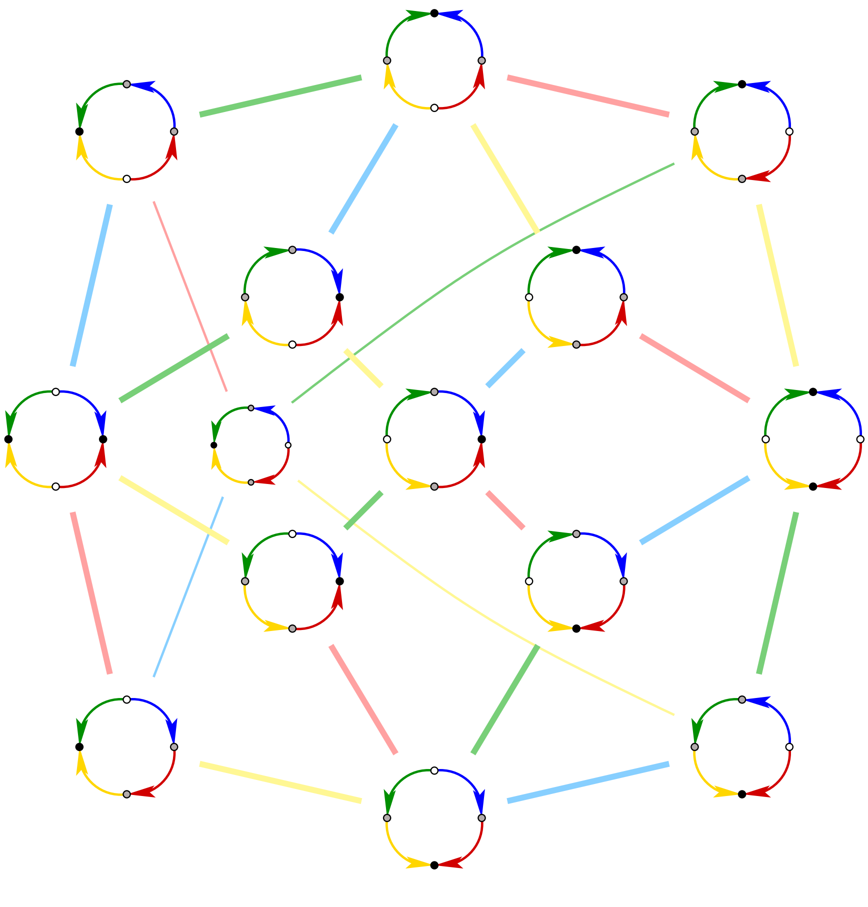
Graphentheorie

Erweiterung des Teilnehmerskriptes

zu einer Vorlesung von
Stefan Felsner
Wintersemester 2013/14
Technische Universitaet Berlin
erweiter fuer die Vorlesung
Wintersemester 2023/24



Die 14 azyklischen Orientierungen des 4-Kreises.

23 Jan. 2018

Inhaltsverzeichnis

Graphentheorie, TU Berlin, WS 2023/24

Dozentin: Stefan Felsner/Alexandra Wesolek
Mitschrift: Mirko Vogel
Datum: 13.11.2023

Vorlesung
8

Extremale Graphentheorie

Ein F-freier graph ist ein Graph der F nicht als Teilgraphen enthält. Zum Beispiel wenn G ein bipartiter Graph ist, dann ist G ein C_{2k+1}-freier Graph für jeden Kreis C_{2k+1} mit $2k+1$ Knoten (Proposition 5.2).

8.1 Satz von Mantel

Eine Fluggesellschaft möchte Flüge anbieten, so dass wenn es einen Direktflug von A nach B gibt und von B nach C gibt, dann gibt es keinen Direktflug von A nach C. Die Fluggesellschaft möchte so viele Flüge wie möglich anbieten. Wie viele Flüge kann sie bei n Städten anbieten

Theorem 8.1 (**Satz von Mantel, 1907**)
Für die Anzahl der Kanten eines dreiecksfreien Graphen G gilt:

$$|E)G)| \leq \left\lfloor \frac{n^2}{4} \right\rfloor$$

und $K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$ *ist (bis auf Isomorphie) der eindeutige maximierende Grap h.*

Beweis. Wir führen den Beweis per Induktion über n. Da wir im Induktionsschritt jeweils von $n-2$ nach n schließen, brauchen wir zwei Induktionsanfänge: Für $n \in \{1,2\}$ ist die Aussage offenbar wahr.

Sei $G=(V,E)$ nun ein dreiecksfreier Graph mit n Knoten und $\{u,v\} \in E$. Da u und v keinen gemeinsamen Nachbarn haben, gilt $(\deg(u)-1)+(\deg(v)-1) \leq n-2$. Für die Anzahl der mit u oder v inzidenten Kanten gilt also:

$$1+(\deg(u)-1)+(\deg(v)-1) \leq n-1.$$

Auf $G' = G-u-v$, also den Graphen der sich durch Löschen der Knoten u und v (sowie aller inzidenten Kanten) ergibt, können wir die Induktionsvorraussetzung anwenden:

$$|E| \leq n-1+\left|E(G')\right| \leq n-1+\left\lfloor \frac{(n-2)^2}{4} \right\rfloor = \left\lfloor \frac{4(n-1)+(n-2)^2}{4} \right\rfloor = \left\lfloor \frac{n^2}{4} \right\rfloor$$

Dass $K_{\lceil \frac{n}{2} \rceil, \lfloor \frac{n}{2} \rfloor}$ der eindeutige kantenmaximale Graph ist, kann man genauso per Induktion zeigen. □

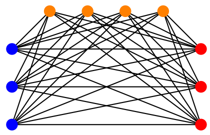


Abbildung 1: $T(10,3)$

8.2 Turán-Zahlen

Für einen Graphen F ist die *Turân-Zahl* $ex(n,F)$ die maximale Anzahl an Kanten in einem F-freien Graphen mit n Knoten.

Wir erinnern an die Definition einer vollständigen Menge: $I \subset V$ ist eine *vollständig Menge* (eine Clique) in G, falls der induzierte Teilgraph $G[I]$ isomorph zum vollständigen Graph ist: $G[I] \simeq K_{|I|}$. (Man sagt G sei K_n-frei, falls G keine Clique der Größe n enthält.) Die *Cliquenzahl* von G ist die Größe der größten vollständigen Menge:

$$\omega(G) = \max\{|I| : I \subset V \text{ vollständig in } G\}$$

Theorem 8.2 (Turán, 1941)
The Turán graph $T(n,k)$ maximizes the number of edges of K_{r+1}-free graphs, i.e.

$$ex(n, K_{k+1}) = e(T(n,k)) \leq \left(1 - \frac{1}{k}\right)\frac{n^2}{2}.$$

Der vollständige k-partite Graph $K_{n_1,\ldots,n_k}$ mit $\sum n_i = n$ und $\left|n_i - n_j\right| \leq 1$, der auch als Turán-Graph $T_k(n)$ bezeichnet wird, ist bis auf Isomorphie der einzige maximierende Graph.

The Turán graph $T(n,r)$ is a *balanced* complete r-partite graph.

Beispiel. Betrachte tripartite Graphen ($k=3$) mit 10 Knoten, siehe Abbildung 1. Die Schranke aus dem Satz ist $\left(1-\frac{1}{3}\right)\frac{10^2}{2} = 33,\bar{3}$. Für die kantenzahl des Turán-Graphen $T_3(10) = K_{4,3,3}$ gilt: $|E(K_{4,3,3})| = 4 \cdot 3 + 3 \cdot 3 + 4 \cdot 3 = 33$.

Wir stellen zuerst fest, dass das Verbieten von Cliquen der Größe $k+1$ äquivalent ist zum Verbieten von Cliquen der Größe $\geq k+1$. Insofern können wir auch nach der maximal möglichen Kantenanzahl eines Graphen, der nur Cliquen der Größe $\leq k$ enthält, fragen:

$$\max\left\{|E| : \text{ es gibt } G(V,E) \text{ mit } \omega(G) \leq k\right\} = ?$$

Intuitiv ist klar, dass inklusionsmaximale Cliquen der Größe $< k$ Kanten „verschwenden“,sodass ein in Bezug auf die Kantenanzahl maximaler Graph nur maximale Cliquen des Größe k haben sollte — eine Forderung, die vollständige k-partite Graphen erfüllen.

Beweis. [Beweis mittels Verschieben von Gewichten]

Sei $G = (V, E)$ ein einfacher Graph mit $V = \{v_1, \ldots, v_n\}$. Für einen Gewichtsvektor $w \in \mathbb{R}^n_{\geq 0}$ mit $\sum_{i=1}^n w_i = 1$ definieren wir den Gewichtsausdruck

$$g(w) = \sum_{\{v_i, v_j\} \in E} w_i w_j.$$

Wir können g nutzen, um eine Aussage über die Kantenanzahl von G zu machen, indem wir den „Einheitsgewichtsvektor“ $w_0 = \frac{1}{n}\mathbb{1}$ einsetzen: $g(w_0) = \frac{1}{n^2}|E|$.

Ein Maximierer $w^{\max}$ von g liefert dann eine obere Schranke für $|E|$:

$$|E| \leq n^2 g(w^{\max}) \tag{$*$}$$

Nun zeigen wir die Existenz eines Maximierers $w^{\max}$, dessen Gewicht auf einer Clique gleichverteilt ist. Dabei verschieben wir das Gewicht iterativ auf die Clique, in dem wir in jedem Schritt aus einem Gewichtsvektor w einen neuen w' erzeugen, s.d. $g(w') \geq g(w)$.

1. *Konzentriere das Gewicht auf einer Clique*
Falls w das Gewicht noch nicht auf einer Clique konzentriert, dann gibt es $v_i, v_j \in V$ mit $w_i, w_j \neq 0$ und $\{v_i, v_j\} \notin E$. Seien

$$s_i := \sum_{v_k \in N(v_i)} w_k \quad \text{und} \quad s_j := \sum_{v_k \in N(v_j)} w_k$$

die Gewichte der Nachbarschaften von v_i und v_j. Wir nehmen an, dass $s_i \geq s_j$ gilt. Erzeuge nun einen neuen Gewichtsvektor w' durch Verschieben des Gewichts von Knoten v_j auf den Knoten v_i:

$$w'_i := w_i + w_j, \quad w'_j := 0, \quad \text{und } w'_k := w_k \,\forall k \neq i, j$$

Es gilt $g(w') \geq g(w)$, da

$$\begin{aligned} g(w') - g(w) &= \left(w'_i s_i + w'_j s_j\right) - \left(w_i s_i + w_j s_j\right) \\ &= s_i\left(w_i + w_j - w_i\right) - s_j w_j \geq 0. \end{aligned}$$

2. *Verteile die Gewichte gleichmäßig auf der Clique*
Sei C eine Clique, auf der sich die Gewichte von w konzentrieren. Falls die Gewichte der Knoten noch nicht gleich sind, so gibt es $v_i, v_j \in C$ mit $w_i > w_j$. Erzeuge nun einen neuen Gewichtsvektor w' durch gleichmäßiges Verteilen des Gewichts $w_i + w_j$ auf beiden Knoten v_i und v_j:

$$w'_i := w'_j := \tfrac{w_i + w_j}{2}, \quad \text{und } w'_k := w_k \,\forall k \neq i, j.$$

Es gilt $g(w') \geq g(w)$, da

$$\begin{aligned} g(w') - g(w) &= 2\left(\frac{w_i + w_j}{2}\right)\left(1 - \frac{w_i + w_j}{2}\right) - w_i(1 - w_i) - w_j(1 - w_j) \\ &= \frac{1}{2}\left(w_i - w_j\right)^2 > 0 \end{aligned}$$

Sei w nun also ein Gewichtsvektor, der seine Masse gleichmäßig auf einer Clique C der Größe t verteilt, d.h $w_i = \frac{1}{t}$ für alle $v_i \in C$. Dann gilt:

$$g(w) = \sum_{v_i \neq v_j \in C} w_i w_j = \sum_{v_i \neq v_j \in C} \frac{1}{t^2} = \frac{t(t-1)}{2} \frac{1}{t^2} = \frac{1}{2}\left(1 - \frac{1}{t}\right) \leq \frac{1}{2}\left(1 - \frac{1}{k}\right)$$

Die letzte Ungleichung folgt hier aus $t \leq \omega(G) \leq k$. Also gilt $g(w^{\max}) \leq \frac{1}{2}\left(1 - \frac{1}{k}\right)$ und mit $(*)$ ist der Satz bewiesen. □

Beweis. [Beweis durch Induktion und Rechnung]

Wir verallgemeinern den Beweis des Satzes von Mantel (Prpp. 8.1): Falls G eine maximale Anzahl von Kanten hat, so enthält G mindestens eine k-Clique A. Nun können wir die Kanten in G aufteilen in die Kanten innerhalb von A, die Kanten innerhalb des Komplements $B = V \backslash A$ sowie die Kanten zwischen A und B. Für die Anzahl der Kanten in B können wir Induktion verwenden. Es gilt:

$$|E[A]| \leq \binom{k}{2}, \quad |E[B]| \leq \left(1 - \frac{1}{k}\right) \frac{(n-k)^2}{2}, \quad |\{vw \in E : v \in A, b \in B\}| \leq (n-k)(k-1)$$

Die letzte Ungleichung ergibt sich aus der Tatsache, dass jeder der $n-k$ Knoten in B maximal $k-1$ Nachbarn in A haben kann, sonst gäbe es eine $k+1$ Clique. Die Behauptung folgt durch Ausrechnen der Summe. □

Beweis. [Beweis, der uns uns ein maximierendes Beispiel schenkt]

Wir zeigen zuerst, dass ein kantenmaximaler Graph G isomorph zu einem vollständigen multipartiten Graphen $K_{n_1,\ldots,n_t}$ ist. Hierfür stellen wir die Knotenmenge V als disjunkte Vereinigung unabhängiger Mengen dar. Hilfsmittel ist die durch

$$u \sim v \iff uv \notin E$$

gegebene Relation auf V. Offenbar ist $\sim$ symetrisch und reflexiv (G ist einfach, enthält also keine Schleifen). Damit $\sim$ eine Zerlegung von V in Äquivalenzklassen liefert, müssen wir Transitivität zeigen. Für allgemeine Graphen ist sie nicht gegeben, wohl aber für $K_1 + K_2$-freie, d.h. falls es keine drei Knoten u, v, w gibt, so dass der induzierte Graph $G[u, v, w]$ genau eine Kante enthält.

Wir zeigen nun, dass jeder kantenmaximale Graph $K_1 + K_2$-frei ist: Angenommen es gäbe $u, v, w \in V$ mit $vw \in E$ und $uv, uw \notin E$. Wir nehmen an, dass $\mathsf{deg}(v) \geq \mathsf{deg}(w)$ und unterscheiden zwei Fälle:

1. Falls $\mathsf{deg}(u) < \mathsf{deg}(v)$: Lösche u und verdopple v (d.h. wir führen einen neuen Knoten v' mit $N(v) = N(v')$ ein). Der neue Graph G' hat mehr Kanten als G und offenbar $\omega(G') \leq k$. Das steht im Widerspruch zur Maximalität von G.

2. Falls $\mathsf{deg}(u) \geq \mathsf{deg}(v)$: Lösche v, w und verdreifache u (d.h. wir führen neue Knoten u', u'' mit $N(u) = N(u') = N(u'')$ ein). Da die Kante vw nur einmal gelöscht wird, hat der neue Graph G' mehr Kanten als G ($3\mathsf{deg}(u) > \mathsf{deg}(u) + \mathsf{deg}(v) + \mathsf{deg}(w) - 1$), was wegen $\omega(G') \leq k$ im Widerspruch zur Maximalität von G steht.

Somit haben wir gezeigt, dass $\sim$ eine Äquivalenzrelation auf V ist. Also lässt sich V als disjunkte Vereinigung unabhängiger Mengen darstellen: $V = \bigcup_{j=1}^{t} I_j$ mit $\left|I_j\right| = n_j$. Ein kantenmaximaler Graph ist also isomorph zu einem $K_{n_1,\ldots,n_t}$.

Aus der Forderung der K_{k+1}-Freiheit folgt $t \leq k$. Man überlegt sich nun noch, dass $t = k$ und $n_i \in \left[\left\lfloor \frac{n}{k} \right\rfloor, \left\lfloor \frac{n}{k} \right\rfloor + 1\right]$ die Kantenanzahl maximiert. Dafür kann man z.B. wieder Argumente aus dem ersten Beweis verwenden (zu $K_{n_1,\ldots,n_t}$ gehört eine Gewichtung mit $w_i = n_i/n$). □

Beweis. [Beweis mittels Cauchy-Schwarz-UG]
Wir verwenden eine untere Schranke für die Cliquenzahl,

$$\omega(G) \geq \sum_{v \in V} \frac{1}{n - \deg(v)}, \tag{$\star$}$$

welche sich direkt aus dem Satz von Wei (siehe VL 12) ergibt, sowie die bekannte Cauchy-Schwarz-Ungleichung

$$\langle a, b\rangle^2 \leq \|a\|^2 \|b\|^2 .$$

Wir definieren

$$a_i = \frac{1}{\sqrt{n - \deg(v_i)}} \quad \text{und} \quad b_i = \sqrt{n - \deg(v_i)}$$

und bekommen wegen $\langle a, b\rangle = n$ die folgende Ungleichungskette:

$$n^2 = \langle a, b\rangle^2 \leq \sum_{i=1}^{n} a_i^2 \cdot \sum_{i=1}^{n} b_i^2 = \sum_{i=1}^{n} \frac{1}{n - \deg(v_i)} \cdot \sum_{i=1}^{n} (n - \deg(v_i)) \leq k\left(n^2 - 2|E|\right)$$

wobei wir für die letzten Ungleichung ($\star$) und die Voraussetzung $\omega(G) \leq k$ verwandt haben. Die Behauptung folgt durch Auflösung der Ungleichung nach $|E|$. □

Graphentheorie, TU Berlin, WS 2023/24

Dozentin: Stefan Felsner/Alexandra Wesolek
Mitschrift: Alexandra Wesolek
Datum: 15.11.2023

Vorlesung
9

Graph Limits

These are preliminary notes. Better notes to come.

9.1 Review

In this whole section we assume an N-vertex graph is a graph on the vertex set $\{1,\ldots,N\}$.

Recall that a homomorphism from a graph H to a graph G is an edge-preserving map. This means it is a map $\phi : V(H) \to V(G)$ such that if uv is an edge in H then $\phi(u)\phi(v)$ is an edge in G.

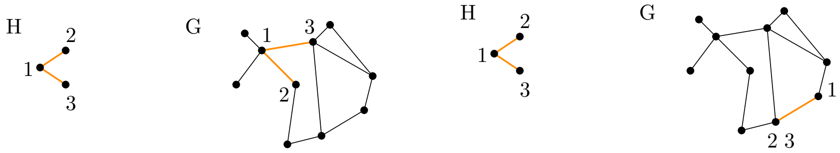


Abbildung 1: Two different homomorphisms from H to G.

9.2 Triangle Removal Lemma

The main goal of this section is to prove the following theorem:

Theorem 9.1 *For all $\varepsilon > 0$ there exists $\varepsilon' > 0$ such that for any N, if G is an N-vertex graph with at most $\varepsilon' N^3$ triangles then there exists a set $E' \subset E(G)$ of εN^2 edges so that after deleting those edges, the remaining graph $G' = G - E'$ has no triangles.*

Beobachtung 1. If G has $\varepsilon' N^3$ triangles, we have to delete more than $\varepsilon' N^2$ edges to make the graph triangle-free: every edge is in at most $N-2$ triangles, so deleting $\varepsilon' N^2$ edges deletes less than $\varepsilon' N^3$ triangles.

Beobachtung 2. If this theorem would be false, then there would exist some $\varepsilon > 0$ such that for every $\varepsilon' > 0$ there exists a graph $G_{\varepsilon'}$ with some number of vertices N (N depends on ε') and at most $\varepsilon' N^3$ triangles such that removing any εN^2 edges from G would keep at least one triangle.

Beobachtung 3. Note that

$$3!|\{K_3 \text{ in } G\}| = |\{\text{homomorphisms from } H = K_3 \text{ to } G\}|$$

since each triangle can be mapped to another triangle in 3! ways.

Definition 9.2 *We define $t(H,G)$ to be the probability that a random map $\phi : V(H) \to V(G)$ preserves edges. That is*

$$t(H,G) = \frac{\#\textit{homomorphisms from } H \textit{ to } G}{\textit{all maps from } H \textit{ to } G} = \frac{\#\textit{homomorphisms from } H \textit{ to } G}{|V(G)|^{|V(H)|}}.$$

Beispiel. Let $H = K_2$ with $V(K_2) = \{1,2\}$. Then $t(K_2,G) = \frac{2|E(G)|}{|V(G)|^2}$, since the edge 12 of K_2 can be mapped in two different ways to an edge $ij \in E(G)$ (either $1 \to i$ and $2 \to j$ or $1 \to j$ and $2 \to i$). Note that $t(K_2,K_N) = \frac{N(N-1)}{N^2} \to 1$.

Beispiel. Let $H = K_3$ and G be an N-vertex graph with $\varepsilon' N^3$ triangles. Then

$$t(K_3,G) \leq \frac{3!\varepsilon' N^3}{N^3} = 6\varepsilon'.$$

Beispiel. Let $H = G = K_3$ be the triangle on $V(K_3) = \{1,2,3\}$. Then

$$t(K_3,K_3) = \frac{3!}{3^3} = \frac{2}{9}.$$

Note that $\frac{2}{9}$ is the probability that mapping $\{1,2,3\}$ randomly to $\{1,2,3\}$ gives a triangle. For example if $1 \to 1, 2 \to 1, 3 \to 1$ then the image of $\{1,2,3\}$ is only the vertex $\{1\}$ and not a triangle whereas $1 \to 2, 2 \to 1, 3 \to 3$ gives a triangle.

9.3 Basics of graph limits

In fact, proving the Triangle Removal Lemma is not the only goal of this section. We also introduce a beautiful recent limit theory for graphs, introduced by Lovász and Szegedy. Let $e_G(S,T)$ be the number of edges in G between a set of vertices S and a set of vertices T.

Definition 9.3 *We define for two N-vertex graphs G, G' the labelled and unlabelled cut distance $d_\square$ and $\delta_\square$ as*

$$d_\square(G,G') = \max_{S,T \subset [N]} \frac{|e_G(S,T) - e_{G'}(S,T)|}{N^2},$$
$$\delta_\square(G,G') = \min_{\pi:[N]\to[N]} \max_{S,T \subset [N]} \frac{|e_G(S,T) - e_{G' \circ \pi}(S,T)|}{N^2},$$

where $G' \circ \pi$ has vertices $V(G' \circ \pi) = V(G)$ and ij is an edge in $G' \circ \pi$ if and only if $\pi^{-1}(i)\pi^{-1}(j)$ is an edge in G'.

Definition 9.4 *Let G be a graph on the vertex set $\{1,\dots,N\}$. For an edge $ij \in E(G)$ we define the square associated to ij as $R_{i,j} = [(i-1)/N, i/N] \times [(j-1)/N, j/N]$. We define a function $W_G : [0,1]^2 \to \{0,1\}$ as*

$$W_G(x,y) = \begin{cases} 1 & \text{if } (x,y) \in R_{i,j} \text{ for some } ij \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

Further we define $I_k = [(k-1)/N, k/N]$.

We consider the set $\mathcal{W} := \{\text{symmetric measurable functions } W : [0,1]^2 \to [0,1]\}$. We define a similar notion of distance for this set.

Definition 9.5 *We define for $W, W' \in \mathcal{W}$,*

$$d_\square(W,W') = \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W(x,y) - W'(x,y)dxdy \right|$$

$$\delta_\square(W,W') = \inf_{\phi:[0,1]\to[0,1],\phi \text{ measure-preserving}} \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W(x,y) - W'(\phi(x),\phi(y))dxdy \right|$$

Further we introduce the following norm on the set of measurable functions on $[0,1]^2$.

$$\|W\|_\square = \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W(x,y)dxdy \right|.$$

We begin by proving the following lemma, which follows from the norm-definition we used above.

Lemma 9.6 *Let $W_i : [0,1]^2 \to [-1,1]$ be such that $\|W_i\|_\square \to 0$ for $i \to \infty$. Then for every $Z \in L_1([0,1]^2)$ also $\|Z \cdot W_i\|_\square \to 0$.*

Beweis. Case 1: Z is the indicator function of a rectangle $S' \times T'$. Then

$$\begin{aligned} \|Z \cdot W_i\|_\square &= \sup_{S,T\subset[0,1]} \left| \int_{S\times T} Z \cdot W_i(x,y)dxdy \right| \\ &= \sup_{S\subset S',T\subset T'} \left| \int_{S\times T} W_i(x,y)dxdy \right| \\ &\le \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W_i(x,y)dxdy \right| = \|W_i\|_\square. \end{aligned}$$

Hence $\|W_i\|_\square \to 0$ for $i \to \infty$ implies $\|Z \cdot W_i\|_\square \to 0$.
Case 2: Z is a step function, i.e. $Z = \sum_{j=1}^k Z_j$ for some rectangle indicator functions Z_j. Then

$$\left\| \sum_{j=1}^k Z_j \cdot W_i \right\|_\square \le \sum_{j=1}^k \left\| Z_j \cdot W_i \right\|_\square.$$

By Case 1, $\left\|Z_j \cdot W_i\right\| \to 0$ for $i \to \infty$. Therefore $\sum_{j=1}^{k} \left\|Z_j \cdot W_i\right\|_\square \to 0$ for $i \to \infty$ and therefore $\left\|\sum_{j=1}^{k} Z_j \cdot W_i\right\|_\square \to 0$.

Case 3: $Z \in L^1([0,1]^2)$. Then Z can be approximated by step functions $Z_1, Z_2, \dots$ such that

$$\int_{[0,1]^2} |Z(x,y) - Z_n(x,y)| dx dy \to 0$$

for $n \to \infty$. But then for m, n large enough

$$\begin{aligned}
&\sup_{S,T\subset[0,1]^2} \left| \int_{S\times T} Z(x,y)\cdot W_n(x,y) dxdy \right| \\
&= \sup_{S,T\subset[0,1]^2} \left| \int_{S\times T} (Z(x,y) - Z_m(x,y) + Z_m(x,y))\cdot W_n(x,y) dxdy \right| \\
&\le \sup_{S,T\subset[0,1]^2} \left| \int_{S\times T} (Z(x,y) - Z_m(x,y))\, W_n(x,y) dxdy \right| + \sup_{S,T\subset[0,1]^2} \left| \int_{S\times T} Z_m(x,y)\cdot W_n(x,y) dxdy \right| \\
&\le \sup_{S,T\subset[0,1]^2} \int_{S\times T} \left| Z(x,y) - Z_m(x,y) \right| \cdot \left| W_n(x,y) \right| dxdy + \|Z_m W_n\|_\square < 2\varepsilon,
\end{aligned}$$

where we take m large enough such that the first term is smaller than ε and then we take n large enough such that the second term is smaller than ε. Note that in the first term we can bound $\left|W_n(x,y)\right| \le 1$. For the second term we reduce to Case 2.

□

The following two theorems are powerful theorems developed in the area of graph limits.

Theorem 9.7 *For any sequence of graph $G_1, G_2, \dots$ there exists a subsequence $G_{i,1}, G_{i,2}, \dots$ and $W \in \mathcal{W}$ such that*

$$W_{G_{i_j}} \to W$$

for $j \to \infty$, where the convergence is in $\delta_\square$-distance (as defined on $\mathcal{W}$). Further,

$$t(H, G_{i_j}) \to t(H, W) := \int_{[0,1]^{|V(H)|}} \left(\prod_{\{k,\ell\}\in E(H)} W(x_k, x_\ell) \right) dx_1, \dots, dx_n$$

for $j \to \infty$.

Beispiel. Let G be an N-vertex graph with a triangle K_3 on $\{1,2,3\}$ and $N-3$ isolated vertices (that is, vertices with no neighbours). We consider W_G. Let $I_1 = [0, 1/N], I_2 =$

$[1/N, 2/N], I_3 = [2/N, 3/N]$. Then $W_G(x, y) = 1$ if and only if $x \in I_i$ and $y \in I_j$ with $i \neq j$ and $1 \leq i, j \leq 3$. Further for $H = K_3$

$$t(K_3, W_G) = \int_{[0,1]^3} W_G(x_1, x_2) W_G(x_2, x_3) W_G(x_3, x_1)\, dx_1, \ldots, dx_n.$$

Note that $W_G(x_1, x_2) W_G(x_2, x_3) W_G(x_3, x_1) = 1$ exactly if $x_1 \in I_i, x_2 \in I_j, x_3 \in I_k$ with i, j, k pairwise different and in $\{1, 2, 3\}$. Hence

$$\begin{aligned} t(K_3, W_G) &= \int_{[0,1]^3} W_{K_3}(x_1, x_2) W_{K_3}(x_2, x_3) W_{K_3}(x_3, x_1)\, dx_1, \ldots, dx_n. \\ &= |I_1 \times I_2 \times I_3| + |I_1 \times I_3 \times I_2| + \cdots + |I_3 \times I_2 \times I_1| \\ &= 3!|I_1 \times I_2 \times I_3| = \frac{6}{N^3} = t(K_3, G). \end{aligned}$$

Theorem 9.8 *If $G_1, G_2, \ldots$ is a sequence of graphs such that*

$$W_{G_i} \to W$$

in $\delta_\square$-distance, then there exist permutation $\pi_1, \pi_2, \ldots$ such that

$$W_{G_i \circ \pi_i} \to W$$

in $d_\square$-distance.

We will use those theorems to proof the triangle removal lemma. For convenience, we restate the theorem.

Theorem 9.1 *For all $\varepsilon > 0$ there exists $\varepsilon' > 0$ such that for any N, if G is an N-vertex graph with at most $\varepsilon' N^3$ triangles then there exists a set $E' \subset E(G)$ of εN^2 edges so that after deleting those edges, the remaining graph $G' = G - E'$ has no triangles.*

Beweis. The proof goes by contradiction. Note that by Beobachtung 3 to Theorem 9.1 for any n there exists a graph G_n with N vertices (N depends on n) with $\frac{1}{n} N^3$ triangles such that removing any εN^2 edges from G_n does not remove all triangles in the graph. By Theorem 9.7 and Theorem 9.8 we can assume that

$$\|W_{G_n} - W\|_\square \to 0 \tag{9.1}$$

and since $t(K_3, G_n) \to 0$ also $t(K_3, W) = 0$.

Main strategy: Let $S = \{(x, y) \in [0, 1]^2 | W(x, y) > 0\}$. We delete an edge $ij \in E(G_n)$ if $\lambda(S \cap R_{i,j}) < \frac{3}{4} \lambda(R_{i,j})$ (recall that $R_{i,j} = J_i \times J_j$ with $J_i = [(i-1)/N, i/N]$ is the square associated to the edge ij).

The idea is : W_{G_n} converges to W, so G_n should not have too many edges whose square $R_{i,j}$ is "outside" of S. We want to show that after deleting the edges, the resulting graph is triangle-free and we deleted at most εN^2 edges, giving us a contradiction.

Behauptung 1. G_n is triangle-free.

Suppose G_n has a triangle. Then there exist i,j,k such that $ij, ik, jk \in E(G_n)$ with $\lambda(S \cap R_{i,j}) \geq \frac{3}{4}\lambda(R_{i,j}) = \frac{3}{4N^2}$, $\lambda(S \cap R_{i,k}) \geq \frac{3}{4N^2}$ and $\lambda(S \cap R_{i,k}) \geq \frac{3}{4N^2}$.

But then

$$\begin{aligned}
&\int_{J_i \times J_j \times J_k} 1_S(x_1,x_2)1_S(x_2,x_3)1_S(x_1,x_3)dx_1dx_2dx_3 \\
&= \int_{J_i \times J_j \times J_k} (1 - 1_{R_{i,j}\setminus S}(x_1,x_2))(1 - 1_{R_{i,k}\setminus S}(x_2,x_3)(1 - 1_{R_{j,k}\setminus S}(x_1,x_3))dx_1dx_2dx_3 \\
&\geq \frac{1}{N^3} - \frac{1}{N}\lambda(R_{i,j} \setminus S) - \frac{1}{N}\lambda(R_{i,k} \setminus S) - \frac{1}{N}\lambda(R_{k,j} \setminus S) \geq \frac{1}{N^3} - \frac{3}{4N^3} \\
&= \frac{1}{4N^3} > 0.
\end{aligned}$$

where we used above that for $x,y,z \in [0,1]$

$$(1-x)(1-y)(1-z) = 1 - x(1-y)(1-z) - y(1-z) - z \geq 1 - x - y - z.$$

But then

$$\int_{[0,1]^3} W(x_1,x_2)W(x_2,x_3)W(x_1,x_3)dx_1dx_2dx_3 > 0,$$

which is a contradiction to $0 = t(K_3, W)$.

Behauptung 2. For n large enough, we removed at most εN^2 edges.

Since W_{G_n} converges to W_G in $d_\square$-distance, it follows

$$\int_{[0,1]^2} (W_{G_n}(x,y) - W(x,y))dxdy \leq \sup_{S \times T \subset [0,1]^2} \int_{S \times T} (W_{G_n}(x,y) - W(x,y))dxdy \to 0$$

By Lemma 9.6, if 1_S is the indicator function on S,

$$\int_{[0,1]^2} (1 - 1_S)(W_{G_n}(x,y) - W(x,y))dxdy \to 0$$

and hence

$$\int_{[0,1]^2} (1 - 1_S)W_{G_n}(x,y)dxdy \to \int_{[0,1]^2} (1 - 1_S)W(x,y)dxdy = 0.$$

Choose n large enough such that

$$\int_{[0,1]^2} (1 - 1_S)W_{G_n}(x,y)dxdy < \frac{\varepsilon}{4}. \tag{9.2}$$

Let S be the set of removed edges from G_n. Let $R_{i,j}$ be such that ij got removed, then

$$\int_{R_{ij}} (1 - 1_S)W_{G_n} = \lambda(R_{i,j} \setminus S) > \frac{1}{4}\lambda(R_{i,j}) \text{ and } \int_{[0,1]^2} (1 - 1_S)W_{G_n} > \frac{1}{4}\lambda(R_{i,j}) \cdot |S|$$

Since $\lambda(R_{i,j}) = \frac{1}{N^2}$, it follows from Equation (9.2) that $|S| < \varepsilon N^2$.

Graphentheorie, TU Berlin, WS 2023/24

Dozentin: Stefan Felsner/Alexandra Wesolek
Mitschrift: Martin Schramm (up to some added material)
Datum: 19.11.2013

Vorlesung 10

Extremale Graphentheorie II

10.1 Many edges, many triangles.

Beobachtung.

- Wenn $|E| > \frac{n^2}{4}$, dann existieren Dreiecke (gemeinsame Nachbarn zweier Knoten).
- Wenn $|E| > \frac{n^2}{4} + \delta$, dann existieren mindestens δ Dreiecke.

Beweis. Der erste Teil ist der Satz von Mantel (Prop. 8.1). Der zweite Teil ergibt sich durch iteratives Löschen einer Kante der (wegen Teil 1) an einem Dreieck beteiligt ist.

Satz 10.1 *Wenn $|E| \geq \left(\frac{1}{4} + c\right)n^2$, dann enthält G mindestens $2c\binom{n}{3}$ Dreiecke.*

Beweis. Wir beginnen mit einfachen Beobachtungen zu Dreiecken.

- Eine Kante (u, v) ist in mindestens $d_u + d_v - n$ Dreiecken enthalten.

Dies gilt weil $d_u + d_v \leq n + |N(u) \cap N(v)|$ und jeder Knoten in $N(u) \cap N(v)$ mit der Kante (u, v) ein Dreieck bildet.

Daraus folgt:

$$\#\text{Dreiecke} \geq \frac{1}{3} \sum_{uv \in E} (d_v + d_u - n) = \frac{1}{3}\left((\sum_v \sum_{u \in N(v)} d_v) - n|E| \right) \stackrel{A}{=} \frac{1}{3}\left(\sum_v d_v^2 - n|E| \right).$$

Wir betrachten nun den Durchschnittsgrad d_{av} in G:

$$d_{\mathsf{av}} = \frac{1}{n}\sum_v d_v \stackrel{B}{=} \frac{2|E|}{n} \geq \left(\frac{1}{2} + 2c\right)n \stackrel{C}{\geq} \frac{n}{2} \qquad \text{und} \qquad d_{\mathsf{av}} - \frac{n}{2} \stackrel{D}{\geq} 2cn.$$

Ausserdem wenden wir die Cauchy-Schwarz-Ungleichung auf die Vektoren $\mathbb{1}$ und d an: $(\sum d_v)^2 \leq n \cdot (\sum d_v^2)$. Multiplikation mit n^{-2} liefert:

$$\frac{1}{n}\sum d_v^2 \geq (\frac{1}{n}\sum d_v)^2 \stackrel{E}{=} d_{\mathsf{av}}^2.$$

Mit diesen Hilfsmitteln erhalten wir:

$$\#\text{Dreiecke} \stackrel{A}{\geq} \frac{1}{3}\left(\sum_v d_v^2 - n|E| \right) \stackrel{E,B}{\geq} \frac{1}{3}\left(n \cdot d_{\mathsf{av}}^2 - \frac{n^2}{2} d_{\mathsf{av}} \right) = \frac{n \cdot d_{\mathsf{av}}}{3}\left(d_{\mathsf{av}} - \frac{n}{2} \right)$$

$$\stackrel{D}{\geq} \frac{n \cdot d_{\mathsf{av}}}{3} \cdot 2cn \stackrel{C}{\geq} 2c\frac{n^3}{3 \cdot 2} > 2c\binom{n}{3}.$$

□

10.2 Füredi stability lemma

Theorem 10.2 (Füredi, 2015)
Sei G ein K_{k+1}-freier Graph mit n Ecken. Sei

$$e(G) = e(T_{n,k}) - t,$$

dann gibt es einen k-partiten Graphen $H \subseteq G$ mit

$$e(H) \geq e(G) - t.$$

Beweis. Wir betrachten den folgenden Algorithmus.
Algorithmus:
Initialisierung: $i := 1$, $H_1 := G$, $V_1 := V(G)$.
While-Schleife: Solange V_i nicht leer ist:
Sei x_i ein Knoten maximalem Grades in H_i.
Setze die neue Knotenmenge als $V_{i+1} := N(x_i) \cap H_i$, d.h. entferne $W_i := H_i \setminus N(x_i)$.
$H_{i+1} = G[V_{i+1}]$;
$i = i + 1$;
Output: $W_1, \ldots, W_s$

Behauptung 1. Der Algorithmus stoppt nach $s \leq k$ Schritten.
Wir zeigen dass die Knoten $x_1, \ldots, x_s$ eine Clique in G bilden. Weil G kein K_{k+1} enthält, gilt dann $s \leq k$.
Es gilt $V_s \subset V_{s-1} \subset \cdots \subset V_1$. Da $x_i \in V_i$, gilt auch $x_2, \ldots, x_s \in V_2 = N(x_1)$. Also ist x_1 Nachbar von $x_2, \ldots, x_t$. Es gilt aber auch, dass $x_3, \ldots, x_t \in V_3 \subset N(v_2)$ und so weiter. Daher bilden $x_1, \ldots, x_s$ eine Clique.

Behauptung 2. Sei $V = W_1 \cup W_2 \cup \cdots \cup W_s$. Wir löschen alle Kanten innerhalb von $W_1, W_2, \ldots, W_s$. Dann haben wir maximal t Kanten von G entfernt, d.h. $\sum_{i=1}^{t} |E(G[W_i])| \leq t$.

Da x_i ein Knoten maximalen Grades in H_i ist, gilt $|N(x) \cap V_i| \leq |V_{i+1}|$ für alle $x \in V(H_i)$ (also insbesondere für alle $x \in W_i$). Sei $deg_{V_i}(x) = |N(x) \cap V_i|$. Dann gilt

$$\sum_{x \in W_i} deg_{V_i}(x) \leq \sum_{x \in W_i} |V_{i+1}| \leq |W_i||V_{i+1}|$$

Andererseits gilt auch

$$\sum_{x \in W_i} deg_{V_i}(x) = \sum_{x \in W_i} deg_{W_i}(x) + deg_{V_{i+1}}(x) = 2|E(G[W_i])| + e(V_{i+1}, W_i),$$

wobei wir das Handshake Lemma in W_i benutzen und die Kanten in dem bipartiten Graphen zwischen W_i und V_{i+1} separat zählen.

Es gilt

$$\sum_{i=1}^{t} |E(G[W_i])| + e(V_{i+1}, W_i) = |E(G)| = |E(T_k(n)| - t.$$

und daher

$$|E(T_k(n)| - t + \sum_{i=1}^{t} |E(G[W_i])| = \sum_{x \in W_i} deg_{V_i}(x) \leq \sum_{i=1}^{t} |W_i||V_{i+1}| \tag{10.3}$$

Da $V_{i+1} = W_{i+1} \cup \cdots \cup W_t$ gilt auch

$$\sum_{i=1}^{t} |W_i||V_{i+1}| = |E(K_{|W_1|,\ldots,|W_t|})| \leq |E(T_k(n))|. \tag{10.4}$$

Aus (10.3) and (10.4) folgern wir, dass $\sum_{i=1}^{t} |E(G[W_i])| \leq t$. Nachdem wir diese Kanten entfernt haben, ist der Graph s-partit, wobei $s \leq k$, also auch k-partit.

10.3 Erdős-Stone-Simonovitz

Satz 10.3 (Erdős-Stone 1948)
Für den Turán-Graphen $T_{k+1}(r)$ gilt:

$$ex(n, T_{k+1}(r)) = (\tfrac{k-1}{k} + o(1))\tbinom{n}{2}$$

Die *Färbungszahl* von G ist die Größe der kleinsten Partition von V in unabhängige Mengen:

$$\chi(G) = \min\{t : \exists \text{Partition } I_1, \ldots, I_t \text{ von } V \text{ so dass alle } I_i \text{ unabhängig in } G \text{ sind}\}$$

Satz 10.4 (Erdős-Simonovits 1966)
Für jeden Graphen H mit $\chi(H) = k+1$ gilt:

$$ex(n, H) = (\tfrac{k-1}{k} + o(1))\tbinom{n}{2}$$

Wir folgern ein etwas schwächeres Ergebnis direkt aus Erdős-Stone. Unser Argument zeigt, dass die Färbungszahl von H die Grössenordnung von $\mathrm{ex}(n, H)$ bestimmt.

Die Färbungszahl von H erzwingt, dass H nicht Teilgraph eines Turán-Graphen $T_k(r)$ sein kann, $\forall r$. Andererseits gibt es ein r, so dass H Teilgraph von $T_{k+1}(r)$ ist. Also gilt für großes n:

$$\mathrm{ex}(n, H) \leq \mathrm{ex}(n, T_{k+1}(r))$$

□

Beispiel. Für bipartite H ist die genaue Bestimmung von $\mathrm{ex}(n, H)$ schwierig.

- $\mathrm{ex}(n, K_{s,t}) \leq c_s n^{2-(1/t)}, \quad s \leq t \leq 1$
- $c_1 n^{5/3} \leq \mathrm{ex}(n, K_{4,4}) \leq c_2 n^{7/4}$

10.4 Eine Schranke für die Unabhängigkeitszahl α

Zur Erinnerung: die *Unabhängigkeitszahl* von G ist die Größe der größten unabhängigen Menge:

$$\alpha(G) = \max\{|I| : I \subset V \text{ unabhängig in } G\}$$

Wenn wir eine maximale unabhängige Menge I betrachten, dann gilt $V = \bigcup_{v\in I} N[v]$ und also $n \le \sum_{v\in I} d_v + 1 \le |I|(\Delta+1)$. Daraus folgt eine einfache untere Schranke an α:

$$\alpha \ge \frac{n}{\Delta+1}.$$

Mit einer kleinen wahrscheinlichkeitstheoretischen Überlegung[I] können wir eine Schranke herleiten die häufig besser ist: Erzeuge eine zufällige Teilmenge S von V indem jeder Knoten v mit Wahrscheinlichkeit p für S ausgewählt wird und zwar unabhängig von anderen Knoten. Die erwartete Größe von S ist $\mathbf{E}(|S|) = p \cdot n$ und für die erwartete Kantenzahl von $G_{[S]}$ gilt

$$\mathbf{E}(|E_{[S]}|) = \sum_{e\in E} \mathbf{Pr}(e \in E_{[S]}) \overset{*}{=} \sum_{(u,v)\in E} \mathbf{Pr}(u\in S)\mathbf{Pr}(v \in S) = p^2 \cdot |E|.$$

Die Gleichung $(*)$ folgt hier aus der Unabhängigkeit. Wenn wir von jeder Kante in $G_{[S]}$ einen der Endknoten löschen bekommen wir eine unabhängige Menge S' mit $|S'| \ge |S| - |E_{[S]}|$. Für den Erwartungswert dieser unabhängigen Menge gilt $\mathbf{E}(|S'|) \ge pn - p^2|E|$. Wegen $2|E| = nd_{\mathsf{av}}$ bekommen wir mit der Wahl von $p = \frac{1}{d_{\mathsf{av}}}$ die Ungleichung $\mathbf{E}(|S'|) \ge \frac{n}{2d_{\mathsf{av}}}$. Aus der unteren Schranke an den Erwartungswert, für die Größe der unabhängigen Menge S', folgt die Existenz einer unabhängigen Menge die mindestens diese Größe hat und also

$$\alpha(G) \ge \frac{n}{2d_{\mathsf{av}}}$$

Achtung: Wahrscheinlichkeiten liegen im Intervall $[0,1]$. Also gilt dieser Beweis nur für $p = \frac{1}{d_{\mathsf{av}}} \le 1$, also $d_{\mathsf{av}} \ge 1$. Dass die Ungleichung sonst nicht immer gilt, sieht man z.B. an $G = ([5], \{\{1,2\},\{3,4\}\})$.

Eine noch etwas stärkere Schranke bekommt man mit dem Satz von Turán. Aus dem Satz folgt sofort $d_{\mathsf{av}} \le (1-\frac{1}{\omega})n$ und also $\omega \ge \frac{n}{n-d_{\mathsf{av}}}$. Diese untere Schranke an die Cliqenzahl übersetzt sich wegen $\alpha(G) = \omega(\overline{G})$ und $d_{\mathsf{av}}(G) = n-1-d_{\mathsf{av}}(\overline{G})$ in die folgende Schranke für die Unabhängigkeitszahl:

$$\alpha(G) \ge \frac{n}{d_{\mathsf{av}}+1}$$

Nun wollen wir aber endlich das Hauptresultat dieses Abschnitts formulieren.

[I] W-Theorie in der Nußschale gibt es im nächsten Kapitel.

Satz 10.5 (Wei 1981)
Für jeden Graphen G gilt:

$$\alpha(G) \geq \sum_v \frac{1}{d_v + 1}.$$

Beweis. Wir betrachten den Algorithmus `MIN`. In diesem und ebenso im Algorithmus `MAX` ist der Grad eines Knotens immer der Grad im aktuell gültigen Graphen.

```
MIN(G)
  while V_G ≠ ∅ do
    x ← Knoten minimalen Grades
    I ← I + x
    G ← G − N[x]
  return I
```

Die Menge I ist offenbar unabhängig. Der Algorithmus gehört zur Familie der Greedy-Algorithmen (greedy=gierig) denn der Knoten x wird jeweils *kurzsichtig* als der Knoten gewählt der für die folgende Iteration möglicht viele Knoten übrig lässt. Diese Vorgehensweise ist einfach, muss aber nicht das optimale Resultat liefern.

Für einen Graphen G definieren wir $A(G) = \sum_v \frac{1}{d_v+1}$. Nun betrachten wir die Differenz $A(G) - A(G - N[x])$:

$$A(G) - A(G - N[x]) \leq \sum_{v \in N[x]} \frac{1}{d_v + 1} \leq (d_x + 1)\frac{1}{d_x + 1} = 1.$$

Die erste Ungleichung gilt, da für jeden Knoten $u \notin N[x]$ der Grad d'_u in $G - N[x]$ kleiner oder gleich dem Grad d_u in G ist und also $\frac{1}{d_u+1} - \frac{1}{d'_u+1} \leq 0$. Für die zweite Ungleichung verwenden wir $d_v \geq d_x$.

Der A-Wert des Graphen reduziert sich in jedem Schleifendurchlauf um höchstens 1. Wenn der Algorithmus endet ist der A-Wert $A(\emptyset) = 0$, also gibt es mindestens $A(G)$ Schleifendurchläufe. Da der unabhängigen Menge in jedem Schleifendurchlauf ein neues Element hinzugefügt wird gilt also $|I| \geq A(G)$. □

Wir geben nun einen zweiten Beweis des Satzes, dafür betrachten wir den folgenden Algorithmus `MAX`:

```
MAX(G)
  while E_G ≠ ∅ do
    x ← Knoten maximalen Grades
    G ← G − x
  return V_G
```

Wiederum ist offensichtlich, dass der Algorithmus eine unabhängige Menge erzeugt. Zur Analyse betrachten wir erneut die Veränderung des A-Wertes in einem

Schleifendurchlauf:

$$\begin{aligned} A(G-x)-A(G) &= \sum_{v\in N(x)} \frac{1}{d_v} - \sum_{v\in N[x]} \frac{1}{d_v+1} \\ &= \sum_{v\in N(x)} \frac{1}{d_v(d_v+1)} - \frac{1}{d_x+1} \geq \sum_{v\in N(x)} \frac{1}{d_x(d_x+1)} - \frac{1}{d_x+1} \\ &= \frac{1}{d_x+1} - \frac{1}{d_x+1} = 0 \end{aligned}$$

Also kann der A-Wert insgesamt nur wachsen. Für eine unabhängige Menge I gilt $A(I) = \sum_{v\in I} 1 = |I|$. Daher gibt der Algorithmus eine unabhängige Menge zurück, deren Größe zumindest $A(G)$ ist. $\square$

Graphentheorie, TU Berlin, WS 2023/24

Dozentin: Stefan Felsner/Alexandra Wesolek
Mitschrift: Alexandra Wesolek
Datum: 21.01.2024

Vorlesung

26

Fast algorithms

26.1 Maximum Independent Set and Maximum Matching for trees

The maximum weight independent set problem as well as the maximum weight matching problem are NP-Hard for general graphs. However, they can be solved in polynomial time for trees using so-called dynamic programming.

Satz 26.1 *The Maximum Weight Independent Set problem can be solved on trees in linear time (with respect to the number of vertices of the tree).*

Beweis. Let T be a tree. Pick an arbitrary vertex v_0 and consider T as a rooted tree with root v_0. We denote by C_v the children of a vertex $v \in T$. The subtree T_v rooted at v consists of all vertices reachable from v through the children relation in T. Let $W[v]$ denote the maximum weight of an independent set of T_v. We want to compute $W[v_0]$. Let $W_+[v]$ denote the maximum weight of an independent set of T_v that contains v, and let $W_-[v]$ denote the maximum weight of an independent set of T_v that does not contain v. Then $W[v] = \max\{W_+[v], W_-[v]\}$. Further, we get the following other relations:

$$W_+[v] = w(v) + \sum_{u \in C_v} W_-[u],$$

and

$$W_-[v] = \sum_{u \in C_v} W[u].$$

Hence we get the following algorithm:

Algorithm 1: Maximum Weight Independent Set Algorithm for trees.

Input: A tree T rooted at v_0.
Result: The size of a maximum weight independent set I.

Compute the set of leafs L of T and for each leaf v set $W_-[v] = 0$ and $W_+[v] = w(v)$.
$S := V(T) \setminus L$
$X := N(L) = \cup_{v \in L} N(v)$.
while $X \neq \emptyset$ **do**
  Choose a vertex $v \in X$.
  if *v is a leaf in $T[S]$* **then**

$$W_+[v] = w(v) + \sum_{u \in C_v} W[u],$$

  and

$$W_-[v] = \sum_{u \in C_v} \max\{W_+[u], W_-[u]\}.$$

  Remove v from S and from X.
  Add $N(w) \cap S$ to X.
  end
  else
  v is not a leaf in $T[S]$
  end
  Remove v from X.
end
Return $\max\{W_+[v_0], W_-[v_0]\}$

For the Initialization, it takes $O(n)$ time to go through all the vertices and compute the set of leaves L. Each while loop removes a vertex from X. If a vertex u is added to X, then because it is a neighbour of v at the step of the while loop. The addition of a vertex to X corresponds therefore to an edge. Hence, we run the while loop at most $O(|E(G)|) = O(n)$ times.
Updated proof:
For each $v \in X$ considered in the while loop that is a leaf in $T[S]$, we perform at most $2\mathbf{deg}(v) - 1$ additions and have to calculate at most $\mathbf{deg}(v)$ maxima. Summing over all $v \in V(G)$, we have at most $\sum_{v \in V(G)} (3\mathbf{deg}(v) + 1) \leq 6|E(G)| + |V(G)|$ operations. □

Thus, although the maximum weight independent set problem is NP-Hard for general graphs, we have just shown that it can be solved in linear time for trees.

Satz 26.2 *The Maximum Weight Matching problem can be solved in linear time on trees.*

Let T be a tree. Pick an arbitrary vertex v_0 and consider T as a rooted tree with root v_0.

For each vertex v, let

$f^+(v)$ = the size of the maximum weight matching in the subtree T_v rooted at v in which v is matched;

$f^-(v)$ = the size of the maximum weight matching in the subtree T_v rooted at v in which v is not matched.

As before we call C_v the set of children of a vertex v. The update equations are as follows

$$f^+(v) = w(v) + \max_{u \in C_v} \left(w(u) + f^-(u) + \sum_{x \in C_v \setminus \{u\}} \max\{f^+(x), f^-(x)\} \right),$$

and

$$f^-(v) = \sum_{u \in C_w} \max\{f^+(w), f^-(w)\}.$$

In the former case, we need to match v to some child w of v. So then we can't match w with any of its children, so we add in $f^-(w)$, the maximum matching we can get in that subtree without matching w to any of its children. For all the other children u of v, we don't care whether we take $f^+(u)$ or $f^-(u)$, so we just take the maximum. Now we want to find the child w that gives us the maximum possible value for the respective sum. The latter case is clear as we don't care whether any child of v is matched or not. With this dynamic programming approach, we can solve the problem in $O(n)$.

26.2 The pathwidth of a graph

An interval graph of width k is an intersection graph of intervals such that every $x \in \mathbb{R}$ is in at most $k+1$ intervals.

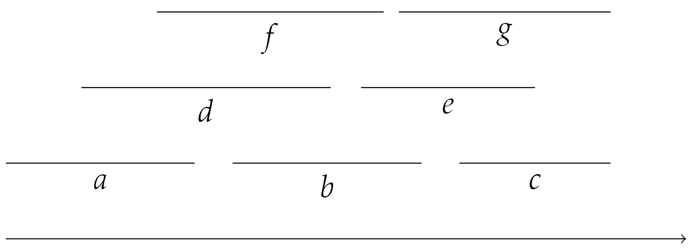


Abbildung 1: An interval graph of width 2.

A graph G *has pathwidth* k if k is the smallest integer such that G is a subgraph of an interval graph of width k. The graph below (Figure 2) is a subgraph of the interval graph in Figure 1.

Beobachtung. The pathwidth of the path is 1 and the pathwidth of a cycle is 2.

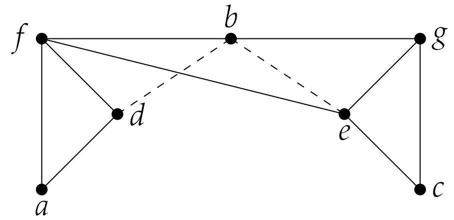


Abbildung 2: A graph of pathwidth 2. In order to see that the pathwidth is at most 2 note that this graph is a subgraph of the interval graph of the intervals in Abbildung 1.

Satz 26.3 *The pathwidth of the complete graph K_n is $n-1$.*

Beweis. Let the intervals be $[a_i, b_i]$ and suppose they pairwise intersect. Take the interval $[a_i, b_i]$ that maximizes a_i. We claim that the point a_i is in every interval. Suppose not: it's not in interval $[a_j, b_j]$. Then, either $a_i > b_j$ or $a_i < a_j$. The latter case cannot happen by the definition of a_i. So $a_i > b_j$. But then $[a_i, b_i]$ and $[a_j, b_j]$ don't intersect, a contradiction.

□

Note that alternatively this follows from Helly's Theorem in one dimension. Helly's theorem states that, given n convex subsets $X_1, \ldots, X_n$ in $\mathbb{R}^d$ for $n > d$ such that the intersection of every $d+1$ of these sets is nonempty. Then the whole collection has a nonempty intersection. (Here we can use $d = 1$ and project all the intervals on the x-axis.)

A *path decomposition* of G is a sequence $(X_1, \ldots, X_N)$ of bags $X_i \subseteq V(G)$ such that

1. Every vertex belongs to at least one bag.
2. For each edge in G, there is a bag containing its endpoints.
3. For every vertex v, the set of bags containing v forms a connected interval of $(X_1, \ldots, X_N)$.

The *width* of the path decomposition is the size of the largest bag X_i. A graph has pathwidth at most k if and only if there exists a path decomposition of width at most $k+1$.

Lemma 26.4 *If $X_i \neq X_{i+1}$ for all $i \leq N$ then $N \leq 2n+1$.*

Beweis. From X_i to X_{i+1} at least one vertex gets added or deleted, i.e. $|X_i \setminus X_{i+1}| \geq 1$ or $|X_{i+1} \setminus X_i| \geq 1$. Each vertex gets added or deleted at most once. □

We define $X_{\leq i}$ to be $X_1 \cup X_2 \cup \cdots \cup X_i$ and $X_{\geq i}$ as $X_i \cup X_{i+1} \cup \cdots \cup X_N$.

Satz 26.5 (Separator Theorem for Path Decompositions)
For each bag X_i, there is no edge between $X_{\leq i-1} \setminus X_i$ and $X_{\geq i+1} \setminus X_i$. We say X_i separates the sets $X_{\leq i-1} \setminus X_i$ and $X_{\geq i+1} \setminus X_i$.

Beweis. Suppose not: there exists an edge uv with $u \in X_{\leq i-1} \setminus X_i$ and $v \in X_{\geq i+1} \setminus X_i$. By property 3. of path decompositions, we know the bags that contain u form some interval, say, $X_{a_1}, X_{a_1+1}, \dots, X_{b_1}$. Also, the bags that contain v form some interval, say, $X_{a_2}, X_{a_2+1}, \dots, X_{b_2}$. By property 2. of path decompositions,there must be a bag containing both u and v. This means that the two intervals must share a bag. The union of the two intervals of bags is also an interval $X_{\min\{s_1,s_2\}}, \dots, X_{\max\{t_1,t_2\}}$. Now, we know that X_i does not contain u or v, so either X_i lies to the left of the joint interval, or X_i lies to the right. But in the former case, u cannot be inside $X_{\leq i-1} \setminus X_i$, because all bags that contain u lie to the right of X_i. Similarly, in the latter case, v cannot be inside $X_{\geq i+1} \setminus X_i$. This is a contradiction.

□

Satz 26.6 *The Maximum Independent Set problem can be solved in time $4^k O(n)$ given a path decomposition of width k of a graph with $X_i \neq X_{i+1}$ for all $i \leq N$.*

Beweis. Let G be a graph and $(X_1, \dots, X_N)$ be a path decomposition of G. We define $I(S, X_{\leq i})$ to be the size of a largest independent set I in $X_1 \cup X_2 \cup \dots \cup X_i$ with $I \cap X_i = S$. Note that if we compute $I(S, X_{\leq N})$ for each independent set $S \subset X_N$, then

$$\alpha(G) = \max_{S \subset X_N, S \text{ independent}} I(S, X_{\leq N}).$$

We say $S' \subset X_{i-1}$ is compatible with $S \subset X_i$ if S' and S agree on $(X_{i-1} \cap X_i)$, that is, $S \cap X_{i-1} \cap X_i = S' \cap X_{i-1} \cap X_i$.
We will proof that the following update rule is valid for an independent set $S \in X_i$:

$$I(S, X_{\leq i}) = |S \setminus X_{i-1}| + \max_{S' \subset X_{i-1}, S' \text{ compatible with S}} I(S', X_{\leq i-1}) \tag{26.5}$$

Every independent set I on the vertices of $X_{\leq i}$ such that $X_i \cap I = S$ induces an independent set $I' = I \cap X_{\leq i-1}$ on $X_{\leq i-1}$. From the separator theorem for path decompositions it follows that $|I| = |S \setminus X_{i-1}| + |I'|$ and in addition $S' = I \cap X_{i-1}$ is compatible with S. Therefore,

$$|I| \leq |S \setminus X_{i-1}| + \max_{S' \subset X_{i-1}, S' \text{ compatible with S}} I(S', X_{\leq i-1}) \tag{26.6}$$

and since the inequality (26.6) holds for all independent sets containing S it holds in particular for a maximum independent set containing S. Hence $|I|$ can be exchanged with $I(S, X_{\leq i})$ in inequality (26.6).

We want to show equality in (26.5). Suppose J' is an independent set in $X_{\leq i-1}$ such that $S' = X_{i-1} \cap J'$ is compatible with S. Let J be of maximum size with respect to this property. Then $J = S \setminus X_{i-1} \cup J'$ is a set in $X_{\leq i-1}$ with

$$|J| = |S \setminus X_{i-1}| + \max_{S' \subset X_{i-1}, S' \text{ compatible with S}} I(S', X_{\leq i-1}).$$

It is left to show that

(a) $X_i \cap J = S$,

(b) and J is an independent set.

For (a) note that $S = (S \setminus X_{i-1}) \cup (S \cap X_{\leq i-1})$ and $S \setminus X_{i-1} = J \setminus X_{i-1}$. But since $S \subset X_i$ it holds that $S \cap X_{\leq i-1} = S \cap X_{i-1} = S \cap X_{i-1} \cap X_i$ which is equal to $S' \cap X_{i-1} \cap X_i$ since S and S' are compatible. But since $S' \cap X_{i-1} \cap X_i = J \cap X_{i-1} \cap X_i$ it follows that $S = (J \setminus X_{i-1}) \cup J \cap X_{i-1} \cap X_i = J \cap X_i$.
Runtime analysis: There are at most 2^k independent sets $S \in X_i$ since each bag X_i has size k. There are at most 2^k independent sets S' in X_{i-1}. Hence the update in Equation (26.5) from bag X_{i-1} to bag X_i needs to consider at most $2^k \cdot 2^k = 4^k$ pairs (S', S). In fact, we can do a pre-processing of time $2 \cdot 2^k \cdot k \leq 4^k$ and classifying S, S' depending on their intersection with $X_i \cap X_{i-1}$ so we only need to check pairs (S, S') which are compatible.