)/2 - k) - (t^2/4)\log n} \cdot \\ &\quad \cdot \left(1 + it\left(M_n'(1) - \frac{\Gamma'(2)}{2}\right) - i\frac{t^3}{12}\log n + \mathcal{O}(t^2 + t^4 \log n)\right). \end{aligned}$$

Parciális integrálással kapjuk, hogy

$$\int_{-\infty}^{\infty} e^{-(t^2/4)\log n}(t^2 + t^4 \log n)dt = 72\sqrt{\pi}(\log n)^{-3/2}.$$

Ugyanúgy becsülve, mint (10) kapjuk, hogy

$$\int_{|t|\geq(\log n)^{-(1-\vartheta)/2}} e^{-t^2 \log n}(1 + t + t^3 \log n) << e^{-(\log n)^{\vartheta}}.$$

Így azt kapjuk, hogy

$$\frac{I_0}{n} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{it((1/2)\log n - k)-(t^2/4)\log n}\cdot$$
$$\cdot\left(1 + it\left(M_n'(1) - \frac{\Gamma'(2)}{2}\right) - i\frac{t^3}{12}\log n\right) dt + \mathcal{O}((\log n)^{-3/2}).$$

Az integrálást elvégezve

$$\frac{I_0}{n} = \frac{2}{\sqrt{\pi \log n}} \exp\left(-\frac{(\frac{1}{2}\log n - k)^2}{\log n}\right)\cdot$$
$$\cdot\left(1 + \frac{\frac{1}{2}\log n - k}{\log n} - \frac{2(\frac{1}{2}\log n - k)^3}{3(\log n)^2} - 2\frac{\frac{1}{2}\log n - k}{\log n}(M_n'(1) - \Gamma'(2)/2)\right) +$$
$$+ \mathcal{O}((\log n)^{-3/2}).$$

Tehát 1 valószínűséggel, k-ban egyenletes hibataggal

$$\frac{X[n,k]}{2n/\sqrt{\pi \log n}} = \exp\left(-\frac{(\frac{1}{2}\log n - k)^2}{\log n}\right)\cdot$$
$$\cdot\left(1 + \frac{\frac{1}{2}\log n - k}{\log n} - \frac{2(\frac{1}{2}\log n - k)^3}{3(\log n)^2} - 2\frac{\frac{1}{2}\log n - k}{\log n}(M_n'(1) - \Gamma'(2)/2)\right) +$$
$$+ \mathcal{O}\left(\frac{1}{\log n}\right).$$

Ebből már következik a tétel állítása, hiszen $|\frac{1}{2}\log n - k| \leq c\sqrt{\log n}$ esetén

$$\frac{X[n,k]}{2n/\sqrt{\pi \log n}} = \exp\left(-\frac{(\frac{1}{2}\log n - k)^2}{\log n}\right)\left(1 + \mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right)\right).$$

Ez az eset a tétel lényeges állítása, azaz ha valamely köz epső szint mérett́ vizsgáljuk. Ha ugyanis k messze van $\frac{1}{2}\log n$-től, akkor az $\mathcal{O}\left(\frac{1}{\sqrt{\log n}}\right)$ tag meghatározó, így az alsó és felsö szintek méretére ad felsö becslést.

Ezzel a bizonyítást befejeztük.

$\square$

Itt szeretnék köszönetet mondani Móri Tamásnak témavezetői munkájáért.

Hivatkozások

[1] Barabási, A.-L., Albert, R., Emergence of scaling in random networks, *Science* **286** (1999), 509-512.

[2] Barabási, A.-L., Albert, R., Hawoong Jeong Mean-field theory for scale-free random networks, *Physica A* **272** (1999) 173-187.

[3] Barabási, A.-L., Albert, R., Hawoong Jeong Scale-free characteristics of random networks: the topology of the world-wide web, *Physica A* **281** (2000) 69-77.

[4] Bollobás, B. Random Graphs, *Academic Press,* London (1985)

[5] Bollobás, B., Riordan, O., The diameter of a scale-free random graph, *Combinatorica* (2003)

[6] Bollobás, B., Riordan, O., Spencer, J., Tusnády, G., The degree sequence of a scale-free random graph process, *Random Structures Algorithms* **18** (2001), 279-290.

[7] Chauvin, B., Drmota, M., Jabbour-Hattab, J., The profile of binary search trees, *The Annals of Applied Probability* **11** (2001), 1042-1062.

[8] Erdős, P., Some remarks on the theory of graphs, *Bull. Res. Council Israel, Section F* **7** (1947) 292-294.

[9] Flajolet, P., Odlyzko, A.M. Singularity analysis of generating functions, *SIAM J. Discrete Math.* **3** (1990), 216-240

[10] Gilbert, E.N., Random graphs, *Ann. Math. Statist.* **30** (1959), 1141-1144.

[11] JOFFE, A. LE CAM, L., NEVEU, J. Sur la loi des grands nombres pour des variables aléatoires attachées à un arbre dyadique, *C. R. Acad. Sci. Paris. Ser. A* **277** (1973), 963-964.

[12] MÓRI, T.F.,On random trees, *Studia Sci. Math. Hungar.* **39** (2002), 143-155.

[13] MÓRI, T.F.,The maximum degree of the Barabási random tree, *Combinatorics, Probability & Computing* (megjelenés alatt)

[14] NEVEU, J., Martingales à temps discret, *Masson and Cie, Paris* (1972)

[15] RÉNYI, A., SZEKERES, G., On the height of trees, *J. Austral. Math. Soc.* **7** (1967) 497-507.

[16] REVUZ, D., YOR, M. Continuous Martingales and Brownian Motion, *Springer, New York* (1990)

[17] SZELE, T., Combinatorial investigations concerning complete directed graphs, *Matematikai és Fizikai Lapok* **50** (1943) 287-293.

[18] WATTS, D.J., STROGATZ, S.H., Collective dynamics of small-worlds' networks, *Nature* **393** (1998), 440-442.