Lecture 8. Spectral proof of regularity lemma

Hong Liu

14th April 2021

1 Chvátal-Rödl-Szemerédi-Trotter theorem

In this section, we present an application of the regularity lemma in graph Ramsey theory. Recall that the Ramsey number $r(G, G)$ for a graph G is the minimum integer N such that any 2-edge-colouring of K_N contains a monochromatic copy of G. The Ramsey number for cliques K_t is exponential in t. It is intuitive to guess that it is easy to force monochromatic copy of sparser graphs, i.e. the Ramsey number for sparser graphs might be smaller. A classical result of Chvátal-Rödl-Szemerédi-Trotter states that bounded degree graphs have linear Ramsey number.

Theorem 1.1. *Let $d \in \mathbb{N}$ and G be a graph with $\Delta(G) \leq d$, then*

$$r(G, G) = O_d(|G|).$$

We will make use of the multicolour version of the Szemerédi Regularity Lemma. For a k-edge-coloured graph G, a partition $V(G) = V_1 \cup \ldots \cup V_r$ is an *ε-regular partition* if

- for all $ij \in \binom{[r]}{2}$, $\big| |V_i| - |V_j| \big| \leq 1$;
- for all but at most $\varepsilon\binom{r}{2}$ choices of $ij \in \binom{[r]}{2}$, the pair (V_i, V_j) is ε-regular in *every* colour.

Lemma 1.2 (Multicolour regularity lemma)**.** *For every real $\varepsilon > 0$ and integers $k \geq 1$ and m, there exists $M = M(\varepsilon, m, k)$ such that every k-edge-coloured graph G with $n \geq m$ vertices admits an ε-regular partition $V(G) = V_1 \cup \ldots \cup V_r$ with $m \leq r \leq M$.*

We can similarly define a reduced graph corresponding to a regular partition, with the only difference that $ij \in E(R)$ if and only if (V_i, V_j) is regular with respect to every colour. The reduced graph inherits a (multi)edge-colouring from G: we can assign each edge $ij \in E(R)$, the set of all colours that is dense in $G[V_i, V_j]$. For the application here, it suffices to just assign the *majority* colour, that is, the densest colour between the pair. Thus, if $ij \in E(R)$ is red, then $G[V_i, V_j]$ has red-density at least $1/k - o(1) = \Omega(1)$.

Let us recall Brook's theorem, which will be needed in the proof.

Theorem 1.3 (Brook's theorem)**.** *Every graph G can be properly vertex-coloured using $\Delta(G)+1$ colours, i.e. $\chi(G) \leq \Delta(G) + 1$.*

Proof of Theorem 1.1. Let $m \geq 5r(K_{d+1}, K_{d+1})$ be sufficiently large, $\varepsilon = 1/m$, and

$$C := 2M/(1/2 - \varepsilon)^d,$$

where $M = M(\varepsilon, m, 2)$ is the constant returned from Lemma 1.2. Let $N \geq C|G|$ and fix an arbitrary 2-edge-colouring of K_N. We shall find a monochromatic copy of G, which then implies $r(G, G) \leq C|G|$ as desired.

Apply multicolour regularity lemma to the given 2-edge-coloured K_N and let R be the corresponding reduced graph with a 2-edge-colouring indicating the majority colour. Note that as there are at most $\varepsilon\binom{r}{2}$ irregular pairs, R is almost complete:

$$e(R) \geq (1-\varepsilon)\binom{r}{2} > \left(1 - \frac{1}{m/3-1}\right)\frac{r^2}{2}.$$

Now, by Turán theorem, R contains a clique $K_{m/3}$. As $m \geq 5r(K_{d+1}, K_{d+1})$, in this 2-edge-coloured clique $K_{m/3}$, there is a monochromatic K_{d+1}. By Brook's theorem, K_{d+1} is a homomorphic image of G, as $\Delta(G) \leq d$. Then by the embedding lemma, the original 2-edge-coloured K_N contains a monochromatic copy of G. □

2 Spectral proof of regularity lemma

Finally, we give a proof of the regularity lemma. The original proof proceeds by refining partition and energy increment strategy. Here, we shall give a proof based on the spectral decomposition of the adjacency matrix given by Tao, and independently by Szegedy. This idea originates from Frieze-Kannon's proof of the weak regularity lemma.

We shall only prove the following weaker version in which we do not require equipartition. One can refine this regular partition further to get a equipartition.

Lemma 2.1. *Let G be an n-vertex graph and let $\varepsilon > 0$. Then there exists a partition $V = V_1 \cup \cdots \cup V_M$, $M \leq M(\varepsilon)$, such that apart from an exceptional set $\Lambda \subseteq \binom{[M]}{2}$ with*

$$\sum_{(i,j)\in\Sigma} |V_i||V_j| = O(\varepsilon|V|^2),$$

we have for every $(i,j) \notin \Lambda$, $A \subseteq V_i$ and $B \subseteq V_j$ that

$$\big|e(A,B) - d_{ij}|A||B|\big| = O(\varepsilon|V_i||V_j|).$$

Before we dive into the details of the proof, let us sketch briefly how it goes. We write the adjacency matrix T of G as the sum of the rank-1 matrices from eigenvectors of T with weights being the associated eigenvalues. Then the structure of T is dictated mostly by the part with large eigenvalues (main term); while the part with small eigenvalues is more like noise (error term). Thus, to capture the behaviour of T, we shall get a partition in which each eigenvector with large eigenvalue is approximately constant in each part.

We need Cauchy-Schwarz inequality for the proof, let us recall it here.

Lemma 2.2 (Cauchy-Schwarz inequality). *Let $u, v \in \mathbb{C}^n$, then*

$$\sum_{i=1}^{n} u_i\overline{v_i} = \langle u, v\rangle \leq \|u\|_2\|v\|_2 = \sqrt{\sum_{i=1}^{n}|u_i|^2\sum_{i=1}^{n}|v_i|^2}.$$

Furthermore, equality holds if and only if u and v are linearly dependent.

Proof of the regularity lemma, Lemma 2.1. Let T be the adjacency matrix of G. As T is a real symmetric matrix, it is self-adjoint and has eigenvalue decomposition:[1]

$$T = \sum_{i=1}^{n} \lambda_i u_i u_i^*,$$

[1] We treat all vectors here as column vectors.

where $u_1, \dots, u_n$ form an orthonormal basis of $\mathbb{C}$ with real eigenvalues ordered as $|\lambda_1| \geq \dots \geq |\lambda_n|$.

Splitting T. As outlined above, we shall splits $T = T_1 + T_2 + T_3$ into main term T_1 and error terms T_2, T_3. To do so, we need a bound on the eigenvalues. Note that the (i,j)-th entry in T^k records the number of v_i, v_j-walks with length k in G. In particular, each diagonal entry of T^2 is the degree of the corresponding vertex. Thus, the trace

$$\mathrm{tr}(T^2) = \sum_i d_i = 2e(G) \leq n^2,$$

and for each $i \in [n]$, we have

$$i \cdot |\lambda_i|^2 \leq \sum_{i=1}^{n} |\lambda_i|^2 = \mathrm{tr}(T^2) \leq n^2 \implies |\lambda_i| \leq \frac{n}{\sqrt{i}}. \tag{1}$$

Let $F = F(\varepsilon) : \mathbb{N} \to \mathbb{N}$ be a function to be chosen later with $F(i) \geq i$. By averaging, for some $J \leq F^{1/\varepsilon^3}(1)^2$, we can take out a piece in the middle with small weight[3]:

$$\sum_{i \in [J, F(J)]} |\lambda_i|^2 \leq \varepsilon^3 n^2. \tag{2}$$

We can now write $T = T_1 + T_2 + T_3$, where

- $T_1 = \sum_{i \leq J} \lambda_i u_i u_i^*$ is the "*structured*" term;
- $T_2 = \sum_{i \in [J, F(J)]} \lambda_i u_i u_i^*$ is the "*small*" term;
- $T_3 = \sum_{i > F(J)} \lambda_i u_i u_i^*$ is the "*pseudorandom*" term.

Partition for the structured term T_1. We now construct a partition of $V(G)$ such that T_1 is approximately constant in most parts. For each $i \leq J$, we partition $V(G)$ into $O_{J,\varepsilon}(1)$ parts in which u_i only fluctuates by $O(\frac{\varepsilon^{3/2}}{J} n^{-1/2})$ apart from an exceptional part of size at most $\frac{\varepsilon}{J} n$ where $|u_i|$ is excessively large (of value at least $\sqrt{\frac{J}{\varepsilon}} n^{-1/2}$). Let $u = u_i$ and write $u(j)$ for the j-th coordinate of u. Recall that $\|u\|_2^2 = \sum_{j \in [n]} u(j)^2 = 1$, so the number of coordinates with value at least $\sqrt{\frac{J}{\varepsilon}} n^{-1/2}$ is at most $\frac{\varepsilon}{J} n$. Thus, for the rest of the coordinates, we can partition it into at most $\sqrt{\frac{J}{\varepsilon}} n^{-1/2} / (\frac{\varepsilon^{3/2}}{J} n^{-1/2}) = O_{J,\varepsilon}(1)$ parts as claimed.

Combining all of these J partitions together, we get $V(G) = V_1 \cup \dots \cup V_{M-1} \cup V_M$, $M = O_{J,\varepsilon}(1)$, where the exceptional part $|V_M| \leq J \cdot \frac{\varepsilon}{J} n = \varepsilon n$, and for any $1 \leq i \leq M-1$, the eigenvectors $u_1, \dots, u_J$ all fluctuate at most $O(\frac{\varepsilon}{J} n^{-1/2})$. We claim that T_1 fluctuates at most $O(\varepsilon)$ on each block $V_i \times V_j$, for $1 \leq i, j \leq M-1$, and consequently, writing d_{ij} for the mean value of entries of T_1 on $V_i \times V_j$, we have for any $A \subseteq V_i$, $B \subseteq V_j$, that

$$\mathbb{1}_A^* T_1 \mathbb{1}_B = d_{ij} |A||B| + O(\varepsilon |V_i||V_j|). \tag{3}$$

Indeed, recall that $|\lambda_i| \leq n/\sqrt{i}$, we see that each $V_i \times V_j$-entry of $T_1 = \sum_{i \leq J} \lambda_i u_i u_i^*$ fluctuates by at most

$$\sum_{i \leq J} \lambda_i \cdot \sqrt{\frac{J}{\varepsilon}} n^{-1/2} \cdot O\left(\frac{\varepsilon^{3/2}}{J} n^{-1/2}\right) = O\left(\frac{\varepsilon}{\sqrt{J}} n^{-1}\right) \cdot \sum_{i \leq J} \frac{n}{\sqrt{i}} = O(\varepsilon).$$

[2] We use $F^k = F \circ \dots \circ F$ for k iteration of F.

[3] Proof: Consider the partition of $[n]$ into intervals $[1, F(1)) \cup [F(1) \cup F^2(1)) \cup [F^2(1), F^3(1)) \cdots$. As $\sum_{i \in [n]} |\lambda_i|^2 \leq n^2$, one of the first $1/\varepsilon^3$ intervals should be at most $\varepsilon^3 n^2$.

Bounding error term T_2. By the choice of T_2 and (2), $\mathrm{tr}(T_2^2) \le \varepsilon^3 n^2$. On the other hand, let x_{ab} be (a,b)-th entry of T_2, as T_2 is self-adjoint, we have $\mathrm{tr}(T_2^2) = \sum_{a,b\in V(G)} |x_{ab}|^2$. Then by Markov inequality, we get

$$\sum_{a\in V_i, b\in V_j} |x_{ab}|^2 \le \varepsilon^2 |V_i||V_j|, \tag{4}$$

for all $1 \le i,j \le M-1$ apart from an exceptional set $\Lambda' \subseteq \binom{[M-1]}{2}$ with $\sum_{(i,j)\in\Lambda'} |V_i||V_j| \le \varepsilon n^2$. Hence, for any $(i,j) \notin \Lambda'$ and $A \subseteq V_i, B \subseteq V_j$, by (4) and Cauchy-Schwarz inequality, we have

$$\mathbb{1}_A^* T_2 \mathbb{1}_B \le \sum_{a\in V_i, b\in V_j} |x_{ab}| \le \Big(\sum_{a\in V_i, b\in V_j} |x_{ab}|^2\Big)^{1/2} \big(|V_i||V_j|\big)^{1/2} = O(\varepsilon |V_i||V_j|). \tag{5}$$

Bounding error term T_3. By the choice of T_3 and (1), the operator norm of T_3 is at most $\|T_3\|_{\mathsf{op}} \le \frac{n}{\sqrt{F(J)}}$. Then by Cauchy-Schwarz inequality, we have

$$\mathbb{1}_A^* T_3 \mathbb{1}_B = \langle \mathbb{1}_A, T_3\mathbb{1}_B\rangle \le \|\mathbb{1}_A\|_2 \cdot \|T_3\mathbb{1}_B\|_2 \le \|\mathbb{1}_A\|_2 \cdot \|T_3\|_{\mathsf{op}} \cdot \|\mathbb{1}_B\|_2 = O\Big(\frac{n^2}{\sqrt{F(J)}}\Big). \tag{6}$$

Set $\Lambda := \Lambda' \cup \{(i,j):\ i \text{ or } j = M\} \cup \{(i,j) : \min\{|V_i|, |V_j|\} \le \varepsilon n/M\}$. Then it is easy to check that $\sum_{(i,j)\in\Lambda} |V_i||V_j| \le O(\varepsilon n^2)$. By (3), (5), (6), we have

$$e(A,B) = \mathbb{1}_A^* T \mathbb{1}_B = \sum_{i=1}^{3} \mathbb{1}_A^* T_i \mathbb{1}_B = d_{ij}|A||B| + O(\varepsilon |V_i||V_j|) + O\Big(\frac{n^2}{\sqrt{F(J)}}\Big).$$

As $|V_i|, |V_j| \ge \varepsilon n/M$, we have $\frac{n^2}{\sqrt{F(J)}} \le \frac{M^2|V_i||V_j|}{\varepsilon^2\sqrt{F(J)}}$. To absorb the 2nd error term into the first one, we need $\frac{1}{\sqrt{F(J)}} = O(\varepsilon^3/M^2)$. □

Remark 2.3. The point of having T_2-term is to have local control on the fluctuation of $e(A,B)$, i.e. $O(\varepsilon|V_i||V_j|)$. For the tail T_3, we only have a global type control $O(\frac{n^2}{\sqrt{F(J)}})$, and we need $\sqrt{F(J)} \ge \frac{M^2}{\varepsilon^3}$, to make it into a local error. Recall that $M \ge J^J$ when we combine the partitions for each u_i, $i \le J$. Thus, we need to create a gap between $F(J)$ and J by splitting out a small term T_2 in the middle.