Graphen und Algorithmen II

Sommersemester 2008

Mitschrift von
Yves Radunz

Inhaltsverzeichnis

Vorlesung am 17.04.2008

Kapitel 0

Einführung

In Graphen und Algorithmen II werden wir uns mit den folgenden Themen beschäftigen:

1. zufällige Graphen

   (a) Eigenschaften

   (b) Algorithmen, Average-Case-Analysis

   Wir betrachten die zufälligen Graphen $G(n,p) = (\mathcal{G}_n, 2^{\mathcal{G}_n}, \mathbf{P})$, wobei $\mathcal{G}$ die Menge aller markierten Graphen mit Knotenmenge $[n]$, $2^{\mathcal{G}_n}$ die Potenzmenge von $\mathcal{G}_n$ und $\mathbf{P}$ eine Verteilung $\mathbf{P} : \mathcal{G}_n \to [0,1]$ ist. (Im Allgemeinen ist $\mathbf{P}$ eigentlich eine Funktion $\mathbf{P} : 2^{\mathcal{G}_n} \to [0,1]$, in unserem Fall (endliches $\mathcal{G}_n$, wegen $n < \infty$) können wir uns jedoch auf eine Abbildung $\mathbf{P} : \mathcal{G}_n \to [0,1]$ beziehen.) Für $\mathbf{P}$ gilt $\sum_{G \in \mathcal{G}_n} \mathbf{P}(G(n,p) = G) = 1$.

   In diesem Modell erzeugen wir einen zufälligen Graphen, indem wir für jedes Knotenpaar mit Wahrscheinlichkeit p eine Kante zwischen diesen Knoten hinzufügen.

   Dann gilt $\mathbf{P}(G(n,p) = G) = p^{e(G)}(1-p)^{\binom{n}{2}-e(G)}$, $e(G) = |E(G)|$.

   Frage: Sei $\mathcal{H}_n \subseteq \mathcal{G}_n$ die Menge aller hamiltonischen Graphen. Wie hoch ist $\mathbf{P}(G(n,p) \in \mathcal{H}_n)$?

   Wir bestimmen diese Wahrscheinlichkeit in Abhängigkeit von p.

   Wenn die Eigenschaft $\mathcal{P}_n \subseteq \mathcal{G}_n$ monoton ist, d.h. sie wird durch Hinzunahme von Kanten nicht zerstört, dann gibt es einen Schwellenwert, d.h. es existiert $\hat{p} = \hat{p}(n)$ sodass

   $\lim_{n\to\infty} \mathbf{P}(G(n,p) \in \mathcal{P}_n) = \begin{cases} 0 & p \ll \hat{p} \\ 1 & p \gg \hat{p} \end{cases}$

   Weitere Frage: Wie groß ist der Bereich um den Schwellenwert, in welchem das Zutreffen der Eigenschaft und das Nicht-Zutreffen ungefähr gleich wahrscheinlich ist?

   Erste Anwendungen:

   - $\mathbf{P}(G(n,p) \in \mathcal{P}_{k,l}) > 0$, $\mathcal{P}_{k,l} = \{G \in \mathcal{G}_n | \chi(G) \geq k, girth(G) \geq l\}$
     Wir haben damit in G&A I gezeigt, dass Graphen mit $\chi(G) \geq k, girth(G) \geq l$ für alle k, l existieren.
   - $\mathbf{P}(G(n,p) \in R_k) > 0$, $R_k = \{G \in \mathcal{G}_n | \alpha(G) \leq k, \omega(G) \leq k\}$
     Wir hatten dies in Graphen und Algorithmen I für $k = 2\log n$ und $p = \frac{1}{2}$ betrachtet.

2. extremale Graphentheorie

   (a) Turín, Erdős, Erdős-Simonovits, Ramseytheorie

   (b) Szemerédis Regularitätslemma

   (c) Testen von Grapheneigenschaften

Betrachten wir einmal $G(n,p)$ und eine Teilmenge $U \subseteq [n]$. Dann ist der Erwartungswert der Anzahl der Kanten X in $G(n,p)[U]$ gerade $P\binom{|U|}{2}$. Laut Chernoff haben wir $X \sim Bi(m,p)$.

$\mathbf{P}(|X - pm| \geq \varepsilon pm) \leq 2e^{-\varepsilon'(\varepsilon)pm}$

$\rightsquigarrow$ Regularitätlemma

$\forall \varepsilon > 0 \forall t \exists T_0 = T_0(\varepsilon, t_0) \approx t_0 \cdot \left. 2^{2^{2^{\cdots^{2}}}} \right\} \frac{1}{\varepsilon^5} \forall G = (V,E) \exists V = V_0 \dot\cup V_1 \dot\cup ... \dot\cup V_t :$

- $t_0 \leq t \leq T_0$
- $|V_0| \leq \varepsilon n, n = |V|$
- $|V_1| = ... = |V_t|$
- $|\{\{i,j\} \in \binom{[t]}{2} | (V_i, V_j)$ ist nicht ε-regulär$\}| \leq \varepsilon t^2$

(V_i, V_j) ist (ε, d)-reguär, falls für alle $U_i \subseteq V_i : |U_i| \geq \varepsilon |V_i|$ und für alle $U_j \subseteq V_j : |U_j| \geq \varepsilon |V_j|$ die Ungleichung $|d_G(U_i, U_j) - d| \leq \varepsilon$ gilt.

(V_i, V_j) ist ε-regulär, falls es (ε, d)-regulär für ein $d \in [0,1]$ ist.

Vorlesung am 19.04.2008

Kapitel 1

Zufällige Graphen

1.1 Allgemeines

Definition (Diskreter Wahrscheinlichkeitsraum)
Ein *diskreter Wahrscheinlichkeitsraum* ist in unserem Fall ein Paar (Ω, p), wobei Ω eine endliche Menge und $p : \Omega \to [0,1]$ eine Abbildung mit $\sum_{\omega\in\Omega} p(\omega) = 1$ ist. Eine Teilmenge $A \subseteq \Omega$ bezeichnen wir als ein Ereignis in Ω. Es gilt $p(A) = \sum_{\omega\in A} p(\omega)$.

Beispiel (6-seitiger Würfel)
Wir haben $\Omega = \{1,2,3,4,5,6\}$ und $p(\omega) = \frac{1}{6}$ für alle $\omega \in \Omega$. Das Ereignis „Wir würfeln eine gerade Zahl" ist $A = \{2,4,6\}$ und hat die Wahrscheinlichkeit $p(A) = \frac{1}{2}$.

Definition ($\mathcal{G}_n$)
Wir setzen $\mathcal{G}_n = \{G | G \text{ ist ein Graph mit } V(G) = [n]\}$ und erhalten $p(G) = \frac{1}{|\mathcal{G}_n|} = \frac{1}{2^{\binom{n}{2}}}$.

Definition (fast sicher)
Wir betrachten Wahrscheinlichkeitsräume $(\Omega_i)_{i\in\mathbb{N}}$ und Ereignisse $A_i \subseteq \Omega_i$.
Falls $\lim_{i\to\infty} p(A_i) = 1$, dann hat (Ω_i) *fast sicher* (die Eigenschaft) (A_i).

Bemerkung (fast sicher $\leftrightarrow$ asymptotisch fast sicher)
In der Stochastik bezeichnet man mit „fast sicher" Ereignisse, deren Wahrscheinlichkeit 1 ist. Die Eigenschaft, dass die Wahrscheinlichkeit gegen 1 konvergiert, wird als „asymptotisch fast sicher" bezeichnet. Um uns die Schreibarbeit zu sparen, werden wir jedoch das „asymptotisch" weglassen und nur „fast sicher" schreiben.

Beispiel
Sei $\mathcal{G} = \bigcup_{i=0}^{\infty} \mathcal{G}_i$ die Menge aller endlichen Graphen.
Es sei $A \subseteq \mathcal{G}$ eine bestimmte Teilmenge, die zum Beispiel durch eine Eigenschaft (bipartit, zusammenhängend, gerade Anzahl an Knoten etc.) gegeben ist. Weiterhin seien $A_i = A \cap \mathcal{G}_i$ und $p_i = p(A_i)$.
$\mathcal{G}_i$ hat die Eigenschaft A_i fast sicher, falls $p_i \xrightarrow{i\to\infty} 1$ gilt.

Interpretation von „fast sicher"
Falls $\mathcal{G}_i$ eine Eigenschaft A_i fast sicher hat, dann bedeutet dies, dass für genügend große i der zufällige Graph $\mathcal{G}_i$ die Eigenschaft A_i mit einer Wahrscheinlichkeit beliebig dicht an 1 hat:
$\forall\varepsilon\exists n_0\forall\mathcal{G}_n, n \geq n_0 : \mathbf{P}(A_n) \geq 1 - \varepsilon$

Proposition 1.1. (Seltenheit isolierter Knoten) *Fast kein Graph hat einen isolierten Knoten.*

Beweis:
Wir haben $|\{G|G \in \mathcal{G}_n, \delta(G) = 0\}| \leq n|\mathcal{G}_{n-1}|$.
$\Rightarrow p(\{G \in \mathcal{G}_n|\delta(G) = 0\}) \leq \frac{n \cdot 2^{\binom{(n-1)}{2}}}{2^{\binom{n}{2}}} = n \cdot 2^{\frac{(n-1)(n-2)}{2} - \frac{n^2-n}{2}} = 2^{\log_2 n + 1 - n} \overset{n\to\infty}{\longrightarrow} 0$ □

Proposition 1.2. (Zusammenhang) *Fast alle Graphen sind zusammenhängend.*

Beweis:
Wenn ein Graph G nicht zusammenhängend ist, existiert eine Komponente mit höchstens $\lfloor \frac{n}{2} \rfloor$ Knoten.
$\Rightarrow$ Die Anzahl der nicht zusammenhängenden Graphen ist höchstens $\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{i} 2^{\binom{i}{2}} \cdot 2^{\binom{n-i}{2}}$.
Wir haben $2^{\binom{i}{2}} \cdot 2^{\binom{n-i}{2}} = 2^{\binom{i}{2}+\binom{n-i}{2}} = 2^{\binom{n}{2}+i(i-n)}$.
$\Rightarrow p(\{G \text{ ist nicht zusammenhängend}\}) \leq \frac{1}{2^{\binom{n}{2}}} \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{i} 2^{\binom{n}{2}+i(i-n)} = \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{i} 2^{i(i-n)}$
(Bemerkung: Für $a > b$ gilt $(\frac{a}{b})^b \leq \binom{a}{b} \leq (\frac{ea}{b})^b$.
Wir verwenden jedoch zunächst $\binom{n}{i} \leq n^i$ und $2^{i(i-n)} = (2^{-(n-i)})^i$.)

$$\begin{aligned} \Rightarrow p(\{G \text{ ist nicht zusammenhängend}\}) &\leq \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{i} 2^{i(i-n)} \leq \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{n}{2^{n-i}}\right)^i \\ &\leq \sum_{i=1}^{\frac{n}{2}} \left(\frac{n}{2^{\frac{n}{2}}}\right)^i \leq \sum_{i=1}^{\frac{n}{2}} \frac{n}{2^{\frac{n}{2}}} \end{aligned}$$

(Die dritte Ungleichung gilt wegen $\forall i \leq \frac{n}{2} : 2^{n-i} \geq 2^{\frac{n}{2}}$ und die letzte Ungleichung folgt aus der Ungleichung $\forall n \neq 3 : n \leq 2^{\frac{n}{2}}$, d.h. $\frac{n}{2^{\frac{n}{2}}} \leq 1$.)
$\Rightarrow p(\{G \text{ ist nicht zusammenhängend}\}) \leq \sum_{i=1}^{\frac{n}{2}} \frac{n}{2^{\frac{n}{2}}} < \sum_{i=1}^{n} \frac{n}{2^{\frac{n}{2}}} = \frac{n^2}{2^{\frac{n}{2}}} \overset{n\to\infty}{\longrightarrow} 0$ □

1.2 $\mathcal{G}(n, \frac{1}{2})$-Modell, Erwartungswert, Markoff

Definition ($\mathcal{G}(n,p)$-Modell)
Mit $\mathcal{G}(n,p)$ bezeichnen wir den Wahrscheinlichkeitsraum mit $\Omega = \mathcal{G}_n$ und der Wahrscheinlichkeitsverteilung $\mathbf{P}(\mathcal{G}(n,p) = G) = p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|}$.

Bemerkung ($\mathcal{G}(n, \frac{1}{2})$)
Wir werden zunächst nur den Fall $p = \frac{1}{2}$ betrachten.

Beispiel (Fragen)
Mit dem Durchmesser $\operatorname{diam}(G)$ eines Graphen G bezeichnen wir die maximale Distanz zwischen zwei Knoten des Graphen, d.h. $\operatorname{diam}(G) = \max_{u,v \in V} \operatorname{dist}_G(u,v)$.
Was ist nun $\operatorname{diam}(\mathcal{G}_n)$?
Besser beantworten lässt sich die Frage nach $p_{n,k} = \mathbf{P}(\{G \in \mathcal{G}_n | \operatorname{diam}(G) = k\})$.
Bzw.: Was ist $\lim_{n\to\infty} p_{n,k}$?

Beobachtung
Für alle $x < y \in [n]$ gilt $p(\{G \in \mathcal{G}_n | \{x,y\} \in E(G)\}) = \frac{1}{2}$.

Definition (Erwartungswert, Zufallsvariable)
Sei (Ω, p) ein diskreter Wahrscheinlichkeitsraum. $X : \Omega \to \mathbb{R}$ heißt *Zufallsvariable*.
Mit $\mathbf{E}[X] = \sum_{\omega \in \Omega} X(\omega)p(\omega)$ bezeichnen wir den *Erwartungswert*.
Spezialfall: $X : \Omega \to 2 = \{0,1\}$
$\Rightarrow \mathbf{E}[X] = \sum_{\omega \in \Omega \cap X^{-1}(1)} p(\omega) = \mathbf{P}(X = 1)$

Proposition 1.3. (Linearität des Erwartungswerts) *Seien X, Y Zufallsvariablen, $\lambda \in \mathbb{R}$ und $Z = \lambda X + Y$.*
Dann gilt $\mathbf{E}[Z] = \lambda\mathbf{E}[X] + \mathbf{E}[Y]$.

Beweis:
$\lambda\mathbf{E}[X]+\mathbf{E}[Y] = \lambda\sum_{\omega\in\Omega} X(\omega)p(\omega)+\sum_{\omega\in\Omega} Y(\omega)p(\omega) = \sum_{\omega\in\Omega} \lambda X(\omega)p(\omega)+Y(\omega)p(\omega) = \mathbf{E}[Z]$ $\square$

Satz 1.1. (Markoff / Markoff'sche Ungleichung) *Sei $X : \Omega \to \mathbb{R}_{\geq 0}$ eine (nicht negative!) Zufallsvariable.*
Dann gilt $\forall t \in \mathbb{R}_{>0} : \mathbf{P}(X \geq t) \leq \frac{\mathbf{E}[X]}{t}$.

Beweis:
$\mathbf{E}[X] = \sum_{\omega\in\Omega} X(\omega)p(\omega) \geq t\sum_{\omega\in\Omega, X(\omega)\geq t} p(\omega) = t\mathbf{P}(X \geq t)$ $\square$

Vorlesung am 23.04.2008

Bemerkung
Eine äquivalente Ungleichung ist $P(X \geq \alpha\mathbf{E}[X]) \leq \frac{1}{\alpha}$ mit $t = \alpha\mathbf{E}[X]$.

1.3 1. Moment-Methode

Korollar 1.1. *Seien $(X_i)_{i\in\mathbb{N}}$ Zufallsvariablen mit $X_i : \Omega_i \to \mathbb{N}$ und $\lim_{n\to\infty} \mathbf{E}[X] \to 0$.*
Dann gilt $\lim_{n\to\infty} \mathbf{P}(X_n > 0) = 0$.

Beweis:
$\lim_{n\to\infty} \mathbf{P}(X_n > 0) = \lim_{n\to\infty} \mathbf{P}(X_n \geq 1) \leq \lim_{n\to\infty} \frac{\mathbf{E}[X_n]}{1} = 0$ $\square$

Satz 1.2. *Es gilt $\mathbf{P}(\mathrm{diam}(\mathcal{G}(n, \frac{1}{2})) = 2) = 1 - o(1)$.*

Beweis:
$\mathbf{P}(\mathrm{diam}(\mathcal{G}(n, \frac{1}{2})) = 1) = \mathbf{P}(\mathcal{G}(n, \frac{1}{2}) = K_n) \leq (\frac{1}{2})^{\binom{n}{2}} = o(1)$
$\mathrm{diam}(G) > 2 \Leftrightarrow \exists u, v \in V(G) : u \neq v, (\{u\} \cup N(u)) \cap N(v) = \emptyset$
Wir definieren die Zufallsvariablen $Y = Y_n : \mathcal{G}_n \to \mathbb{N}, Y_n(G) = |\{\{u, v\} \in \binom{[n]}{2} | N_G(u) \cap N_G(v) = \emptyset\}|$ und $Y_n^{u,v}(G) = \begin{cases} 1 & N_G(u) \cap N_G(v) = \emptyset \\ 0 & \text{sonst} \end{cases}$.
Dann gilt $Y_n = \sum_{\{u,v\}\in\binom{[n]}{2}} Y_n^{u,v}$. Zu zeigen ist $\mathbf{P}(Y_n > 0) \stackrel{n\to\infty}{\longrightarrow} 0$.
Für beliebiges $\{u, v\} \in \binom{[n]}{2}$ gilt $\mathbf{P}(Y_n^{u,v} > 0) = \mathbf{P}(Y_n^{u,v} = 1) = (\frac{3}{4})^{n-2}$.

$$\begin{aligned} \Rightarrow \mathbf{E}[Y_n] &= \mathbf{E}\left[\sum_{\{u,v\}\in\binom{[n]}{2}} Y_n^{u,v}\right] = \sum_{\{u,v\}\in\binom{[n]}{2}} \mathbf{E}[Y_n^{u,v}] \\ &= \sum_{\{u,v\}\in\binom{[n]}{2}} \mathbf{P}(Y_n^{u,v} = 1) = \binom{n}{2}(\tfrac{3}{4})^{n-2} \stackrel{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

$\Rightarrow \mathbf{P}(\mathrm{diam}(\mathcal{G}(n, \frac{1}{2})) \neq 2) \leq o(1) + \mathbf{P}(\mathrm{diam}(\mathcal{G}(n\frac{1}{2})) > 2) = o(1) + \mathbf{P}(Y_n > 0) \stackrel{n\to\infty}{\longrightarrow} 0$ $\square$

Definition (k-erweiterbar)
Ein Graph $G = (V, E), k \in \mathbb{N}$ ist *k-erweiterbar*, falls für alle k-elementigen Teilmengen $S \in \binom{V}{k}$ und für alle $T \subseteq S$ ein $x \in V \setminus S$ mit $N(x) \cap S = T$ existiert.

Satz 1.3. *Für alle $k \in \mathbb{N}$ gilt $\mathbf{P}(\mathcal{G}(n, \frac{1}{2})$ ist k-erweiterbar$) = 1 - o(1)$.*

Beweis:
Für beliebiges $n \geq k$ betrachten wir die Zufallsvariable

$$Y_S = \begin{cases} 0 & \forall T \subseteq S : \exists x \in V \setminus S : N(x) \cap S = T \\ 1 & \text{sonst} \end{cases}.$$

Weiterhin sei Y die Anzahl der nicht erweiterbaren Mengen in $\binom{V}{k}$, d.h. $Y = \sum_{S\in\binom{V}{k}} Y_S$.

Es gilt $P(Y_S = 1) \leq \underbrace{2^k}_{\forall T\subseteq S} \cdot \left(\underbrace{\frac{2^k - 1}{2^k}}_{\mathbf{P}(N(x)\cap S\neq T)}\right)^{n-k,\ \text{weil}\ \forall x\in V\setminus S}$.

$\Rightarrow \mathbf{E}[Y] = \sum_{S\in\binom{V}{k}} \mathbf{E}[Y_s] = \sum_{S\in\binom{V}{k}} \mathbf{P}(Y_S = 1) \leq \underbrace{\binom{n}{k}}_{=O(n^k)} \cdot 2^k \underbrace{(1 - \frac{1}{2^k})}_{<1}{}^{n-k} \overset{n\to\infty}{\longrightarrow} 0$

$\Rightarrow \mathbf{P}(\mathcal{G}(n, \frac{1}{2})$ ist nicht k-erweiterbar$) = \mathbf{P}(Y > 0) = o(1)$ □

Korollar 1.2. *Für alle Graphen F gilt $\mathbf{P}(F \leq \mathcal{G}(n, \frac{1}{2})) = 1 - o(1)$. (Mit $\leq$ bezeichnen wir induzierte Teilgraphen.)*
Genauer: $\forall F, \varepsilon \exists n_0 \forall n \geq n_0 : \mathbf{P}(F \leq \mathcal{G}(n, \frac{1}{2})) \geq 1 - \varepsilon$

Beweis:
Wir führen eine Induktion über $l = v(F) = |V(F)|$ durch:
Der Induktionsanfang $l = 1$ ist klar.
Es bleibt der Induktionsschritt $l \to l + 1$:
Wir haben bereits gezeigt: $\mathbf{P}(\mathcal{G}(n, \frac{1}{2})$ ist l-erweiterbar$) = 1 - o(1)$.
Laut Induktionsvoraussetzung haben wir $\mathbf{P}(F' \leq \mathcal{G}(n, \frac{1}{2})) = 1 - o(1)$ für $F' = F \setminus \{v\}$, $v \in V(F)$ beliebig.
Wir zeigen, dass F-freie Graphen F' frei oder nicht l-erweiterbar sind: Wenn ein Graph G l-erweiterbar ist, F' als induzierten Teilgraphen enthält und $n \geq l + 1$ gilt, so folgt $F \leq G$. (Wir wählen $S = V(F') \subseteq V(G)$ und $T = \{x \in S | x$ ist Nachbar von v in $F\}$. Dann erhalten wir durch die l-Erweiterbarkeit, dass ein $y \in V(G) \setminus S$ existiert, welcher das gleiche Nachbarschaftsmuster wie v erzeugt.)
$\Rightarrow G[S \cup \{y\}] \cong F$
$\Rightarrow \mathbf{P}(F \not\leq \mathcal{G}(n, \frac{1}{2})) \leq \mathbf{P}(F' \not\leq \mathcal{G}(n, \frac{1}{2})$ oder $G(n, \frac{1}{2})$ ist nicht l-erweiterbar$) \leq o(1) + o(1) = o(1)$ □

Korollar 1.3. *1. $\forall k \in \mathbb{N} : \mathbf{P}(\chi(\mathcal{G}(n, p)) \geq k) \geq 1 - o(1)$, weil $K_{k-1} \leq \mathcal{G}(n, \frac{1}{2})$.*

2. $\mathbf{P}(\mathcal{G}(n, \frac{1}{2})$ ist planar$) = o(1)$ *„Kuratovski“*

3. $\mathbf{P}(\mathcal{G}(n, \frac{1}{2})$ ist perfekt$) = o(1)$ *„Berge“*

Beweis:
Die Beweise werden hier nicht angegeben. □

Vorlesung am 25.04.2008

1.4 Varianz, Chebyshev (bzw. Tschebyscheff oder Čebyšev), Anzahl der Kanten

Wieviele Kanten hat $\mathcal{G}(n, \frac{1}{2})$?
Antwort: $\mathbf{E}[e(\mathcal{G}(n, \frac{1}{2}))] = \frac{n(n-1)}{4} = \frac{1}{2}\binom{n}{2}$
Eine Begründung / Herleitung dieser Antwort:
$X_{u,v} = \begin{cases} 0 & \{u, v\} \notin G \\ 1 & \{u, v\} \in G \end{cases}$
$\Rightarrow \mathbf{E}[X_{u,v}] = \frac{1}{2}$
$\Rightarrow \mathbf{E}[e(\mathcal{G}(n, \frac{1}{2}))] = \mathbf{E}\left[\sum_{\{u,v\}\in\binom{[n]}{2}} X_{u,v}\right] = \sum_{\{u,v\}\in\binom{[n]}{2}} \mathbf{E}[X_{u,v}] = \sum_{\{u,v\}\in\binom{[n]}{2}} \frac{1}{2} = \frac{1}{2}\binom{n}{2}$

Eine weitere Frage: Wie wahrscheinlich (bzw. wie weit) weicht $e(G(n, \frac{1}{2}))$ vom Erwartungswert ab?

Definition (Varianz)
Sei X eine Zufallsvariable. Dann ist $\mathrm{Var}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2]$ die *Varianz*.

Lemma 1.1. *Es gilt* $\mathrm{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2$.

Beweis:
$$\begin{aligned}\mathrm{Var}[X] &= \mathbf{E}[(X - \mathbf{E}[X])^2] = \mathbf{E}[X^2 - 2X\mathbf{E}[X] + (\mathbf{E}[X])^2] = \mathbf{E}[X^2] - 2\mathbf{E}[X] \cdot \mathbf{E}[X] + \mathbf{E}[(\mathbf{E}[X])^2] \\ &= \mathbf{E}[X^2] - 2(\mathbf{E}[X])^2 + (\mathbf{E}[X])^2 = \mathbf{E}[X^2] - (\mathbf{E}[X])^2\end{aligned}$$ □

Bemerkung
Für eine Zufallsvariable $X : \Omega \to \{0,1\} = 2$ gilt $X^2 : \Omega \to 2$ und insbesondere $X^2 = X$.

Satz 1.4. (Tschebyscheff(-Ungleichung)) *Sei X eine Zufallsvariable. Dann gilt für alle $t > 0$:* $\mathbf{P}(|X - \mathbf{E}[X]| \geq t) \leq \frac{\mathrm{Var}[X]}{t^2}$

Beweis:
Sei $Y = |X - \mathbf{E}[X]|$.

$$\begin{aligned} \Rightarrow \mathbf{P}(|X - \mathbf{E}[X]| \geq t) &= \mathbf{P}(Y \geq t) = \mathbf{P}(Y^2 \geq t^2) \\ &\leq \frac{\mathbf{E}[Y^2]}{t^2} = \frac{\mathbf{E}[|X-\mathbf{E}[X]|^2]}{t^2} = \frac{\mathbf{E}[(X-\mathbf{E}[X])^2]}{t^2} \\ &= \frac{\mathrm{Var}[X]}{t^2} \end{aligned}$$
□

Satz 1.5. (Abweichung der Kantenanzahl vom Erwartungswert) *Sei $X_n = e(\mathcal{G}(n, \frac{1}{2}))$ und $\omega(n)$ eine beliebige Funktion mit $\omega(n) \stackrel{n\to\infty}{\longrightarrow} \infty$.*
Dann gilt $\mathbf{P}(|X_n - \frac{1}{2}\binom{n}{2}| > n\omega(n)) \stackrel{n\to\infty}{\longrightarrow} 0$.

Beweis:
Nach Tschebyscheff gilt $\mathbf{P}(|X_n - \frac{1}{2}\binom{n}{2}| > n\omega(n)) \leq \frac{\mathrm{Var}[X_n]}{(\omega(n)n)^2}$.
Wir zerlegen nun X_n wieder in einfache Zufallsvariablen $X_{u,v}$ wie oben, d.h. $X = \sum_{\{u,v\}\in\binom{[n]}{2}} X_{u,v}$.
Damit ergibt sich:

$$\begin{aligned} \mathbf{E}[X^2] &= \mathbf{E}\left[\left(\sum_{\{u,v\}\in\binom{[n]}{2}} X_{u,v}\right)^2\right] = \mathbf{E}\left[\sum_{\{u,v\}\in\binom{[n]}{2}} X_{u,v}^2 + \sum_{\{u,v\}\in\binom{[n]}{2}\ni\{x,y\}\neq\{u,v\}} X_{u,v}X_{x,y}\right] \\ &= \sum_{\{u,v\}\in\binom{[n]}{2}} \mathbf{E}[X_{u,v}^2] + \sum_{\{u,v\}\in\binom{[n]}{2}\ni\{x,y\}\neq\{u,v\}} \mathbf{E}[X_{u,v}X_{x,y}] \\ &= \sum_{\{u,v\}\in\binom{[n]}{2}} \mathbf{E}[X_{u,v}] + \sum_{\{u,v\}\in\binom{[n]}{2}\ni\{x,y\}\neq\{u,v\}} \mathbf{P}(X_{u,v}X_{x,y} = 1) \\ &= \sum_{\{u,v\}\in\binom{[n]}{2}} \frac{1}{2} + \sum_{\{u,v\}\in\binom{[n]}{2}\ni\{x,y\}\neq\{u,v\}} \mathbf{P}[X_{u,v} = 1, X_{x,y} = 1] \\ &= \frac{1}{2}\binom{n}{2} + \sum_{\{u,v\}\in\binom{[n]}{2}\ni\{x,y\}\neq\{u,v\}} \mathbf{P}[X_{u,v} = 1]\cdot\mathbf{P}[X_{x,y} = 1] \\ &= \frac{1}{2}\binom{n}{2} + \sum_{\{u,v\}\in\binom{[n]}{2}\ni\{x,y\}\neq\{u,v\}} \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2}\binom{n}{2} + 2\binom{\binom{n}{2}}{2}\cdot\frac{1}{4} = \frac{1}{2}\binom{n}{2} + \frac{1}{2}\binom{\binom{n}{2}}{2} \end{aligned}$$

$$\begin{aligned} \Rightarrow \mathrm{Var}[X_n] &= \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \frac{1}{2}\binom{n}{2} + \frac{1}{2}\binom{\binom{n}{2}}{2} - (\frac{1}{2}\binom{n}{2})^2 \\ &= \frac{1}{4}(n^2-n) + \frac{1}{4}(\frac{n^2-n}{2})(\frac{n^2-n}{2} - 1) - \frac{1}{4}(\frac{n^2-n}{2})^2 \\ &= \frac{1}{8}(n^2-n) + \frac{1}{4}(\frac{n^2-n}{2})^2 - \frac{1}{4}(\frac{n^2-n}{2})^2 = \frac{1}{8}(n^2-n) \end{aligned}$$

$$\Rightarrow \frac{\mathrm{Var}[X_n]}{(\omega(n)n)^2} = \frac{n^2-n}{8(\omega(n)n)^2} \stackrel{n\to\infty}{\longrightarrow} 0$$
□

1.5 Schwellenwertfunktion, $\mathcal{G}(n,p)$

Definition ($\mathcal{G}(n,p)$-Modell)
Wir definieren das $\mathcal{G}(n,p)$-*Modell* analog zum $\mathcal{G}(n, \frac{1}{2})$-Modell, wobei die Wahrscheinlichkeit, dass eine Kante existiert, genau p (statt bisher $\frac{1}{2}$) ist.

Bemerkung
Es gilt $\mathbf{P}(\mathcal{G}(n,p) = G) = p^{e(G)}(1-p)^{\binom{n}{2}-e(G)}$ mit $V(G) = [n]$.
Typischerweise betrachten wir $0 \leq p = p(n) \leq 1$, d.h. p ist eine Funktion, die von n abhängt.

Definition (Schwellenwertfunktion)
Sei $\mathcal{P} \subseteq \{G | G \text{ ist ein Graph}\}$ und $\mathcal{P}_n = \{G \in \mathcal{P} | v(G) = n\}$.
Wir sagen $\hat{p} = \hat{p}(n) : \mathbb{N} \to [0,1]$ ist eine *Schwellenwertfunktion* von $\mathcal{P}$, falls gilt:

$$\lim_{n\to\infty} \mathbf{P}(\mathcal{G}(n,p) \in \mathcal{P}) = \begin{cases} 0 & \text{falls } p(n) \ll \hat{p}(n) \\ 1 & \text{falls } p(n) \gg \hat{p}(n) \end{cases}$$

Bemerkung
$\hat{p}(n)$ ist nicht eindeutig, da mit $\hat{p}$ auch $\tilde{p}(n) = c\hat{p}(n)$ für alle $0 < c \leq 1$ eine Schwellenwertfunktion für $\mathcal{P}$ ist.

Satz 1.6. *Sei $\mathcal{P} = \{G | \delta(G) > 0\}$ die Eigenschaft der Graphen ohne isolierte Knoten. Es gilt $\lim_{n\to\infty} \mathbf{P}(\mathcal{G}(n,p) \in \mathcal{P}) \to 1$, falls $p(n) \gg \frac{\ln n}{n}$.*

Beweis:
Sei Y die Anzahl der isolierten Knoten und $Y_u = \begin{cases} 0 & \deg(u) \geq 1 \\ 1 & \deg(u) = 0 \end{cases}$.
$\Rightarrow \mathbf{E}[Y] = \sum_{u\in[n]} \mathbf{E}[Y_u] = \sum_{u\in[n]} \mathbf{P}(u \text{ ist isoliert}) = n \cdot (1-p)^{n-1} = \frac{n}{1-p}(1-p)^n$
Wir haben die Ungleichung $\forall x \in [0,1] : (1-x) \leq e^{-x}$ (folgt direkt aus der Reihendarstellung von e^x).
$\Rightarrow \mathbf{E}[Y] = \frac{n}{1-p}(1-p)^n \leq \frac{n}{1-p}e^{-pn}$
Zu zeigen ist nun $\mathbf{P}(Y > 0) \to 0$, d.h. $\mathbf{E}[Y] \to 0$.
Wir haben $\mathbf{E}[Y] \leq \frac{n}{1-p}e^{-pn} \leq \frac{n}{1-p}e^{-\ln n\omega(n)} \leq \frac{n}{(1-p)n^{\omega(n)}} \overset{n\to\infty}{\longrightarrow} 0$.
(Die dritte Ungleichung gilt wegen $p(n) = \omega(n)\frac{\ln n}{n}$.) □

Bemerkung
Im obigen Beweis würde $\omega(n) = (1+\varepsilon)$ ausreichen.

1.6 2. Moment Methode

Frage: Wann gilt $\mathbf{P}(X_n = 0) \overset{n\to\infty}{\longrightarrow} 0$?

Satz 1.7. (2. Moment Methode)
*Seien $(X_n)_{n\in\mathbb{N}}$ Zufallsvariablen mit $\mathbf{E}X_n \to \infty$ und $\frac{\mathrm{Var}[X_n]}{(\mathbf{E}[X_n])^2} \overset{n\to\infty}{\longrightarrow} 0$.
Dann gilt $\mathbf{P}(X_n = 0) \overset{n\to\infty}{\longrightarrow} 0$.*

Beweis:
$\mathbf{P}(X_n = 0) \leq \mathbf{P}(|X_n - \mathbf{E}[X_n]| \geq \mathbf{E}[X_n]) \leq \frac{\mathrm{Var}[X_n]}{(\mathbf{E}[X_n])^2} \overset{n\to\infty}{\longrightarrow} 0$ □

Bemerkung
Es reicht aus, wenn $\mathbf{E}[X_n] > c$ für alle $n \geq n_0$ und eine Konstante $c > 0$ gilt.

Satz 1.8. *Seien $\mathcal{P}$ und $\mathcal{P}_n$ wie im vorletzten Satz.
Falls $p(n) \ll \frac{\ln n}{n}$, dann gilt $\lim_{n\to\infty} \mathbf{P}(\mathcal{G}(n,p) \in \mathcal{P}_n) = 0$.*

Beweis:
Seien Y und Y_n wie im vorletzten Satz definiert. Dann gilt $\mathbf{E}[Y] = n(1-p)^{n-1}$.
Wir haben $(1-x) \geq e^{-x(x+1)}$ für $x < 0.68$ (Beweis durch Reihendarstellung der Exponentialfunktion).

$$\begin{aligned}\Rightarrow \mathbf{E}[Y] &= n(1-p)^{n-1} \geq ne^{-p(p+1)(n-1)} \geq ne^{-\frac{\ln n}{\omega(n)n}(1+\frac{\ln n}{\omega(n)n})n} = nn^{-\frac{1}{\omega(n)} - \frac{\ln n}{\omega(n)^2 n}} \\ &\geq \frac{n}{n^{\frac{1}{\omega(n)} + \frac{n}{\omega(n)^2 n}}} = \frac{n}{n^{\frac{1}{\omega(n)} + \frac{1}{\omega(n)^2}}} \geq n^{1-\varepsilon} \overset{n\to\infty}{\longrightarrow} \infty\end{aligned}$$

Zu zeigen bleibt also noch $\frac{\mathrm{Var}[X_n]}{(\mathbf{E}[X_n])^2} \overset{n\to\infty}{\longrightarrow} 0$, d.h. $\mathrm{Var}[Y] \ll (\mathbf{E}[Y])^2 = n^{2-\varepsilon}$ zu zeigen.

$$Y^2 = \sum_{u\in[n]} Y_u^2 + \sum_{(u,v)\in[n]^2, u\neq v} Y_u Y_v = \underbrace{\sum_{u\in[n]} Y_u}_{=Y} + \sum_{(u,v)\in[n]^2, u\neq v} Y_u Y_v$$

$$\begin{aligned}\Rightarrow \mathbf{E}[Y^2] &= \mathbf{E}[Y] + n(n-1)(1-p)^{2n-3} = n(1-p)^{n-1}(1+(n-1)(1-p)^{n-2}) \\ &= \mathbf{E}[Y] \cdot (1+(n-1)(1-p)^{n-2}) \\ \Rightarrow \frac{\mathrm{Var}[y]}{(\mathbf{E}[Y])^2} &= \frac{\mathbf{E}[Y^2] - (\mathbf{E}[Y])^2}{(\mathbf{E}[Y])^2} = \frac{\mathbf{E}[Y^2]}{(\mathbf{E}[Y])^2} - 1 = \frac{1+(n-1)(1-p)^{n-2}}{\mathbf{E}[Y]} \cdot \frac{\mathbf{E}[Y]}{\mathbf{E}[Y]} - 1 = \frac{1}{\mathbf{E}[Y]} + \frac{(n-1)(1-p)^{n-2}}{n(1-p)^{n-1}} - 1 \\ &= \frac{1}{\mathbf{E}[Y]} + \frac{n-1}{n(1-p)} - 1 \overset{n\to\infty}{\longrightarrow} 0 + 1 - 1 = 0\end{aligned}$$

$\Rightarrow \mathbf{P}(Y=0) \to 0$
$\Rightarrow \mathbf{P}(\mathcal{G}(n,p) \in \mathcal{P}) \to 0$ $\square$

Vorlesung am 30.04.2008

Beispiel (Schwellenwert für $K_3 \subseteq G$)
Sei $X = |\{K_3 \subseteq G\}|$ die Anzahl der Kopien von K_3 in G.
$\Rightarrow \mathbf{E}X = \binom{n}{3}p^3$
Wegen $\binom{n}{3} \sim n^3$ erhalten wir $n^3p^3 \to \infty$, wenn $p \gg \frac{1}{n}$.
Wir versuchen also $\hat{p}(n) = \frac{1}{n}$ nachzuweisen.

Satz 1.9. (Schwellenwert von $K_3 \subseteq G$) *Der Schwellenwert von $K_3 \subseteq G$ ist $\frac{1}{n}$.*

Beweis:

1. Sei $p(n) = \frac{1}{\omega(n)n}$ mit $\omega(n) \to \infty$.

   Zu zeigen ist $\mathbf{P}(\mathcal{G}(n,p) \subseteq K_3) \to 0$.

   Sei X die Anzahl der Dreiecke in G. Dann gilt $\mathbf{E}X = \binom{n}{3}p^3 \leq n^3p^3 \leq \frac{1}{\omega(n)^3} \to 0$.

   Mit Markoff folgt $\mathbf{P}(\mathcal{G}(n,p) \subseteq K_3) \to 0$.

2. Sei $p(n) = \frac{\omega(n)}{n}$ mit $\omega(n) \to \infty$.

   Zu zeigen ist

   (a) $\mathbf{E}X \to \infty$

   (b) $\frac{\mathrm{Var}(X)}{(\mathbf{E}X)^2} \to 0$

   Zu den Beweisen für diese Behauptungen:

   (a) $\mathbf{E}X = \binom{n}{3}p^3 \geq (\frac{n}{3})^3p^3 = \frac{\omega(n)^3}{27} \to \infty$

   (b) $\mathrm{Var}\, X = \mathbf{E}X^2 - (\mathbf{E}X)^2$

   Wir definieren $X_T = \begin{cases} 0 & T \text{ spannt nicht } K_3 \\ 1 & T \text{ spann } K_3 \end{cases}$. Dann gilt $X = \sum_{T \in \binom{[n]}{3}} X_T$.

   Wir haben also $\mathbf{E}X^2 = \sum_{T \in \binom{[n]}{3}} X_T^2 + \sum_{T \neq S \in \binom{[n]}{3}} X_T X_S$.

   $\Rightarrow \mathbf{E}X^2 = \mathbf{E}X + \sum_{T \neq S} \underbrace{\mathbf{E}[X_T X_S]}_{=\mathbf{P}(X_T=1, X_S=1)}$

   $\mathbf{P}(T \text{ und } S \text{ sind } K_3) = \begin{cases} p^6 & T \cap S = \emptyset \\ p^6 & |T \cap S| = 1 \\ p^5 & |T \cap S| = 2 \end{cases}$

   Der erste Fall tritt $\binom{n}{3}\binom{n-3}{3}$ mal auf,
   der zweite $3\binom{n}{3}\binom{n-3}{2}$ mal und
   der dritte $3\binom{n}{3}(n-3) = 12\binom{n}{4}$ mal.

   $\Rightarrow \mathbf{E}X^2 = \mathbf{E}X + \binom{n}{3}\binom{n-3}{3}p^6 + 3\binom{n}{3}\binom{n-3}{2}p^6 + 3\binom{n}{3}(n-3)p^5$

   $$\begin{aligned} \frac{\mathrm{Var}\, X}{(\mathbf{E}X)^2} &= \frac{1}{\mathbf{E}X} + \frac{\binom{n}{3}\binom{n-3}{3}p^6}{\binom{n}{3}^2p^6} + \frac{3\binom{n}{3}\binom{n-3}{2}p^6}{\binom{n}{3}^2p^6} + \frac{3\binom{n}{3}(n-3)p^5}{\binom{n}{3}^2p^6} - \frac{(\mathbf{E}X)^2}{(\mathbf{E}X)^2} \\ &\leq \frac{1}{\mathbf{E}X} + \frac{\binom{n-3}{3}}{\binom{n}{3}} + \frac{n^5}{(\frac{n}{3})^6} + \frac{n^4}{(\frac{n}{3})^6 p} - 1 \stackrel{n\to\infty}{\longrightarrow} 0 + \lim_{n\to\infty} \frac{\binom{n-3}{3}}{\binom{n}{3}} + 0 + 0 - 1 \\ &= 1 - 1 = 0 \end{aligned}$$

$\square$

Implikationen in $\mathcal{G}(n,p)$
Gilt in $\mathcal{G}(n,p)$ die Implikation $\delta(G) > 0 \Rightarrow G$ ist zusammenhängend?
Genauer: Haben „$\delta(G) > 0$" und „G zusammenhängend" den selben Schwellenwert?

Satz 1.10. $\frac{\ln n}{n}$ *ist der (oder besser: ein) Schwellenwert für „G ist zusammenhängend".*

Beweis:

1. $p \ll \frac{\ln n}{n} \Rightarrow \mathbf{P}(\delta(\mathcal{G}(n,p)) = 0) \to 1 \Rightarrow \mathbf{P}(\mathcal{G}(n,p)$ ist zusammenhängend$) \to 0$

2. Sei $p(n) = \frac{\omega(n)\ln n}{n}$ und X die Anzahl der Zusammenhangskomponenten mit höchstens $\frac{n}{2}$ Knoten.

   Sei X_k die Anzahl der Zusammenhangskomponenten der Größe k, d.h. es gilt $X = \sum_{k=1}^{\frac{n}{2}} X_k$.

   Wir wollen zeigen, dass $\mathbf{P}(X > 0) \overset{n\to\infty}{\longrightarrow} 0$.

   Idee: Wir verwenden die 1.M.M., indem wir zeigen, dass $\mathbf{E}X \to 0$.

   Wegen $\mathbf{E}X = \sum_{k=1}^{\frac{n}{2}} \mathbf{E}X_k$ werden wir versuchen zu zeigen, dass für alle $k = 1, ..., \frac{n}{2}$ die Ungleichung $\mathbf{E}X_k \ll \frac{1}{n}$ gilt.

   Sei $K \subseteq \binom{[n]}{k}$.

   Dann gilt $\mathbf{P}(K$ ist Zusammenhangskomponente in $\mathcal{G}(n,p)) \leq (1-p)^{k(n-k)}$, weil keine Kante die Zusammenhangskomponente verlassen darf.

   $\Rightarrow \mathbf{E}X_k \leq \binom{n}{k}(1-p)^{k(n-k)} \leq n^k e^{-pk(n-k)} \leq n^k e^{-\frac{\omega(n)\ln n}{n}k\frac{n}{2}} = n^k n^{-\omega(n)\frac{k}{2}} = n^{-k(\frac{\omega(n)}{2}-1)}$
   $\leq n^{-k} \leq n^{-1} = \frac{1}{n} \to 0$ □

Vorlesung am 07.05.2008

Satz 1.11.
Es existiert ein $l_0 : \mathbb{N} \to \mathbb{N}$ mit $l_0 \sim 2\log n$ und $\mathbf{P}(l_0(n) - 1 \leq \omega(\mathcal{G}(n, \frac{1}{2})) \leq l_0(n)) \overset{n\to\infty}{\longrightarrow} 1$.

Beweis:
Sei X_l die Anzahl der K_l in G.
Wir wollen zeigen:

1. $\mathbf{P}(X_{l_0(n)+1} = 0) \overset{n\to\infty}{\longrightarrow} 1$

2. $\mathbf{P}(X_{l_0(n)-1} = 0) \overset{n\to\infty}{\longrightarrow} 0$

3. $l_0(n) \sim 2\log n$

Zu den Beweisen:

1. Idee: Setze l_0 so, dass $\mathbf{E}(X_{l_0}) \geq \frac{1}{\sqrt{n}}$.

   Wegen $n = \mathbf{E}X_1 \gg \frac{1}{\sqrt{n}}$ und $\mathbf{E}X_n = (\frac{1}{2})^{\binom{n}{2}} \ll \frac{1}{\sqrt{n}}$ existiert ein l_0 mit $\mathbf{E}X_{l_0} \geq \frac{1}{\sqrt{n}}$ und $\mathbf{E}X_{l_0+1} < \frac{1}{\sqrt{n}}$.

   $\Rightarrow \mathbf{P}(X_{l_0+1} > 0) \leq \mathbf{E}X_{l_0+1} < \frac{1}{\sqrt{n}} \overset{n\to\infty}{\longrightarrow} 0$ (Markoff)

   Damit haben wir 1. bewiesen.

3. Sei $l \in \mathbb{N}$ und $X_S = \begin{cases} 1 & S \text{ induziert eine Clique} \\ 0 & \text{sonst} \end{cases}$.

   Wir haben $\mathbf{E}X_l = \mathbf{E}\sum_{S\in\binom{[n]}{l}} X_S = \sum_S \mathbf{E}X_S = \binom{n}{l}(\frac{1}{2})^{\binom{l}{2}}$.

   Wegen $(\frac{a}{b})^b \leq \binom{a}{b} \leq (\frac{ea}{b})^b$ gilt:

$b(\log_2(a) - \log_2(b)) \leq \log_2 \binom{a}{b} \leq b(\log_2(a) + \log_2(e) - \log_2(b))$

$\Rightarrow \log_2 \binom{a}{b} = b\log_2(a) - b\log_2(b) + O(b)$

$\Rightarrow \mathbf{E}X_l = 2^{l\log_2(n) - l\log_2(l) + O(l) - \frac{l^2}{2} + \frac{l}{2}} = 2^{l\log_2(n) - \frac{l^2}{2} + o(l^2)}$

Für $l = c\log n$ erhalten wir $\mathbf{E}X_l = 2^{c(1-\frac{c}{2})(\log n)^2 + o((\log n)^2)}$.

Für $c = 2+\varepsilon$ gilt $\mathbf{E}X_l = 2^{(2+\varepsilon)(1-1-\frac{\varepsilon}{2})(\log n)^2 + o((\log n)^2)} = 2^{-(\varepsilon+\frac{\varepsilon^2}{2})(\log n)^2 + o((\log n)^2)} \ll \frac{1}{n}$.

Für $c = 2-\varepsilon$ gilt $\mathbf{E}X_l = 2^{(2-\varepsilon)(1-1+\frac{\varepsilon}{2})(\log n)^2 + o((\log n)^2)} = 2^{(\varepsilon-\frac{\varepsilon^2}{2})(\log n)^2 + o((\log n)^2)} \gg n$.

Da $\mathbf{E}X_{l_0+1} < \frac{1}{\sqrt{n}} \leq \mathbf{E}X_{l_0}$ gilt $l_0 \sim 2\log n$.

2. $\frac{\mathbf{E}X_{l+1}}{\mathbf{E}X_l} = \frac{\binom{n}{l+1}(\frac{1}{2})^{\binom{l+1}{2}}}{\binom{n}{l}(\frac{1}{2})^{\binom{l}{2}}} = \frac{n!l!(n-l)!}{(l+1)!(n-l-1)!n!}(\frac{1}{2})^{\binom{l+1}{2}-\binom{l}{2}} = \frac{n-l}{l+1}(\frac{1}{2})^l < n2^{-l}$

$l \sim 2\log n \wedge l \geq 2\log n \Rightarrow l = (2+o(1))\log n$

$\Rightarrow \frac{\mathbf{E}X_{l+1}}{\mathbf{E}X_l} < n2^{-(2+o(1))\log n} = \frac{1}{n^{1+o(1)}}$

$\Rightarrow \mathbf{E}X_{l_0-1} \geq \mathbf{E}X_{l_0} n^{1+o(1)} \geq \frac{1}{\sqrt{n}} n^{1+o(1)} = n^{\frac{1}{2}+o(1)}$

$\Rightarrow \mathbf{E}X_{l_0-1} \to \infty$

Es ist zu zeigen, dass $\mathrm{Var}(X_{l_0-1}) \ll (\mathbf{E}X_{l_0-1})^2$ gilt, da dann 2. aus Tschebyscheff folgt.

Wir setzen $r = l_0 - 1$.

$$\begin{aligned}
\Rightarrow \mathbf{E}X_r^2 &= \mathbf{E}(\textstyle\sum_{R\in\binom{[n]}{r}} X_R)^2 = \mathbf{E}X_r + \sum_{R,R'\in\binom{[n]}{r}, R\neq R'} \mathbf{E}X_R X_{R'} \\
&= \mathbf{E}X_r + \textstyle\sum_{R,R'\in\binom{[n]}{r}, R\neq R'} \mathbf{P}(R, R' \text{ induzieren Cliquen}) \\
&= \mathbf{E}X_r + \textstyle\sum_{k=0}^{r} \underbrace{\binom{n}{r}}_{R} \underbrace{\binom{r}{k}}_{R\cap R'} \underbrace{\binom{n-r}{r-k}}_{R'\setminus R} (\frac{1}{2})^{2\binom{r}{2}-\binom{k}{2}}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \frac{\mathrm{Var}\, X_r}{(\mathbf{E}X_r)^2} &= \frac{1}{\mathbf{E}X_r} + \frac{\sum_{k=0}^{r-1}\binom{r}{k}\binom{n-r}{r-k}2^{\binom{k}{2}}}{\binom{n}{r}} - 1 \\
&= \frac{1}{\mathbf{E}X_r} + \frac{\sum_{k=1}^{r-1}\binom{r}{k}\binom{n-r}{r-k}2^{\binom{k}{2}}}{\binom{n}{r}} + \underbrace{\frac{\binom{r}{0}\binom{n-r}{r}2^0}{\binom{n}{r}} - 1}_{\xrightarrow{n\to\infty} 0}
\end{aligned}$$

Nun ist $\frac{\sum_{k=1}^{r-1}\binom{r}{k}\binom{n-r}{r-k}2^{\binom{k}{2}}}{\binom{n}{r}} \to 0$ zu zeigen.

Wir setzen $g(k) = \frac{\binom{r}{k}\binom{n-r}{r-k}2^{\binom{k}{2}}}{\binom{n}{r}}$. Gesucht ist $\sum_{k=1}^{\frac{r}{2}} g(k) + \sum_{k=1} \frac{r}{2} g(r-k)$.

Für $1 \leq k \leq \frac{r}{2}$ haben wir $g(k) = \binom{r}{k}\frac{(n-r)_{r-k} r!}{(r-k)!(n)_r}2^{\binom{k}{2}}$ mit $(a)_b = \frac{a!}{(a-b)!}$. Damit folgt:

$g(k) \leq r^k \frac{(n)_{r-k}}{(n)_r}(r)_k 2^{\binom{n}{2}} \leq \frac{r^{2k}2^{\binom{k}{2}}}{(n-r)^k} \leq (\frac{r^2 2^{\frac{k}{2}}}{n-r})^k$

Wegen $k \leq \frac{r}{2}$ und $r = l_0 - 1 = (2+o(1))\log n$ gilt:

$\frac{r^2 2^{\frac{k}{2}}}{n-r} \leq \frac{r^2 2^{\frac{r}{4}}}{n-r} \leq \frac{8(\log n)^2}{n^{\frac{1}{2}} - o(1)} \leq \frac{1}{n^{\frac{1}{4}}}$

$\Rightarrow \sum_{k=1}^{\frac{r}{2}} g(k) \leq \sum_{k=1}^{\frac{r}{2}} (\frac{1}{n^{\frac{1}{4}}})^k \leq \sum_{k=1}^{\infty} (\frac{1}{n^{\frac{1}{4}}})^k = \frac{1}{n^{\frac{1}{4}} - 1} \xrightarrow{n\to\infty} 0$

(Für $\frac{r}{2} \leq k \leq r$ gilt $g(r-k) \leq (\frac{1}{n^{\frac{1}{4}}})^k \Rightarrow \sum_{k=1}^{\frac{r}{2}} g(r-k) \xrightarrow{n\to\infty} 0$) $\square$

Bemerkung: Es gilt $l_0(n) = \lfloor 2\log_2(n) - 2\log_2\log_2(n) + 2\log_2(e) - \frac{3}{2} + o(1) \rfloor$.

Bemerkung
Eigentlich hätten wir $\mathbf{P}(X_{l_0(n)} = 0) \overset{n\to\infty}{\longrightarrow} 0$ zeigen müssen.
(Bisher wurde nur $P(\omega(G(n, \frac{1}{2})) \in \{l_0(n), l_0(n) \pm 1\}) \overset{n\to\infty}{\longrightarrow} 1$ gezeigt.)

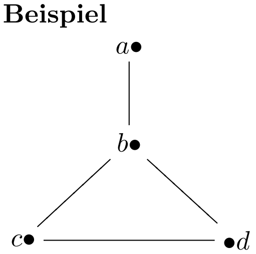
Vorlesung am 09.05.2008

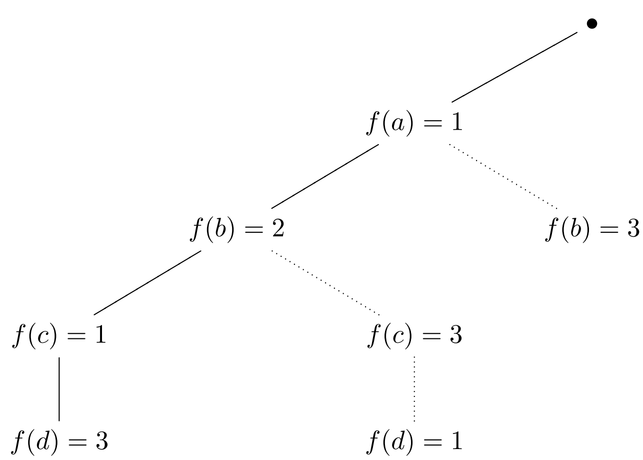
1.7 Chromatische Zahl von $\mathcal{G}(n, \frac{1}{2})$

Bemerkung
Für jeden Graphen gilt $\omega(\mathcal{G}(n, \frac{1}{2})) \le \chi(\mathcal{G}(n, \frac{1}{2}))$ und (was besser ist) $\frac{n}{\alpha(\mathcal{G}(n,\frac{1}{2}))} \le \chi(\mathcal{G}(n, \frac{1}{2}))$.

Back-Tracking-Algorithmus
Dieser Algorithmus probiert rekursiv alle möglichen k-Färbungen aus.

Beispiel





Satz 1.12. (Wilt) *Für alle $k \in \mathbb{N}$ ist die erwartete Laufzeit des Back-Tracking-Algorithmus konstant.*

Beweis:
Sei $\beta(G)$ die Anzahl der Schritte des BTA. Dann gilt $\mathbf{E}[\text{Laufzeit des BTA}] = 2^{-\binom{n}{2}} \sum_{G\in\mathcal{G}_n} \beta(G)$.
Zunächst einmal: Wie viele k-färbbare Graphen mit n Knoten gibt es?
Idee: Die Anzahl der k-partiten Graphen ist höchstens die Anzahl der k-Partitionen $V_1 \dot\cup ... \dot\cup V_k$ mal die Anzahl der Untergraphen des vollständigen $K(V_1, ..., V_k)$. Die Anzahl der Kanten in $K(V_1, ..., V_k)$ ist
$e(K(V_1, ..., V_k)) = \sum_{1\le i<j\le k} |V_i||V_j| = \frac{1}{2}(\sum_{i=1}^k |V_i|)^2 - \frac{1}{2}\sum_{i=1}^k |V_i|^2 \le \frac{n^2}{2} - \frac{k}{2}(\frac{n}{k})^2 = \frac{n^2}{2}(1 - \frac{1}{k})$
Damit ist die Anzahl der k-färbbaren Graphen in $\mathcal{G}_n$ höchstens $k^n \cdot 2^{\frac{n^2}{2}(1-\frac{1}{k})} \le 2^{\frac{n^2}{2}(1-\frac{1}{k}+o(1))}$.

$$
\begin{aligned}
\sum_{G\in\mathcal{G}_n} 2^{-\binom{n}{2}}\beta(G) &\le 2^{-\binom{n}{2}} \sum_{i=1}^n \sum_{H\in\mathcal{G}_i} \sum_{G\in\mathcal{G}_n, G[\{1,...,i\}]=H} \#(k\text{-Färb.}(H)) \\
&\le 2^{-\binom{n}{2}} \sum_{i=1}^n \underbrace{k^i}_{\#k-\text{färbb. pro Graph}} \cdot \underbrace{2^{\frac{i^2}{2}(1-\frac{1}{k}+o(1))}}_{\#k-\text{färbb. Graphen}} \cdot \underbrace{2^{\binom{n}{2}-\binom{i}{2}}}_{G[\{1,...,i\}]=H} \\
&= \sum_{i=1}^n k^i \cdot 2^{\frac{i^2}{2}(1-\frac{1}{k}+o(1))} \cdot 2^{-\binom{i}{2}} \\
&= \sum_{i=1}^n k^i \cdot 2^{\frac{i^2}{2}(1-\frac{1}{k}+o(1))-\binom{i}{2}} \\
&= \sum_{i=1}^n k^i \cdot 2^{\frac{i^2}{2}-\frac{i^2}{2k}+o(i^2))-\frac{i^2}{2}+\frac{i}{2}} \\
&= \sum_{i=1}^n k^i \cdot 2^{-\frac{i^2}{2k}+o(i^2))+\frac{i}{2}} \\
&= \sum_{i=1}^n \left(k \cdot 2^{-\frac{i}{2k}+o(i))+\frac{1}{2}}\right)^i \\
&= \sum_{i=1}^n \left(k \cdot 2^{-\frac{i}{2k}+o(i))} \cdot \sqrt{2}\right)^i
\end{aligned}
$$

$\forall k \exists i_k : \forall i > i_k : \frac{k\sqrt{2}}{\left(2^{-\frac{1}{2k}+o(1))}\right)^i} < 1$

$\Rightarrow \mathbf{E}[\text{Laufzeit des BTA für } k] \leq \sum_{i=1}^{i_k} k\sqrt{2}\left(2^{-\frac{1}{2k}+o(1)}\right)^i + \sum_{i=i_k+1}^{\infty} k\sqrt{2}\left(2^{-\frac{1}{2k}+o(1)}\right)^i$ $\square$
$= O(1) + O(1)$

Für $k = 3$ und $i = 19$ gilt $\frac{k\sqrt{2}}{\left(2^{-\frac{1}{2k}+o(1))}\right)^i} < \frac{1}{2}$. Die erwartete Laufzeit ist kleiner als 479 für $k = 3$ und 34881 für $k = 4$.

Vorlesung am 14.05.2008

$\mathbf{E}\chi(\mathcal{G}(n, \frac{1}{2}))$
Was ist $\mathbf{E}\chi(\mathcal{G}(n, \frac{1}{2}))$?
Wir wissen, dass $\alpha(\mathcal{G}(n, \frac{1}{2})) = (2 + o(1)) \log n$ bzw. $\omega(\mathcal{G}(n, \frac{1}{2})) = (2 + o(1)) \log n$ gilt.
Damit erhalten wir auch $\chi(\mathcal{G}(n, \frac{1}{2})) \geq (2 + o(1)) \log n$.
Wegen $\chi(G) \cdot \alpha(G) \geq n$ erhalten wir $\chi(\mathcal{G}(n, \frac{1}{2})) \geq \frac{n}{(2+o(1)) \log n}$.

Satz 1.13. *Es gilt* $\lim_{n\to\infty} \mathbf{P}\left(\frac{n}{(2+o(1)) \log n} \leq \chi(\mathcal{G}(n, \frac{1}{2})) \leq \frac{n}{(1-o(1)) \log n}\right) = 1.$

Beweis:
Wegen der vorherigen Bemerkung reicht der Beweis der oberen Schranke.
Als Idee für diesen Teil des Beweises verwenden wir die Johnson-Heuristik, d.h. wir suchen große stabile Mengen als Farbklassen und entfernen diese aus dem Graphen.

Färbung „Johnson“-Greedy-Coloring($G = (V, E)$){
  $k = 0$; $n = |V|$; $G' = G$;
  while($|V(G')| \geq \frac{n}{(\log n)^2}$){
    $S = \emptyset$;
    $H = G'$;
    while($H \neq \emptyset$){
      $x = \text{select}(\{V(H)|d(x) \text{ minimal}\})$;
      $S = S \cup \{x\}$;
      $H = H \setminus \{x\} \cup N(x)$;
    }
    $G' = G' \setminus S$;
    $k{+}{+}$;
    color(S)= k;
  }
  color(G')=jeder Knoten anders;
  return color;
}

Wir zeigen nun, dass fast sicher in jedem Durchlauf der inneren while-Schleife eine stabile Menge mindestens der Größe $(1-o(1)) \log n$ gefunden wird (d.h. nicht nur in jedem Durchlauf *fast sicher*, sondern fast sicher *in allen* Durchläufen). Dann gilt nämlich fast sicher:
$\chi(\mathcal{G}(n, \frac{1}{2})) \leq \frac{n - \frac{n}{(\log n)^2}}{(1-o(1)) \log n} + \frac{n}{(\log n)^2} \leq \frac{n}{(1-o(1)) \log n} - \frac{n}{(\log n)^3} + \frac{n}{(\log n)^2} = \frac{n}{(1-o(1)) \log n}$
Zeigen wir also die obige Behauptung:
Wir betrachten einen beliebigen Durchlauf der inneren while-Schleife. Sei $v(G') = r \geq \frac{n}{(\log n)^2}$. Für ein beliebiges $\varepsilon > 0$ und ein S (nach Abschluss eines while-Schleifendurchlaufs) bestimmen wir $\mathbf{P}(|S| \leq (1-\varepsilon) \log n)$. S ist maximal in G', d.h. $\forall v \in V(G') \setminus S : N(v) \cap S \neq \emptyset$.
Sei $v \in V(G') \setminus S$ beliebig. Dann gilt $\mathbf{P}(v \text{ erweitert } S \text{ nicht}) = 1 - 2^{-|S|}$.
Sei S eine stabile Menge. Dann gilt $|V(G') \setminus S| = r - |S|$.

$\Rightarrow \mathbf{P}(S \text{ ist maximal in } G') = (1-2^{-|S|})^{r-|S|}$.
$\Rightarrow p = \mathbf{P}(\text{Algorithmus findet maximales stabiles } S \text{ mit } |S| \leq (1-\varepsilon)\log n)$
$\quad \leq \binom{r}{|S|}\mathbf{P}(S \text{ ist maximal}) \leq \binom{r}{|S|}(1-2^{-|S|})^{r-|S|}$
Sei $|S| = s = (1-\varepsilon)\log n$.
$\Rightarrow p \leq r^s(1-2^{(\varepsilon-1)\log n})^{r-(1-\varepsilon)\log n}$ mit $r-(1-\varepsilon)\log n \geq \frac{r}{2}$
$\Rightarrow p \leq r^s(1-2^{(\varepsilon-1)\log n})^{\frac{r}{2}}$ wegen $(1-2^{(\varepsilon-1)\log n}) < 1$
$\Rightarrow p \leq r^s(1-n^{\varepsilon-1})^{\frac{r}{2}} \leq r^s(1-r^{\varepsilon-1})^{\frac{r}{2}} \leq r^{\log r}e^{-r^{\varepsilon-1}\cdot\frac{r}{2}}$
$\Rightarrow p \leq e^{\ln r\cdot\log r-\frac{r^\varepsilon}{2}}$
Es gilt $r^\varepsilon \gg \ln r\log r$ für alle $\varepsilon > 0$. Wegen $n \geq r \geq \frac{n}{(\log n)^2}$ gilt $\log r = (1\pm o(1))\log n$.
$\Rightarrow \mathbf{P}(\exists \text{ Schleifendurchlauf mit } |S| < (1-\varepsilon)\log n) \leq ne^{\ln r\cdot\log r-\frac{r^\varepsilon}{2}} = e^{\ln n+\ln r\log r-\frac{r^\varepsilon}{2}}$
$\quad \leq e^{\ln n+\ln n\log n-\frac{n^\varepsilon}{2(\log n)^{2\varepsilon}}} \xrightarrow{n\to\infty} 0$ □

Korollar 1.4. *Es existiert ein polynomieller Algorithmus, welcher die chromatische Zahl mit hoher Wahrscheinlichkeit (d.h. für fast alle Graphen) bis auf den Faktor* 2 *genau approximiert.*

Beweis:
Hmm ... naja ... trivial?
Schließlich haben wir ja oben einen Algorithmus im Beweis verwendet. □

Vorlesung am 16.05.2008

1.8 Chernoff

Definition (Unabhängigkeit zweier Ereignisse)
Zwei Ereignisse A und B sind *unabhängig*, falls $P(A\cap B) = P(A)\cdot P(B)$ gilt.

Beispiel
Wir würfeln zwei mal. Die Ereignisse „Erster Wurf ist 3“ und „Summe ist 8“ sind abhängig. Die Ereignisse „Erster Wurf ist 6“ und „Summe ist 7“ sind unabhängig.

Bemerkung
Die Unabhängigkeit von Ereignissen haben wir bereits im $\mathcal{G}(n,p)$-Modell ausgenutzt, wo die Existenz von verschiedenen Kanten unabhängig ist.

Definition (Unabhängigkeit von Zufallsvariablen)
Zufallsvariablen $X_1,...,X_n$ heißen *unabhängig*, falls für alle $x_1,...,x_n$ gilt:
$P(X_1 = x_1,...,X_n = x_n) = \prod_{i=1}^n P(X_i = x_i)$.

Bemerkung (Aus paarweiser Unabhängigkeit folgt nicht Unabhängigkeit)
Wenn X_1 und X_2, X_1 und X_3 sowie X_2 und X_3 unabhängig sind, so sind X_1, X_2 und X_3 nicht notwendigerweise insgesamt unabhängig.

Beispiel (Aus paarweiser Unabhängigkeit folgt nicht Unabhängigkeit)
Wir betrachten die Ereignisse „Erster Wurf ist 1“, „Zweiter Wurf ist 1“ und „Summe ist 7“.

Definition (unabhängigkeit mehrerer Ereignisse, k-fache Unabhängigkeit)
Ereignisse $A_1, ..., A_n$ sind unabhängig, falls für alle $I \subseteq [n]$ die Gleichung $P(\bigcap_{i\in I} A_i) = \prod_{i\in I} P(A_i)$ gilt.
Die Ereignisse sind *k-fach unabhängig*, falls wir nur Indexmengen mit $|I| = k$ betrachten.

Lemma 1.2. *Seien $X_1, ..., X_n$ unabhängige Zufallsvariablen. Dann gilt* $\mathbf{E}(\prod_{i=1}^n X_i) = \prod_{i=1}^n \mathbf{E}(X_i)$.

Beweis:
Wir zeigen die Aussage nur für $n = 2$. Für $n > 2$ erfolgt der Beweis analog.
Seien X und Y unabhängig. Dann gilt:
$$\begin{aligned}\mathbf{E}(XY) &= \sum_x \sum_y P(X = x, Y = y)xy = \sum_x \sum_y P(X = x)P(Y = y)xy \\ &= \sum_x P(X = x)x \sum_y P(Y = y)y = \mathbf{E}X\mathbf{E}Y\end{aligned}$$
□

Lemma 1.3. *Wenn X und Y zwei unabhängige Zufallsvariablen sind, so gilt $Cov(X, Y) = 0$. Weiterhin gilt* $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Beweis:
Rechnerei. □

Bemerkung
Bisher haben wir die Abschätzung $\mathbf{P}(|X - \mathbf{E}X| \geq t) \leq \frac{\mathrm{Var}\,X}{t^2}$. Diese Schranke ist gut, wenn $t > \sqrt{\mathrm{Var}\,X}$ gilt.
Wir wollen nun eine Schranke der Form $\mathbf{P}(|X - \mathbf{E}X| \geq \varepsilon\mathbf{E}X) \leq e^{-\Omega(\varepsilon)\mathbf{E}X}$ finden. Dies ist allgemein nicht möglich, da die Tschebyscheff-Ungleichung scharf ist. Wir suchen diese Schranke daher nur für spezielle X.

Beschränkung der Abweichung vom Erwartungswert
Wie gleichmäßig sind die Kanten in $\mathcal{G}(n, p)$ verteilt?
Mit welcher Wahrscheinlichkeit gilt $e(U) \approx p\binom{U}{2}$ für alle (nicht zu kleinen!) U?
Sei X_U die Anzahl der Kanten in U. Wir wissen, dass $\mathbf{E}X_U = p\binom{|U|}{2}$ gilt.
Die obige Frage ist also äquivalent zu der Frage nach der Wahrscheinlichkeit, dass ein U mit $|X_U - \mathbf{E}X_U| > \varepsilon\mathbf{E}X_U$ existiert.
Wir wissen, dass $\mathbf{P}(\exists U : |X_U - \mathbf{E}X_U| > \varepsilon\mathbf{E}X_U) \leq 2^n \max \mathbf{P}(|X_U - \mathbf{E}X_U| > \varepsilon\mathbf{E}X_U)$ gilt.
Hoffnung: $\mathbf{P}(|X_U - \mathbf{E}X_U| > \varepsilon\mathbf{E}X_U) \ll 2^{-n}$ für geeignet große U.
$\mathbf{P}(X \geq \lambda) = \mathbf{P}(e^{tX} \geq e^{t\lambda}) \leq \frac{\mathbf{E}e^{tX}}{e^{t\lambda}}$
Analog gilt $\mathbf{P}(X \leq -\lambda) = \mathbf{P}(-tX \geq t\lambda) = \mathbf{P}(e^{-tX} \geq e^{t\lambda}) \leq \frac{\mathbf{E}e^{-tX}}{e^{t\lambda}}$.
Beobachtung: $\mathbf{E}[e^{tX}] = \sum_{n=0}^{\infty} \frac{t^n\mathbf{E}[X^n]}{n!}$.

Lemma 1.4. *Sei X eine Zufallsvariable mit $|X| \leq 1$ und $\mathbf{E}X = 0$.
Dann gilt $\mathbf{E}[e^{tX}] \leq e^{t^2\,\mathrm{Var}\,X}$ für alle $t \in [-1, 1]$.*

Beweis:
Es gilt $|tX| \leq 1$, d.h. $e^{tX} \leq 1 + tX + t^2X^2$, wegen $\sum_{n=3}^{\infty} \frac{1}{n!} \leq \frac{1}{2}$.
$\Rightarrow \mathbf{E}[e^{tX}] \leq 1 + t\mathbf{E}X + t^2\mathbf{E}[X^2] = 1 + t^2\,\mathrm{Var}\,X$ wegen $\mathbf{E}X = 0$
$\Rightarrow \mathbf{E}[e^{tX}] \leq 1 + t^2\,\mathrm{Var}\,X \leq e^{t^2\,\mathrm{Var}\,X}$ □

Satz 1.14. (Chernoff) *Seien $X_1, ..., X_n$ unabhängige Zufallsvariablen mit $|X_i - \mathbf{E}X_i| \leq 1$ für alle $i = 1, ..., n$, $X = \sum_{i=1}^n X_i$ und $\sigma = \sqrt{\mathrm{Var}\,X}$.
Für alle $\lambda > 0$ gilt* $\mathbf{P}(|X - \mathbf{E}X| \geq \lambda\sigma) \leq 2\max\{e^{-\frac{\lambda^2}{4}}, e^{-\frac{\lambda\sigma}{2}}\}$.

Beweis:
O.B.d.A. sei $\mathbf{E}X_i = 0$.
Dann gilt $\mathbf{P}(|X| \geq \lambda\sigma) = \mathbf{P}(X \geq \lambda\sigma) + \mathbf{P}(X \leq -\lambda\sigma)$.
Zu zeigen ist $\mathbf{P}(X \geq \lambda\sigma) \leq e^{-\frac{t\lambda\sigma}{2}}$ mit $t = \min\{\frac{\lambda}{2\sigma}, 1\}$. Der andere Fall wird analog gezeigt.

$\mathbf{P}(X \geq \lambda\sigma) \leq \mathbf{P}(e^{tX} \geq e^{t\lambda}) \leq \frac{\mathbf{E}e^{tX}}{e^{t\lambda}}$.
Wir haben $\mathbf{E}e^{tX} = \mathbf{E}\left[e^{t\sum_{i=1}^n X_i}\right] = \mathbf{E}\prod_{i=1}^n e^{tX_i}$. Weil mit $X_1, ..., X_n$ auch $e^{tX_1}, ..., e^{tX_n}$ unabhängig sind, gilt:
$\mathbf{E}e^{tX} = \prod_{i=1}^n \mathbf{E}[e^{tX_i}] \leq \prod_{i=1}^n e^{t^2 \operatorname{Var} X_i} = e^{t^2 \sum_{i=1}^n \operatorname{Var} X_i} = e^{t^2 \operatorname{Var} X} = e^{(t\sigma)^2}$
$\Rightarrow \mathbf{P}(X \geq \lambda\sigma) \leq e^{t^2\sigma^2 - t\lambda\sigma}$
$t \leq \frac{\lambda}{2\sigma} \Rightarrow t^2\sigma^2 \leq \frac{t\lambda\sigma}{2} \Rightarrow \mathbf{P}(X \geq \lambda\sigma) \leq e^{-\frac{t\lambda\sigma}{2}}$ □

Korollar 1.5.
Seien $X_1, ..., X_n$ unabhängige Zufallsvariablen mit $\mathbf{P}(X_i = 1) = p_i$ und $\mathbf{P}(x_i = 0) = (1 - p_i)$. Weiterhin sei $X = \sum_{i=1}^n X_i$.
Dann gilt $\mathbf{P}(|X - \mathbf{E}X| \geq \varepsilon\mathbf{E}X) \leq 2e^{-\min\{\frac{\varepsilon^2}{4}, \frac{\varepsilon}{2}\}\mathbf{E}X}$.

Beweis:
Wir setzen $\lambda = \frac{\varepsilon\mathbf{E}X}{\sigma}$ und $\sigma = \sqrt{\sum_{i=1}^n \operatorname{Var} X_i}$. Wegen $\operatorname{Var} X_i \leq \mathbf{E}X_i$ und $\sigma \leq \sqrt{\mathbf{E}X}$ folgt die Ungleichung. □

Vorlesung am 21.05.2008

Kapitel 2

Extremale Graphentheorie

2.1 Einführung

Keine gute Frage
Was ist das?
Gegeben ist eine Graphenklasse $\mathcal{G}$, z.B. die Menge aller dreiecksfreien Graphen, und ein Graphenparameter τ, z.B. die Anzahl der Kanten, oder die Anzahl bestimmter Untergraphen, etc....
Gesucht ist das Minimum und/oder Maximum von $\tau(G)$, sodass $G \in \mathcal{G}, |V(G)| = n$.

Beispiel (kein gutes Beispiel)
Sei $\mathcal{G}_{ham}$ die Klasse der hamiltonischen Graphen und $\tau(G) = \delta(G)$.
Dann gilt $\max \delta(G) = n - 1$.

Bessere Frage
Welche Bedingung an τ impliziert $G \in \mathcal{G}$ (bzw. $G \notin \mathcal{G}$)?

Beispiel (besseres Beispiel)
Es gilt $\delta(G) \geq \frac{n}{2} \Rightarrow G \in \mathcal{G}_{ham}$.

Definition ($Forb$)
Für einen Graphen F sei $Forb(F) = \{G | F \not\subseteq G\}$.
Weiterhin sei $Forb_n(F) = \{G \in Forb(F) | v(G) = n\}$.

Definition
Sei $n \in \mathbb{N}$ und F ein Graph mit $e(F) \geq 1$.
Wir definieren $ex(n, F) = \max\{e(G) | G \in Forb(F), v(G) = n\}$.

Beispiel

1. $F = K_2 \Rightarrow ex(n, F) = 0$

2. $F = K_3 \Rightarrow ex(n, F) \geq (\frac{n}{2})^2$

3. $F = K_4 \Rightarrow ex(n, F) \geq 3(\frac{n}{3})^2$

4. $F = K_k \Rightarrow ex(n, F) \geq \binom{k-1}{2}(\frac{n}{k-1})^2 = (1 - \frac{1}{k-1})\frac{n^2}{2}$

Definition (Turán-Graph)
Zu $n, k \in \mathbb{N}$ definieren wir $T_{n,k}$ als den vollständigen k-partiten Graphen mit Knotenklassen der Größe $\lfloor \frac{n}{k} \rfloor$ oder $\lceil \frac{n}{k} \rceil$.

Satz 2.1. (Turán (1943)) *Es gilt* $\forall k, n \in \mathbb{N} : ex(n, K_k) = e(T_{n,k-1})$. *Der Turán-Graph ist der einzige extremale Graph, d.h.* $e(H) < e(T_{n,k})$ *für alle* $H \in Forb_n(K_k) \setminus \{T_{n,k-1}\}$.

Beweis:
Sei $G = (V, E) \in Forb_n(K_k)$ mit maximaler Kantenzahl und $H = \{x_1, ..., x_{k-1}\}$ spanne einen vollständigen Graphen K_{k-1}. (Weil G mit maximaler Kantenzahl gewählt wurde, muss G einen K_{k-1} enthalten.) Es gilt $e(G) \leq A + B + C$ mit $A = e(G \setminus H), B = e(G[H]) = \binom{k-1}{2}$ und $C = e(H, V \setminus H) \leq (n-k+1)(k-2)$ (jeder Knoten aus $G \setminus H$ hat höchstens $k-2$ Nachbarn in H).
$\Rightarrow e(G) \leq e(T_{n-k+1,k-1}) + \binom{k-1}{2} + (n-k+1)(k-2) = e(T_{n,k-1})$
Beweis der Eindeutigkeit: Übungsaufgabe. □

Frage
Was ist $ex(n, F)$ für $F \neq K_k$?
$ex(n, F) \geq e(T_{n,k})$, $k = \chi(F) - 1$

Antwort
$ex(n, F) = (1 - \frac{1}{\chi(F)-1} + o(1))\binom{n}{2} = (1 + o(1))e(T_{n,k}), k = \chi(F) - 1$ löst das Problem für Graphen F mit $\chi(F) \geq 3$.
Für $\chi(F) = 2$ haben wir das folgende Lemma:

Lemma 2.1. (Kövari, Turán, Sós) $\chi(F) = 2 \Rightarrow ex(n, F) = o(n^2)$

Beweis:
Wir zeigen $\forall t \in \mathbb{N}, c > 0 : \exists n_0$: Falls $e(G) \geq cn^2, |V(G)| = n$, dann $T_{2t,2} = K_{t,t} \subseteq G$.

1. Schritt: $\forall c, t : \exists c', n_0 : \#\{K_{1,t} \subseteq G\} \geq c'n^{t+1}$

   Beweis:

   $$\begin{aligned}\#\{K_{1,t} \leq G\} &= \textstyle\sum_{v \in V(G)} \binom{d(v)}{t} \geq n \left(\begin{matrix} \frac{\sum_{v \in V(G)} d(v)}{n} \\ t \end{matrix} \right) \qquad \text{(Jensen)} \\ &= n\binom{2e(G)}{t} \geq n\binom{cn^2}{t} \geq nc'n^t = c'n^{t+1}\end{aligned}$$

2. Schritt: $\forall c', t : \exists c'', n_0 : \#\{K_{1,t} \subseteq G\} \geq c'n^{t+1} \Rightarrow \#\{K_{t,t} \subseteq G\} \geq c''n^{2t}$

   Beweis:

   $\#\{K_{t,t} \subseteq G\} = \sum_{T \in \binom{V(G)}{t}} \binom{\deg(T)}{t}$, wobei $\deg(T) = \{x \in V(G) \setminus T | (x, T)$ ist $K_{1,t}$ in G mit Zentrum $x\}$.

   Nach Jensen folgt $\#\{K_{t,t} \subseteq G\} \geq \binom{n}{t} \cdot \left(\begin{matrix} \frac{\sum \deg(T)}{\binom{n}{t}} \\ t \end{matrix} \right)$ mit $\frac{\sum \deg(T)}{\binom{n}{t}} = \#\{K_{1,t} \subseteq G\} \geq c'n^{t+1}$.

   $\Rightarrow \#\{K_{t,t} \subseteq G\} \geq c_0 n^t c_2 (\frac{c'n^{t+1}}{c_1 n^t})^t \geq c''n^{2t}$

Vorlesung am 23.05.2008

Lemma 2.2. *Für alle Graphen* H *existiert der Grenzwert* $\lim_{n\to\infty} \frac{ex(n,H)}{\binom{n}{2}}$.

Beweis:
Wir zeigen, dass die Folge $a_n = \frac{ex(n,H)}{\binom{n}{2}}$ monoton (fallend) und (nach unten) beschränkt ist.
Es gilt offensichtlich $\frac{ex(n,H)}{\binom{n}{2}} \geq 0$, womit eine untere Schranke gefunden ist.
Sei $G \in Forb_n(H)$ extremal (d.h. mit $ex(n, H) = e(G)$).
Wir betrachten $G_1, ..., G_n$, wobei $G_i = G \setminus \{v_i\}$ und $V(G) = \{v_1, ..., v_n\}$. Weil jede Kante in den Graphen $G_1, ..., G_n$ genau zweimal nicht vorkommt, gilt $\sum_{i=1}^{n} e(G_i) = (n-2)ex(n, H)$. Weiterhin enthält kein $G_i \subseteq G$ den Graphen H, d.h. $ex(n-1, H) \geq e(G_i)$.
$\Rightarrow ex(n-1, H)n \geq \sum_{i=1}^{n} e(G_i) = ex(n, H)(n-2)$
$\Rightarrow \frac{ex(n,H)}{n} \leq \frac{ex(n-1,H)}{n-2}$
$\Rightarrow \frac{ex(n,H)}{\binom{n}{2}} = \frac{ex(n,H)}{\frac{n(n-1)}{2}} \leq \frac{ex(n-1,H)}{\frac{(n-2)(n-1)}{2}} = \frac{ex(n-1,H)}{\binom{n-1}{2}}$ □

Definition
Für alle Graphen F definieren wir $\pi(F) = \lim_{n\to\infty} \frac{ex(n,F)}{\binom{n}{2}}$. (Mit dem vorhergehenden Lemma existiert dieser Grenzwert.)

Was ist $\pi(F)$?
Wir wissen bereits $\pi(K_k) = \lim_{n\to\infty} \frac{e(T_{n,k-1})}{\binom{n}{2}} = \frac{k-2}{k-1} = 1 - \frac{1}{k-1}$ (Turán).
Weiterhin wissen wir $\pi(F) = 0$, wenn F bipartit ist (Kövar, Turán, Sós).
Offensichtlich gilt $\pi(F) \geq 1 - \frac{1}{\chi(F)-1}$, da $F \not\subseteq T_{n,\chi(F)-1}$ ($T_{n,\chi(F)-1}$ und alle seine Untergraphen sind $\chi(F) - 1$-färbbar, F selbst jedoch nicht).
Wir zeigen nun, dass an dieser Stelle sogar die Gleichheit gilt, d.h. $\pi(F) = 1 - \frac{1}{\chi(F)-1}$. Dies ist der folgende Satz von Erdős-Stone.

Lemma 2.3. (Erdős, Stone (1947)) *Für alle $\varepsilon > 0$ und $k, t \in \mathbb{N}$ existiert ein $n_0 \in \mathbb{N}$ sodass: Falls G ein Graph auf $n \geq n_0$ Knoten mit $\delta(G) \geq (1 - \frac{1}{k-1} + \varepsilon)n$ ist, dann enthält G eine Kopie von $K_k(t) = T_{tk,k}$, d.h. dem vollständigen k-partiten Graphen mit Knotenklassengröße t.*

Beweis:
Wir führen eine Induktion über k durch.

1. Induktionsanfang $k = 2$:

   Dies folgt aus dem Satz von Kövari, Turán und Sós, da $\delta(G) \geq \varepsilon n \Rightarrow e(G) \geq \varepsilon \frac{n^2}{2} \geq \varepsilon \binom{n}{2}$.

2. Induktionsschritt $k \to k+1, k \geq 2$:

   Wir setzen $s = \lceil \frac{t}{\varepsilon} \rceil$. Für genügend großes n gilt $K_{k-1}(s) \subseteq G$, falls $\delta(G) \geq (1 - \frac{1}{k-2} + \varepsilon)n$. Sei also G ein Graph mit n Knoten und $\delta(G) \geq (1 - \frac{1}{k-1} + \varepsilon)n > (1 - \frac{1}{k-2} + \varepsilon)n$.

   $\Rightarrow$ Es existieren disjunkte Teilmengen $U_1, ..., U_{k-1} \subseteq V(G)$, welche einen $K_{k-2}(s)$ aufspannen (Induktionsvoraussetzung).

   Wir setzen $\overline{W} = \bigcup_{i=1}^{k-1} U_i, W = V(G) \setminus \overline{W}$ und $X = \{w \in W \big| \forall i = 1, ..., k-1 : |N(w) \cap U_i| \geq t\}$. Zwischen W und $\overline{W}$ fehlen mindestens $|W \setminus X|(s-t)$ Kanten, da für alle Knoten $w \in W \setminus X$ ein i existiert, sodass $|N(w) \cap U_i| < t$.

   $\Rightarrow |W \setminus X|(s-t) \geq |W \setminus X|(1-\varepsilon)s = (n - (k-1)s - |X|)(1-\varepsilon)s$

   Andererseits fehlen jedem Knoten in G auf Grund der Minimalgradeigenschaft höchstens $(\frac{1}{k-1} - \varepsilon)n$ Kanten.

   Wegen $|\overline{W}| = (k-1)s$, erhalten wir als Anzahl der fehlenden Kanten zwischen den Knotenmengen W und $\overline{W}$ höchstens $(\frac{1}{k-1} - \varepsilon)n(k-1)s = (1 - \varepsilon(k-1))sn$.

   $\Rightarrow (n - (k-1)s - |X|)(1-\varepsilon)s \leq e_{\overline{G}}(W, \overline{W}) \leq (1 - \varepsilon(k-1))sn$

   $\Rightarrow n - (k-1)s - |X| - \varepsilon n + (k-1)s\varepsilon + |X|\varepsilon \leq n - \varepsilon(k-1)n$

$\Rightarrow -(k-1)s - |X| - \varepsilon n + (k-1)s\varepsilon + |X|\varepsilon \leq -\varepsilon(k-1)n$
$\Rightarrow -\varepsilon n + (k-1)s(\varepsilon - 1) + |X|(\varepsilon - 1) \leq -\varepsilon(k-1)n$
$\Rightarrow (k-1)s(\varepsilon-1) + |X|(\varepsilon-1) \leq -\varepsilon(k-1)n + \varepsilon n = (2-k)\varepsilon n$
$\Rightarrow |X|(\varepsilon-1) \leq (2-k)\varepsilon n - (k-1)s(\varepsilon-1)$
$\Rightarrow |X| \geq \frac{(2-k)\varepsilon n}{\varepsilon-1} - (k-1)s$
$\Rightarrow |X| \geq \frac{(k-2)\varepsilon}{1-\varepsilon}n - (k-1)s$
$\Rightarrow |X| \geq (t-1)\binom{s}{t}^{k-1} + 1$ für ausreichend großes n
$\Rightarrow \exists x_1, ..., x_t : \forall i_1, i_2, j \in \{1, ..., t\} : \forall j \in \{1, ..., k-1\} : N(x_{i_1}) \cap U_j = N(x_{i_2}) \cap U_j$
$\Rightarrow$ Es gibt t Knoten in X, welche das gleiche Nachbarschaftsmuster in $U_1, ..., U_{k-1}$ haben.
$\Rightarrow U_1 \cap N(x_1), ..., U_{k-1} \cap N(x_1), U_k = \{x_1, ..., x_t\}$ spannen $K_k(t)$ auf. □

Satz 2.2. $\forall k > 2, t \geq 1 : \pi(K_k(t)) = 1 - \frac{1}{k-1}$

Beweis:
Es seien $\varepsilon > 0$ beliebig (aber klein), n ausreichend groß und G ein Graph mit n Knoten und $e(G) \leq (1 - \frac{1}{k-1} + \varepsilon)\frac{n^2}{2}$. Wir entfernen iterativ alle Knoten aus G, die weniger als $(1 - \frac{1}{k-1} + \frac{\varepsilon}{2})n$ Nachbarn haben. Damit erhalten wir den Graphen H mit $\delta(H) \geq (1 - \frac{1}{k-1} + \frac{\varepsilon}{2})v(H)$.
Falls $V(H)$ groß ist, dann gilt mit dem Lemma $K_{k-1}(t) \subseteq H \subseteq G$.
Es stellt sich also die Frage, wie viele Kanten bei der Konstruktion von H entfernt werden, bzw. wie groß $e(G) - e(H)$ ist.
Mit $v(H) = N$ erhalten wir

$$\begin{aligned}
e(G) - e(H) &\leq \textstyle\sum_{i=N+1}^{n} i(1 - \frac{1}{k-1} + \frac{\varepsilon}{2}) = \left(\binom{n+1}{2} - \binom{N+1}{2}\right)(1 - \frac{1}{k-1} + \frac{\varepsilon}{2}) \\
&= \left(\binom{n}{2} + n - \binom{n}{2} - N\right)(1 - \frac{1}{k-1} + \frac{\varepsilon}{2}) \\
&\leq \left(\binom{n}{2} - \binom{n}{2}\right)(1 - \frac{1}{k-1} + \frac{\varepsilon}{2}) + n - N
\end{aligned}$$

$\Rightarrow (1 - \frac{1}{k-1} + \varepsilon)\binom{n}{2} \leq e(G) \leq (1 - \frac{1}{k-1} + \frac{\varepsilon}{2})\left(\binom{n}{2} - \binom{n}{2}\right) + n - N + \binom{n}{2}$
$\Rightarrow \frac{\varepsilon}{2}\binom{n}{2} \leq (\frac{1}{k-1} - \frac{\varepsilon}{2})\binom{n}{2} + n - N$
$\Rightarrow \frac{\varepsilon}{2}\binom{n}{2} - n \leq (\frac{1}{k-1} - \frac{\varepsilon}{2})\binom{n}{2} - N$
$\Rightarrow N \overset{n\to\infty}{\longrightarrow} \infty$
$\Rightarrow N$ kann beliebig große Werte annehmen. □

Korollar 2.1. *Für alle Graphen F gilt $\pi(F) = 1 - \frac{1}{\chi(F)-1}$.*

Beweis:
Wir setzen im obigen Satz $k = \chi(F)$. □

Vorlesung am 28.05.2008

Definition ((ε, d)-Regularität, ε-Regularität)
Seien $G = (V, E), X, Y \subseteq V, X \cap Y = \emptyset, d \geq 0$ und $\varepsilon > 0$.
(X, Y) heißt (ε, d)-*regulär*, falls $\forall X' \subseteq X, Y' \subseteq Y$ mit $|X'| \geq \varepsilon|X|, |Y'| \geq \varepsilon|Y|$ gilt:
$|d(X', Y') - d| \leq \varepsilon$, wobei $d(X', Y') = \frac{e(X', Y')}{|X'||Y'|}$.
(X, Y) ist ε-*regulär*, falls es ein $d \geq 0$ gibt, sodass (X, Y) (ε, d)-regulär ist.

Lemma 2.4. (Counting Lemma) *Für alle $F, \gamma > 0$ und $d_0 > 0$ existieren $\varepsilon > 0$ und m_0 sodass gilt:*
Falls $G = (V_1, ..., V_l, E)$ ein l-partiter Graph ist, $V(F) = \{1, ..., l\}$, für alle $\{i, j\} \in E(F)$ ist (V_i, V_j) (ε, d_{ij}) regulär mit $d_{i,j} \geq d_0$ und $\forall i \in [l] : |V_i'| = m \geq m_0$ dann gilt:

$$|\{f : V(F) \to V(G) | \forall i \in [l] : f(i) \in V_i, \forall \{i, j\} \in E(F) : \{f(i), f(j)\} \in E\}| = (1 \pm \gamma)m^l \prod_{\{i,j\} \in E(F)} d_{i,j}$$

Beweis:
Beweis durch Induktion über $e(F)$:

1. $e(F) = 0 \Rightarrow |\{F \subseteq_P G\}| = \prod_{i=1}^{l} V_i = m^l = \prod_{\{i,j\}\in E(F)} d_{i,j} m^l = m^l$

2. Gegeben seien F, γ und d_0.

   Setze $\gamma' = \frac{\gamma}{2}$.

   Seien ε', m_0' gegeben durch das Counting Lemma für F^-, γ' und d_0.

   Setze $\varepsilon = \min\{\varepsilon', \gamma' d_0^{e(F)}\}$ und $m_0 = m_0'$.

   Sei $F^- = F \setminus \{f\}$, wobei $f \in E(F)$ beliebig.

   $\Rightarrow |\{F \subseteq_P G\}| = \sum_{F' \in \{F^- \subseteq_P G\}} \mathbf{1}_{E(G)}(\eta(F'))$ mit $\eta(F') = \{v_1, v_2\}$, $v_1 \in V_1 \cap V(F)$ und $v_2 \in V_2 \cap V(F)$

   $$\mathbf{1}_{E(G)}(\eta(F')) = \begin{cases} 0, \text{wenn} \eta(F') \notin E(G) \\ 1, \text{wenn} \eta(F') \in E(G) \end{cases}$$

   $\Rightarrow |\{F \subseteq_P G\}| = \sum_{F' \in \{F^- \subseteq_P G\}} (d_{1,2} + \mathbf{1}_{E(G)}(\eta(F')) - d_{1,2})$

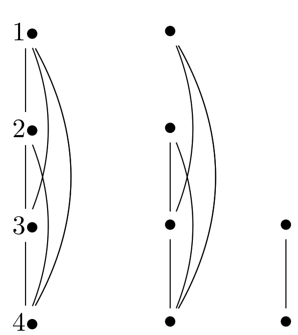
   $$\Rightarrow |\{F \subseteq_P G\}| = \underbrace{\sum_{F' \in \{F^- \subseteq_P G\}} d_{1,2}}_{=A} + \underbrace{\sum_{F' \in \{F^- \subseteq_P G\}} (\mathbf{1}_{E(G)}(\eta(F')) - d_{1,2})}_{=B}$$

   $A = d_{1,2} |\{F^- \subseteq_P G\}| \overset{I.V.}{=} d_{1,2}(1 \pm \gamma') \prod_{\{i,j\}\in E(F^-)} d_{i,j} m^l = d_{1,2}(1 \pm \gamma') \prod_{\{i,j\}\in E(F^-)} d_{i,j} m^l$

   Mit $F^* = F - \{1, 2\}$ (Die Knoten werden entfernt, nicht nur die Kanten zwischen ihnen.) erhalten wir $B = \sum_{F'' \in \{F^* \subseteq_P G \setminus \{V_1 \cup V_2\}\}} \sum_{(v_1,v_2) \in (V_1 \times V_2 \cap EXT(F''))} \mathbf{1}_{E(G)}(\{v_1, v_2\}) - d_{1,2}$, $EXT(F'') = \{(v_1, v_2) \in V_1 \times V_2 | v_1 \in \bigcap_{\{1,i\}\in E(F^-)} N(V_i) \cap V_1, v_2 \in \bigcap_{\{2,i\}\in E(F^-)} N(V_i) \cap V_2$ und $V(F'') = \{v_3, ..., v_l\}, v_i \in V_i\}$.

   Beispiel:

   $F = K_4$ F^- F^*

   $$|B| \le \sum_{F'' \in \{...\}} \underbrace{\left| \sum_{(v_1,v_2) \in EXT(F'')} \mathbf{1}_{E(G)}(\{v_1, v_2\}) - d_{1,2} \right|}_{\le \varepsilon m^2}:$$

   $\left| \sum_{(v_1,v_2) \in EXT(F'')} \mathbf{1}_{E(G)}(\{v_1, v_2\}) - d_{1,2} \right|$

$$= \left| \left\{ e \in G \left[\underbrace{\bigcap_{\{1,i\} \in E(F^-)} N(v_i) \cap V_1}_{=X}, \underbrace{\bigcap_{\{2,i\} \in E(F^-)} N(v_i) \cap V_2}_{=Y} \right] \right\} - d_{1,2}|X||Y| \right|$$
$$= |e(X,Y) - d_{1,2}|X||Y||$$

(a) 1. Fall: $|X|, |Y| \geq \varepsilon m$
Aus der $(\varepsilon, d_{1,2})$-regularität von (V_1, V_2) folgt $|\frac{e(X,Y)}{|X||Y|} - \frac{d_{1,2}|X||Y|}{|X||Y|}| \leq \varepsilon$
$\Rightarrow |e(X,Y) - d_{1,2}|X||Y|| \leq \varepsilon|X||Y| \leq \varepsilon m^2$

(b) 2. Fall: Sei $|X| \leq \varepsilon m$.
$\Rightarrow |X||Y| \leq \varepsilon m^2$
$\Rightarrow |e(X,Y) - d_{1,2}|X||Y|| \leq |X||Y| \leq \varepsilon m^2$
(Die erste Ungleichung gilt für alle Graphen!)

$|\{F^* \subseteq_P G \setminus \{v_1, v_2\}\}| \leq m^{l-2}$
$\Rightarrow |\{F \subseteq_P G\} = A \pm B = (1 \pm \gamma') \prod_{\{i,j\} \in E(F)} d_{i,j} m^l \pm \varepsilon m^l$
Falls $\varepsilon \leq \gamma' d_0^{e(F)}$, dann $|\{F \subseteq_P G\}| = (1 \pm 2\gamma') \prod_{\{i,j\} \in E(F)} d_{i,j} m^l = (1 \pm \gamma) \prod_{\{i,j\} \in E(F)} d_{i,j} m^l$.

□

Lemma 2.5. $\forall \gamma > 0, d > 0 \exists \varepsilon, m_0$:
Falls $G = (X, Y, Z, E)$ ein tripartiter Graph ist, sodass (X, Y) (ε, d_{XY})-regulär ist, (X, Z) ist (ε, d_{XZ})-regulär, (Y, Z) ist (ε, d_{YZ})-regulär und $d_{XY}, d_{XZ}, d_{YZ} \geq d$.
$\Rightarrow \#\{K_3 \subseteq G\}(1 \pm \gamma) d_{XY} d_{XZ} d_{YZ} |X||Y||Z|$

Beweis:
Kein Beweis. □

Lemma 2.6. (Regularitätslemma (Szenerédi: 1978))
$\forall \varepsilon > 0, t_0 \in \mathbb{N} : \exists T_0, n_0 : \forall G = (V, E), |V| \geq n_0 : \exists V_0 \dot\cup ... \dot\cup V_t = V$:

1. $t_0 \leq t \leq T_0$
2. $|V_0| \leq \varepsilon n$, $|V_1| = ... = |V_t|$
3. $|\{\{i, j\} | 1 \leq i < j \leq t, (V_i, V_j)$ *ist nicht* ε*-regulär*$\}| \leq \varepsilon t^2$

Beweis:
Später: 2.1. □

Satz 2.3. (Removal Lemma) *Für alle Graphen F und $\eta > 0$ existieren $\delta > 0$ und n_0 sodass für alle Graphen $G = (V, E)$ mit $|V| = n \geq n_0$ gilt: Falls G höchstens $\delta n^{v(F)}$ Kopien von F enthält, dann existert $E' \subseteq E$ sodass $G' = (V, E')$ F-frei ist und $|E \setminus E'| \leq \eta n^2$ gilt.*

Beweis:
Beweisidee

1. meditiere über Konstanten
2. Sei G mit $o(n^l)$ Kopien von F gegeben.
3. regularisiere G
4. lösche Kanten innerhalb der $V_1, ..., V_l$
   (wir brauchen also $t_0 \gg \frac{1}{\eta}$)

5. lösche Kanten e mit $e \cap V_0 \neq \emptyset$
   (brauchen also $\varepsilon \ll \eta$)

6. lösche Kanten von nicht-regulären Paaren
   (wieder wird $\varepsilon \ll \eta$ benötigt)

7. lösche Kanten von Paaren mit Dichte $< d_0$
   (es muss also $d_0 \ll \eta$ gelten)

8. Stelle sicher, dass nur ηn^2 Kanten gelöscht wurden

9. Zu zeigen: G ist nach dem Löschen der Kanten F-frei.

10. Annahme: Es existiert noch eine Kopie von F.

    Diese kann nur noch Kanten aus dichten und regulären Paaren haben.

11. Mit Hilfe des Counting-Lemmas erhalten wir die Existenz von vielen, d.h. $\frac{1}{2}d_0^{e(F)}\left(\frac{(1-\varepsilon)n}{tl}\right)^l$, Kopien von F im ursprünglichen Graphen G.

12. Mit $\delta \ll \frac{1}{2}d_0^{e(F)}\left(\frac{1-\varepsilon}{T_0 l}\right)^l$ ist dies ein Widerspruch zu der Annahme, dass G nur δn^l Kopien von F enthält.

Beweis

1. Seien F, $\eta > 0$ und $l = v(F)$ gegeben.

   Wir setzen $d_0 = \frac{\eta}{4}$, ε_{CL} als das ε vom Counting Lemma angewandt mit $F, \gamma = \frac{1}{2}$ und d_0. Weiterhin fixieren wir $\varepsilon = \min\{\frac{\eta}{4}, \frac{\varepsilon_{CL}}{l}\}$ und $\frac{4}{\eta}$.

   Sei T_0 die Konstante des Regularitäts Lemmas angewandt auf ε und t_0.

   Wir wählen $\delta = \frac{d_0^{e(F)}}{4}\left(\frac{1-\varepsilon}{lT_0}\right)^l$ und $n_0 = \max\{n_{0,RL}, \frac{T_0 l m_0}{1-\varepsilon}\}$, wobei $n_{0,RL}$ das n_0 des Regularitätslemmas angewandt mit ε und t_0 und m_0 das m_0 des Counting Lemmas angewandt mit $F, \gamma = \frac{1}{2}$ und d_0 sind.

2. Sei $G = (V, E)$ mit $|V| = n \geq n_0$ und höchstens δn^l Kopien von F gegeben.

   Wir müssen nun alle Kopien von F zerstören und dürfen dafür höchstens ηn^2 Kanten löschen.

3. Sei $V_0 \dot\cup ... \dot\cup V_t$ die Partition, welche das Regularitätslemma angewandt mit ε und t_0 garantiert.

4. Lösche alle Kanten innerhalb der $V_1, ..., V_1$.

   Wir löschen also höchstens $t\binom{\frac{n}{t}}{2} \leq \frac{n^2}{2t} < \frac{n^2}{t_0} = \frac{\eta}{4}n^2$ (durch die Wahl von t_0 und $t \geq t_0$) Kanten.

5. Lösche alle Kanten e mit $\cap V_0 \neq \emptyset$.

   Wir löschen höchstens $\varepsilon n^2 \leq \frac{\eta}{4}n^2$ Kanten.

6. Lösche alle Kanten von nicht-regulären Paaren.

   Wir löschen höchstens $\varepsilon t^2(\frac{n}{t})^2 \leq \varepsilon n^2 \leq \frac{\eta}{4}n^2$.

7. Lösche alle Kanten von Paaren mit Dichte $< d_0$.

   Wir löschen höchstens $\binom{t}{2}d_0(\frac{n}{t})^2 \leq \frac{d_0}{2}n^2 < d_0 n^2 \leq \frac{\eta}{4}n^2$.

8. Sei $G' = (V, E')$ der durch das Löschen entstehende Graph. Dann gilt $|E \setminus E'| \leq 4 \cdot \frac{\eta}{4}n^2 = \eta n^2$.

9. Behauptung: G' ist F-frei.

10. Annahme: Es existiert noch eine Kopie von F in G', d.h. eine Menge $W\{v_1, ..., v_l\} \subseteq V$ spannt eine Kopie von F in G' auf, d.h. $\{v_i, v_j\} \in E'$, falls $\{i, j\} \in F$.

    Falls $|V_j \cap W| \leq 1$ für alle $j \in [t]$ kann o.B.d.A. $v_i \in V_i$ angenommen werden. (Wir nummerieren einfach die Klassen V_i neu durch.) Weiterhin wissen wir, dass (V_i, V_j) $(\varepsilon, d_{i,j})$-regulär mit $d_{i,j} \geq d_0$ für alle $\{i, j\} \in E(F)$ ist. (Dies gilt, weil $\{v_i, v_j\} \in E'$ ist und jede Kante aus E' in einem regulären Paar mit Dichte $\geq d_0$ liegt - alle anderen Kanten wurden ja gelöscht.)

    $\Rightarrow G[V_1, ..., V_l]$ erfüllt die Voraussetzung des Counting Lemmas.

    $\Rightarrow G[V_1, ..., V_l]$ enthält mindestens $(1-\gamma)\prod_{\{i,j\}\in E(F)} d_{i,j} \prod_{i=1}^{l} |V_i| \geq \frac{1}{2} d_0^{e(F)} \left(\frac{(1-\varepsilon)n}{t}\right)^l > \delta n^l$ Kopien von F, was einen Widerspruch zur vorgegebenen Eigenschaft von G, wenige Kopien von F zu enthalten, steht.

    Anderenfalls:

    Seien o.B.d.A. $\{V_1, ..., V_k\}$ die Partitionsklassen, die $\{v_1, ..., v_k\}$ enthalten und $k < l$.

    Wir zerlegen $V_i, i = 1, ..., k$ in l gleichgroße Klassen $V_{i,1}, ...V_{i,k}$. (Ggf. muss die Klassengröße abgerundet werden.)

    Behauptung: (V_i, V_j) ist $(\varepsilon, d_{i,j})$-regulär $\Rightarrow (V_{i,x}, V_{j,y})$ ist $(l\varepsilon, d_{i,j})$-regulär für alle $x, y \in [l]$.

    Beweis:

    Seien $X \subseteq V_{i,x}$ und $Y \subseteq V_{j,y}$ mit $|X| \geq l\varepsilon|V_{i,x}| \geq |V_i|$ (*Ungleichung ggf. korrigieren!*) und $|Y| \geq l\varepsilon|V_{j,y}| \geq \varepsilon|V_j|$.

    $\Rightarrow |d(X, Y) - d_{i,j}| \leq \varepsilon < l\varepsilon$

    Da ein Homomorphismus von F nach

    $$(\{V_1, ..., V_k\}, \{\{V_i, V_j\}|(V_i, V_j) \text{ ist } (\varepsilon, d_{i,j})\text{−regulär mit } d_{i,j} \geq d_0\})$$

    existiert und $|V(F)| = l$ gilt, enthält der Graph

    $$(\{V_{i,x}|i = 1 \in [t], x \in [l]\}, \{\{V_{i,x}, V_{j,y}\}|(V_i, V_j) \text{ ist } (\varepsilon, d_{i,j}\text{−regulär mit } d_{i,j} \geq d_0, x, y \in [l]\})$$

    eine Kopie von F.

    Seien $W_1, ..., W_l$ die Klassen in $\{V_{i,x}|i = 1, ..., k, x = 1, ..., l\}$, welche die Kopie von F aufspannen.

    $\Rightarrow$ Die Voraussetzungen des Counting Lemmas sind erfüllt.

    $\Rightarrow (1-\gamma)d_0^{e(F)} \left(\frac{(1-\varepsilon)n}{T_0 l}\right)^l$ Kopien von F sind in $G' \subseteq G$ enthalten. Dies ist ebenfalls ein Widerspruch zu der vorausgesetzten maximalen Anzahl an Kopien von F in G. □

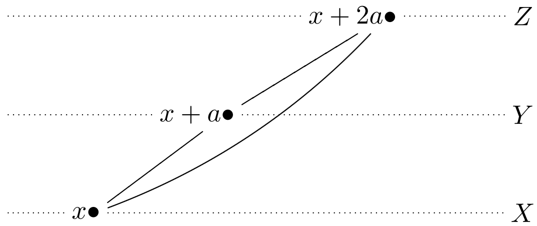
Bemerkung (Kurzform des Removal Lemmas)
Wenige $(o(n^{v(F)}))$ Kopien von F können einfach $(o(n^2)$ Kanten) zerstört werden.

Korollar 2.2.
Falls jede Kante von $G = (V, E)$ in genau einem K_3 enthalten ist, dann gilt $|E| = o(|V|^2)$.

Beweis:
Es sei $n = |V|$.
Es gilt $\#(K_3 \subseteq G) = \frac{|E|}{3} \leq n^2 = o(n^3) \leq \delta n^{v(K_3)}$ für alle $\delta > 0$ und für ausreichend großes n.
Mit Hilfe des Removal Lemmas erhalten wir $\exists E' \subseteq E : G' = (V, E')$ K_3-frei, $|E \setminus E'| = o(n^2)$.
Jede Kante in $E \setminus E'$ zerstört höchstens einen K_3.
$\Rightarrow \frac{|E|}{3} = \#\{K_3 \subseteq G\} \leq |E \setminus E'| = o(n^2)$
$\Rightarrow |E| = o(n^2)$

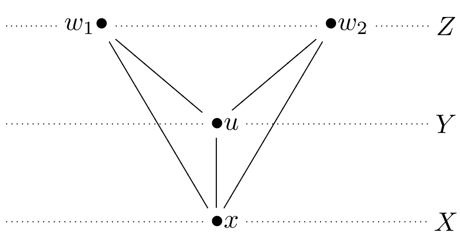
Satz 2.4. (Satz von Roth) $\forall\varepsilon > 0\exists n_0 : \forall n \geq n_0\forall A \subseteq [n], |A| \geq \varepsilon n : A$ *enthält* AP_3*, wobei* AP_3 *eine arithmetische Progression der Länge* 3*, d.h.* $a, b, c \in A : \frac{a+c}{2} = b$.

Beweis:
Übergang von den AP_3s zu Graphen: Sei $A \subseteq [n]$ eine AP_3-freie Menge.
Wir definieren einen tripartiten Graphen $G = (X, Y, Z, E)$ durch $X = [n], Y = [2n], Z = [3n]$ und $E = \bigcup_{x \in X} \bigcup_{a \in A} \{\{x, x+a\}, \{x, x+2a\}, \{x+a, x+2a\}\}$:



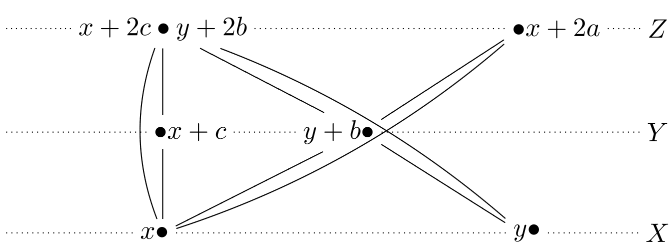
Es gilt $|V(G_A)| = 6n$ und $|E| = 3n|A|$.
Falls $|E| = o(n^2)$, dann $|A| = o(n)$, d.h. $r_3(n) = o(n)$.
Wir wissen, dass jede Kante aus E in einem Dreieck enthalten ist, da E die Vereinigung von Dreiecken ist. Weiterhin beobachten wir, dass sogar jede Kante in *genau* einem Dreieck enthalten ist:
Beweis dieser Beobachtung:

1. Fall



   $\exists a, b \in A : u = x + a = x + b, w_1 = x + 2a, w_2 = x + 2b \Rightarrow a = b \Rightarrow w_1 = w_2$

2. Fall



   $y + 2b = x + 2c$

   $x + a = y + b$

   $\Rightarrow \frac{y-x}{2} = c - b, y - x = a - b$

   $\Rightarrow b, c = b + \frac{y-x}{2}, a = b + 2\frac{y-x}{2}$ ist AP_3, was einen Widerspruch zur AP_3-freiheit von A darstellt.

$\Rightarrow$ Jede Kante ist in genau einem K_3. Nun verwenden wir das Removal Lemma und erhalten, dass der Graph höchstens $o(|V|^2)$ Kanten hat.
$\Rightarrow |A| = o(n)$ $\square$

Bemerkung (Geschichte des obigen Satzes)

- Es war einmal im Jahre 1936, wo Erdős und Turán vermuteten, dass

  $\forall k, \varepsilon > 0 : \exists n_0 : \forall n \geq n_0, A \subseteq [n]$ mit $|A| \geq \varepsilon n : A$ enthält AP_k.

- Motivation: Satz (van der Waerden): $\forall k \geq 3, r \geq 2 : \exists n_0 : \forall n \geq n_0 \forall C_1 \dot\cup ... \dot\cup C_r = [n] \exists i : C_i$ enthält AP_k.

- Viele, viele Jahre später (im Jahre 1947) zeigte Behrend: $\exists A \subseteq [n] : A$ ist AP_3-frei und $|A| \geq \frac{n}{e^{C\sqrt{\log n}}}$, d.h. für alle $\varepsilon > 0, C > 0, k > 0$ gilt asymptotisch $\frac{n}{n^\varepsilon} \ll \frac{n}{e^{C\sqrt{\log n}}} \ll \frac{n}{\log^k n}$.

- Wir definieren $r_k(n) = \max\{|A| \big| A \subseteq [n], A \text{ ist } AP_k\text{-frei}\}$.

  Die ET-Vermutung ist nun $r_k(n) = o(n)$.

- Behrend: $r_k(n) = \Omega(\frac{n}{e^{C\sqrt{\log n}}})$ für ein $C > 0$

- Roth: $r_s(n) = o(n)$, genauer: $r_3(n) \leq O(\frac{n}{\log\log n})$

- Der aktuelle Rekord liegt bei $r_3(n) \leq O\left(\frac{n\log\log n}{(\log n)^{\frac{2}{3}}}\right), C > 0$ (Bourgain (2007)).

- Strenge Alien- ähhm ... ET(Erdős-Turán)-Vermutung:

  Für $(a_i)_{i\in\mathbb{N}} \subseteq \mathbb{N}$ gilt: $\sum a_i = \infty \Rightarrow \{a_i | i \in \mathbb{N}\}$ enthält AP_k.

- Szemeréd (1969): $r_4(n) = o(n)$

  Szemeréd (1975): $r_k(n) = o(n)$ für alle $k \geq 3$

- Fürstenberg (1977): Alternativer Beweis mit ergodischer Theorie

- Gowers (2000): Beweis mit Fourier-Analysis

- Gseen-Tao: $\mathbb{P}$ enthält AP_k für alle k

Vorlesung am 30.05.2008

Vorlesung am 04.06.2008

Vorlesung am 06.06.2008

Vorlesung am 11.06.2008

Lemma 2.7. (1)

1. *$\forall C, D \subseteq V, C \cap D = \emptyset, \mathcal{C}$ Partition von C und $\mathcal{D}$ Partition von D : $ind(C, D) \leq ind(\mathcal{C}, \mathcal{D})$*
2. *$\forall \mathcal{P}' \prec \mathcal{P} : ind(\mathcal{P}) \leq ind(\mathcal{P}')$*

Lemma 2.8. (2) *Für alle $\varepsilon > 0$ und $C, D \subseteq V$ mit $C \cap D = \emptyset$ gilt:*
Falls (C, D) nicht ε-regulär ist, dann existieren Partitionen $\mathcal{C} = \{C_1, C_2\}$ und $\mathcal{D} = \{D_1, D_2\}$ von C bzw. D sodass $ind(\mathcal{C}, \mathcal{D}) \geq ind(C, D) + \frac{\varepsilon^4|C||D|}{n^2}$.

Lemma 2.9. (3)
Sei $\mathcal{P} = \{C_0, ..., C_k\}$ eine Partition von V mit $|C_0| \leq \varepsilon n$ und $|C_1| = ... = |C_k| = c$. Falls $\mathcal{P}$ nicht ε-regulär ist, d.h. $|\{\{i, j\} \in \binom{k}{2} | (C_i, C_j)\ \neg\varepsilon\text{-regulär}\}| > \varepsilon k^2$, dann existiert eine Partition $\mathcal{P}' = \{c_0', ..., c_l'\}$ mit $k \leq l \leq k4^k$, $|C_0'| \leq |C_0| + \frac{n}{2^k}$, $|C_1'| = ... = |C_l'|$ und $ind(\mathcal{P}') \geq ind(\mathcal{P}) + \frac{\varepsilon^5}{2}$.

Beweis:
Für alle $1 \leq i < j \leq k$ definieren wir eine Partition $\mathcal{C}_{i,j}$ von C_i und eine Partition $\mathcal{C}_{j,i}$ von C_j wie folgt:

- Falls (C_i, C_j) ε-regulär ist, dann setzten wir $\mathcal{C}_{i,j} = \{C_i\}$ und $\mathcal{C}_{j,i} = \{C_j\}$.

- Falls (C_i, C_j) nicht ε-regulär ist, dann sind $\mathcal{C}_{i,j}$ und $\mathcal{C}_{j,i}$ die entsprechenden Partitionen aus dem Lemma 2.

  Es gilt dann $ind(\mathcal{C}_i, \mathcal{C}_j) \geq ind(C_i, C_j) + \frac{\varepsilon^4 |C_i||C_j|}{n^2} = ind(\mathcal{C}_i, \mathcal{C}_j) + \frac{\varepsilon^4 c^2}{n^2}$.

Für jedes $i = 1, ..., k$ sei $\mathcal{C}_i$ die minimale Partition, die alle Partitionen $\mathcal{C}_{i,j}$ mit $i \neq j$ verfeinert.
$\Rightarrow |\mathcal{C}_i| \leq 2^{n-1}$
Wir definieren $\mathcal{Q} = \{C_0\} \cup \bigcup_{i=1}^{k} \mathcal{C}_i$. Dann gilt $k \leq |\mathcal{Q}| \leq 1 + 2k^{k-1} \leq k2^k$.
Sei $\mathcal{C}_0 = \{\{v\} | v \in C_0\}$.

$$\begin{aligned}
\Rightarrow ind(\mathcal{Q}) &= \sum_{1 \leq i < j \leq k} ind(\mathcal{C}_i, \mathcal{C}_j) + \sum_{1 \leq i \leq k} ind(\mathcal{C}_0, \mathcal{C}_i) + \sum_{i=0}^{k} ind(\mathcal{C}_i) \\
&\geq \sum_{1 \leq i < j \leq k} ind(\mathcal{C}_{i,j}, \mathcal{C}_{j,i}) + \sum_{i=1}^{k} ind(\mathcal{C}_0, \{C_i\}) \\
&\geq \sum_{1 \leq i < j \leq k} ind(C_i, C_j) + \varepsilon k^2 \cdot \frac{\varepsilon^4 c^2}{n^2} + \sum_{i=1}^{k} ind(\mathcal{C}_0, \{C_i\}) + ind(\mathcal{C}_0) \\
&= ind(\mathcal{P}) + \frac{\varepsilon^5 k^2 c^2}{n^2} > ind(\mathcal{P}) + \frac{\varepsilon^5}{2}
\end{aligned}$$

Die letzte Ungleichung gilt wegen $c = \frac{n - |C_0|}{k} \geq \frac{3n}{4k} \Rightarrow \frac{k^2 c^2}{n^2} \geq \frac{9}{16} \geq \frac{1}{2}$ und $c_0 \leq \varepsilon n \leq \frac{n}{4}$.

Jetzt verfeinern wir $\mathcal{Q}$, um eine Partition mit gleichgroßen Klassen $C'_1, ..., C'_l$ zu erhalten. Sei $d = \lfloor \frac{c}{4^k} \rfloor$ und $C'_1, ..., C'_l$ die maximale Familie von disjunkten Mengen der Größe d, sodass für alle $i \in l$ eine Partitionsklasse in $\mathcal{Q}$ existiert, die C'_i enthält. Sei $C'_0 = V \setminus \bigcup_{i=1}^{l} C'_i$.
$\Rightarrow |C'_0| \leq |C_0| + |\mathcal{Q}| d \leq |C_0| + \frac{k 2^k c}{4^k} \leq |C_0| + \frac{n}{2^k}$ (wegen $c \leq \frac{n}{k}$)
Weiterhin kann C_i höchstens 4^k Klassen der Größe $\lfloor \frac{c}{4^k} \rfloor$ enthalten.
Sei $\mathcal{P}' = \{C'_0, ..., C'_l\}$. Dann ist $\tilde{\mathcal{P}}'$ eine Verfeinerung von $\tilde{\mathcal{Q}}' = \mathcal{C}_0 \cup \bigcup_{k=1}^{k} \mathcal{C}_i$.
$\Rightarrow ind(\mathcal{P}') = ind(\tilde{\mathcal{P}}') \geq ind(\tilde{\mathcal{Q}}) = ind(\mathcal{Q}) \geq ind(\mathcal{P}) + \frac{\varepsilon^5}{2}$ □

Beweis des Regularitätslemmas
Seien ε und t_0 gegeben.
Wir wählen $T_0 = (x \mapsto x4^x)^{\frac{1}{\varepsilon^5}}(t_0)$ und $n_0 = 2^{\frac{t_0}{\varepsilon}}$.
Sei nun $G = (V, E)$ mit $|V| = n \geq n_0$ gegeben.
Wir wählen nun eine beliebige Partition $\mathcal{P}_1 = \{C_0^1, ..., C_{t_0}^1\}$ mit $|C'_0| \leq \frac{1}{2}\varepsilon n$ und $|C_1^1| = ... = |C_{t_0}^1|$.
Solange $\mathcal{P}_i$ nicht ε-regulär ist, wenden wir das vorherige Lemma auf $\mathcal{P}_i$ an und setzen $\mathcal{P}_{i+1} = \mathcal{P}'_i$.

Da $ind(\mathcal{P}) \leq \frac{1}{2}$, terminiert der Algorithmus nach höchstens $\frac{\frac{1}{2}}{\frac{\varepsilon^5}{2}} = \frac{1}{\varepsilon^5}$ Iterationen.
Sei $\mathcal{P} = \{C_0, ..., C_k\}$ die letzte dieser Partitionen.
Dann gilt $|C_0| \leq |C_0^1| + \sum_{k=t_0}^{t} \frac{n}{2^k} \leq \varepsilon n$ für ausreichend großes t_0 ($t_0 \geq \frac{1}{\varepsilon}$).
Weiterhin gilt $t_0 \leq t \leq T_0 = (x \mapsto x4^x)^{\frac{1}{\varepsilon^5}}(t_0)$. □

Vorlesung am 13.06.2008

Lemma 2.10. (Einbettungslemma) *Für alle $d, \Delta > 0$ und Graphen F existieren $\varepsilon, c > 0$ und $m_0 \in \mathbb{N}$ sodass:*
Seien $V(F) = \{1, ..., l\}$ und $G = (V_1 \dot\cup ... \dot\cup V_l, E)$ ein l-partiter Graph mit $|V_1| = ... = |V_l| = m \geq m_0$.
Weiterhin für alle $\{i, j\} \in E(F)$ das Paar(V_i, V_j) (ε, d)-regulär.
$H = (W_1 \dot\cup ... \dot\cup W_l, E_H)$ sei ein l-partiter Graph mit $|W_i| \leq cm$ für alle $i = 1, ..., l$ mit $\Delta(H) \leq \Delta$.
Für alle $w_i \in W_i, i = 1, ..., l$ sei $\varphi_i : V(H) \to V(F), \varphi(w_i) = i$ ein Graphenhomomorphismus.
Dann gilt $H \subseteq G$.

Beweis:
Beweisidee: Wir betten iterativ die Knoten von H in G ein und bestimmen nach jeder Einbettung (des Knotens t) für jeden noch nicht eingebetteten Knoten i eine Kandidatenmenge $C_i^{(t)}$ von Knoten, die diesem Knoten entsprechen könnten.

Nun zum Beweis:

Seien $d > 0, F$ und Δ gegeben. und $[l] = V(F)$.
Wir bestimmen nun $\varepsilon = \varepsilon_i \left(\frac{d}{2}\right)^{\Delta-1}, \tilde{\varepsilon} = \frac{d}{2}, \varepsilon_i = \min\{\frac{1}{\Delta}, \frac{d}{2}\}, c = \frac{1}{2\Delta}\left(\frac{d}{2}\right)^{\Delta-1}$ und $m_0 = 1$.
Seien nun G und H gegeben. O.B.d.A. sei $V(H) = [M]$. Sei $\varphi : F(H) \to F(F)$ ein Homomorphismus. Wir werden die folgende Aussagen (S_i) per Induktioin beweisen:
(S_i): Es existieren eine partielle Einbettung $\psi_i[i] \to V(G), [i] \subseteq V(H)$ und $C^i_{i+1}, ..., C^i_M$ mit:

1. $\forall j = i+1, ..., M : C^i_j \subseteq V_{\varphi(j)} \cap \bigcap_{w \in N_H(j) \cap \{1,...,i\}} N_G(\psi_i(w))$
2. $\forall j = i+1, ..., M : |C^i_j| \geq (d - \tilde{\varepsilon})^{|N_H(S) \cap \{1,...,i\}|} \cdot m$

Weiterhin folgt aus (S_M) das Einbettungslemma.

Induktionsanfang:

1. ψ_0 ist trivial.
2. $C^0_j = V_{\varphi(j)}$

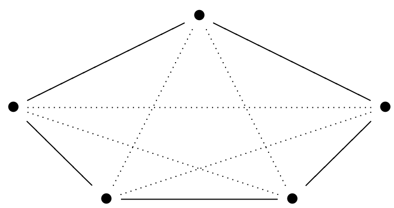
Induktionsschritt: $(S_i) \to (S_{i+1})$
Sei $k \in N_H(i+1) \cap \{i+2, ..., M\}$. Dann ist (C^i_{i+1}, C^i_k) (ε_i, d)-regulär mit $\varepsilon_i \geq \frac{\varepsilon}{(d-\tilde{\varepsilon})^{\Delta-1}}$ (wegen 2. von (S_i)).
$\Rightarrow B^-_{i+1,k} = \left\{v \in C^i_{i+1} \middle| |N_G(v) \cap C^i_k| < (d - \varepsilon_i)|C^i_k|\right\}, |B^-_{i+1,k}| \leq \varepsilon |C^i_{i+1}|$
Aus $B^- = \bigcup_{k \in N_H(i+1) \cap \{i+2,...,M\}} B^-_{i+1,k}$ folgt $|B^-| \leq (\Delta - 1)\varepsilon_i C^i_{i+1}$.
Wir setzen nun $\tilde{C}_{i+1} = C^i_{i+1} \setminus (B^- \cup \psi_i(\{1, ..., i\}))$. Die Knoten $\psi_(\{1, ..., i\})$ müssen wir ebenfalls entfernen, weil schon verwendete Knoten nicht mehr in der Kandidatenmenge enthalten sein dürfen.
Als Größenabschätung erhalten wir:
$|\tilde{C}_{i+1}| \geq |C^i_{i+1}| - (\Delta - 1)\varepsilon_i |C^i_{i+1}| - \underbrace{cm}_{\text{folgt aus } |W_i| \leq cm} \geq ((1 - (\Delta - 1)\varepsilon_i)(d - \tilde{\varepsilon})^{\Delta-1} - c)m$
Es reicht also, wenn wir die Konstanten so wählen, dass der erste Faktor des letzten Terms positiv ist, weil in diesem Fall ein $v \in \tilde{C}_{i+1}$ existiert. Wir definieren dann:
$\psi_{i+1} = \begin{cases} \psi_i(x) & x \leq i \\ v & x = i+1 \end{cases}$
und
$C^{i+1}_j = \begin{cases} C^i_j & j \notin N_H(i+1) \\ N_G(v) \cap C^i_j & j \in N_H(i+1) \end{cases}$ mit $j = i+2, ..., M$.
Zeigen wir zunächst, dass ψ_{i+1} und C^{i+1}_j die geforderten Eigenschaften haben:
ψ_{i+1} ist eine Einbettung von $H[[i+1]]$, da ψ_i eine Einbettung von $H[[i]]$ ist und v in C^i_{i+1} eingebettet wird. Damit ist auch die Nachbarschaft von $\psi_{i+1}(i+1)$ korrekt.
2.1 folgt aus der Definition von C^{i+1}_j und von 1. aus (S_i).
2.2. folgt aus $v \in C^i_{i+1} \setminus \bigcup_{k \in N_H(i+1), k \geq i+2} B^-_{i+1,k}$ und $\varepsilon_i \leq \tilde{\varepsilon}$.
Nun noch zu den Konstanten:
$\tilde{\varepsilon} \leq \frac{d}{2}$
$\frac{2^{\Delta-1}\varepsilon}{d^{\Delta-1}} \leq \varepsilon_i \leq \tilde{\varepsilon}$
$\Rightarrow (1 - (\Delta - 1)\varepsilon_i)(\frac{d}{2})^{\Delta-1} - c > 0$ □

Vorlesung am 18.06.2008

2.2 Burr-Erdős-Vermutung

Definition ($\mathcal{R}(H_1, H_2)$)
Mit $\mathcal{R}(H_1, H_2)$ bezeichnen wir die kleinste natürliche Zahl $n \in \mathbb{N}$, sodass jede 2-Färbung der Kanten des K_n eine mit 0 gefärbte Kopie von H_1 oder eine mit 1 gefärbte Kopie von H_2 enthält.

Beispiel ($\mathcal{R}(K_3, K_3)$)
Es gilt $\mathcal{R}(K_3, K_3) = 6$, denn:
Für $n = 5$ gibt es eine dreicksfreie 2-Färbung:



Es bleibt also noch zu zeigen, dass dies für $n = 6$ nicht mehr möglich ist: Wir wählen einen beliebigen Knoten v des K_6 und betrachten die Kanten zu dessen Nachbarschaft. Nach dem Schubfachprinzip gibt es eine Farbe (o.B.d.A. sei dies die Farbe 0), welche mindestens 3 mal auftritt. Wenn die Knoten, zu denen diese Kanten führen, durch eine Kante der Farbe 0 verbunden sind, so erhalten wir ein mit 0 gefärbtes Dreieck, welches diese Kante und den Knoten v enthält. Anderenfalls sind alle Kanten in diesem Teil der Nachbarschaft von v mit 1 gefärbt. Damit erhalten wir jedoch ein Dreieck der Farbe 1. □

Burr-Erdős-Vermutung

1. $\forall \Delta : \exists C_\Delta : \forall H, \Delta(H) \leq \Delta : \mathcal{R}(H, H) \leq C_\Delta |V(H)|$
2. $\forall k : \exists D_k : \forall H, \hat{\delta}(H) \leq k : \mathcal{R}(H, H) \leq D_k v(H)$

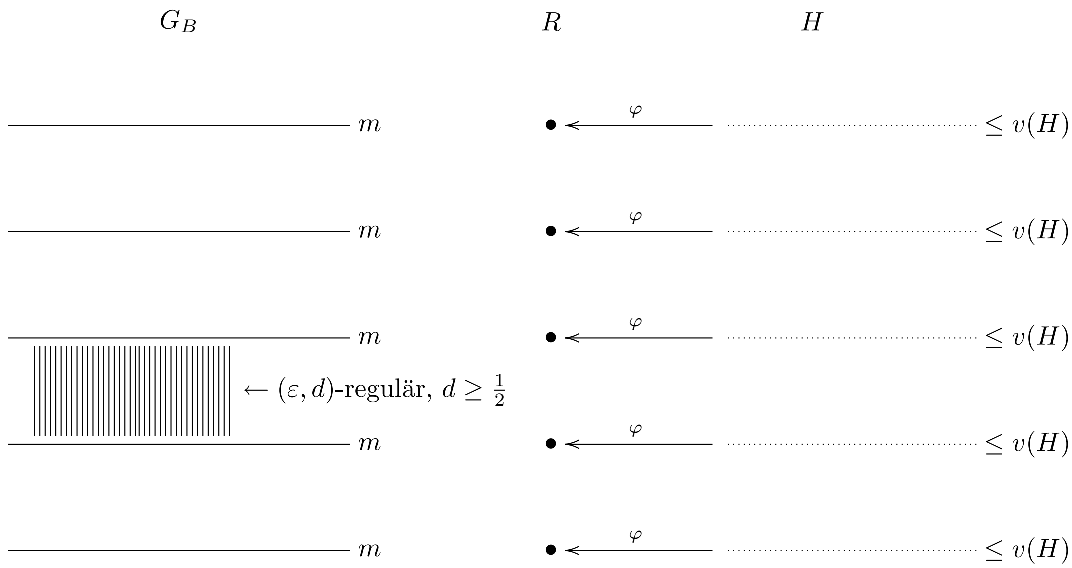
Satz 2.5. (Chvátal, Rödl, Szenerédi, Trotter (1983))
$\forall \Delta : \exists C_\Delta : \forall H, \Delta(H) \leq \Delta : \mathcal{R}(H, H) \leq C_\Delta v(H)$

Beweis:
Beweisidee: Wir regularisieren den roten (oder blauen) Teilgraphen.

Beweis:
Seien $\Delta \geq 1$ und $r = \mathcal{R}(K_{\Delta+1,\Delta+1})$.
Mit Turán erhalten wir $e(R) \geq (1 - \frac{1}{r-1} + o(1))\binom{v(R)}{2} \Rightarrow K_r \subseteq R$.
Wir setzen nun $d_0 = \frac{1}{2}$. Weiterhin erhalten wir aus Δ und F durch das Einbettungslemma ε_0 und C. Nun verwenden wir $\varepsilon \leq \min\{1 \frac{3(r-1)}{,} \varepsilon_0\}$.
Aus ε und $t_0 = r$ erhalten wir mit Hilfe des Regularitätslemmas T_0 und n_0. Jetzt setzen wir $C_\Delta = \frac{2T_0}{c}$. Sei H ein Graph mit n Knoten und $\Delta(H) \leq \Delta$. Sei $G_R \dot\cup G_B = K_n$ mit $N = C_\Delta n$ eine 2-Färbung von $E(K_n)$ mit den Farben rot und blau. Durch das Regularitätslemma erhalten wir eine Partition $V_0 \cot \cup V_1 \dot\cup ... \dot\cup V_t = V = V(G_R) = V(G_B) = V(K_n)$ mit:

1. $|V_0| \leq \varepsilon N, |V_1| = ... = |V_t| = m \geq \frac{(1-\varepsilon)N}{T_0}$
2. Höchstens εt^2 Paare (V_i, V_j) sind irregulär.

Sei $R = ([t], E_R)$ definiert vermöge $\{i, j\} \in E_R \Leftrightarrow (V_i, V_j)$ ist ε-regulär in G_R.
Aus der 2. Aussage des Regularitätslemmas folgt $|E_R| \geq \binom{t}{2} - \varepsilon t^2 > ex(t, K_r)$. Da $t \geq r$ folgt mit Turán $K_r \subseteq R$.
Seien $I \subseteq [t]$ die Knoten der Kopie von K_r in R. Wir definieren eine 2-Färbung der Kanten von $R[I]$ durch $\gamma(i, j) = \begin{cases} \text{rot} & d_{G_R}(V_i, V_j) \geq \frac{1}{2} \\ \text{blau} & \text{sonst} \end{cases}$. Wegen $|I| = r = R(K_{\Delta+1}, K_{\Delta+1})$ existiert nach Ramsey eine monochrome Clique der Größe $\Delta + 1$ in $R[I]$. Sei $J \subseteq I$ die Knotenmenge dieser Clique.
Wir wissen, dass für ein (ε, d)-reguläres Paar (V_i, V_j) in G_R das selbe Paar $(\varepsilon, 1-d)$-regulär in G_B ist. Demnach gilt für alle $\{i, j\} \in \binom{J}{2}$: (V_i, V_j) ist (ε, d)-regulär für G_R mit $d_{i,j} \geq \frac{1}{2}$ oder alle Paare sind (ε, d)-regulär für G_B mit $d_{i,j} \geq \frac{1}{2}$. O.B.d.A. gelte diese Aussage für G_B.



Die Knotenmengen aus H, welche durch einen Homomorphismus φ auf Knoten von R abgebildet werden, sind stabile Mengen. Wegen $\Delta(H) \leq \Delta$ folgt weiterhin $\chi(H) \leq \Delta + 1$.
Durch das Einbettungslemma erhalten wir $H \subseteq G_B$, falls $v(H) \leq cm$.
Wegen $cm \geq c\frac{(1-\varepsilon)N}{T_0} \geq \frac{cN}{2T_0} = \frac{N}{C_\Delta} = n = v(H)$ ist diese Ungleichung erfüllt. $\square$

Vorlesung am 20.06.2008

2.3 Algorithmisches Regularitätslemma

Satz 2.6. *Das folgende Entscheidungsproblem ist co-NP-vollständig:*
Gegeben sind ein Graph G, $t \geq 1$ und $\varepsilon > 0$ mit $V(G) = V_0 \dot\cup ... \dot\cup V_t$.
Gefragt ist, ob die Partition ε-regulär ist, d.h. ob höchstens $\varepsilon\binom{t}{2}$ Paare nicht ε-regulär sind.

Bemerkung
Wir werden sogar zeigen, dass das Problem bereits bei $t = 2$ und $\varepsilon = \frac{1}{2}$ co-NP-vollständig ist.

Satz 2.7. *Das folgende Problem ist co-NP-vollständig.*
Gegeben sind $\varepsilon > 0$ und $G = (A \dot\cup B)$ mit $|A| = |B| = n$.
Gefragt ist, ob (A, B) ε-regulär ist.

Beweis:
Sei $G = (A \dot\cup B, E)$ eine Instanz des speziellen `BiClique`-Problems aus dem Lemma 2.14, d.h. mit $|A| = |B| = n, |E| = \frac{n^2}{2} - 1$.
$\Rightarrow d(A, B) = \frac{1}{2} - \frac{1}{n^2}$
Zu zeigen ist nun: $\exists A' \subseteq A, B' \subseteq B, |A'| \geq \frac{n}{2}, |B'| \geq \frac{n}{2}, |d(A', B') - d(A, B)| > \frac{1}{2} \Leftrightarrow K_{\frac{n}{2},\frac{n}{2}} \subseteq G$

1. "$\Rightarrow$"

   $|d(A', B') - \frac{1}{2} + \frac{1}{n^2}| > \frac{1}{2}$

   Wegen $|0 - \frac{1}{2} + \frac{1}{n^2}| < \frac{1}{2}$, muss $d > \frac{1}{2} - \frac{1}{n^2} + \frac{1}{2} = 1 - \frac{1}{n^2} = \frac{n^2-1}{n^2}$ gelten.

   $\Rightarrow d(A', B') = 1$

   $\Rightarrow (A', B')$ spannt einen vollständigen bipartiten Graphen.

   $\Rightarrow K_{\frac{n}{2},\frac{n}{2}} \subseteq G$

2. "$\Leftarrow$"

   Wenn $K_{\frac{n}{2},\frac{n}{2}} \subseteq G$, dann existieren $A' \subseteq A, B' \subseteq B$ mit $|A'| \geq \frac{n}{2}, |B'| \geq \frac{n}{2}$ und $d(A', B') = 1$.

   $\Rightarrow |d(A', B') - d(A, B)| > \frac{1}{2}$

Definition (`Clique`)
Gegeben sind ein Graph G und $k \in \mathbb{N}$.
Gefragt ist, ob $\omega(G) \geq k$ gilt.

Definition (`BiClique`)
Gegeben ist ein bipartiter Graph $G = (A\dot{\cup}B, E), |A| = |B| = n, k \in \mathbb{N}$.
Gefragt ist, ob $K_{k,k} \subseteq G$ gilt.

Lemma 2.11. *Das folgende Problem ist NP-vollständig:*
Gegeben ist ein Graph G mit $n \in 2\mathbb{N} + 1$ Knoten.
Gefragt ist, ob $K_{\frac{n+1}{2}} \subseteq G$ gilt.

Beweis:
Sei (G, k) eine Instanz von Clique.

1. Fall: $k \leq \frac{v(G)+1}{2}$

   Sei G^* der Graph, welcher G, $v(G) + 1 - 2k$ zusätzliche Knoten und alle Kanten innerhalb dieser neuen Knoten und zwischen den neuen Knoten und den Knoten aus G enthält.

   $\Rightarrow \omega(G^*) \geq \frac{v(G)+1}{2} \Leftrightarrow K_k \subseteq G$

2. Fall: $k > \frac{v(G)+1}{2}$

   Sei nun $G^* = G\dot{\cup}(2k - v(G) - 1)K_1$. Dann gilt $K_k \subseteq G \Leftrightarrow K_{\frac{v(G^*)+1}{2}} \subseteq G^*$. □

Lemma 2.12. *`BiClique` ist NP-vollständig.*

Beweis:
Sei G eine Instanz des Problems des letzten Lemmas.

Wir definieren $H = (A\dot{\cup}B, F)$ mit:
$$\begin{aligned} A &= \{\alpha_{i,j} | 1 \leq i, j \leq n\} \\ B &= \{\beta_{i,j} | 1 \leq i, j \leq n\} \\ F &= \{(\alpha_{i,j}, \beta_{k,l} | i = k \vee (\{i, k\} \in E \wedge l \neq i \wedge j \neq k)\} \end{aligned}$$

Behauptung: $\omega(G) \geq \frac{n+1}{2}, v(G) = n \Leftrightarrow K_{(\frac{n+1}{2})^2,(\frac{n+1}{2})^2} \subseteq H$
Beweis dieser Behauptung:

1. Wir definieren $A' = \{\alpha_{i,j} | i \in W, j \notin W\} \cup \{\alpha_{i,j} | j \in W\}$
   und $B' = \{\beta_{i,j} | i \in W, j \notin W\} \cup \{\beta_{i,j} | i \in W\}$

   Es gilt $|A'| \cdot |B'| = (\frac{n+1}{2})(\frac{n-1}{2}) + \frac{n+1}{2} = (\frac{n+1}{2})^2$.

   Sei $(\alpha_{i,j}, \beta_{k,l}) \in A' \times B'$.

   (a) Fall $i = k$:
       Hier gilt $(\alpha_{i,j}, \beta_{k,l}) \in F$.

   (b) Fall $i \neq k, i \neq j, k \neq l$:
       $\Rightarrow \{i, k\} \in E(G)$, weil $i, k \in W$
       Weiterhin gilt $i \neq l$, wegen $i \in W$ und $l \notin W$ und analog $j \neq k$, wegen $j \notin W$ und $k \in W$.
       $\Rightarrow (\alpha_{i,j}, \beta_{k,l}) \in F$

(c) Fall $i \neq k, i = j, k \neq l$:
    $\{i,k\} \in E(G)$ und $i \neq l$ erhalten wir wie im Fall (b).
    Wegen $j = i \neq k$ folgt nun auch $(\alpha_{i,j}, \beta_{k,l}) \in F$.

(d) Fall $i \neq k, i \neq j, k = l$:
    Dieser Fall kann analog zum Fall (c) behandelt werden.

(e) Fall $i \neq k, i = j, k = l$:
    Auch hier gilt (wie in (b)) $\{i,k\} \in E(G)$.
    Weiterhin haben wir $i \neq k = l$ und $j = i \neq k$ und damit $(\alpha_{i,j}, \beta_{k,l}) \in F$.

Damit ist die Hinrichtung erledigt.

2. "$\Leftarrow$"

Sei $H' = (A'\dot{\cup}B', F')$ ein größter (bezüglich der Anzahl der Kanten) vollständiger bipartiter Untergraph ivon H. O.B.d.A. sei $|A'| \geq |B'| \geq (\frac{n+1}{2})^2$.

Weiterhin seien $s(i) = |\{j \in [n] | \alpha_{i,j} \in A'\}|, t(k) = |\{l \in [n] | \beta_{k,l} \in B'\}|$, $A^* = \{i | s(i) > 0\}$ und analog $B^* = \{k | t(k) > 0\}$.

Fakt:

(a) $s(i) = n - |B^*|$ für $i \in A^* \setminus C^*$
(b) $s(i) = n - |B^*| + 1$ für $i \in C^*$

Analog gilt dies für $t(k)$.

Beweis dieses Fakts:

(a) Falls $i \in A^* \setminus C^*$, dann gilt $\forall k \in B^* : \alpha_{i,k} \notin A'$, denn sonst hätten wir:
    Es gilt: $i \in A^* \setminus C^* \Rightarrow \exists j : \alpha_{i,j} \in A' \wedge \neg \exists k : \beta_{i,k} \in B'$
    mit der Annahme $\alpha_{i,k} \in A'$ und $k \in B^*$ erhalten wir $\exists j : \beta_{k,j} \in B'$. Damit ergibt sich $(\alpha_{i,k}, \beta_{k,j}) \in F$.
    Wegen $k \in B^*$ und $i \in A^* \setminus C^*$, d.h. $i \notin B^*$ gilt $i \neq k$. Dies ist jedoch ein Widerspruch zur der Definition von F, da der zweite Index von α und der erste Index von β gleich sind.

(b) Falls $i \in A^* \setminus C^*$, dann gilt: $\forall k \in B^* : \alpha_{i,k} \in A'$:
    Wegen $i \in A^*$ existiert ein k' mit $\alpha_{i,k'} \in A'$. Sei $\beta_{j,k} \in B'$. Dann erhalten wir $(\alpha_{i,k'}, \beta_{j,l}) \in F' \subseteq F$.
    $\Rightarrow \{i,j\} \in E(G), l \neq i, j \neq k', i \neq j$, da $i \notin C^*$ und $j \in B^*$
    Genauso ergibt sich $j \neq k \mp (\alpha_{i,k}, \beta_{j,l}) \in F \Rightarrow \alpha_{i,k} \in A'$, da sonst H' um $\alpha_{i,k}$ erweitert werden könnte, was einen Widerspruch zur maximalen Wahl von H' darstellte.

(c) Falls $i \in C^*$ erhalten wir $\forall k \in B^* \setminus \{i\} : \alpha_{i,k} \notin A'$ (analog zu (a)).

(d) Falls $i \in C^*$ erhalten wir weiterhin $\forall k \notin B^* \vee k = i : \alpha_{i,k} \in A'$ (analog zu (b)).

Mit (a) ergibt sich $s(i) \leq n - |B^*|$ und mit (b) die umgekehrte Ungleichung $s(i) \geq n - |B^*|$, d.h. es gilt $s(i) = n - |B^*|$.

Mit (c) ergibt sich $s(i) \leq n - |B^*| + 1$ und mit (d) $s(i) \geq n - |B^*| + 1$, d.h. $s(i) = n - |B^*| + 1$.

Damit sind die Bahuptungen des Fakts bewiesen.

Setzen wir nun $x = |A^*|, y = |B^*|$ und $z = |C^*|$.

Damit erhalten wir $|A'| = \sum_{i \in A^*} s(i) = x(n-y) + z$ und $|B'| = \sum_{i \in B^*} t(i) = y(n-x) + z$.

Wegen $|A'| \geq |B'| \geq \left(\frac{n+1}{2}\right)^2$ erhalten wir $x(n-y) + z \geq y(n-x) + z \geq \left(\frac{n+1}{2}\right)^2$, $n \geq x \geq z$, $n \geq y \geq z \geq 0$.

Um diese Ungleichungen zu erfüllen muss $y(n-x)+y \geq \left(\frac{n+1}{2}\right)^2$ gelten (wir ersetzen z durch $y \geq z$). Analog können wir eine ähnliche Ungleichung auch durch die Ersetzung mit $x \geq z$ erhalten. Die erste dieser beiden Ungleichungen gilt für $y \geq x$. Die zweite für $x \geq x$. Wir erhalten also $x = y = \frac{n+1}{2}$.

$\Rightarrow \frac{n+1}{2} \cdot \frac{n-1}{2} + z = x(n-y) + z \geq \left(\frac{n+1}{2}\right)^2$
$\Rightarrow z \geq \frac{n+1}{2}$
Wegen $y \geq z$ folgt nun $z = \frac{n+1}{2}$.

Also haben wir $A^* = B^* = C^*$, d.h. C^* spant einen $K_{\frac{n+1}{2}}$ in G auf: Für $i \neq j$ in C^* gilt $\{i,j\} \in E(G)$.

Damit ist die Rückrichtung ebenfalls erledigt.

Es folgt also die Behauptung.
Damit folgt auch sofort, dass die angegebene Konstruktion eine polynomielle Reduktion ist, d.h. `BiClique` ist NP-vollständig. □

Lemma 2.13. (Übung (C_4-Zählung))
$\forall \gamma, d > 0 : \exists n, n_0 : (X,Y)$ ε-regulär $\Rightarrow |\{C_4 \subseteq G\}| \leq (1+\gamma)d^4n^4$

Beweis:
Seien $\varepsilon \ll \gamma, d$ und (X,Y) ε-regulär.
$\Rightarrow \exists X' \subseteq X, |X'| < 2\varepsilon n : \forall x \in X \setminus X' : |N(x) \cap Y| \in [(d-\varepsilon)n, (d+\varepsilon)n]$
$\Rightarrow \forall x : \exists X_x : |X_x| < 2\varepsilon n \wedge \forall \tilde{x} \in X \setminus X_x : |N(x) \cap N(\tilde{x}) \cap Y| \in [(d-\varepsilon)^2 n, (d+\varepsilon)^2 n]$

$$\Rightarrow |\{Z_4 \subseteq G\}| \leq \underbrace{n}_{\text{Knoten in } X} \cdot \underbrace{n}_{\tilde{x}} \cdot \underbrace{((d+\varepsilon)^2 n)^2}_{\text{gemeinsame Nachbarn}} + |X'|n^3 + n \cdot \max_{x \in X \setminus X'}\{|X_x|\}n^2$$
$$\leq (d+\varepsilon)^4 n^4 + 2\varepsilon n^4 + 2\varepsilon n^4 = ((d+\varepsilon)^4 + 4\varepsilon)n^4 \leq (1+\gamma)d^4n^4$$

Wir müssen lediglich $4\varepsilon \leq \frac{\gamma}{2}d^4$ und $\left(1+\frac{\varepsilon}{d}\right)^4 \leq 1+\gamma 2$ haben. □

Vorlesung am 25.06.2008

Lemma 2.14. (Spezielles `BiClique`-Problem) *Das folgende Problem ist NP-vollständig:*
Gegeben ist ein bipartiter Graph $G = (A\dot{\cup}B, E)$ mit $|A| = |B| = n$ und $|E| = \frac{n^2}{2} - 1$.
Gefragt ist, ob $K_{\frac{n}{2},\frac{n}{2}} \subseteq G$ gilt.

Beweis:
Wir reduzieren `BiClique` auf das Problem dieses Lemmas.
Sei $(G = (A\dot{\cup}B, E), k)$ eine Instanz von `BiClique`.

1. Fall: $k = \frac{n}{2}$

   Hier sind wir so weit erst einmal fertig.

2. Fall: $k \geq \frac{n}{2}$

   Wir fügen jeweils $2k-n$ isolierte Knoten zu A und B hinzu.

   $\Rightarrow k = \frac{n+(2k-n)}{2} = \frac{n'}{2}$

3. Fall: $k < \frac{n}{2}$

   Wir fügen jeweils $n-2k$ Knoten zu A und B hinzu und verbinden diese Knoten mit allen Knoten der jeweils anderen Klasse, womit wir den Graphen G' erhalten.

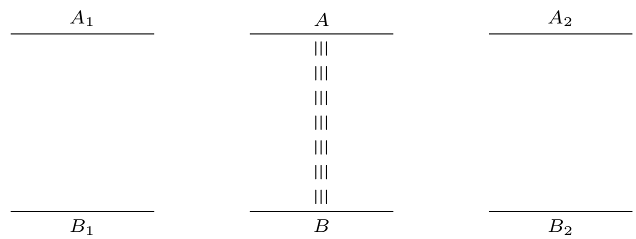
   $\Rightarrow k + (n-2k) = n-k = \frac{2n-2k}{2} = \frac{n+(n-2k)}{2} = \frac{n'}{2}$

   $\Rightarrow K_{k,k} \in G \Leftrightarrow K_{n-k,n-k} \subseteq G'$

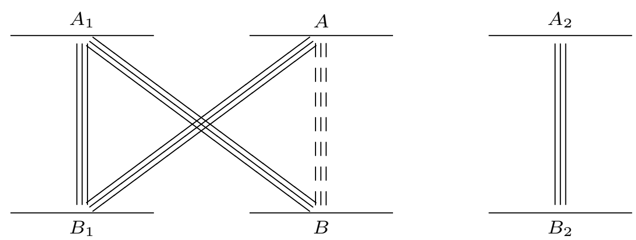
Demnach können wir annehmen, dass $n \in 2\mathbb{N}$ und $k = \frac{n}{2}$ gilt.

1. Fall: $|E| < \frac{n^2}{2} - 1$

   - Falls $|E| < \frac{n^2}{4}$, dann gilt $K_{\frac{n}{2},\frac{n}{2}} \not\subseteq G$, weil zu wenig Kanten vorhanden sind.
   - Sonst fügen wir zu A und B jeweils zwei Knotenklassen A_1 und A_2 bzw. B_1 und B_2 der Größe $\frac{n}{2}$ hinzu.

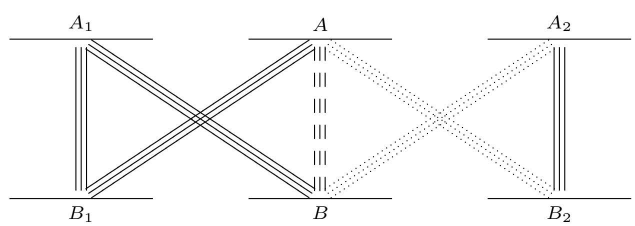
     $G \quad =$ 

     Zwischen $(A_1, B_1), (A_2, B_2), (A_1, B), (B_1, A)$ fügen wir vollständige bipartite Graphen ein.

     $G^* \quad =$ 

     Sei $G^* = (A^* \dot\cup B^*, E^*)$ der modifizierte Graph.
     $\Rightarrow |E^*| = |E| + 2(\frac{n}{2})^2 + 2 \cdot n \cdot \frac{n}{2} \leq \frac{n^2}{2} - 1 + \frac{n^2}{2} + \frac{2n^2}{2} = \frac{(2n)^2}{2} - 1 = \frac{(n^*)^2}{2} - 1$, wobei $n^* = |A^*| = |B^*|$.
     Falls $K_{\frac{n}{2},\frac{n}{2}} \subseteq G \Rightarrow G^*[A_1 \dot\cup A_2, B_1 \dot\cup B_2] \supseteq K_{n,n}$. Falls andererseits $G^* \supseteq K_{n,n}$, dann enthält dieser $K_{n,n}$ keinen Knoten aus $A_2 \cup B_2 \Rightarrow G \supseteq K_{\frac{n}{2},\frac{n}{2}}$.

     Wir haben jetzt bereits $|E^*| \leq \frac{(n^*)^2}{2} - 1$. Um auch $|E^*| = \frac{(n^*)^2}{2} - 1$ zu erreichen, fügen wir nun Kanten zwischen A_2 und B und B_2 und A ein, sodass keinem Knoten aus $A_2 \cup B_2$ mehr als $\frac{n}{2} - 1$ neue Nachbarn hinzugefügt werden.

     $\Rightarrow |E| + 2(\frac{n}{2})^2 + n^2 + n(\frac{n}{2} - 1) \geq \frac{7}{4}n^2 + \frac{n^2}{2} - n \geq \frac{(2n)^2}{2} - 1 = \frac{(n^*)^2}{2} - 1$
     Dies gilt für alle $n \in \mathbb{N}$, d.h. wir können unter der angegebenen Nebenbedingung einen Graphen $G^{**} = (A^* \dot\cup B^*, E^{**})$ konstruieren, für den gilt:

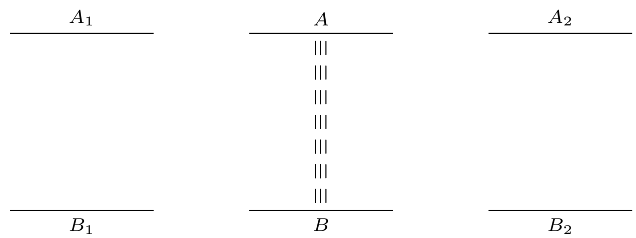
     - $|E^{**}| = \frac{(n^*)^2}{2} - 1$
     - $\forall a \in A_2 : \deg(a) \leq n - 1$
     - $\forall b \in B_2 : \deg(b) \leq n - 1$

     Durch die letzten beiden Eigenschaften erhalten wir $G^{**} \supseteq K_{\frac{n}{2},\frac{n}{2}} \Leftrightarrow G^{**} \supseteq K_{n,n}$ (Beweis wie oben).

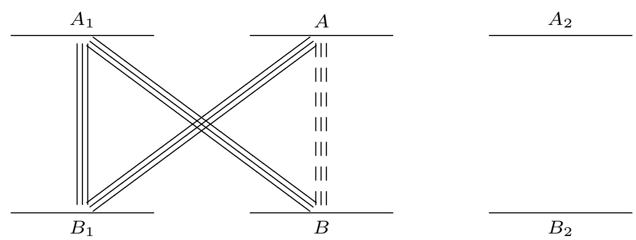
2. Fall: $|E| = \frac{n^2}{2} - 1$

   Hier sind wir fertig.

3. Fall: $|E| > \frac{n^2}{2} - 1$

$G \quad = $ 

Erweitere G zu $G^* = (A^* \dot\cup B^*, E^*)$ mit Knotenklassen A_1, A_2, B_1, B_2 der Größe n und vollständigen bipartiten Graphen zwischen $(A_1, B_1), (A_1, B)$ und (A, B_1).

$G^* \quad = $ 

$\Rightarrow |E^*| = |E| + 3n^2 < \frac{(3n)^2}{2} - 1$ (wie im ersten Fall)

$G^* \supseteq K_{\frac{3n}{2}, \frac{3n}{2}} \Leftrightarrow G \supseteq K_{\frac{n}{2}, \frac{n}{2}}$

Damit können wir nun wie im Fall 1 fortfahren. (Wir verwenden G^* statt G.) $\square$

Vorlesung am 27.06.2008

Idee des algorithmischen Regularitäts-Lemmas - Alon, Duke, Lefmann, Rödl, Yuster (1992)

Für gegebene G und $\varepsilon > 0$ berechne $0 \leq \varepsilon' < \varepsilon$, sodass:

1. Falls G nicht ε-regulär ist, dann wird der Algorithmus dies erkennen und ein Beweisstück ausgeben, welches zeigt, dass G nicht ε'-regulär ist.
2. Falls G ε'-regulär ist, dann erkennt der Algorithmus, dass G ε-regulär ist.

Wenn G ε-regulär, aber nicht ε'-regulär ist, dann kann die Ausgabe des Algorithmus nicht vorhergesagt werden.

Definition

Für einen bipartiten Graphen $G = (X \dot\cup Y, E)$ und $y \neq y' \in Y$ sei $\sigma(y, y') = \big||N(y) \cap N(y')| - d^2|X|\big|$.
Für $Y' \subseteq Y$ sei $\sigma(Y') = \frac{\sum_{(y,y') \in Y'^2, y \neq y'} \sigma(y,y')}{|Y'|^2}$.

Lemma 2.15. *Seien $H = (A \dot\cup B, E)$ bipartit, $|A| = |B| = n$, $\frac{|E|}{|A||B|} = d$ und $0 < \varepsilon < \frac{1}{16}$.*
Falls ein $Y \subseteq B$ mit $|Y| \geq \varepsilon n$ und $\sigma(Y) \geq \frac{\varepsilon^3 n}{2}$ existiert, dann gilt mindestens eine der folgenden Aussagen:

1. $d < \varepsilon^3$
2. $\exists B' \subseteq B, |B'| \geq \frac{\varepsilon^4}{8} n$ *und* $\forall b \in B' : \big||N(b)| - dn\big| \geq \varepsilon^4 n$
3. $\exists A' \subseteq A, B' \subseteq B, |A'| \geq \frac{\varepsilon^4}{4} n \leq |B'|, |d(A', B') - d| \geq \varepsilon^4$

Darüber hinaus existiert ein Algorithmus, der für gegebenes H und Y (wie oben) genau eines der folgenden ausgibt:

1. *„Die erste Aussage ist wahr.“*

2. *„Die zweite Aussage ist wahr.“ (Wobei B' mit angegeben wird.)*

3. *„Die dritte Aussage ist wahr.“ (Wobei A' und B' ebenfalls angegeben werden.)*

Dieser Algorithmus hat eine Laufzeit von $O(M(n))$, wobei $M(n) = O(n^{2{,}376})$ die Laufzeit für das Multiplizieren zweier Matrizen aus $\{0,1\}^{n\times n}$ ist.

Beweis:
Angenommen, 1. und 2. gelten nicht.
Wir zeigen, dass dann 3. gilt. Sei $Y' = \{y \in Y \big| ||N(y)| - dn| \le \varepsilon^4 n\}$. Weil 2. nicht gilt, haben wir $Y' \neq \emptyset$.
$\Rightarrow \sum_{y'\in Y'} \sum_{y\in Y\setminus\{y'\}} \sigma(y',y) = \sigma(Y)|Y|^2 - \sum_{y'\in Y\setminus Y'} \sum_{y\in Y\setminus\{y'\}} \sigma(y',y) \ge \frac{\varepsilon^3 n}{2}|Y|^2 - \frac{\varepsilon^4 n}{8}|Y|n$
Wegen $|Y'| \le |Y|$ existiert ein $y_0 \in Y'$ mit:

$$\sum_{y\in Y\setminus\{y_0\}} \sigma(y_0,y) \ge \frac{\varepsilon^3 n}{2}\frac{|Y|^2}{|Y'|} - \frac{\varepsilon^4 n}{8}\frac{|Y|}{|Y'|}n \ge \frac{\varepsilon^3 n}{2}|Y| - \frac{\varepsilon^4 n^2}{8} \overset{|Y|\ge\varepsilon n}{\ge} \frac{3}{8}\varepsilon^3 n|Y| \qquad (*)$$

Sei $Y_0 = \{y \in Y \setminus \{y_0\} | \sigma(y_0,y) \ge 2\varepsilon^4 n\}$.
Behauptung: $|Y_0| \ge \frac{\varepsilon^4}{4}n$
Beweis: Falls $|Y_0| < \frac{\varepsilon^4}{4}n$, gilt:

$$\sum_{y\in Y} \sigma(y_0,y) \le \underbrace{\frac{\varepsilon^4}{4}n^2}_{\ge\sum_{y\in Y_0}\sigma(y_0,y)} + \underbrace{|Y|\cdot 2\varepsilon^4 n}_{\ge\sum_{y\in Y\setminus Y_0}\sigma(y_0,y)} \overset{|Y|\ge\varepsilon n}{\le} \frac{\varepsilon^3}{4}|Y|n + 2\varepsilon^4|Y|n \overset{2\varepsilon<\frac{1}{8}}{<} \frac{3}{8}\varepsilon^3 n|Y|$$

Dies ist ein Widerspruch zu (*).
$\Rightarrow B' = Y_0$ und $\forall b \in B' : \big||N(b)\cap N(y_0)| - d^2 n\big| > 2\varepsilon^4 n$
Setze $A' = N(y_0)$. Wegen $y_0 \in Y'$ gilt $|N(y_0)| \ge dn - \varepsilon^4 n > \varepsilon^3 n - \varepsilon^4 n > \frac{\varepsilon^4}{4}n$, weil 1. nicht gilt.
Weiterhin gilt $e(A',B') = \sum_{b\in B'} |N(y_0)\cap N(b)|$.
$\Rightarrow \big|e(A',B'') - d|A'||B''|\big| \overset{(*)}{\ge} \frac{3}{8}\varepsilon^3|Y|\frac{n}{2}$ mit $B'' \subseteq B'$, $|B''| \ge \frac{1}{2}|B'|$ und ohne Beschränkung der Allgemeinheit $\sum_{b\in B''} |N(y_0)\cap N(b)| - d^2 n \ge \frac{3}{8}\varepsilon^3|Y|\frac{n}{2}$
(Wir haben als Summe der Beträge $\frac{3}{8}\varepsilon^3 n|Y|$. Damit ist die Summe der positiven Werte mindestens $\frac{1}{2}\cdot\frac{3}{8}\varepsilon^3 n|Y|$ oder die Summe der negativen Werte ist mindestens $\frac{1}{2}\cdot\frac{3}{8}\varepsilon^3 n|Y|$: Es gilt entweder $\sum_{b\in B''} |N(y_0)\cap N(b)| - d^2 n \ge \frac{3}{8}\varepsilon^3|Y|\frac{n}{2}$ oder $\sum_{b\in B'\setminus B''} |N(y_0)\cap N(b)| - d^2 n \le -\frac{3}{8}\varepsilon^3|Y|\frac{n}{2}$.)
$\Rightarrow |d(A',B'') - d| \ge \frac{3}{16}\varepsilon^3\frac{n}{|B''|} \ge \varepsilon^4$
(B'' ist zwar kleiner als B', aber wir haben immer noch $|B''| \ge \frac{\varepsilon^4}{8}n$.)

Algorithmus:

1. Alle Kanten zählen. $\Rightarrow O(n^2)$

2. Alle Nachbarschaften prüfen. $\Rightarrow O(n^2)$

   Dabei kann dann auch gleich B' erzeugt werden.

3. Wir betrachten gemeinsame Nachbarschaften, welche den Wegen der Länge 2 entsprechen. Die Anzahl dieser Wege kann von der quadrierten Adjazenzmatrix abgelesen werden.

   Weil Matrixmultiplikationen mit der gewünschten Laufzeit möglich sind, hat der Gesamtaufwand des Algorithmus die geforderte Komplexität.

Vorlesung am 02.07.2008

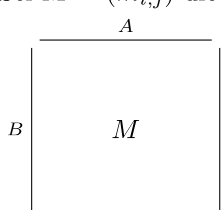
Lemma 2.16. *Sei $H = (A\dot\cup B, E)$ mit $|A| = |B| = n$ und $dn^2 = |E|$ ein bipartiter Graph und $2n^{-\frac{1}{4}} < \varepsilon < \frac{d}{16}$. Falls $\left|\{x \in A\dot\cup B \big| ||N(x)| - dn| > \varepsilon^4 n\}\right| < \frac{1}{8}\varepsilon^4 n$, dann existiert $Y \subseteq B$ mit $|Y| \ge \varepsilon n$ und $\sigma(Y) \ge \frac{\varepsilon^3 n}{2}$, falls H nicht ε-regulär ist. Dabei ist $\sigma(Y) = \sum_{y,y'\in Y}\frac{\sigma(y,y')}{|Y|^2}$ und $\sigma(y,y') = |N(y)\cap N(y')| - d^2 n$.*

Beweis:
Angenommen: $\forall Y \subseteq B, |Y| \geq \varepsilon n : \sigma(Y) < \frac{\varepsilon^3 n}{2}$.
Wir zeigen: H ist ε-regulär.
Dann reicht es aus $|d(A,B) - d(X,Y)| \leq \varepsilon$ für alle $Y \subseteq B, x \subseteq A$ mit $|X| = |Y| = \lceil \varepsilon n \rceil$ zu zeigen. (*)
(Angenomen, es gäbe zunächst nur ein größeres Mengenpaar, welches eine zu große oder zu kleine Dichte hat. Dann können wir einfach die Teilmengen der gewünschten Größe betrachten. Mindestens eine dieser Teilmengen muss dann weit genug von dem Sollwert abweichen.)

Behauptung:
Für alle $X \subseteq A, Y \subseteq B, |X| = |Y| = \lceil \varepsilon n \rceil$ gilt

$$\sum_{x \in X} (|N(x) \cap Y| - d|Y|)^2 \leq e(A,Y) + |Y|^2 \sigma(Y) + \frac{2}{5}\varepsilon^5 n^3.$$

Beweis der Behauptung:
Sei $M = (m_{i,j})$ die (bipartite) Adjazenzmatrix von H der größe $n \times n$:



Dann gilt:
$$\begin{aligned}\sum_{x \in X} (|N_y(x)| - d|Y|)^2 &\leq \sum_{x \in A} (|N_y(x)| - d|Y|)^2 = \sum_{x \in A} ((\sum_{y \in Y} m_{x,y}) - d|Y|)^2 \\ &= \sum_{x \in A} ((\sum_{y \in Y} m_{x,y}^2) + d^2|Y|^2 + \sum_{y \neq y' \in Y} m_{x,y} m_{x,y'} - 2 \sum_{y \in Y} m_{x,y} d|Y| \\ &= e(A,Y) + nd^2|Y|^2 + \sum_{y \neq y'} |N(y) \cap N(y')| - 2e(A,Y)d|Y| \\ &= e(A,Y) + nd^2|Y|^2 + \sum_{y \neq y'} (\sigma(y,y') + d^2 n) - 2e(A,Y)d|Y| \\ &\leq e(A,Y) + \sigma(Y)|Y|^2 + 2nd^2|Y|^2 - 2e(A,Y)d|Y|\end{aligned}$$
Es reicht also zu zeigen, dass $nd^2|Y|^2 - e(A,Y)d|Y| \leq \frac{\varepsilon^5 n^3}{5}$ gilt.
Diese Ungleichung ist äquivalent zu $d - \underbrace{\frac{e(A,Y)}{n|Y|}}_{=d(A,Y)} \leq \frac{\varepsilon^5 n^2}{5d|Y|^2}$ bzw. $d(A,Y) \geq d - \frac{\varepsilon^5 n^2}{5d|Y|}$.
Es gilt $d(A,Y) = \frac{e(A,Y)}{n|Y|} \geq \frac{((d-\varepsilon^4)n)(|Y| - \frac{1}{8}\varepsilon^4 n)}{n|Y|} = d - \varepsilon^4 - \frac{d\varepsilon^4 n}{|Y|} + \frac{\varepsilon^8 n}{8|Y|} \geq d - \varepsilon^4 - \frac{\varepsilon^3}{8}$.
Da $|Y| \leq \varepsilon n + 1$, $d < 1$ und $1 < \varepsilon^4 n$, gilt $\frac{\varepsilon^5 n^2}{5d|Y|^2} \geq \frac{\varepsilon^5 n^2}{5d(1+\varepsilon n)^2} \geq \frac{\varepsilon^5}{5(\varepsilon^4 + \varepsilon)^2} \geq \varepsilon^4 + \frac{1}{8}\varepsilon^3$.
Damit ist die Behauptung beweisen.

Weiterhin gilt (wahlweise nach Jensen oder CS):

$$\sum_{x \in X} (|N_y(x)| - d|Y|)^2 \geq \frac{1}{|X|}((\sum_{x \in X} |N_y(x)|) - d|X||Y|)^2.$$

Mit der Behauptung folgt außerdem:
$((\sum_{x \in X} |N_y(x)|) - d|X||Y|)^2 \leq |X|(e(A,Y) + |Y|^2 \sigma|Y| + \frac{2}{5}\varepsilon^5 n^3)$
$\Leftrightarrow (d(X,Y) - d)^2 \leq \frac{1}{|X||Y|^2}(e(A,Y) + |Y|^2\sigma|Y| + \frac{2}{5}\varepsilon^5 n^3)$ (Division durch $|X|^2|Y|^2$)
Es reicht nun, zu zeigen dass die rechte Seite kleiner oder gleich ε^2 ist:
$$\begin{aligned}\frac{1}{|X||Y|^2}((d + \varepsilon^4 n)|Y| + \frac{1}{8}\varepsilon^4 n^2 + \frac{\varepsilon^3 n}{2}|Y|^2 + \frac{2}{5}\varepsilon^5 n^3) &\leq \frac{1+\varepsilon^4 n}{\varepsilon^2 n^2} + \frac{\frac{1}{8}\varepsilon^4 n^2}{\varepsilon^3 n^3} + \frac{\varepsilon^3 n}{2\varepsilon n} + \frac{\frac{2}{5}\varepsilon^5 n^3}{\varepsilon^3 n^3} \\ &\leq \frac{1}{\varepsilon^2 n^2} + \frac{\varepsilon^2}{n} + \frac{\varepsilon}{8n} + \frac{\varepsilon^2}{2} + \frac{2\varepsilon^2}{5} \leq \varepsilon\end{aligned}$$
Die vorletzte Ungleichung gilt für ausreichend großes n.
Damit folgt nun das Lemma. $\square$

Lemma 2.17. *Sei $H = (A\dot{\cup}B, E)$ bipartit, $|A| = |B| = n$, $2n^{-\frac{1}{4}} < \varepsilon < \frac{d}{16}$ und $dn^2 = |E|$. Dann existiert ein $O(M(n))$ Algorithmus, der entweder verifiziert, dass H ε-regulär ist oder Mengen $A' \subseteq A$ und $B' \subseteq B$ mit $|A'|, |B'| \geq \frac{\varepsilon^4}{16}n$ ausgibt, sodass $|d(A', B') - d| \geq \varepsilon^4$.*

Beweis:
Die vorhergehenden Lemmas. □

Bemerkung
Durch diesen Algorithmus kann nun der Beweis des Regularitätslemmas realisiert werden, d.h. für alle $\varepsilon > 0$ existiert ein Algorithmus mit Laufzeit $O(M(n))$, der eine ε-reguläre Partition eines Eingabegraphen $G = (V, E)$ findet.

Anwendung
Betrachten wir das folgende Problem: Gegeben ist ein Graph $G = (V, E)$ und gesucht ist $V' \subseteq V$ sodass $e(V', V \setminus V')$ maximal ist.
Dieses Problem ist NP-vollständig und nicht beliebig genau approximierbar.
Aber:
Für alle $\varepsilon > 0$ existiert ein Approximationsalgorithmus, der einen Cut der Größe $MC(G) - \varepsilon n^2$ findet, wobei $MC(G)$ die Lösung von MaxCut bezeichnet. Diese Näherung ist also besonders für dichte Graphen sinnvoll.

Vorlesung am 04.07.2008

Problem
Gegeben sind F mit l Knoten, $\mu > 0$ und ein Graph G mit n Knoten. Gesucht ist die Anzahl der Kopien von F in G.
Dieses Problem ist für festes l in polynomieller Zeit trivial lösbar. Selbst wenn l kein Teil der Eingabe ist, ist die Laufzeit des trivialen Algorithmus für große l jedoch sehr hoch. Daher werden wir die Anzahl der Vorkommen von F nur mit einer Genauigkeit von μn^l approximieren.

Satz 2.8. *Für alle $\mu > 0, F, V(F) = l$ existiert ein $O(n^{2,3\cdots})$-Algorithmus, der für gegebenen Graphen $G = (V, E)$ mit $|V| = n$ die Anzahl $|\{F \subseteq G\}|$ bis auf einen Fehler von μn^l approximiert.*

Beweis:
Idee:

1. Wir regularisieren G und bilden den reduzierten Graphen R, dessen Größe von n unabhängig und damit konstant ist. Der Graph R hat für jede Knotenklasse des regularisierten Graphen G einen Knoten und die Kanten, welche reguläre Paare von G repräsentieren, haben als Gewicht die Dichte der Kanten zwischen den entsprechenden Knotenklassen in G.

2. Dann berechnen wir $\sum_{F' \cong F, F' \subseteq R} \prod_{\{i,j\} \in E(F')} w(i, j) = \#_{gew}\{F \subseteq R\}$.

3. Behauptung: $|\{F \subseteq G\}| \approx \#_{gew}\{F \subseteq R\}(\frac{n}{t})^l$

Ausführlicher:
Seien $\mu > 0, F$ und l gegeben. Wir wählen d_0 und ε, sodass $\min\{\frac{1}{l}, \mu\} \gg d_0 \gg \varepsilon > 0$.
Nun regularisieren wir den Graphen G mit $t_0 \ll \min\{\frac{1}{l}, \mu\}$ und erhalten $V(G) = V = V_0 \dot{\cup} V_1 \dot{\cup} ... \dot{\cup} F_t$ mit $t_0 \leq t \leq T_0 = T_0(\varepsilon, t_0)$.
Wir betrachten den reduzierten Graphen $R = R(d_0, \varepsilon, G)$ mit $V(R) = \{1, ..., t\}$ und:
$E(R) = \{\{i, j\} \in \binom{V(R)}{2} | (V_i, V_j)$ ist ε-regulär, $d(V_i, V_j) = \frac{e(V_i,V_j)}{|V_i||V_j|} \geq d_0\}$.
Jetzt zählen wir die Anzahl der markierten F in G. (Wir können anschließend einfach durch die Anzahl der Automorphismen von F dividieren, wenn wir die Anzahl der unmarkierten Kopien

bestimmen wollen.) Sei $\#_{gew}\{F \subseteq R\}$ die Summe aller gewichteten (markierten) Kopien von F in R, d.h. $\sum_{F' \in \binom{R}{F}} \prod_{\{i,j\} \in E(F')} w(i,j)$, wobei $\binom{R}{F}$ die Menge aller Kopien von F im vollständigen Graphen mit Knotenmenge $\{1, ..., t\}$. Dabei sei $w(i,j) = \begin{cases} 0 & (V_i, V_j) \text{ nicht } \varepsilon-\text{regulär} \\ 0 & d(V_i, V_j) < d_0 \\ d(V_i, V_j) & \text{sonst} \end{cases}$,
d.h. $w(i,j)$ ist genau dann 0, wenn $\{i,j\} \notin E(R)$.
Behauptung: $\#_{gew}\{F \subseteq R\}m^l = \#\{F \subseteq G\} \pm \mu n^l$, wobei $m = |V_1| = ... = |V_t|$.
Offensichtlich folgt der Satz direkt aus dieser Behauptung, da in $O(n^{2.3\cdots})$ die Partitionen bestimmt werden kann. Anschließend kann in $O(n^{2.3})$ R berechnet werden. Zuletzt ist es in konstanter Zeit möglich, $\#_{gew}\{F \subseteq R\}$ zu berechnen und mit m^l zu multiplizieren.

Beweisen wir nun die Behauptung:
Betrachte $G' \subseteq G$, definiert durch $V(G') = V(G) \setminus V_0$ und $E(G') = \bigcup_{\{i,j\} \in E(R)} E(V_i, V_j)$.
$\Rightarrow \#\{F \subseteq G'\} \geq \#\{F \subseteq G\} - |E(G) \setminus E(G')|\binom{n}{l-2}|Aut(F)| \geq \#\{F \subseteq G\} - l^2|E(G) \setminus E(G')|n^{l-2}$
(wegen $|Aut(F)| \leq l!$)
Es gilt $|E(G) \setminus E(G')| \leq \frac{\mu}{2l^2}n^2$, da:

$$|E(G) \setminus E(G')| \leq \underbrace{\underbrace{|V_0|n}_{Kante\ berührt\ V_0}}_{\leq \varepsilon n^2 \ \rightsquigarrow\ \varepsilon \leq \frac{\mu}{8l^2}} + \underbrace{\underbrace{d_0\binom{t}{2}\left(\frac{n}{t}\right)^2}_{zu\ dünn}}_{\leq \frac{d_0}{2}n^2 \ \rightsquigarrow\ d_0 \leq \frac{\mu}{4l^2}} + \underbrace{\underbrace{\varepsilon t^2\left(\frac{n}{t}\right)^2}_{irregulär}}_{\leq \varepsilon n^2} + \underbrace{\underbrace{t\binom{\frac{n}{t}}{2}}_{Kante\ innerhalb\ von\ V_i}}_{\leq \frac{n^2}{t} \ \rightsquigarrow\ t \geq t_0 \geq \frac{8l^2}{\mu}}$$

$\Rightarrow \#\{F \subseteq G'\} \geq \#\{F \subseteq G\} - \frac{\mu}{2}n^l$
Weiterhin gilt $\#\{F \subseteq G'\} = (1 \pm \gamma(\varepsilon)) \cdot \#_{gew}(F \subseteq R) \cdot m^l \pm \frac{\mu}{4}n^l$ wobei $\gamma(\varepsilon) \to 0$ mit $\varepsilon \to 0$ für festes d_0. (**)

Sei F' eine Kopie von F in G'.

1. Fall: $\exists i : |V(F') \cap V_i| \geq 2$

   Die Anzahl solcher Kopien ist höchstens $t\binom{\frac{n}{t}}{2} \cdot \binom{n}{l-2}|Aut(F)| \leq \frac{l^2 n^l}{2t} \leq \frac{\mu}{4}n^l$ mit $t \geq t_0 \geq \frac{2l^2}{\mu}$.

2. Fall: $\forall i : |V(F') \cap V_i| \leq 1$

   Weil $\forall \{x,y\} \in E(F') : (V_i, V_j)$ ε-regulär ist und $d(V_i, V_j) \geq d_0$ mit $\{x,y\} \in E(V_i, V_j)$ gilt, folgt:

   Seien $V_{i_1}, ..., V_{i,l}$ die Knotenklassen, die $V(F')$ schneiden. Dann erfüllt $G[V_{i_1}, ..., V_{i_l}]$ die Voraussetzungen des Zähllemmas.

   $\Rightarrow |\{F \subseteq G[V_{i_1}, ..., V_{i_l}]\}| = (1 \pm \gamma(\varepsilon)) \prod_{1 \leq j < k \leq l, E(F') \cap E(V_{i,j}, V_{i,k}) \neq \emptyset} d(V_{i_j}, V_{i_k})m^l$

   Daher liegt die Anzahl der Kopien F' von F in G, die diesen Fall erfüllen, im Bereich

   $$(1 \pm \gamma(\varepsilon)) \underbrace{\sum_{\tilde{F} \subseteq \binom{R}{F}} \prod_{\{i,j\} \in E(\tilde{F})} w(i,j)}_{= \#_{gew}\{F \subseteq R\}} \cdot m^l.$$

Damit folgt (**).

Aus (*) und (**) folgt $\#\{F \subseteq G\} = \#_{gew}\{F \subseteq R\} \pm \frac{3\mu}{4}n^l + \underbrace{\gamma(\varepsilon)\#_{gew}\{F \subseteq R\}m^l}_{\leq \gamma(\varepsilon)t^l m^l \leq \gamma(\varepsilon)n^l}$. Wir wählen also $\gamma(\varepsilon) \leq \frac{\mu}{4}$.
Sei $\varepsilon > 0$ klein genug, sodass $\gamma(\varepsilon) \leq \mu 4$. □

Konstanten:
Für l und $\mu > 0$ setzen wir $d_0 = \frac{\mu}{4l^2}$.

$\varepsilon > 0$ sei die Konstante, welches das Countinglemma für d_0, F und $\gamma = \frac{\mu}{4}$ garantiert.
Jetzt wählen wir $\varepsilon = \min\{\varepsilon_0, \frac{\mu}{8l^2}\}$ und setzen n_0' groß genug sowie $t_0 = \frac{8l^2}{\mu}$. Mit dem durch das Regularitätslemma (angewandt auf t_0 und ε) gegebene T_0 bestimmen wir nun $n_0 \geq 2T_0 n_0'$.

Vorlesung am 09.07.2008

Problem: MaxCut
Seien $G = (V, E)$ und $U \subseteq V$. Dann heißt $(U, \overline{U} = V \setminus U)$ Cut bzw. Schnitt von G.
$e(U, \overline{U}) = \{\{x, y\} \in E | x \in U, y \in \overline{U}\}$ heißt Wert des Schnittes.

MaxCut: $mc(G) = \max_{U \subseteq V} e(U, \overline{U})$

MinCut: $\min_{U \subseteq V} e(U, \overline{U})$
$\rightsquigarrow$ MaxFlow-MinCut, MinCut $\in$ P

Bemerkung
MaxCut is NP-vollständig.
Es existiert kein polynomieller Approximationsalgorithmus, der eine $\frac{16}{17} + \varepsilon$ Approximationsgüte hat, falls P$\neq$NP.
Aber: Es existiert ein $\frac{16}{17} - \varepsilon$-Approximationsalgorithmus.

Satz 2.9. (Frieze-Kannen (1996)) *Für alle $\gamma > 0$ existiert ein polynomieller Algorithmus, der für einen gegebenen Graphen $G = (V, E)$ mit $|V| = n$ einen Cut $(U, \overline{U})$ mit $e(U, \overline{U}) \geq mc(G) - \gamma n^2$ berechnet.*

Beweis:
Idee: Zunächst wird der Graph regularisiert. Dann betrachten wir wieder den reduzierten Graphen R mit $w(i, j) = d(V_i, V_j)$ und lösen für diesen Graphen das folgende Optimierungsproblem:
Gegeben ist $R = ([t], E_r, w : E_R \to [0, 1])$.
Gesucht ist $(\alpha_1, ..., \alpha_t) \in [0, 1]^t$ sodass $\sum_{i=1}^t \sum_{j=1}^t \alpha_i(1-\alpha_j)w(i, j)$ maximal ist.Dabei sei $w(i, i) = 0$ für $i = 1, ..., t$.
Dieses Problem diskretisieren wir und lösen es in $\text{\Huge O}(1)$. (Ja, das ist ein gaaaaanz großes O.)

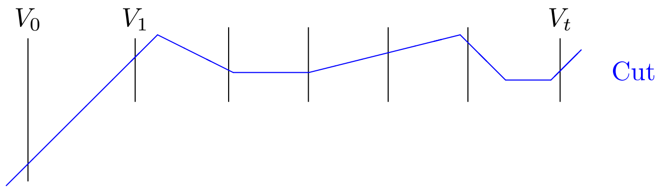
Behauptung: $mc(R) \cdot m^2 \geq mc(G) - \gamma n^2$ mit $m = |V_1| = ... = |V_t|$.
Beweis: Für gegebenes $\gamma > 0$ wählen wir $0 < \varepsilon \ll \gamma$. Durch das Regularitätslemma mit $t_0 \gg \frac{1}{\gamma}$ erhalten wir $T_0 = T_0(\varepsilon, t_0)$ und $n_0 = n_0(\varepsilon, t_0)$.
Sei $G = (V, E)$ gegeben. Falls $e(G) \leq \gamma n^2$ geben wir irgendeine Knotenmenge aus.
Sei also $e(G) > \gamma n^2$. Mit dem Regularitätslemma erhalten wir $V_0 \dot\cup ... \dot\cup V_t$ mit $t_0 \leq t \leq T_0, |V_0| \leq \varepsilon n$ und $|V_1| = ... = |V_t| = m \geq \frac{(1-\varepsilon)n}{t}$. ($\rightsquigarrow$ Aufwand von $O(M(n))$) Sei $R = ([t], E_R, w : R_R \to [0, 1])$ der gewichtete reduzierte Graph, wobei:
$\{i, j\} \in E_R \Leftrightarrow (V_i, V_j)$ ist ε-regulär, $w(i, j) = d(V_i, V_j)$ für alle $\{i, j\} \in E_R$ und $w(i, j) = 0$ für $\{i, j\} \notin E_R$. ($\rightsquigarrow$ Aufwand von $O(n^2)$)
Jetzt suchen wir eine maximierende Zuordnung $(\alpha^*, ..., \alpha_t^*) \in \{\underbrace{0, \varepsilon, 2\varepsilon, ..., 1}_{\frac{1}{\varepsilon}+1 \text{ Elemente}}{}^t$. ($\rightsquigarrow O(1)$)

Für $\underline{\alpha}^* = (\alpha_1^*, ..., \alpha_t^*)$ definieren wir einen cut $U_{\underline{\alpha}^*}$ in G wie folgt:
Für $i = 1, ..., t$ sei $U_{\underline{\alpha}^*, i} \subseteq V_i$ von der Größe $\lfloor \alpha_i^* m \rfloor$ (wir werden jedoch in der folgenden Rechnung der Einfachheit halber nicht mehr abrunden) und $U_{\underline{\alpha}^*} \bigcup_{i=1}^t U_{\underline{\alpha}^*, i}$.

$$\begin{aligned}
\Rightarrow e(U_{\underline{\alpha}^*, 1}, \overline{U}_{\underline{\alpha}^*}) &\geq \sum_{\{i,j\} \in E_R} \left|U_{\underline{\alpha}^*, i}\right| \left|V_j \setminus U_{\underline{\alpha}^*, j}\right| (d(V_i, V_j) - \varepsilon) + \left|V_i \setminus U_{\underline{\alpha}^*, i}\right| \left|U_{\underline{\alpha}^*, j}\right| (d(V_i, V_j) - \varepsilon) \\
&\geq m^2 \sum_{\{i,j\} \in E_R} (\underbrace{\alpha_i^*(1 - \alpha_j^*) + (1 - \alpha_i^*)\alpha_j}_{\leq 1})(w(i, j) - \varepsilon) \\
&\geq cut(R, \underline{\alpha}^*) m^2 - \varepsilon n^2
\end{aligned}$$

($\rightsquigarrow$ Aufwand von $O(n)$)

Zu zeigen ist nun: $cut(R,\underline{\alpha}^*)m^2 \geq mc(G) - \frac{\gamma}{2}n^2$
Idee:



Sei $U \subseteq U$ sodass $e(U,\overline{U}) = mc(G)$.
Seien $U_i = U \cap V_i$ für $i = 1,...,t$, $\beta_i = \frac{|U_i|}{m}$ und $k_i \in \mathbb{N} : k_i\varepsilon \leq \beta_i < k_i\varepsilon + \varepsilon$ mit $k_i \in \{0,...,\frac{1}{\varepsilon}\}$ sowie $\beta_i^* = k_i\varepsilon$. Zuletzt definieren wir $\underline{\beta}^* = \{\beta_i^*,...,\beta_t^*\}$.

1. $cut(R,\underline{\beta}^*) \leq cut(R,\underline{\alpha}^*)$, da $\underline{\beta}^* \in \{0,\varepsilon,...,1\}^t$

2. $cut(R,\underline{\beta}^*) \geq cut(R,\underline{\beta}) - 2\varepsilon t^2$ mit $\underline{\beta} = (\beta_1,...,\beta_t)$, weil:
$$\beta_i^*(1-\beta_j^*) \geq (\beta_i - \varepsilon)(1-\beta_j-\varepsilon) \geq \beta_i(1-\beta_j) - \varepsilon\underbrace{(1-\beta_j)}_{\leq 1} + \varepsilon^2 - \varepsilon\underbrace{\beta_i}_{\leq 1} \geq \beta_i(1-\beta_j) - 2\varepsilon$$

3. $cut(R,\underline{\beta}) \cdot m^2 \geq e(U,\overline{U}) - (5\varepsilon + \frac{1}{t_0}n^2$
$$\begin{aligned} e(U,\overline{U}) \leq\ & \underbrace{|\{\{x,y\} \in E | x \in V_0\}|}_{\leq \varepsilon n^2} + \underbrace{\sum_{1\leq i<j\leq t, (V_i,V_j)\text{nicht } \varepsilon-\text{regulär}} e(V_i,V_j)}_{\leq \varepsilon n^2} \\ & + \underbrace{\sum_{i=1,...,t,\beta_i<\varepsilon} \varepsilon|V_i|n}_{\leq \varepsilon n^2} + \underbrace{\sum_{i=1,...,t,1-\beta_i<\varepsilon} \varepsilon|V_i|n}_{\leq \varepsilon n^2} + \underbrace{\sum_{i=1,...,t} e(V_i)}_{\leq \frac{n^2}{t_0}} \\ & + \sum_{\substack{1\leq i<j\leq t \\ \beta_i \geq \varepsilon \\ 1-\beta_i \geq \varepsilon \\ \beta_j \geq \varepsilon \\ 1-\beta_j \geq \varepsilon}} \beta_i(1-\beta_j)m^2(d(V_i,V_j)+\varepsilon) + (1-\beta_i)\beta_j m^2(d(V_i,Vj)+\varepsilon) \\ \leq\ & 4\varepsilon n^2 + \frac{n^2}{t_0} + cut(R,\underline{\beta}) + \varepsilon n^2 \\ \leq\ & 5\varepsilon + \frac{1}{t_0}n^2 + cut(R,\underline{\beta}) \end{aligned}$$

$\Rightarrow m^2 cut(R,\underline{\alpha}^*) \geq m^2 cut(R,\underline{\beta}^*) \geq m^2 cut(R,\underline{\beta}) - 2\varepsilon t^2 m^2 \geq e(U,\overline{U}) - 7\varepsilon n^2 - \frac{n^2}{t_0} \geq e(U,\overline{U}) - \frac{\gamma}{2}n^2$,
wobei die letzte Ungleichung für $\varepsilon \leq \frac{\gamma}{28}$ und $t_0 \geq \frac{4}{\gamma}$ gilt.
$\Rightarrow e(U_{\underline{\alpha}^*}\overline{U}_{\underline{\alpha}^*}) \geq mc(G) - \gamma n^2$ $\square$

Bemerkung
Es existiert sogar ein randomisierter Algorithmus, der den Wert des Schnittes in konstanter Zeit bestimmt. Um den Schnitt selbst ausgeben zu können, ist jedoch ein linearer Zeitaufwand nötig.

Vorlesung am 11.07.2008

2.4 Generalisierungen des Removal Lemmas

Erinnerung: Removal Lemma
$\forall F, \eta > 0 : \exists c > 0, n_0 \in \mathbb{N} : \forall G = (V,E), |V| = n \geq n_0 :$
$|\{F \subseteq G\}| \leq cn^{v(F)} \Rightarrow \exists E' \subseteq E, |E'| \leq \eta n^2, G' = (V, E \setminus E')$ ist F-frei.

Definition
$\mathcal{P}$ sei eine *Grapheneigenschaft*, d.h. eine unter Isomorphie abgeschlossene Familie von Graphen. Weiterhin definieren wir $\mathcal{P}_n = \{H \in \mathcal{P} | v(H) = n\}$. Für $\eta > 0$ ist $G = (V, E)$ mit $|V| = n$ *η-weit* von $\mathcal{P}$ falls alle Graphen $G' = (V, E')$ mit $|E' \vartriangle E| \leq \eta n^2$ nicht in $\mathcal{P}_n$ liegen.

Bemerkung
Mit dieser Definition lässt sich das Removal Lemma wie folgt darstellen:
Wenn G η-weit von F-frei ist, dann gilt $|\{F \subseteq G\}| > cn^{v(F)}$.

Ein Korollar hiervon ist: $\exists k = k(c) : \mathbf{P}(G[U] \supseteq F) \geq 0.99$ für $U \in \binom{[n]}{k}$ (zufällig und gleichverteilt)
Es existiert also ein randomisierter Algorithmus mit konstanter Laufzeit und den folgenden Eigenschaften:

1. Falls G F-frei ist, dann gibt der Algotithmus dies mit Wahrscheinlichkeit 1 aus.
2. Falls G η-weit von F-frei ist, dann gibt der Algorithmus „G ist nicht F-frei“ mit Wahrscheinlichkeit 0.99 aus.
3. In allen übrigen Fällen gibt es keine Garantie für die Ausgabe des Algorithmus.

Definition
Algorithmen mit den Eigenschaften 1. bis 3. heißen *einseitige Tester* oder auch *Tester mit einseitigem Fehler*.
(Graphen)Eigenschaften, für die es Tester (mit einseitigem Fehler) für jedes $\eta > 0$ gibt, heißen *testbar* (*mit einseitigem Fehler*).

Frage
Welche (entscheidbaren) Grapheneigenschaften sind testbar?

- „$|E|$ gerade“ ist nicht testbar.
- „$\mathcal{P} = Forb(F)$“ ist testbar.

Definition (monotone Eigenschaften)
Grapheneigenschaften, die abgeschlossen unter Untergraphen sind, d.h. $G \in \mathcal{P} \Rightarrow G' \in \mathcal{P} \forall G' \subseteq G$, nennt man *monoton*.

Beispiel

- Für alle F ist „$G \in Forb(F)$“ monoton.
- „$\{G | \chi(G) \leq k\}$“ ist ebenfalls monoton.
- „$|E|$ gerade“ ist nicht monoton.

Bemerkung
Für alle monotonen Eigenschaften $\mathcal{P}$ existiert eine Familie $\mathcal{F} = \mathcal{F}_{\mathcal{P}}$ sodass $\mathcal{P} = Forb(\mathcal{F})$. (Wie wählen einfach $\mathcal{F} = \mathcal{G} \setminus \mathcal{P}$.) Sei $\mathcal{F}_{\min}$ die inklusionsminimale Teilfamilie von $\mathcal{F}$, d.h.

1. $\forall F, F' \in \mathcal{F}_{\min}$ gilt $F \not\subset F'$
2. $\forall F \in \mathcal{F} : \exists F' \in \mathcal{F}_{\min} : F' \subseteq F$

$\Rightarrow \mathcal{P} = Forb(\mathcal{F}) = Forb(\mathcal{F}_{\min})$

Im Allgemeinen gilt jedoch nicht $|\mathcal{F}_{\min}| < \infty$.
Beispiel: $\mathcal{P} = \{G \in \mathcal{G} | \chi(G) \leq 2\}, \mathcal{F} = \{H \in \mathcal{G} | \chi(H) \geq 3\}, \mathcal{F}_{\min} = \{C_{2k+1} | 0 \neq k \in \mathbb{N}\}$

Lemma 2.18. (Removal Lemma für monotone Eigenschaften (Alon, Shapira (2006))) *Für alle $\mathcal{F}$ (möglicherweise unendlich) und $\eta > 0$ existieren $c > 0, L, n_0 \in \mathbb{N}$ sodass für alle $G = (V, E)$ mit $|V| = n \geq n_0$ gilt:*
Wenn G η-weit von $Forb(\mathcal{F})$ ist, dann existiert ein $F \in \mathcal{F}, v(F) \leq L : |\{F \subseteq G\}| > cn^{v(F)}$.

Korollar 2.3. *Entscheidbare, monotone Grapheneigenschaften sind (mit einseitigem Fehler) testbar.*

Erinnerung: Beweis des Removal Lemmas
Wir regularisieren den gegebenen Graphen G und löschen alle Kanten, die innerhalb der Klassen liegen, die zu irregulären Paaren gehören, die zu zu dünnen regulären Paaren gehören und die in V_0 beginnen.
Wenn G η-weit von $Forb(F)$ entfernt ist, dann existiert ein $F \in Forb(F) : F \subseteq G'$.
$\Rightarrow \exists F', \varphi : F \to F'$ (Homomorphismus): $F' \subseteq R(d_0, \varepsilon)$
$\Rightarrow cn^{v(F)}$ Kopien von F sind in $G' \subseteq G$ enthalten (Counting Lemma).

Lemma 2.19. (Iteriertes Reg. Lemma) *Für alle $t_0, \varepsilon : \mathbb{N} \to]0, 1], \delta > 0$ existieren T_0 und n_0 sodass für alle $G = (V, E)$ mit $|V| = n \geq n_0$ eine Partition $V = V_0 \dot\cup ... \dot\cup V_t$ und eine Verfeinerung $W_0 \dot\cup \bigcup_{k=1}^{t} \bigcup_{i=1}^{s} W_{k,i}$ mit folgenden Eigenschaften:*

1. *$ts \leq T_0$*
2. *$\forall i \in [t], j \in [s] : |V_1| = ... = |V_t|, |W_{i,j}| = \frac{|V_i|}{s}$*
3. *Alle bis auf höchstens δt^2 Paare (V_i, V_j) erfüllen:*
   (a) *(V_i, V_j) ist δ-regulär.*
   (b) *Alle bis auf höchstens δs^2 Unterpaare $(W_{i,k}, W_{j,l})$ erfüllen:*
   $|d(V_i, V_j) - d(W_{i,k}, W_{i,l})| \leq \delta$
4. *Alle bis auf $\varepsilon(t) d^2 s^2$ Paare $(W_{i,k}, W_{j,l}), i, j \in [t], k, l \in [s]$ sind $\varepsilon(t)$-regulär.*

Beweis:
Idee: Dies erreicht man, indem man zunächst mit δ und anschließend mit $\varepsilon(t)$ regularisiert. Wenn die Eigenschaft 3b erfüllt ist, ist man fertig, sonst iteriert man dieses Vorgehen. □

Beweisskizze des Removal Lemmas für monotone Eigenschaften
Der obige Beweis des Removal Lemmas funktioniert analog für endliche Mengen $\mathcal{F}$ statt F. Lediglich im letzten Schritt müssen wir bei der Verwendung des Counting Lemmas zur Regularisierung das kleinste der zu den Graphen aus $\mathcal{F}$ durch das Counting Lemma bestimmten ε auswählen.
Zu einem verbotenen Graphen F (der durchaus $\log n$ oder $\sqrt{n}$ Knoten haben kann) bilden wir die Projektion in dem reduzierten Graphen der Regularisierung von G. Anschließend suchen wir nach dem kleinsten verbotenen Graphen, der diese Projektion als homomorphes Bild haben kann und betten diesen durch das Counting Lemma ein.

- $d_0, \frac{1}{t_0}, \delta \ll \eta$
- $\mathcal{F}_t = \{F \in \mathcal{F} | v(F) \leq t\}$
- Für $F' \subseteq K_t$ sei $\mathcal{F}(F') = \{F \in \mathcal{F} | \exists$ Hom. $\varphi : F \to F'\}$
- $p_t = \max_{F' \subseteq K_t} \min_{F \in \mathcal{F}(F')} |V(F)|$
- $\varepsilon(t) \ll \varepsilon(CountingLemma(\gamma = \frac{1}{2}, d_0, p_t))$
- $L = \max_{t=1,...,T_0} p_t$, wobei T_0 aus dem iterierten Regularitätslemma stammt.

- $c = \frac{d_0^{\binom{L}{2}}}{4T_0^L}$

Wir wählen $U_i \in \{W_{i,1}, ..., W_{i,s}\}$ zufällig.
Durch die Eigenschaft 3d gilt $\left|\left\{\{i,j\} \in \binom{[t]}{2} \middle| |d(U_i,U_j) - d(V_i,V_j)| \geq \delta\right\}\right| \leq 4\delta\binom{[t]}{2}$ mit einer Wahrscheinlichkeit von $> \frac{1}{2}$.
Durch die 4. Eigenschaft des iterierten Regularitätslemmas und $\varepsilon(t) \ll \frac{1}{t^4}$ sind alle Paare (U_i, U_j) mit Wahrscheinlichkeit $> \frac{1}{2}$ $\varepsilon(t)$-regulär. ($\mathbf{E}(|\{\text{irreguläre Paare in } U_i, ..., U_t\}|) \approx \varepsilon(t)t^4 \ll 1$)

Nun fixieren wir $U_1, ..., U_t$ mit diesen Eigenschaften und löschen:

- alle Kanten, die mindestens ein Ende in V_0 haben
- (V_i, V_j), falls $d(U_i, U_j) < d_0$
- alle Kanten innerhalb der V_i

Wir haben also höchstens $\delta n^2 + \left(\underbrace{4\delta\binom{t}{2}\left(\frac{n}{t}\right)^2}_{|d(U_i,U_j)-d(V_i,V_j)|>\delta} + \underbrace{(d_0+\delta)\binom{t}{2}\left(\frac{n}{t}\right)^2}_{d(V_i,V_j)\leq d(U_i,U_j)|+\delta}\right) + \frac{n^2}{t_0}$ Kanten gelöscht.
$\Rightarrow \exists \tilde{F} \in \mathcal{F} : \tilde{F} \subseteq G' \subseteq G$

Der reduzierte Graph ist $R = ([t], E_R)$ mit $\{i,j\} \in E_R \Leftrightarrow d(U_i, U_j) \geq d_0$.
$\Rightarrow \exists$ Hom $\varphi : \tilde{F} \to R$
$\Rightarrow \exists F \in \mathcal{F} : F \to R$ mit $v(F) \leq p_t \leq L$
Wegen $\varepsilon(t) \ll \frac{1}{p_t}$ folgt:
$|\{F \subseteq G'[U_1, ..., U_t]\} \geq \frac{1}{2} d_0^{e(F)} \left(\frac{(1-\delta)n}{ts}\right)^{v(F)} \geq cn^{v(F)}$ □

Bemerkung
Es existieren Familien $\mathcal{F}$, sodass p_t nicht berechenbar ist.

Definition (hereditär)
Eine Grapheneigenschaft heißt *hereditär*, wenn sie abgeschlossen unter induzierten Untergraphen ist.

Lemma 2.20. (Removal Lemma für hereditäre Eigenschaften (Alon, Shapira - 2008))
Für alle $\mathcal{F}$ (möglicherweise unendlich) und $\eta > 0$ existieren $c > 0, L, n_0 \in \mathbb{N}$ sodass für alle $G = (V, E)$ mit $|V| = n \geq n_0$ gilt:
Wenn G η-weit von $Forb_{induziert}(\mathcal{F})$ ist, dann existiert ein $F \in \mathcal{F}, v(F) \leq L : |\{F \leq G\}| > cn^{v(F)}$.
Dabei steht $F \leq G$ für $G[V(F)] = F$.

Vorlesung am 16.07.2008

Kapitel 3

Letzte Woche - letzte Vorlesung

3.1 Monty-Hall-Problem

Das Problem
Es gibt drei Tore. Hinter einem liegt ein Gewinn und die beiden anderen Tore sind leer.
Der Spieler wählt nun ein Tor aus. Dann ist noch mindestens eines der nicht ausgewählten Tore leer und wird geöffnet.
Die Frage ist nun, ob sich der Spieler umentscheiden oder bei seiner Wahl bleiben sollte.

Lösung
Bei einer Umentscheidung gewinnt er mit Wahrscheinlichkeit von $\frac{2}{3}$ (nämlich genau dann, wenn er sich zunächst falsch entschieden hatte), sonst nur mit einer Wahrscheinlichkeit von $\frac{1}{3}$.

3.2 Geburtstagsparadoxon

Das Problem
Wie hoch ist die Wahrscheinlichkeit, dass unter $n \in \mathbb{N}$ Personen zwei am gleichen Tag des Jahres Geburtstag haben?

Lösung
Bei $n \geq 23$ ist diese Wahrscheinlichkeit größer als $\frac{1}{2}$.

3.3 Aktienwetten

Das Problem
Eine Firma X habe heute einen Aktienwert von 100 €, welcher innerhalb eines Jahres mit einer Wahrscheinlichkeit von 80% auf 120 € ansteigt und mit einer Wahrscheinlichkeit von 20% auf 80 € abfällt. Die Bank verspricht nun einem Kunden, am Ende dieses Jahres den Betrag max{0,Aktienwert−100 €} auszuzahlen. Wie viel kann dieses Angebot einem Kunden wert sein, wenn dieser zuvor für dieses Versprechen bezahlen muss?
Der Erwartungswert ist $\mathbf{E}(Gewinn) = 16\,€ - Preis$.
Sind wir also glücklich, solange der Preis geringer als 16 € ist?
Wie wir sehen werden, gewinnt die Bank bei einem Angebot von 14 € auf jeden Fall 4 €.

Verhalten der Bank

- Kaufe $\frac{1}{2}$ Aktie von X.

- 1. Fall: Aktienkurs ist nach einem Jahr 120 €
     $\Rightarrow$ 10 € Gewinn der halben Aktie, 14 € durch den Verkauf des Angebots und −20 € für die Auszahlung an den Kunden.
     $\Rightarrow$ 4 € Gewinn.
  2. Fall: Aktienkurs ist nach einem Jahr 80 €
     $\Rightarrow$ −10 € Verlust durch die halbe Aktie und +14 € durch den Verkauf des Angebots.
     $\Rightarrow$ 4 € Gewinn.

3.4 Ein Spiel

100 Gefangene dürfen sich erst eine Strategie zu dem folgenden Spiel überlegen und müssen anschließend ohne Kommunikation spielen: Sie werden einzeln in einen Raum kommen wo sie das Licht an- bzw. ausschalten können. Die Reihenfolge ist nicht bekannt. Es ist jedoch bekannt, dass jeder noch beliebig oft in diesem Raum geführt wird. Irgendwann darf einer „Stop“ sagen. Wenn zu diesem Zeitpunkt alle in dem Raum waren, haben sie gewonnen. Zu Beginn ist das Licht aus.

Lösung
Ein Spieler zählt, wie oft das Licht an war, als er in den Raum gekommen ist, schaltet es aus und sagt stop, wenn er das Licht 99 mal eingeschaltet wurde. Jeder andere schaltet das Licht ein, wenn es aus ist und er es noch nicht eingeschaltet hatte.

3.5 Noch mal Gefängnis

Gegeben sind n Gefangene, deren Nummern 1 bis n permutiert in n Töpfen versteckt sind. Sie dürfen einzeln in $\frac{n}{2}$ Töpfe sehen. Wenn jeder Gefangene dabei seine Nummer entdeckt hat, dann haben sie gewonnen.

Frage
Wie groß kann $\mathbf{P}(\forall x : x \text{ findet sich})$ mit der besten Strategie sein?
Antwort: Man erreicht sogar eine Wahrscheinlichkeit von mindestens $1 - \ln 2$.

Strategie
Setze $t = Meine\_Nummer = s$ und prüfe den Topf a_t. Falls $a_t = s$, dann wurde die Nummer gefunden, sonst setze $t = a_t$ und mache weiter, bis $\frac{n}{2}$ Töpfe geprüft wurden.
x findet sich genau dann, wenn die Nummer von x in einem Zykel der Länge $\leq \frac{n}{2}$ der Permutation $(a_1, ..., a_n)$ enthalten ist.
$\Rightarrow \mathbf{P}(\forall x \text{ finden sich}) = 1 - \frac{N}{n!}$ wobei N die Anzahl der Permutationen mit einem Zykel der Länge $\geq \frac{n}{2}$ ist.
Als Abschätzung für N erhalten wir:
$N \leq \sum_{i=\frac{n}{2}+1}^{n} \binom{n}{i} \cdot \frac{i!}{(}i-1)! \cdot (n-i)! = \sum_{i=\frac{n}{2}+1}^{n} \frac{n!}{i} \leq n!(H_n - H_{\frac{n}{2}})$ mit $H_k = \sum_{i=1}^{k} \frac{1}{k} \approx \ln k$
$\Rightarrow N \approx n!(\ln n - \ln(\frac{n}{2})) = n! \ln 2$
$\Rightarrow$ Die Erfolgswahrscheinlichkeit ist mindestens $1 - \ln 2$. $\square$

3.6 Ein letztes Problemchen

Wir wählen $0 < n \in \mathbb{N}$ mit Wahrscheinlichkeit $\frac{1}{2^k}$.
In zwei Umschläge legen wir 10^k und 10^{k+1} Euro. Nach dem Öffnen eines Umschlags gibt es die Möglichkeit noch einmal zu tauschen.
(Der Erwartungswert sagt: Tauschen.)

Index