AN INTRODUCTION TO THE REGULARITY LEMMA

CASSIE DESKUS

ABSTRACT. The purpose of this paper is to introduce basic concepts relating to random graphs and to provide an introduction to and proof of the Regularity Lemma.

CONTENTS

1. INTRODUCTION

Szemeredi's Regularity Lemma shows that any large enough graph can be broken into a restricted number of equally sized subsets that behave for the most part fairly randomly: that is, a large enough graph partitions, for a suitably nice partition, in the same way as a random graph. This result is important in graph theory because it is often simpler to state and prove results for random graphs with probabilistic methods than for general graphs. Before making this more explicit we will give some basic vocabulary.

Definition 1.1. A *simple graph* is a pair $G = (V, E)$, where V is a finite set, called the *vertices* of G, and E is a subset of the set of two-element subsets of V, called the *edges* of V.

For example, one might let V be the set of U.S. telephone numbers, and let E be the set of all pairs of distinct telephone numbers where one has called the other at least once. One might think of a graph pictorially as a finite set of points (the vertices) connected by lines (the edges). If we do, the definition of simple graph excludes pictures with *loops* — lines connecting a vertex to itself — and *multiple edges* — more than one line connecting two vertices. In particular, the pair ({U.S. cities}, {U.S. roads}), where we identify a road with the two cities it connects, is not a simple graph: there are cities connected to each other by more than one road, and roads that connect a city to itself. However, if $C \subset$ {U.S. cities}

Date: September 24th, 2011.

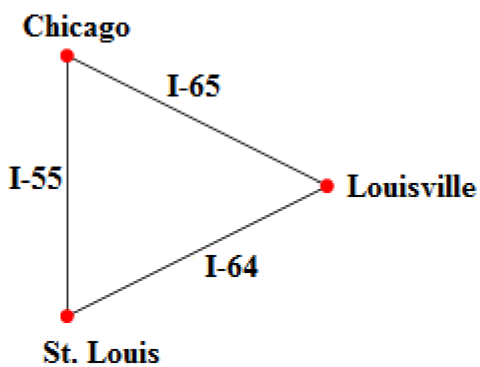

FIGURE 1. This diagram shows the three vertices (Louisville, Chicago, St. Louis) and the three edges (I-55, I-64, I-65) as in the graph (C, R) previously described

and $R \subset \{\text{U.S. roads}\}$, it is possible for (C, R) to be a simple graph. One such example is given by

$$C_v = \{\text{Louisville, Chicago, St. Louis}\}; \quad R_v = \{\text{I-55, I-64, I-65}\}.$$

The admittedly very small graph (C_v, R_v) is topologically a triangle, which we can see from the graph below.

We will use vertical bars to denote the cardinality of a set, so that if $X \subset V$, $|X|$ is the number of vertices in the set X. In the example above, $|C_v| = |R_v| = 3$.

Definition 1.2. Let $G = (V, E)$ be a simple graph, and let $X, Y \subset V$. The *number of edges between X and Y*, denoted $e(X, Y)$, is

$$e(X, Y) = |\{(e_1, e_2) \in E \ : \ \text{either } e_1 \in X \text{ and } e_2 \in Y \text{ or } e_2 \in X \text{ and } e_1 \in Y\}|.$$

In the example above, if $X = \{\text{Chicago}\}$ and $Y = \{\text{Louisville}\}$, then

$$e(X, Y) = |\{\text{I-65}\}| = 1.$$

In a simple graph, if $|X| = |Y| = 1$, then $e(X, Y) \in \{0, 1\}$ (either the vertices are connected by an edge or they are not).

We did not require in the above definition that the sets X and Y be disjoint, but suppose now that this is the case. There are potentially $|X||Y|$ edges between X and Y, i.e.

$$0 \leq e(X, Y) \leq |X||Y|.$$

We would like some measure of how connected the sets X and Y are to each other, and the simplest way to measure this is to compare the actual number of edges with the potential (greatest) number of edges. Our next definition clarifies this:

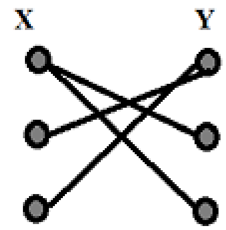

FIGURE 2. This graph shows an example of a graph having density $\frac{4}{9}$

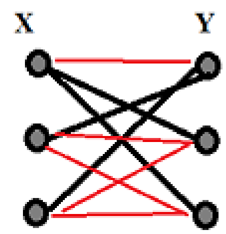

FIGURE 3. This graph contains all possible edges between the two sets, therefore it has a density 1.

Definition 1.3. Let $G = (V, E)$ be a simple graph, and let X, $Y \subset V$ be disjoint and nonempty. The *density between X and Y*, denoted $d(X, Y)$, is

$$d(X, Y) = \frac{e(X, Y)}{|X||Y|}.$$

Of course, we have $0 \leq d(X, Y) \leq 1$. Returning again to our example (C_v, R_v), the density between any two disjoint nonempty subsets of C_v is 1; that is to say, every city is connected to every other city by a road. On the other hand, in the picture below, there are four edges connecting the subsets X and Y (drawn in black) with a possible total of 9; hence $d(X, Y) = 4/9$ for this picture.

Definition 1.4. Let $G = (V, E)$ be a simple graph, and let $X, Y \subset V$ be disjoint and nonempty. We call the pair (X, Y) *ϵ-regular* if for any $A \subseteq X$ and $B \subseteq Y$ such that $|A| \geq \epsilon|X|$ and $|B| \geq \epsilon|Y|$, it is the case that $|d(A, B) - d(X, Y)| < \epsilon$. We call a pair *$\epsilon$-irregular* if it is not ϵ-regular.

Roughly, an ϵ-regular pair (X, Y) is a pair whose density is close to the density of any two subsets A and B that are not too small. Let us explain some of the intuition behind these definitions - why should we care whether subsets of X and Y resemble the pair itself?

Definition 1.5. Intuitively, a random variable is a function that obtains its values through some kind of random process. In this paper we will be looking specifically

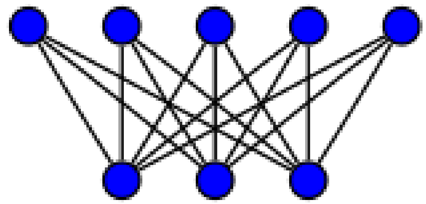

FIGURE 4. An example of a bipartite graph between two sets–one with 3 vertices, the other with 5

at random variables that take on a finite number of discrete values. That is, a random variable X may be seen as a finite set of values x_i and a set of probabilities $\mathbb{P}\{X = x_i\}$ where the probabilities sum to 1.

For example, we might consider flipping a fair coin. The random variable

$$X = \left\{ \begin{array}{ll} 1 & \text{if heads} \\ 0 & \text{if tails} \end{array} \right\}$$

is given by the information

$$\mathbb{P}\{X = 1\} = \mathbb{P}\{X = 0\} = 1/2.$$

The process of obtaining a value for X in practice is random and depends on the experiment of flipping a coin, but the information is deterministic.

We see here that the probability of outcomes 0 and 1 are 0.5 and the probability of any other outcome is 0. All discrete random variables with finitely many possible values have a well-defined*expected value*, written $\mathbb{E}(X)$ and given by

$$\mathbb{E}(X) = \sum_{i \in I} x_i P(X = x_i)$$

where I indexes the set of possible values of X. In the example above, X has expected value $1 \cdot 0.5 + 0 \cdot 0.5 = 0.5$.

Definition 1.6. A random graph, $G(n, p)$ is a graph on n vertices, where two vertices are joined by an edge with probability $p \in [0, 1]$. In a random graph, edges between vertices are constructed randomly, so the probablility of any two vertices being connected is always the same. Constructing a random graph can be very similar to assigning a value to a random variable. For instance, the above variable could instead be

$$e(X, Y) = \left\{ \begin{array}{ll} \text{exists} & \text{if heads} \\ \text{does not exist} & \text{if tails} \end{array} \right\}.$$

This method would allow us to construct edges randomly by flipping a (possibly) biased coin that comes up heads with probability p and tails with probability $1-p$, producing a random graph.

Definition 1.7. A complete bipartite graph $G(V_1 + V_2, E)$ is a graph for which the set of vertices V it divided into two subsets V_1 and V_2 and for which every $v_n \in V_1$ shares an edge with every $w_k \in V_2$. An example of a bipartite graph would be the one pictured in the previous section. The picture below depicts a complete bipartite graph in which there are 3 vertices in V_1 and 5 vertices in V_2.

Since all possible edges between V_1, V_2 exist in a complete bipartite graph, it is not difficult to see that the density of a complete bipartite graph is 1. Also, for a bipartite graph, $G(X, Y)$ where $X \subseteq V_1$ and $Y \subseteq V_2$ it is easy to see that the density of G is merely the ratio of the number of edges in G to the number of edges in the complete bipartite graph.

If we take the ϵ-regular pair (X, Y) as previously defined to be a pair with density close to the density of any large enough subsets $A \subset X$ and $B \subset Y$, we see that because $d(X, Y)$ is close to $d(A, B)$, the probability of any edge existing between X, Y is close to $d(A, B)$. This shows that the edges in the ϵ-regular pair act like edges constructed through a random process, and therefore, the graph is very similar to a random graph: ϵ-regularity is a substitute for randomness.

2. Some inequalities

The proof of the Regularity Lemma requires that we first prove a "defect" form of the Cauchy-Schwarz inequality, which we recall here:

Lemma 2.1. *(Cauchy-Schwarz) Let $\{u_i\}_{1\le i\le n}$, $\{v_i\}_{1\le i\le n}$ be real numbers. Then*

$$\left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right) \ge \left(\sum_{i=1}^{n} u_i v_i\right)^2.$$

We will need a more refined version of this inequality, which we will prove and then explain.

Lemma 2.2. *Let $\{a_i\}_{1\le i\le n}$, $\{b_i\}_{1\le i\le n}$ be nonnegative, and let $\mu > 0$, with*

$$\sum_{i=1}^{n} a_i = 1; \qquad \sum_{i=1}^{n} a_i b_i = b.$$

Suppose there exists $j < n$ such that

$$\sum_{i=1}^{j} a_i b_i \ge b \sum_{i=1}^{j} a_i + \mu.$$

Then

$$\sum_{i=1}^{n} a_i b_i^2 \ge b^2 + \frac{\mu^2}{a(1-a)},$$

where $a := \sum_{i=1}^{j} a_i$.

Proof. We have

$$\begin{aligned}\sum_{i=1}^{n} a_i b_i^2 - b^2 &= \sum_{i=1}^{n} a_i (b_i - b)^2 \\ &= \sum_{i=1}^{j} a_i (b_i - b)^2 + \sum_{i=j+1}^{n} a_i (b_i - b)^2 \\ &\geq \frac{1}{a}\left(\sum_{i=1}^{n} a_i (b_i - b)\right)^2 + \frac{1}{1-a}\left(\sum_{i=j+1}^{n} a_i (b_i - b)\right)^2 \\ &\geq \frac{\mu^2}{a} + \frac{\mu^2}{1-a} \\ &\geq \frac{\mu^2}{a(1-a)},\end{aligned}$$

where the first and last lines both use the fact that $\sum_{i=1}^{n} a_i = 1$.

The third line (i.e. the first inequality) follows from an application of Cauchy-Schwarz to the numbers $u_i = \sqrt{a_i}$ and $v_i = \sqrt{a_i}(b_i - b)$, and the fourth line follows from the hypothesis and the fact that

$$\sum_{i=1}^{n} a_i (b_i - b) = 0.$$

□

There is a probabilistic interpretation of this lemma that we will keep in mind for the rest of the paper. Suppose that we have a random variable X satisfying, for $1 \leq i \leq n$,

$$\mathbb{P}\{X = a_i\} = b_i.$$

Then the quantity b above is precisely $\mathbb{E}[X]$, the expectation of the random variable, and the inequality is a statement about the *variance* $\mathrm{Var}[X] := \mathbb{E}[X^2] - \mathbb{E}[X]^2$. If we want to examine how "random" a graph is, we can use this inequality to bound the variance of, say, the number of edges in the graph.

Theorem 2.3 (Defect C-S)**.** *Suppose $m \leq n$ are such that*

$$\sum_{k=1}^{m} X_k = \frac{m}{n}\sum_{k=1}^{n} X_k + \delta.$$

Then

$$\sum_{k=1}^{n} X_k^2 \geq \frac{1}{n}\left(\sum_{k=1}^{n} X_k\right)^2 + \frac{\delta^2 n}{m(n-m)}.$$

Proof. Apply the previous lemma with $a_i = \frac{1}{n}$ and $b_k = X_k$. With the previous notation, this implies that $\mu = \delta/m$, $b = \frac{1}{n}\sum_{i=1}^{n} X_i$, and $a = m/n$.

This gives us

$$\sum_{k=1}^{n}\frac{1}{n}X_k^2 \geq \left(\frac{1}{n}\sum_{k=1}^{n}X_k\right)^2 + \frac{(\frac{\delta}{m})^2\frac{m}{n}}{1-\frac{m}{n}},$$

which becomes

$$\sum_{k=1}^{n}X_k^2 \geq \frac{1}{n}(\sum_{k=1}^{n}X_k)^2 + \frac{\delta^2 n}{m(m-n)}.$$

□

3. Partitions of graphs

Our goal is to develop a result about subsets of large graphs. To do this we need one more definition:

Definition 3.1. Let $G = (V, E)$ be a simple graph, and let $X \subset V$. A *partition* $P = \{C_0, C_1, \ldots, C_k\}$ of X is a finite set of pairwise disjoint subsets of X, called *classes*, whose union is X. A partition is called *equitable* if $|C_1| = |C_i|$ for $1 \leq i \leq k$. The class C_0, which may have a different cardinality, is called the *exceptional class*.

We measure a partition's size not by the number of elements in the smallest class but by the number of classes in the partition; hence $|P| = k + 1$.

Definition 3.2. Let $G = (V, E)$ be a simple graph, let $X \subset V$, and let P be an equitable partition of X into classes $C_0, C_1, ..., C_k$. P is called *ϵ-regular* if $|C_0| \leq \epsilon k$ and if at most ϵk^2 of the pairs of classes (C_i, C_j) with $1 \leq i, j \leq k$ are ϵ-irregular.

Note that the condition of ϵ-regularity places constraints not just on the classes $C_1, \ldots, C_k$ but also on the exceptional class. Because of this, it is not clear a priori that, for a given ϵ and a given G, any ϵ-regular equitable partition should even exist. We will for the moment take the existence of such objects on faith. When we proceed with the proof of the regularity lemma, it will be necessary to use not only ϵ-regular partitions but refinements of them, so we will need some measure of the 'size' of a partition other than its cardinality.

Definition 3.3. Let $G = (V, E)$ be a simple graph, and let P be an equitable partition of V into classes $C_0, C_1, ..., C_k$. The *index of* P, written $\operatorname{Ind} P$, is the number

$$\operatorname{Ind} P = \frac{1}{k^2}\sum_{1\leq i\leq j\leq k} d(C_i, C_j)^2.$$

(There is a more general definition of index that applies to non-equitable partitions, but it reduces to the expression above). Note that for any equitable partition P, we have $0 \leq \operatorname{Ind} P \leq 1$. Furthermore, if Q is a refinement of P, then $\operatorname{Ind} Q \geq \operatorname{Ind} P$.

Here is the central lemma in the proof of the regularity theorem:

Lemma 3.4. *Let $G = (V, E)$ be a graph on n vertices (i.e. $|V| = n$), and let P be an equitable partition of V into classes $C_0, C_1, ..., C_k$. Choose ϵ such that $4^k > 600\epsilon^{-5}$. Then, if there are more than ϵk^2 pairs (C_i, C_j) such that $0 < i < j \leq k$ are ϵ-irregular, there is another equitable partition Q of V, with $1 + k4^k$ classes, such that the cardinality of the exceptional class of Q is at most $|C_0| + \frac{n}{4^k}$, and such that*

$$\operatorname{Ind} Q \geq \operatorname{Ind} P + \frac{\epsilon^5}{20}.$$

Proof. Suppose (C_i, C_j), $i \neq j$, is an ϵ-irregular pair. Then by definition there exist subsets $X = X(i,j) \subset C_i$ and $Y = Y(i,j) \subset C_j$ such that $|X| \geq \epsilon|C_i|$, $|Y| \geq \epsilon|C_j|$, and

$$|d(X,Y) - d(C_i, C_j)| > \epsilon.$$

Note that, for a fixed i, there are at most $k-1$ sets $X(i,j)$, so the set of subsets of $\{X(i,j)\}_{i\neq j}$ has cardinality 2^{k-1}. Order this power set, writing, say,

$$2^{\{X(i,j)\}_{i\neq j}} = \{\emptyset = A_1, \ldots, A_{2^{k-1}}\}.$$

For a vertex $v \in C_i$, we say $v \in A_m$ if and only if $v \in X(i,j)$ for every $X(i,j) \in A_m$. It is clear that this gives an equivalence relation on the vertices of C_i with at most 2^{k-1} classes (note that some vertices, of course, may not belong to any of the sets $X(i,j)$, i.e. they are associated to A_1). The resulting partition

$$C_i = V_{i,1} \sqcup V_{i,2} \sqcup \cdots \sqcup V_{i,2^{k-1}}$$

is called *atomic*, each piece being called an *atom*, and we have

$$\bigcup_{j=1}^{k} \bigcup_{l=1}^{2^k-1} V_{j,l} = V \setminus C_0.$$

Now for all $1 \leq i \leq k$ set

$$m = \left\lfloor \frac{1}{4^k}|C_i| \right\rfloor.$$

Choose a set Q of pairwise-disjoint subsets of $V \setminus C_0$ such that:

- For each subset B_i of Q, $|B_i| = m$;
- Each atom A of the partition $\{V_{j,l}\}$ has exactly $\left\lfloor \frac{|A|}{m} \right\rfloor$ members of Q;
- Every class C_i has exactly $\left\lfloor \frac{|C_i|}{m} \right\rfloor$ members of Q.

By the definition of m, $|C_i| = m4^k$ for all $i > 1$; thus by the third item above we may assume every class C_i contains 4^k members of Q. Since there are k nonexceptional classes of P, there must then be a total of $k4^k$ members of Q. Q together with C_0 is almost an equitable partition of V. The issue is that Q does not necessarily partition $V \setminus C_0$. Each time we examine an atom A of $\{V_{j,l}\}$, there may be $|A| - \lfloor |A|/m \rfloor$ members not included in Q. If we simply add those left-over vertices to C_0, though, we will obtain an equitable partition of V; what is the size of this new exceptional class C_0'? Certainly the excess vertices in each atom number at most m, and there are k classes C_i, so that

$$|C_0|' \leq |C_0| + km \leq |C_0| + \frac{n}{4^k},$$

as claimed. It remains to show the bound on the index of Q.

Label the $k4^k$ members of Q that are contained in some C_i (i.e. the non-exceptional classes) as $C_i(j)$ where $1 \leq i \leq k$, $1 \leq j \leq 4^k$, and $C_i(j) \subset C_i$ for all i, j. For each i, set

$$C_i^* = \bigcup_{j=1}^{4^k} C_i(j).$$

As in the argument giving the size of C_0' above,

$$|C_i^*| > |C_i| - m > |C_i|\left(1 - \frac{\epsilon^5}{600}\right),$$

since by hypothesis $4^k > 600\epsilon^{-5}$. Thus

$$|d(C_i^*, C_j^*)^2 - d(C_i, C_j)^2| < \frac{\epsilon^5}{50}$$

for all $i < j$. In particular, by Cauchy-Schwarz

$$\begin{aligned}
\frac{1}{(4^k)^2}\sum_{i=1}^{4^k}\sum_{j=i+1}^{4^k} d(C_m(i), C_n(j))^2 &= \frac{1}{(4^k)^2}\sum_{i=1}^{4^k}\sum_{j=i+1}^{4^k} \frac{e(C_m(i), C_n(j))^2}{|C_m(i)|^2|C_n(j)|^2} \\
&= \frac{4^{-2k}}{|C_m(1)|^2|C_n(1)|^2}\sum_{i=1}^{4^k}\sum_{j=i+1}^{4^k} e(C_m(i), C_n(j))^2 \\
&= \frac{4^{2k}}{|C_m^*|^2|C_n^*|^2}\sum_{i=1}^{4^k}\sum_{j=i+1}^{4^k} e(C_m(i), C_n(j))^2 \\
&\geq d(C_m^*, C_n^*)^2 \\
&\geq d(C_m, C_n)^2 - \frac{\epsilon^5}{50}.
\end{aligned}$$

This is almost what we need to give the bound on the index of Q, but we will strengthen it using Theorem 2.3. Fix $i \neq j$, and suppose the pair (C_i, C_j) is ϵ-irregular. Recall that we obtained two sets $X(i,j)$ and $Y(i,j)$ by a process described earlier. Let $X_0 = X(i,j)_0$, resp. $Y_0 = Y(i,j)_0$ be the largest subset of $X(i,j)$, resp. $Y(i,j)$ that is a union of members of Q. Then

$$|X_0| \geq |X(i,j)| - 2^k m > |X(i,j)|(1 - \epsilon/100);$$

here we have used the hypothesis that $|X| \geq \epsilon|C_i|$ and again that $4^k > 600\epsilon^5$.

Let ν be the smallest integer greater than or equal to $\frac{|X|}{m}\left(1 - \frac{\epsilon}{100}\right)$. We may assume without loss of generality that

$$X^* := \bigcup_{n=1}^{\nu} C_i(n), \qquad Y^* := \bigcup_{n=1}^{\nu} C_j(n)$$

are subsets of X, respectively Y, where the $C_i(n), C_j(n)$ are as defined earlier. Having assumed this,

$$|X| \geq |X^*| \geq |X|\left(1 - \frac{\epsilon}{100}\right) \text{ and } |Y| \geq |Y^*| \geq |Y|\left(1 - \frac{\epsilon}{100}\right).$$

Hence the sets $X \setminus X^*$ and $Y \setminus Y^*$ contain at most $\epsilon|X|/100$, respectively $\epsilon|Y|/100$ elements, and we have

$$\begin{aligned}
e(X,Y) &= e(X^*, Y^*) + e(X^*, Y \setminus Y^*) + e(Y^*, X \setminus X^*) + e(X \setminus X^*, Y \setminus Y^*) \\
&\leq e(X^*, Y^*) + \left(\frac{\epsilon}{50} + \frac{\epsilon^2}{100^2}\right)|X||Y|,
\end{aligned}$$

so that

$$\frac{e(X,Y)}{|X||Y|} - \left(\frac{\epsilon}{50} + \frac{\epsilon^2}{100^2}\right) \leq \frac{e(X^*,Y^*)}{|X||Y|} \leq \frac{e(X^*,Y^*)}{|X^*||Y^*|}.$$

For sufficiently small ϵ, we deduce

$$|d(X^*,Y^*) - d(X,Y)| < \frac{\epsilon}{4}.$$

(In fact, this is a specific instance of a more general idea that the density behaves continuously, examined in Szemeredi's original paper). Now, we recall that $|d(X,Y) - d(C_i,C_j)| > \epsilon$ (as well as the definition of C_i^* and C_j^*) to deduce

$$|d(X^*,Y^*) - d(C_i^*,C_j^*)| > \frac{\epsilon}{2}.$$

Finally, we apply 2.3 with $n = 4^{2k}$, $m = \nu^2$, and $\delta = \nu^2(d(X^*,Y^*) - d(C_i^*,C_j^*))$ to find

$$\frac{1}{4^{2k}}\sum_{n=1}^{4^k}\sum_{m=1}^{4^k} d(C_i(n),C_j(m))^2 \geq d(C_i^*,C_j^*)^2 + \frac{\epsilon^2}{4}\cdot\frac{\nu^2}{4^{k^2}-\nu^2} > d(C_i,C_j)^2 - \frac{\epsilon^5}{50} + \frac{\epsilon^4}{16},$$

where we have used the Cauchy-Schwarz argument given earlier in the proof. Then

$$\begin{aligned}
\operatorname{Ind} Q &= \frac{1}{k^2}\sum_{s=1}^{k}\sum_{t=s+1}^{k}\left(\frac{1}{4^{2k}}\sum_{i=1}^{4^k}\sum_{j=1}^{4^k} d(C_s(i),C_t(j))^2\right) \\
&\geq \frac{1}{k^2}\sum_{s=1}^{k}\sum_{t=s+1}^{k}\left(d(C_s,C_t)^2 - \frac{\epsilon^5}{50} + \frac{\epsilon^4}{16}\right) \\
&> \operatorname{Ind} P + \frac{\epsilon^5}{20},
\end{aligned}$$

which finishes the proof. □

4. Regularity Lemma

Having done the brunt of the work in the previous section, we conclude by proving the Regularity Lemma in the same brief way that Szemeredi did in his original paper.

Theorem 4.1. *For every $\epsilon > 0$ and $m \in \mathbb{N}$, there exist positive integers $N(\epsilon,m)$ and $M(\epsilon,m)$ such that for every graph $G(E,V)$ with at least N vertices, there is an ϵ-regular partition of G into $k+1$ classes such that $m \leq k \leq M$.*

Proof. We follow the notation of the previous section. Let s be the smallest integer satisfying $4^s > 600\epsilon^{-5}$, $s \geq m$, and $s > 2/\epsilon$. Construct an iterative sequence $f(k)$ by setting $f(0) = s$ and $f(k+1) = f(k)4^{f(k)}$ for all positive integers k; we think of $f(k)$ as giving the number of classes in the refinement of a partition produced by Lemma 3.4. Let t be the largest integer such that there exists an equitable partition P of V into $1 + f(t)$ classes such that

$$\operatorname{Ind} P \geq \frac{t\epsilon^5}{20}$$

and
$$|C_0| \leq \epsilon n \left(1 - \frac{1}{2^{t+1}}\right).$$
(Such an integer is well-defined, as this partition exists tautologically for $t = 0$ and since the index of P is bounded above by 1). By the maximality of t, Lemma 3.4 implies this partition is ϵ-regular. Setting $M = f(\lfloor 10\epsilon^{-5} \rfloor)$, we have produced the partition we wanted. □

Acknowledgments. I would like to thank my mentor, Zachary Madden for all of his help in writing and editing this paper and Maryanthe Malliaris for Logic lectures.

References

[1] János Komlós, Ali Shokoufandeh, Miklós Simonovits, Endre Szemerédi. The Regularity Lemma and its Applications in Graph Theory
[2] Endre Szemerédi Regular Partitions of Graphs
[3] Fan Chung Graham The Combinatorics of Patterns in Subsets and Graphs