Das Regularitätslemma von Szemerédi

Gerrit Gruben

12. November 2009

Das Regularitätslemma

Grundbegriffe und Formulierung

Das Regularitätslemma gestattet es uns die Eckenmenge eines hinreichend großen Graphen so zu partionieren, dass die zwischen diesen Eckenmengen verlaufenden Kanten nahezu gleichmäßíg verteilt sind, also wie wir es erwarten würden.

Zuerst einige Grundbegriffe und -tatsachen. Im folgendem sei $G = (V, E)$ ein einfacher Graph und A, B disjunkte Teilmengen von V.

Definition 1 (Dichte eines Paares). Es ist

$$d(A, B) := \frac{\|A, B\|}{|A||B|} \in [0, 1]$$

die Dichte des Eckenpaares (A, B). Wobei wir mit $\|A, B\|$ die Anzahl der A-B-Kanten in G bezeichnen.

Die Dichte gibt also an, wieviele der möglichen Kanten zwischen A und B tatsächlich existieren.

Definition 2 (ϵ-reguläre Eckenpaare und ϵ-reguläre Partitionen). Wir nennen für ein $\epsilon > 0$ (A, B) ϵ-regulär, wenn für alle $A' \subseteq A$, $B' \subseteq B$ mit $(|A'|, |B'|) \geq \epsilon(|A|, |B|)$ gilt:

$$|d(A, B) - d(A', B')| \leq \epsilon$$

Eine Partition $\{V_0, V_1, \ldots, V_k\}$ von V mit Ausnahmemenge V_0 heißt ϵ-reguläre Partition, wenn gilt

i) $|V_0| \leq \epsilon |V|$.

ii) $|V_i| = |V_j|$ für $i, j \in [k]$.

iii) Maximal ϵk^2 Paare $(V_i, V_j)_{1 \leq i < j \leq k}$ sind nicht ϵ-regulär. [1]

Je näher das ϵ bei 0 ist, desto gleichmäßiger (regulärer) sind die Kanten zwischen (A, B) verteilt. Wir lassen den Parameter ϵ im folgendem fort. Die Ausnahmemenge V_0 agiert als Abfalleimer, der nicht zu groß werden darf. Man kann jedoch auch fordern, dass $||V_i| - |V_j|| \leq 1$ gilt. Die letzte Eigenschaft sichert, dass nicht zuviele Paare irregulär sind.

Zuerst werden einige Eigenschaften von ϵ-regulären Eckenpaaren gezeigt.

Satz 1. Sei für ein $\epsilon > 0$ (A, B) ϵ-regulär und $d := d(A, B)$. Dann gilt für alle $B' \subseteq B$ mit $|B'| \geq \epsilon |B|$, dass die Menge

$$A' := \{v \in A \mid |N(v) \cap B'| < (d - \epsilon)|B'|\}$$

weniger als $\epsilon |A|$ Elemente enthält. Das heisst mehr als $\epsilon |A|$ viele Elemente aus A haben mehr als $(d - \epsilon)|B'|$ viele Nachbarn in B'.

[1] Man könnte auch anschaulicher $\epsilon\binom{k}{2}$ fordern.

Beweis Annahme: $|A'| \geq \epsilon|A|$, dann gilt $|d - d(A', B')| < \epsilon$, also $d(A', B') \geq d - \epsilon \Leftrightarrow \|A', B'\| \geq (d - \epsilon)|A'||B'|$, womit die Existenz eines $v \in A'$ mit $|N(v) \cap B'| \geq (d - \epsilon)|B'|$ folgt. ↯ ■

Es gibt also bei hinreichend großen Teilmengen einer Menge eines regulären Eckenpaares verschwindend wenige Nachbarn, deren Grad von der Erwartung abweicht. Eine weitere Eigenschaft der Regularität ist es, dass sie unter Teilmengenbildung invariant bleibt:

Satz 2 (Slicing Lemma). Sei (A, B) ein ϵ-reguläres Paar, $d := d(A, B)$. Für $\alpha > \epsilon$, seien $A' \subseteq A$, $B' \subseteq B$ mit $|A'| \geq \alpha|A|$ und $|B'| \geq \alpha|B|$. Dann ist (A', B') ein ϵ'-reguläres Paar mit $\epsilon' := \max\{\epsilon/\alpha, 2\epsilon\}$.

Beweis Seien $A'' \subseteq A', B'' \subseteq B'$ mit $|A''| \geq \epsilon'|A'| \geq \frac{\epsilon}{\alpha}\alpha|A| = \epsilon|A|$ und $|B''| \geq \epsilon'|B'|$. Es folgt mit der Dreiecksungleichung

$$|d(A'', B'') - d(A', B')| \leq |d(A'', B'') - d(A, B)| + |d(A, B) - d(A', B')| = 2\epsilon \leq \epsilon'$$

Womit (A', B') ϵ'-regulär ist. ■

Nun kann das Hauptresultat dieser Arbeit wohlformuliert werden:

Lemma 1 (Regularitätslemma - Szemerédi 1976). Für alle $\epsilon > 0$ und $m \geq 1$ existiert ein M und ein $N \in \mathbb{N}$, sodass sich alle Eckenmengen der Graphen mit $n \geq N$ Ecken in die ϵ-reguläre Partition $\{V_0, \ldots, V_k\}$ partitionieren lassen. Dabei ist k aus $[m, M]$.

Die Parameter ϵ und eine Mindestanzahl an Partitionsmengen m, welches höhere Dichten zwischen den Paaren und weniger Kanten innerhalb der Eckenmengen V_i zur Folge hat, können frei gewählt werden. Das Regularitätslemma liefert dann eine ϵ-reguläre Partition mit einer beschränkten Anzahl an Partitionsmengen für alle hinreichend großen Graphen. Für den Beweis bedarf es einiger Vorarbeit, die im folgendem Abschnitt geleistet wird.

Konzeptionell definiert man eine durch eine Konstante beschränkte(!) Funktion von Partitionen von V. Man zeigt daraufhin, dass bei Verfeinerungen von Partition der Wert unter dieser Funktion nicht kleiner wird. Ausgestattet mit diesen Wissen ermittelt man eine Abschätzung für eine Verfeinerung von irregulären Eckenpaaren. Dieses Ergebnis erweitert man dann auf eine Abschätzung wie eine Verfeinerung einer irregulären Partionierung mit Ausnahmemenge den Wert unter Funktion mindestens um eine Konstante hebt. Da die Funktion mit einer Konstante beschränkt ist, muss eine Endlosschleife irgendwann terminieren und das ist dann, wenn die Partitionierung regulär wird (weil das die Voraussetzung für das Verfahren ist). Technische Probleme dabei sind das Wachstum der Ausnahmemenge und die Anzahl der Partitionsmengen (das Lemma liefert ja eine obere Schranke für die Anzahl dieser).

Beweis des Regularitätslemmas

Die folgende Ungleichung wird in diesem Abschnitt benötigt:

Satz 3. Für reelle $\mu_1, \dots, \mu_k > 0$, $e_1, \dots, e_k \geq 0$ gilt

$$\frac{\left(\sum e_i\right)^2}{\sum \mu_i} \leq \sum \frac{e_i^2}{\mu_i} \tag{1}$$

Beweis Dies folgt aus der Cauchy-Schwarzschen Ungleichung

$$\left(\sum a_i b_i\right)^2 \leq \sum a_i^2 \sum b_i^2$$

mit $a_i = \sqrt{\mu_i}$ und $b_i = \frac{e_i}{\sqrt{\mu_i}}$. ■

Es ist nun $G = (V, E)$ ein einfacher Graph.

Für disjunkte $A, B \subset V$ definieren wir

$$q(A, B) := \frac{|A||B|}{n^2} d^2(A, B) = \frac{\|A, B\|^2}{n^2|A||B|}$$

Für Partitionen $\mathcal{A}$ von A und $\mathcal{B}$ von B erweitern wir:

$$q(\mathcal{A}, \mathcal{B}) := \sum_{A \in \mathcal{A},\, B \in \mathcal{B}} q(A, B)$$

sowie für $V = \dot{\bigcup} V_i$

$$q(\{V_1, \dots, V_k\}) := \sum_{i<j} q(V_i, V_j)$$

betrachten wir reguläre Partitionen von V, so wird die Ausnahmemenge in ihre 1-elementigen Teilmengen partioniert. Wir können q als Maß für Regularität einer Partition deuten. Eine weitere wichtige Beobachtung ist, dass für Partitionen $V_1, \dots, V_k$ von V gilt:

$$\begin{aligned} q(\{V_1, \dots, V_k\}) &= \sum_{i<j} q(V_i, V_j) \\ &= \sum_{i<j} \frac{|V_i||V_j|}{n^2} \underbrace{d^2(V_i, V_j)}_{\leq 1} \\ &\leq 1 \end{aligned}$$

Unser Vorgehen ist nun wie folgt: Wir werden eine irreguläre Partition schrittweise verfeinern, sodass der Wert unter q mindestens um einen Wert in Abhängigkeit nur von ϵ steigt. Da er nach oben beschränkt ist durch eine Konstante, werden wir sie nach einer beschränkten Anzahl an Verfeinerungen ϵ-regulär machen.

Als erstes zeigen wir, dass eine die Verfeinerung einer Partition deren Wert unter q nicht senkt.

Satz 4. Seien $A, B \subseteq V$ disjunkt, $\mathcal{A}$ und $\mathcal{B}$ Partitionen von A bzw. B, so gilt

$$q(A, B) \leq q(\mathcal{A}, \mathcal{B}) \tag{2}$$

und für eine Partitionen $\mathcal{P}$ von V und einer Verfeinerung $\mathcal{P}'$ dieser gilt

$$q(\mathcal{P}) \leq q(\mathcal{P}') \tag{3}$$

Beweis (2): $\mathcal{A} = \{A_1, A_2\}$, $\mathcal{B} = \{B_1, B_2\}$, $A = \dot\cup A_i$, $B = \dot\cup B_i$.

$$\begin{aligned} q(\mathcal{A}, \mathcal{B}) &= \sum_{i,j} q(A_i, B_j) \\ &= \frac{1}{n^2} \sum_{i,j} \frac{\|A_i, B_j\|^2}{|A_i||B_j|} \\ &\overset{(1)}{\geq} \frac{1}{n^2} \frac{\overbrace{\left(\sum_{i,j} \|A_i, B_j\| \right)^2}^{=\|A,B\|^2}}{\underbrace{\sum_{i,j} |A_i||B_j|}_{=(\Sigma A_i)(\Sigma B_j)}} \\ &= q(A, B) \end{aligned}$$

(3): O.B.d.A. $\mathcal{P} = \{X, V_1, \ldots, V_k\}$ und $\mathcal{P}' = \mathcal{P} - X + \{X_1, X_2\}$ mit $X = X_1 \dot\cup X_2$.

$$\begin{aligned} q(X, V_i) &\overset{(2)}{\leq} q(\{X_1, X_2\}, V_i) \\ &= q(X_1, V_i) + q(X_2, V_i) \end{aligned}$$

nun folgt

$$\begin{aligned} q(\mathcal{P}') &= \underbrace{q(X_1, X_2)}_{\geq 0} + \sum_{i<j} q(V_i, V_j) + \underbrace{\sum_i q(X_1, V_i) + \sum_i q(X_2, V_i)}_{\geq \sum_i q(X_i, V_i)} \\ &\geq \sum_{i<j} q(V_i, V_j) + \sum_i q(X, V_i) \\ &= q(\mathcal{P}) \end{aligned}$$

■

Satz 5. Sei $\epsilon > 0$, $A, B \subseteq V$ disjunkt. Ist (A, B) ϵ-irregulär, so existieren A_1, A_2, B_1, B_2 mit $A = A_1 \dot\cup A_2$, $B = B_1 \dot\cup B_2$ mit

$$q(A, B) + \epsilon^4 \frac{|A||B|}{n^2} \leq q(\{A_1, A_2\}, \{B_1, B_2\})$$

Beweis Seien (A,B) ein ϵ-irreguläres Paar, also gibt es $|A_1| \geq \epsilon|A|$, $|B_1| \geq \epsilon|B|$ mit

$$|\underbrace{d(A,B) - d(A_1,B_1)}_{=:\eta}| > \epsilon$$

Setze $a := |A|$, $a_i := |A_i|$, $b := |B|$, $b_i = |B_i|$, $e_{ij} := \|A_i, B_j\|$, $e := \|A, B\|$. Mit (1) erhalten wir:

$$\begin{aligned}
q(\{A_1,A_2\},\{B_1,B_2\}) &= \frac{1}{n^2}\sum_{i,j}\frac{e_{ij}^2}{a_ib_j} \\
&= \frac{1}{n^2}\left(\frac{e_{11}^2}{a_1b_1} + \sum_{i+j\geq 2}\frac{e_{ij}^2}{a_ib_j}\right) \\
&\overset{(1)}{\geq} \frac{1}{n^2}\left(\frac{e_{11}^2}{a_1b_1} + \frac{\left(\sum_{i+j\geq 2} e_{ij}\right)^2}{\sum_{i+j\geq 2} a_ib_j}\right) \\
&= \frac{1}{n^2}\left(\frac{e_{11}^2}{a_1b_1} + \frac{(e-e_{11})^2}{ab - a_1b_1}\right)
\end{aligned}$$

Es gilt ferner

$$e_{11} = \frac{a_1b_1e}{ab} + \eta a_1b_1$$

was mit der Definition von η leicht nachprüfbar ist. Also:

$$\begin{aligned}
n^2q(\{A_1,A_2\},\{B_1,B_2\}) &\geq \frac{1}{a_1b_1}\left(\frac{a_1b_1e}{ab} + \eta a_1b_1\right)^2 + \frac{1}{ab - a_1b_1}\left(\left(1 - \frac{a_1b_1}{ab}\right)e - \eta a_1b_1\right)^2 \\
&\geq \frac{a_1b_1e^2}{a^2b^2 + \eta^2a_1b_1}{}^2 \\
&\geq \frac{e^2}{ab} + \epsilon^4ab = n^2q(A,B) + \epsilon^4ab
\end{aligned}$$

Wobei $|\eta| > \epsilon$, $a_1 \geq \epsilon a$, $b_1 \geq \epsilon b$ benutzt wurde. ■

Satz 6. Seien $0 < \epsilon \leq 1/4$, $\mathcal{P} = \{V_0, \dots, V_k\}$ Partition von V, $|V_0| \leq \epsilon n$ Ausnahmemenge und $|V_i| = |V_j| = r$ für $(i \neq j)$. Ist $\mathcal{P}$ ϵ-irregulär, dann existiert eine Partitition $\mathcal{P}' = \{V_0', \dots, V_l'\}$ von V mit Ausnahmemenge V_0', $l \in [k, k4^k]$, $|V_0'| \leq |V_0| + n2^{-k}$, $|V_i'| = |V_j'|$ und

$$q(\mathcal{P}) + \epsilon^5/2 \leq q(\mathcal{P}')$$

Beweis Sei $\mathcal{P} = \{V_0, \dots, V_k\}$ eine irreguläre Partition von V.
Für $1 \leq i < j \leq k$ definieren wir Partitionen $\mathcal{P}_{ij}$ von V_i und $\mathcal{P}_{ji}$ von V_j wie folgt:

[2]Ausmultiplizieren, der lineare Teil beider Terme kürzt sich heraus, dann werden die quadratischen (und daher nichtnegativen) Teile des zweiten Terms nach unten abgeschätzt.

Ist (V_i, V_j) ϵ-regulär dann $\mathcal{P}_{ij} := \{V_i\}$ und $\mathcal{P}_{ji} := \{V_j\}$.
Ansonsten können wir (V_i, V_j) nach dem vorherigen Satz in 2-er Mengen $\mathcal{P}_{ij}$ und $\mathcal{P}_{ji}$ zerlegen, sodass gilt:

$$q(V_i, V_j) + \epsilon^4 \frac{|V_i||V_j|}{n^2} \leq q(\mathcal{P}_{ij}, \mathcal{P}_{ji})$$

Für $i \in [k]$ sei $\mathcal{P}_i$ die kleinste Partition von V_i, welche alle V_{ij} verfeinert $(i \neq j)$, es gilt dann $|\mathcal{P}_i| \leq 2^{k-1}$. Setze

$$\widetilde{\mathcal{P}} = \{V_0\} \cup \bigcup_{i=1}^{k} \mathcal{P}_i$$

Dann ist $|\widetilde{\mathcal{P}}| \in [k, k2^k]$. Sei $\mathcal{P}_0 := \{\{v\} \mid v \in V_0\}$.
Es sind mehr als ϵk^2 der (V_i, V_j) und damit V_{ij} irregulär, daher gilt:

$$\begin{aligned} q(\widetilde{\mathcal{P}}) &= \sum_{1 \leq i < j \leq k} q(\mathcal{P}_i, \mathcal{P}_j) \\ &\geq \sum_{1 \leq i < j \leq k} q(\mathcal{P}_{ij}, \mathcal{P}_{ji}) \\ &\geq \sum_{1 \leq i < j \leq k} q(V_i, V_j) \\ &= q(\mathcal{P}) + \epsilon^5 \left(\frac{kr}{n}\right)^2 \\ &\geq q(\mathcal{P}) + \frac{\epsilon^5}{2} \end{aligned}$$

■

Anwendungen des Regularitätslemmas

Ursprünglich stieß Szemerédi beim Beweis dieses Satzes auf das Regularitätslemma:

Satz 7 (Szemerédi 1975). Für alle natürlichen $k \geq 3$ und $\epsilon > 0$ existiert ein $n_0 \in \mathbb{N}$, sodass für alle natürlichen $n \geq n_0$ gilt: Wenn $A \subseteq \{1, \ldots, n\}$ mit $|A| > \epsilon n$, dann enthält A eine arithmetische Progression der Länge k.

Ein wichtiges Lemma, welches aus dem Regularitätslemma folgt, ist das Einbettungslemma. Der folgende Begriff wird für die Formulierung dieses Lemmas benötigt:

Definition 3 (Regularitätsgraph). Zu gegebenem $\epsilon > 0$, $d \in [0, 1], m \in \mathbb{N}$ und $V_1, \ldots, V_k$ Partition einer Eckenmenge V mit $|V_i| = m$ konstruieren wir den Graphen R mit Eckenmenge $V_1, \ldots, V_k$ und Kanten zwischen V_i und V_j genau dann wenn (V_i, V_j) ein ϵ-reguläres Paar mit Dichte $\geq d$ ist. R heißt Regularitätsgraph [3]. Für ein $s \in \mathbb{N}$ entsteht R_s aus R in dem die Ecken durch s Ecken ersetzt werden und vormals verbundene Ecken nach dieser Transformation vollständige bipartite Graphen $K_{s,s}$ werden.

[3] Auch Cluster Graph und Reduced Graph genannt.

Lemma 2 (Einbettungslemma). $\forall d \in (0,1]$ und $\Delta \geq 1$ existiert ein $\epsilon_0 > 0$, sodass für einfache Graphen G und H mit $\Delta(H) \leq \Delta$ gilt: Für den Regularitätsgraphen R von G mit den Parametern $\epsilon \leq \epsilon_0$, m und d, sowie $s \in \mathbb{N} : m \geq \frac{2s}{d^\Delta}$, dann folgt aus $H \subseteq R_s$ dass H Teilgraph von G ist.

Beweis Sei $d \in (0,1]$ und $\Delta \geq 1$ gegeben. Da gilt

$$(d-\epsilon)^\Delta - \Delta\epsilon \xrightarrow{\epsilon \to 0} d^\Delta$$

können wir wegen des Zwischenwertsatzes und der Stetigkeit des Ausdruckes in ϵ $\epsilon_0 \in (0,d)$ mit

$$(d-\epsilon_0)^\Delta - \Delta\epsilon_0 \geq \frac{1}{2} d^\Delta$$

wählen. Jetzt halten wir $s \in \mathbb{N}$, $m \geq \frac{2s}{d^\Delta}$, G, H und R von G mit den obigen Eigenschaften fest. Sei $V(R_s) = \{V_1^s, \ldots, V_k^s\}$ und $H \subseteq R_s$. Nummerieren wir die h Ecken $u_1, \ldots, u_h$ von H so erhalten wir eine Zuordnung $\sigma : [h] \to [k]$ mit $u_i \in V_{\sigma(i)}^s$ für $i = 1, \ldots, h$. Nun wollen wir H in G einbetten, dafür wählen wir für jedes u_i ein $v_i \in V_{\sigma(i)}$, sodass aus $u_i u_j \in E(H)$ folgt, dass $v_i v_j \in E(G)$.
Ziel: Für $i \in [h]$:

$$V_{\sigma(i)} =: V_i^0 \supseteq V_i^1 \supseteq \ldots \supseteq V_i^i = \{v_i\}$$

Sodass $u_i \mapsto v_i$ H in G einbettet.
Algorithmus:
Dabei wird $v_i \in V_i^{i-1}$ so gewählt, dass für die maximal Δ vielen $j > i$ mit

---

Algorithm 1 Iterative Einbettung

$V_i^0 := V_{\sigma(i)}$ $\forall i \in [h]$
for all $i = 1 \ldots h$ **do**
  Wähle $v_i \in V_i^{i-1}$
  for all $j = i+1 \ldots h$ **do**
    if $u_i u_j \in E(H)$ **then**
      $V_j^i := V_j^{i-1}$
    else
      $V_j^i := V_j^{i-1} \cap N(v_i)$
    end if
  end for
end for

$u_i u_j \in E(H)$ gilt

$$|V_j^i| \geq (d-\epsilon)|V_j^{i-1}|$$

Das geht nach Satz (1) mit den Mengen $V_{\sigma(i)}$ und $V_j^{i-1} \subseteq V_{\sigma(j)}$. Damit fallen maximal $\Delta\epsilon m$ Wahlen von v_i weg. Folgendes muss noch gezeigt werden:

1. Es müssen in jedem V_i^{i-1} noch s mögliche v_i übrig bleiben (maximal $s-1$ Ecken aus $V_{\sigma(i)}$ können schon eingebettet worden sein). Das ist äquivalent zu

$$|V_i^{i-1}| - \Delta\epsilon m \geq s$$

2. $|V_i^{j-1}| \geq \epsilon m$ damit wir den Satz (1) anwenden können ($j > i$ und $u_i u_j \in E(H)$)

Erinnern wir uns an $\epsilon \leq \epsilon_0$, $m \geq 2s/d^\Delta$, $\Delta(H) \leq \Delta$ und der Wahl von ϵ_0 so gilt:

$$\begin{aligned}
|V_i^{j-1}| - \Delta\epsilon m &\geq (d-\epsilon)^\Delta m - \Delta\epsilon m \\
&\geq (d-\epsilon_0)^\Delta m - \Delta\epsilon_0 m \\
&\geq \frac{1}{2} d^\Delta m \\
&\geq s
\end{aligned}$$

so folgen beide Behauptungen. ∎

Literatur

[1] Bondy, A. ; Murty, U.S.R.: *Graph Theory.* Springer, 2008 (Graduate Texts in Mathematics). – ISBN 978–1–84628–969–9

[2] Diestel, R.: *Graph Theory.* Springer, 2006 (Graduate Texts in Mathematics)

[3] Komlós, János ; Simonovits, Miklós: *Szemerédi's Regularity Lemma and Its Applications in Graph Theory.* 1996