Topics in Probabilistic Method

Selected mainly from Introduction to Graph Ramsey Theory

by Y. Li and W. Zang

Summer School 2017, Supported by NSFC

Jiangsu Normal University

Yusheng Li

Tongji University

Contents

Chapter 1

Semi-random method

1.1 An example for random method

Let $\alpha(G)$ be its independence number. We shall show the following result, called Turán bound, by random method. However, the problem itself is nothing on randomness.

Theorem 1.1 *Let $G = (V, E)$ be a graph of order N and average degree d. Then*

$$\alpha(G) \geq \sum_{v \in V} \frac{1}{1 + d(v)} \geq \frac{N}{1 + d}.$$

Proof. Label all vertices in $V(G)$ randomly by $\{1, 2, \cdots, N\}$. Define a set

$$I = \{v \in V : \ell(v) < \ell(w) \text{ for any } w \in N(v)\},$$

where $\ell(v)$ is the label of v. Note that I is a random set determined by ℓ. Let X_v be the indicator random variable for $v \in I$ and $X = \sum_{v \in V} X_v$, clearly $X = |I|$ and

$$E(X) = \sum_v E(X_v) = \sum_v \Pr[v \in I] = \sum_v \frac{1}{1 + d(v)},$$

since $v \in I$ if and only if v is the least element among v and its neighbors. So there exists a specific labeling such that $|I| \geq E(X)$. Clearly I is always an independent set thus $\alpha(G) \geq |I|$. Then the first inequality follows. The second comes from the convexity of the function $f(x) = 1/(1 + x)$. □

1.2 Semi-random method

For graphs F and H, Ramsey number $r(F,H)$ is defined to be the smallest N such that if G is a graph of order N, then either G contains F or $\overline{G}$ contains H. Let us denote $r(K_k, K_n)$ by $r(k,n)$, called as the *classic Ramsey number*. Thus $r(k,n)$ is the smallest N such that either $\omega(G) \geq k$ or $\alpha(G) \geq n$ for any graph G of order N. Note that $\omega(G) = \alpha(\overline{G})$. Good bounds for $\alpha(G)$ are expected to give good estimation on $r(k,n)$. It is well known that $r(k,n) \leq \binom{k+n-2}{k-1}$. Ajtai, Komlós and Szemerédi (1980) proved that $r(k,n) \leq (5000)^k \frac{n^{k-1}}{(\log n)^{k-2}}$ for fixed $k \geq 2$. Let us gain an overview of the technique employed by them.

A greedy algorithm to obtain an independent set is to put a vertex v into the independent set and then delete all neighbors of v, and repeat the process.

In order to produce a larger independent set by above algorithm, we hope to delete less vertices and more edges in each step so that the remaining graph is sparser. What v should be chosen? To obtain some criterion, we define $Q(v)$ to be the number of edges that incident with a neighbor of v, and define

$$Q_0(v) = \sum_{u \in N(v)} \deg(u).$$

Note that if we delete a vertex v and its neighbors, we delete exactly $Q(v)$ edges.

Lemma 1.1 *Let v be a vertex of a graph G, then*

$$Q(v) \leq Q_0(v),$$

and the equality holds if and only if $N(v)$ contains no edge.

We shall establish a criterion $R(v) \geq 0$ for choosing v. Before this, let us have a property of $Q_0(v)$.

Lemma 1.2 *Let $G = (V,E)$ be any graph with average degree d. Then the average value of $Q_0(v)$ over $v \in V$ is at least d^2.*

Proof. Let N denote the order of G. Then

$$\begin{aligned}\frac{1}{N}\sum_{v\in V} Q_0(v) &= \frac{1}{N}\sum_{v\in V}\sum_{u\in N(v)} \deg(u) = \frac{1}{N}\sum_{u\in V}\deg(u)\sum_{v\in N(u)} 1 \\ &= \frac{1}{N}\sum_{u\in V}\deg^2(u) \geq \left(\frac{1}{N}\sum_{u\in V}\deg(u)\right)^2 = d^2,\end{aligned}$$

where we have used the convexity of function $f(x) = x^2$. □

Ajtai, Komlós and Szemerédi defined a vertex v to be a *groupie* if the average degree of its neighbors is at least the average degree of G. Then every graph has a groupie. This comes from an easy fact that there is some vertex $v \in V$ so that $Q_0(v) - d\deg(v) \geq 0$ as the inequality holds on average from Lemma 1.2. By deleting groupie and its neighbors recursively, they proved that for any triangle-free graph G of order N and average degree d,

$$\alpha(G) \geq cN\frac{\log d}{d}, \tag{1.1}$$

where $c = 1/100$, and $\log x$ is the natural logarithmic function. Ajtai, Erdős, Komlós and Szemerédi conjectured in 1981 that the best constant is $1 - o(1)$, where the small term $o(1)$ tends to zero as $d \to \infty$. By adding the details to the proof for (1.1), Griggs (1983) improved the constant to 1/2.4. These results considerably improved the Turán bound when d is large. Now, let $N = r(3, n) - 1$, there is a triangle-free graph G on N vertices with independence number at most $n - 1$. Since each neighborhood of a triangle-free graph is an independent set, thus $\alpha(G) \geq \Delta(G) \geq d$, and $n - 1 \geq cN\frac{\log(n-1)}{n-1}$, it follows by

$$r(3, n) = N + 1 \leq \left(\frac{1}{c} + o(1)\right)\frac{n^2}{\log n},$$

as $n \to \infty$. This bound is much stronger than the bound $r(3, n) \leq \binom{n+1}{2} = \frac{1}{2}n(n + 1)$. A deep result of Kim (1995) proved that

$$r(3, n) \geq \left(\frac{1}{162} - o(1)\right)\frac{n^2}{\log n}.$$

This fact and the argument just mentioned imply that one cannot expect to improve lower bound (1.1) more than a multiplicative constant even just for triangle-free graphs.

The method of Ajtai, Komlós and Szemerédi is now called "semi-random method" or "nibble method" initialized by Rödl (1985), in which they selected their objects in many small "nibbles" rather than a big "bite", and then analyzed how the nibbles change the structure of the remainder. It was the method that Kim made an elaborate use in his random structure to obtain his lower bound of $r(3, n)$.

In order to find a larger independent set, the key step is to determine the vertex, which together with its neighbors, will be deleted. We now look how Shearer (1983) determined this vertex for triangle-free graphs. Suppose that $f(x)$ is the function for which we want to prove

$$\alpha(G) \geq Nf(d)$$

for a triangle-free graph G. Naturally, we *assume that $f(x)$ is positive, decreasing*, and more importantly, $f(x) \geq c \log x/x$ for some constant $c > 1/100$ when x is sufficiently large. Let $P(v) = d(v) + 1$ and $Q(v)$ denote the number of edges incident with v or one of neighbors of v. Let H denote the graph obtained from G by deleting v and its neighbors. Note that we delete exactly $P(v)$ vertices and $Q(v) = Q_0(v)$ edges since G is triangle-free. Then H has $N - P(v)$ vertices and $Nd/2 - Q(v)$ edges. So its average degree is

$$d_H = \frac{Nd - 2Q(v)}{N - P(v)}.$$

By induction,

$$\alpha(G) \geq 1 + \alpha(H) \geq 1 + [N - P(v)]f(d_H).$$

We do not know which of d and d_H is bigger. That is why the algorithm of Ajtai et. al. deletes a groupie and its neighbors. However we can swap $f(d_H)$ with $f(d)$, if we *further assume that $f(x)$ is convex* so that we can use the fact $f(x) \geq f(d) + f'(d)(x - d)$. Thus we have

$$\begin{aligned} & 1 + [N - P(v)]f(d_H) \\ \geq\ & 1 + (N - P(v))\Big(f(d) + f'(d)\Big(\frac{Nd - 2Q(v)}{N - P(v)} - d\Big)\Big) \\ =\ & Nf(d) + R(v) \end{aligned}$$

where

$$R(v) = 1 - P(v)f(d) - [2Q(v) - dP(v)]f'(d).$$

The only thing left to prove is $R(v) \geq 0$ for some v. To find such a vertex v, let us look the average of $R(v)$ as follows.

$$\begin{aligned} \frac{1}{N}\sum_v R(v) &\geq 1 - (d+1)f(d) - [2d^2 - d(d+1)]f'(d) \\ &= 1 - (d+1)f(d) - d(d-1)f'(d), \end{aligned}$$

where we used the fact that the average of $Q(v)$ is at least d^2 by Lemma 1.2 and the assumption that $f'(x) \leq 0$. So we need that the function $f(x)$ satisfies the following differential equation

$$x(x-1)f'(x) + (x+1)f(x) = 1.$$

Solving this differential equation, Shearer thus obtained that

$$f(x) = \frac{x\log x - x + 1}{(x-1)^2}.$$

Luckily enough, $f(x)$ is positive, decreasing and convex. Moreover, $f(x) \sim \log x/x$ as $x \to \infty$.

1.3 New functions

We shall generalize Shearer's result to improve upper bound of $r(k,n)$ for general k.

Lemma 1.3 *For $m \geq 1$ and $x \geq 0$, the function*

$$f_m(x) = \int_0^1 \frac{(1-t)^{1/m}}{m + (x-m)t}dt$$

satisfies the differential equation

$$x(x-m)f_m'(x) + (x+1)f_m(x) = 1. \tag{1.2}$$

Moreover, $f_m(x)$ is completely monotonic on $[0,\infty)$, that is to say, $(-1)^k f_m^{(k)}(x) > 0$ for all $k \geq 0$ and $x \geq 0$. In particular, $f_m(x)$ is positive, decreasing, and convex.

Proof. By differentiating under the integral and then integrating by parts, we have

$$\begin{aligned} & x(x-m)f'_m(x) \\ = & -x(x-m)\int_0^1 \frac{(1-t)^{1/m}t}{(m+(x-m)t)^2}dt \\ = & x\int_0^1 (1-t)^{1/m}t\frac{d}{dt}\left(\frac{1}{m+(x-m)t}\right)dt \\ = & -x\int_0^1 \left(1-\frac{t}{m(1-t)}\right)\frac{(1-t)^{1/m}}{m+(x-m)t}dt \\ = & -xf_m(x)+\int_0^1 \frac{(1-t)^{1/m}}{m}\left[\frac{1}{1-t}-\frac{m}{m+(x-m)t}\right]dt \\ = & -xf_m(x)+1-f_m(x). \end{aligned}$$

Hence (1.2) follows. The complete monotonicity of $f_m(x)$ can be seen by repeated differentiating under the integral. □

Corollary 1.1 *For $0 \le x \le m$, $f_m(x) \le 1/(1+x)$, and for $m \ge 1$, $f_m(x) \ge \frac{\log(x/m)-1}{x}$.*

Proof. The first statement comes from the differential equation in Lemma 1.3 immediately since $f'_m(x) < 0$. To just the statement for the case $x > m$, note that

$$f_m(x) \ge \int_0^1 \frac{(1-t)dt}{m+(x-m)t} = \frac{x\log(x/m)-(x-m)}{(x-m)^2} > \frac{\log(x/m)-1}{x}.$$

The last inequality holds since it amounts to

$$(2x-m)\log(x/m) > x-m \quad \text{or} \quad (2t-1)\log t > t-1$$

for $t > 1$, which is easy to prove. □

1.4 Independence number of sparse graph

Theorem 1.2 *Let G be a graph with N vertices and average degree d. If any subgraph induced by a neighborhood has the maximum degree at most a, then*

$$\alpha(G) \ge N f_{a+1}(d).$$

Proof. We prove by induction on N, the number of vertices of G. If $N \leq a+2$, then $d \leq a+1$. By the corollary, we have $1/(d+1) \geq f_{a+1}(d)$. It follows from Turán's theorem that $\alpha(G) \geq \frac{N}{d+1} \geq N f_{a+1}(d)$. So we suppose $N > a+2$ hereafter. By the preceding argument, we may also assume $d > a+1$. We shall let G_v stand for the subgraph of a graph G induced by the neighborhood of v. In case some vertex v of G has degree $N-1$, by virtue of Turán's theorem, we have $\alpha(G_v) \geq \frac{N-1}{a+1}$ as the maximum degree of G_v is at most a. It follows that

$$\alpha(G) \geq \alpha(G_v) \geq \frac{N-1}{a+1} \geq \frac{N}{a+2} = N f_{a+1}(a+1) \geq N f_{a+1}(d).$$

So we suppose henceforth that

$$\text{the maximum degree of } G \text{ is at most } N-2.$$

For each $v \in V(G)$, let $P(v) = d(v)+1$ and let $Q(v)$ denote the number of edges of G that are incident with either v or one of its neighbors. Since the average degree of G_v is at most a, G_v contains at most $\frac{a}{2}d(v)$ edges. It follows that

$$Q(v) \geq \sum_{u \in N(v)} d(u) - \frac{a}{2}d(v).$$

Consequently, the average value of Q satisfies

$$\frac{1}{N}\sum_{v \in V} Q(v) \geq d^2 - \frac{ad}{2}.$$

Set

$$R(v) = 1 + [P(v)d - 2Q(v)]f'_{a+1}(d) - P(v)f_{a+1}(d).$$

Note that the coefficient of $Q(v)$ is positive since $f'_{a+1}(d) < 0$. Then

$$\begin{aligned} \frac{1}{N}\sum_{v \in V} R(v) &\geq 1 + [(d+1)d - 2d^2 + ad]f'_{a+1}(d) - (d+1)f_{a+1}(d) \\ &= 1 - d(d-a-1)f'_{a+1}(d) - (d+1)f_{a+1}(d) = 0. \end{aligned}$$

Hence there exists a vertex $v_0 \in V(G)$ such that $R(v_0) \geq 0$. Let$R(v_0) = \hat{R}$, $P(v_0) = \hat{P}$ and $Q(v_0) = \hat{Q}$. Then

$$\hat{R} = 1 + (\hat{P}d - 2\hat{Q})f'_{a+1}(d) - \hat{P}f_{a+1}(d) \geq 0.$$

Delete v_0 and its neighbors from G, in view of that the maximum degree of G is at most $N-2$, we obtain a nontrivial graph H with $N-\hat{P}$ vertices and $Nd/2-\hat{Q}$ edges. Note that any subgraph induced by a neighborhood of H has a maximum degree at most a, so by induction hypothesis,

$$\alpha(H) \geq (N-\hat{P}) f_{a+1}\left(\frac{Nd-2\hat{Q}}{N-\hat{P}}\right).$$

Clearly $\alpha(G) \geq 1+\alpha(H)$. Since f_{a+1} is convex, $f_{a+1}(x) \geq f_{a+1}(d) + f'_{a+1}(d)(x-d)$ for all $x \geq 0$. Combining these two facts, we obtain

$$\begin{aligned}
\alpha(G) &\geq 1+\alpha(H) \geq 1+(N-\hat{P}) f_{a+1}\left(\frac{Nd-2\hat{Q}}{N-\hat{P}}\right) \\
&\geq 1+(N-\hat{P})\left\{f_{a+1}(d)+f'_{a+1}(d)\left(\frac{\hat{P}d-2\hat{Q}}{N-\hat{P}}\right)\right\} \\
&= 1+(N-\hat{P}) f_{a+1}(d)+(\hat{P}d-2\hat{Q}) f'_{a+1}(d) \\
&= N f_{a+1}(d)+\hat{R} \geq N f_{a+1}(d).
\end{aligned}$$

This completes the proof. □

Theorem 1.3 *Let $k \geq 2$ be fixed. Then for all large n,*

$$r(k,n) \leq (1+o(1))\frac{n^{k-1}}{(\log n)^{k-2}}.$$

Theorem 1.4 *Let $k \geq 2$ be fixed. Then for all large n,*

$$r(k,n) \leq (1+o(1))\frac{n^{k-1}}{(\log n)^{k-2}}.$$

Proof We will prove the assertion by induction on k. For $k=2$, it follows the trivial case $r(2,n)=n$.

For the case $k=3$. Let G be graph of order $N=r(3,n)-1$ which contains no triangles and $\alpha(G) \leq n-1$. Since G is triangle-free, each subgraph induced by a neighborhood is empty thus its average degree is zero. Let d be average degree of G. Since each neighborhood is an independence set, $n-1 \geq \alpha(G) \geq d$. By the inequality that f_1 satisfies,

$$n-1 \geq N f_1(d) \geq N f_1(n-1) \geq N\frac{\log(n-1)-1}{n-1}.$$

Then $r(3,n)-1 < \frac{(n-1)^2}{\log n - 1}$, and it follows by $r(3,n) \le \frac{n^2}{\log(n/e)}$ for large n.

Suppose the statement holds for $2,3,\ldots,k$. We proceed to the induction step. Let G be a graph of order $N = r(k+1,n)-1$ such that G contains no K_{k+1} and that $\alpha(G) \le n-1$. Then for each vertex v of G, we have

- the degree of v is at most $r(k,n)-1$, and
- the maximum degree of $\langle N(v)\rangle$ is at most $r(k-1,n)-1$,

where $\langle N(v)\rangle$ is the subgraph induced by $N(v)$. For any $\epsilon > 0$, let

$$d = (1+\epsilon)\frac{n^{k-1}}{(\log n)^{k-2}}, \quad \text{and} \quad m = \left\lfloor (1+\epsilon)\frac{n^{k-2}}{(\log n)^{k-3}} \right\rfloor.$$

From induction hypothesis, $r(k,n) < d$ and $r(k-1,n) < m$ for large n. Thus Theorem 1.2 implies that

$$n > \alpha(G) \ge N f_m(r(k,n)-1) > N f_m(r(k,n)),$$

We may assume that $r(k,n) > m$ as otherwise we are done. By the property of the function $f_m(x)$ in last section as it is decreasing on either of x and m, we obtain

$$n > N\frac{\log(r(k,n)/m)-1}{r(k,n)} \ge N\frac{\log(n/\log n)-1}{(1+\epsilon)n^{k-1}/(\log n)^{k-2}},$$

implies that for large n,

$$r(k+1,n) = N+1 \le (1+2\epsilon)\frac{n^k}{(\log n)^{k-1}},$$

as asserted for case $(k+1)$. □

Chapter 2

Basic probabilistic method

2.1 Random graphs

Probabilistic method basing on mathematical expectation is called basic probabilistic method. We have an example in Lecture 1. To use this method, we should establish a appropriate probabilistic space. The space of random graphs is used often and the graph Ramsey theory is the birthplace of random graphs.

Every probability space whose points are graphs gives a notion of a random graph. For a family of graphs $\mathcal{G} = \{G_1, G_2, \cdots\}$ with probabilities $\Pr(G_i)$ such that $0 \leq \Pr(G_i) \leq 1$ and $\sum_{i\geq 1} \Pr(G_i) = 1$, we have a probability space of random graphs. Each G_i is called a *random graph* of $\mathcal{G}$ with probability $\Pr(G_i)$. We shall consider the probability space that consists of graphs on a fixed set $V = [n] = \{1, 2, \ldots, n\}$, where the vertices in V are *distinguishable*, so edges on V are distinguishable, too. Note that the complete graph K_n on vertex set V has

$$\binom{n}{1} + \binom{n}{2}2 + \cdots + \binom{n}{k}2^{\binom{k}{2}} + \cdots + \binom{n}{n}2^{\binom{n}{2}}$$

subgraphs. The general term corresponds the subgraphs that have exactly k vertices, and the last term $2^{\binom{n}{2}}$ corresponds all spanning subgraphs.

Let us label all edges of K_n on vertex set $V = [n]$ as $e_1, e_2, \ldots, e_m$, where $m = \binom{n}{2}$. Note that the number of graphs on vertex set $[n]$

is 2^m since the edges are distinguishable. The space $\mathcal{G}(n; p_1, \ldots, p_m)$ is defined for $0 \leq p_i \leq 1$ as follows. To get a random element of this space, one selects the edge e_i independently, with probability p_i. Putting it another way, the ground set of $\mathcal{G}(n; p_1, \ldots, p_m)$ is the set of all 2^m graphs on $V = [n]$. For a specific graph H in the space with $E(H) = \{e_j : j \in S\}$, where $S \subseteq \{1, \ldots, m\}$ is the index set of edges of H, the probability that H appears is

$$\prod_{j \in S} p_j \prod_{j \notin S} (1 - p_j).$$

That is to say, each of the edges of H has to be selected and none of $\overline{H}$ is allowed to be selected. Write $q_j = 1 - p_j$ and $G(p_1, \ldots, p_m)$ for a random element in $\mathcal{G}(n; p_1, \ldots, p_m)$, then

$$\Pr(G(p_1, \ldots, p_m) = H)) = \left(\Pi_{j \in S} p_j\right) \left(\Pi_{j \notin S} q_j\right).$$

Since the vertices (and hence edges) are distinguishable, the event $G(p_1, \ldots, p_n) = H$ is different from that $G(p_1, \ldots, p_m)$ is isomorphic to H. To see that $\mathcal{G}(n; p_1, \ldots, p_m)$ is truly a probability space, let us verify that

$$\begin{aligned} \sum_H \Pr(G(p_1, \ldots, p_n) = H)) &= \sum_{S \subseteq [m]} \left(\Pi_{j \in S} p_j\right) \left(\Pi_{j \notin S} q_j\right) \\ &= \Pi_{j=1}^m (p_j + q_j) = 1. \end{aligned}$$

We shall concentrate on the case $p_1 = p_2 = \cdots = p_m = p$, for which the probability space $\mathcal{G}(n; p_1, \ldots, p_m)$ is written as $\mathcal{G}(n, p)$.

In space $\mathcal{G}(n, p)$ the probability of a specific graph H in the space with k edges is $p^k(1-p)^{m-k}$: each of the k edges of H has to be selected and none of $\overline{H}$ is allowed to be selected. Write G_p for a random element of $\mathcal{G}(n, p)$, then

$$\Pr(G_p = H) = p^{e(H)} q^{m - e(H)}.$$

In the space $\mathcal{G}(n, 0)$, the probability that the empty graph $\overline{K_n}$ appears is one, and the probability that any other graph appears is zero. Similarly, in the space $\mathcal{G}(n, 1)$, the only graph that appears is K_n. Other than these two extremal cases, $0 < p < 1$, any graph on vertex set $[n]$ can appear with a positive probability. As p increases from 0 to 1, random graph G_p *evolves* from empty to full.

Another point of view may be convenient in which one colors all edges of the complete graph K_n with probability p, randomly and independently. Thus random graph G_p is viewed as a random coloring of edge set of K_n. The coloring of edge set of K_n is also said a coloring of K_n in short. Recalling the definition of Ramsey numbers, we see why *the relation between random method and Ramsey theory is so natural and tight.*

It is worth remarking that $p = p(n)$ is often a function. The space $\mathcal{G}(n,p)$ is of great interest for fixed values of p as well; in particular, $\mathcal{G}(n,1/2)$ could be viewed as the space: it consists of all 2^m graphs on $V = [n]$, and the probability of any graph is equiprobable. This is just a classical probability space. Thus $G_{n,1/2}$ is also obtained by picking any of the 2^m graphs on $V = [n]$ at random with probability 2^{-m}. No matter how p is fixed or a function, we tend to be interested in what happens as $n \to \infty$.

Now we have obtained a space of random graphs, every graph invariant becomes a random variable; the nature of such a random variable depends crucially on the space. For instant, the number $X_k(G)$ of complete graphs of order k in G is a random variable on our space of random graphs.

To be proficient in the probabilistic method one must have a feeling for asymptotic calculation. For the sake of convenience, we state some simple inequalities that will be used in the calculations. The following precise formula is called *Stirling formula.*

Lemma 2.1 *For all $n \geq 1$,*

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\{\frac{1}{12n+\theta}\},$$

where $0 < \theta = \theta_n < 1$. Thus

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n < n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/12n},$$

and

$$n! = (1+o(1))\sqrt{2\pi n}\left(\frac{n}{e}\right)^n > \left(\frac{n}{e}\right)^n.$$

□

Lemma 2.2 *For any positive integers $N \geq n$,*

$$\binom{N}{n} \leq \frac{N^n}{n!} \leq \left(\frac{eN}{n}\right)^n.$$

If $n = o(\sqrt{N})$ as $n \to \infty$, then

$$\binom{N}{n} \sim \frac{N^n}{n!}.$$

Proof. The first two inequalities are immediate from Stirling's formula, and then it suffices to see

$$\begin{aligned} \frac{N(N-1)\cdots(N-n+1)}{N^n} &= \exp\left[\sum_{i=1}^{n-1} \log\left(1 - \frac{i}{N}\right)\right] \\ &= \exp\left[-\sum_{i=1}^{n-1} \frac{i}{N} - O\left(\frac{n^2}{N}\right)\right], \end{aligned}$$

which goes to 1, and the desired asymptotical formula follows. □

The following simple fact from calculus is often used.

Lemma 2.3 *For any $0 \leq x \leq 1$ and $n \geq 0$, $(1-x)^n \leq e^{-nx}$. If $x = x(n) \to 0$ and $x^2 n \to 0$ as $n \to \infty$, then $(1-x)^n \sim e^{-nx}$.*

2.2 Elementary Examples

In the original proof of the exponent lower bound for $r(n,n)$ in 1947, Erdős did not use the formal probabilistic language. So his paper has been considered as an informal starting point of random graphs. But in two papers published in 1959 and 1961, in which he gave a lower bound $c(n/\log n)^2$ for $r(3,n)$, he even wrote probabilities in the titles.

Theorem 2.1 *For $n \geq 1$,*

$$r(n,n) > \frac{n}{e\sqrt{2}} 2^{n/2}.$$

Proof. Consider the random graphs in $\mathcal{G}(N, 1/2)$, or color K_N randomly and independently with probability $p = 1/2$, where N is a positive integer to be chosen. Let S be a set of n vertices and let A_S be the event that S is monochromatic. Then

$$\Pr[A_S] = 2\left(\frac{1}{2}\right)^{\binom{n}{2}} = 2^{1-\binom{n}{2}},$$

as for S to hold all $\binom{n}{2}$ edges must be colored the same. Consider the event $\bigcup A_S$ over all n-sets on $[N]$. We use the simple fact that the probability of a disjunction is at most the sum of the probability of the events. Thus

$$\Pr[\cup A_S] \le \sum \Pr[A_S] = \binom{N}{n} 2^{1-\binom{n}{2}}.$$

If this probability is less than one, then the event $B = \bigcap \overline{A}_S$ has positive probability. Therefore B is not the null event. Thus there is a point in the probability space for which B holds. But a point in the probability space is precisely a coloring of the edges of K_N. And the event B is precisely that under this coloring there is no monochromatic K_n. Hence $r(n, n) > N$.

We need to find the maximum possible N such that $\Pr[\cup A_S] < 1$. From Stirling formula, we have

$$\binom{N}{n} 2^{1-\binom{n}{2}} \le \frac{N^n}{n!} 2^{1-\binom{n}{2}} < \frac{2}{\sqrt{2\pi n}} \left(\frac{e\sqrt{2}N}{n2^{n/2}}\right)^n.$$

This can be ensured by setting $N = \left\lfloor \frac{n}{e\sqrt{2}} 2^{n/2} \right\rfloor$ such that the fraction in the parenthesis is at most one. Therefore $r(n, n) \ge N + 1 > \frac{n}{e\sqrt{2}} 2^{n/2}$. □

Remark. The original proof of Erdős used the counting argument. He in fact used the space $\mathcal{G}(N, 1/2)$, which is a classical probability space as mentioned. It is interesting to see that this space is the only one that counting argument works! For a property Q of graphs, if the probability of graphs G_p of $\mathcal{G}(n, p)$ satisfying Q tends to 1 as $n \to \infty$, we say that *almost all* graphs G_p satisfies Q. The above argument is an "almost all" argument, but nobody can construct one (family) of them.

Theorem 2.2 *Let m, n and N be positive integers. If for some $0 < p < 1$,*

$$\binom{N}{m} p^{\binom{m}{2}} + \binom{N}{n}(1-p)^{\binom{n}{2}} < 1,$$

then $r(m,n) > N$.

Proof. Consider random graphs G_p in $\mathcal{G}(N,p)$. Let S be a set of m vertices and let A_S be the event that S induces a complete graph. Then $\Pr[A_S] = p^{\binom{m}{2}}$, and

$$\Pr[\cup A_S] \leq \sum \Pr[A_S] = \binom{N}{m} p^{\binom{m}{2}}.$$

Let T be a set of n vertices and let B_T be the event that T induces an independent set. Then

$$\Pr[\cup B_T] \leq \sum \Pr[B_T] = \binom{N}{n} p^{\binom{n}{2}}.$$

Thus

$$\Pr[(\cup A_S) \cup (\cup B_T)] < 1.$$

So there exists a graph on N vertices such that there is neither an induced K_m nor an induced $\overline{K_n}$, thus $r(m,n) > N$. □

The above result is ineffective in bounding $r(3,n)$. We now examine the lower bound of $r(4,n)$. We shall give details in calculation for choosing a suitable value of p, and that of N as large as possible for large n. To make the condition in Theorem 2.2 satisfied, we roughly estimate $\binom{N}{n}$ as $(eN/n)^n$, and $(1-p)^{\binom{n}{2}}$ as $e^{-p\binom{n}{2}}$, hence $\binom{N}{n}(1-p)^{\binom{n}{2}}$ as

$$\left(\frac{eN}{n}\right)^n \exp\left\{-p\binom{n}{2}\right\} = \left(\frac{eN}{ne^{p(n-1)/2}}\right)^n.$$

We have known that $r(4,n) \leq (1+o(1))n^3/(\log n)^2$ in the last chapter, thus $e^{p(n-1)/2} = n^{a+o(1)}$ for some constant a, so we take $p = (c_1 \log n)/(n-1)$. Then

$$\binom{N}{4} p^6 \sim \frac{1}{24} N^4 \left(\frac{c_1 \log n}{n}\right)^6 \sim \frac{c_1^6}{24}\left(N\left(\frac{\log n}{n}\right)^{3/2}\right)^4 < 1,$$

so $N \sim c_2(n/\log n)^{3/2}$ for some constant c_2.

Formally, we let $p = (c_1 \log n)/(n-1)$ and $N = \lfloor c_2(n/\log n)^{3/2} \rfloor$, where c_1 and c_2 are positive constants to be chosen satisfying that $c_1^6 c_2^4 < 24$. Then

$$\binom{N}{4} p^6 < \frac{N^4}{24} p^6 \le \frac{c_1^6 c_2^4}{24} \left(\frac{n}{n-1}\right)^6 \le c_3 < 1$$

for large n, where c_3 is a constant. For the second term, we estimate that $(1-p)^{\binom{n}{2}} < e^{-pn(n-1)/2} = n^{-c_1 n/2}$ and

$$\binom{N}{n}(1-p)^{\binom{n}{2}} < \left(\frac{eN}{n}\right)^n n^{-c_1 n/2} = \left(\frac{eN}{n^{1+c_1/2}}\right)^n.$$

In order to make the above tending to zero, we have to take $c_1 \ge 1$. On the other hand, in order to take c_2 as large as possible with $c_1^6 c_2^4 < 24$, we have to take c_1 as small as possible. So it has to be $c_1 = 1$.

Now, we may hope to optimize the constant c_2. Since we need only $c_2 < 24^{1/4}$, so $c_2 = 24^{1/4} - \epsilon$ will be ok. Thus we have

$$r(4,n) \ge (24^{1/4} - o(1)) \left(\frac{n}{\log n}\right)^{3/2}.$$

Hereafter we will choose p with some foresight. For general $m \ge 4$, by taking $p = (m-3)\log n/(n-1)$, the similar calculation yields

$$r(m,n) \ge c \left(\frac{n}{\log n}\right)^{(m-1)/2}.$$

It is often to replace $\log n/(n-1)$ in the expression of p by $\log n/n$.

We have seen that the property of G_p is sensitive with the value of p. To ensure that G_p contains no K_m (with a positive probability, more precisely, or $\binom{N}{m} p^{\binom{m}{2}}$ is small), it is better to take smaller p. But it is better to take a bigger p to ensure that there is no induced $\overline{K_n}$ (i.e, $\binom{N}{n}(1-p)^{\binom{n}{2}}$ is small). Our task is to balance both sides to obtain a larger N as possible.

We shall improve the obtained lower bounds for $r(n,n)$ and $r(m,n)$ by the proofs so called *deletion method*.

Theorem 2.3 *As $n \to \infty$,*

$$r(n,n) \geq (1-o(1))\frac{n}{e}2^{n/2}.$$

Proof. Consider the random graphs in $\mathcal{G}(N,1/2)$. Let X be the number of clique or independent set of size n. Then

$$X = \sum X_S,$$

the sum over all $n-$set S, where X_S is the indicator random variable of the event A_S that S is a clique or independent. That is

$$X_S = \begin{cases} 1 & \text{if } S \text{ induces } K_n \text{ or } \overline{K_n} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$E[X_S] = \Pr[A_S] = 2\left(\frac{1}{2}\right)^{\binom{n}{2}}.$$

By linearity of expectation

$$E[X] = \sum E[X_S] = \binom{N}{n}2^{1-\binom{n}{2}}.$$

There is a point in the probability space for which X does not exceed its expectation. That is, there is a graph with at most

$$\binom{N}{n}2^{1-\binom{n}{2}}$$

S that induces a K_n or a $\bar{K}_n$. Fix that graph. For each such S select a point $x \in S$ arbitrarily and delete it from the vertex set. The remaining point V^* have neither K_n nor $\bar{K}_n$. Thus

$$r(n,n) > |V^*| \geq N - E(X).$$

The rest of the proof is to find N such that $|V^*|$ as large as possible. By taking $N = \lfloor \frac{n2^{n/2}}{e} \rfloor$, from the Stirling formula, we have

$$\binom{N}{n}2^{1-\binom{n}{2}} < \left(\frac{eN}{n}\right)^n 2^{1-\binom{n}{2}} < 2\left(\frac{e\sqrt{2}N}{n2^{n/2}}\right)^n \leq 2^{n/2+1},$$

which is $o(N)$. Thus $r(n,n) \geq (1-o(1))N$. □

Theorem 2.4 *For any positive integer m, n and N, and any real number $0 < p < 1$,*

$$r(m,n) > N - \binom{N}{m} p^{\binom{m}{2}} - \binom{N}{n}(1-p)^{\binom{n}{2}}.$$

Consequently,

$$r(m,n) \geq c\left(\frac{n}{\log n}\right)^{m/2}$$

for all large n.

Proof. The first assertion is obvious. For the second, setting $N = a(n/\log n)^{m/2}$ and $p = (m-2)\log n/(n-1)$ such that $a - \frac{(m-2)^{\binom{m}{2}}}{m!}a^m > 0$, this is possible since $a^m = o(a)$ as $a \to 0^+$. Then

$$N_1 = \binom{N}{m} p^{\binom{m}{2}} \sim \frac{(m-2)^{\binom{m}{2}} a^m}{m!}\left(\frac{n}{\log n}\right)^{m/2},$$

and

$$N_2 = \binom{N}{n}(1-p)^{\binom{n}{2}} < \left(\frac{eN}{n}\right)^n e^{-pn(n-1)/2} = \left(\frac{eN}{n^{m/2}}\right)^n \to 0.$$

So if $c < a - \frac{(m-2)^{\binom{m}{2}} a^m}{m!}$

$$r(m,n) \geq N - N_1 - N_2 > c\left(\frac{n}{\log n}\right)^{m/2}.$$

□

2.3 Label Vertices Randomly

We have proved the Turán bound $\alpha(G) \geq \sum_v 1/(1+d(v))$ in Lecture 1, where we label vertices randomly. Let us have another result proven in similar way. Let us introduce some notions before presenting results. Given a graph $G = (V, E)$, set

$$N_i(v) = \{w \in V : d(w,v) = i\},$$

which is the set of all the vertices of distance i from vertex v in G, and set $d_i(v) = |N_i(v)|$. Thus $d_0(v) = 1$ and $d_1(v) = d(v)$. We do not distinguish the subset $N_i(v)$ and the subgraph of G induced by $N_i(v)$ when there is no danger of confusion. The graph G is called (m, k)-colorable if $N_i(v)$ is $k-$colorable for any vertex v and any $i \leq m$, that is, there is an assignment of k colors on vertices of $N_i(v)$ so that no two adjacent vertices receive the same color.

Theorem 2.5 *Let $m \geq 2$ and $k \geq 1$ be integers and let $G = (V, E)$ be an (m, k)-colorable graph. Then*

$$\alpha(G) \geq c \left(\sum_{v \in V} d(v)^{1/(m-1)} \right)^{(m-1)/m},$$

where $c = \frac{1}{k2^{(m-1)/m}}$. So if G is d-regular, then $\alpha(G) \geq c|V|^{1-1/m} d^{1/m}$.

Lemma 2.4 *Let $G = (V, E)$ be a $(1, k)$-colorable graph. Then*

$$\alpha(G) \geq \frac{1}{k} \sum_{v \in V} \frac{d_1(v)}{1 + d_1(v) + d_2(v)}.$$

Proof. Randomly label the vertices of G with a permutation of integers $1, 2, \ldots, N$, where $N = |V|$. Let X be the set of all the vertices v such that the minimum label on the vertices in $\{v\} \cup N_1(v) \cup N_2(v)$ is on a vertex in $N_1(v)$. Then the probability that X contains a vertex v is $\frac{d_1(v)}{1+d_1(v)+d_2(v)}$. So the expected size of X is $\sum_{v \in V} \frac{d_1(v)}{1+d_1(v)+d_2(v)}$. It follows that for certain fixed permutation of integers from 1 to N, we have $|X| \geq \sum_{v \in V} \frac{d_1(v)}{1+d_1(v)+d_2(v)}$. We aim to prove that there is an independent set in this X with size at least $|X|/k$.

To this end, we define a relation R on X as follows. Let $u, v \in X$. Call u and v satisfy the relation R if the minimum label on $\{u\} \cup N_1(u) \cup N_2(u)$ is precisely the same as that on the vertices in $\{v\} \cup N_1(v) \cup N_2(v)$. Clearly R is an equivalence relation, and thus X can be partitioned into certain equivalence classes $X_1, X_2, \ldots, X_p$ for some positive integer p. For each $1 \leq i \leq p$, by the definition of relation R, all vertices in X_i share a neighbor v_i in common, in which for any $u \in X_i$, the label of v_i is the minimum label on vertices in $\{u\} \cup N_1(u) \cup N_2(u)$. Hence

$X_i \subseteq N_1(v_i)$ and clearly $v_i \neq v_j$ for $i \neq j$. We claim that there is no edge between X_i and X_j whenever $1 \leq i \neq j \leq p$. To justify it, assume the contrary: some $w_i \in X_i$ is adjacent to some $w_j \in X_j$. Since $w_i \in X$ and since $v_j \in N_1(w_i) \cup N_2(w_i)$, by the definition of X we see that the label on v_i is less than that on v_j. Similarly, by considering w_j, we conclude that the label on v_j is less than that on v_i, yielding a contradiction.

Since $X_i \subseteq N_1(v_i)$ for each $1 \leq i \leq p$ is $k-$colorable there is an independent set Y_i in X_i with $|Y_i| \geq |X_i|/k$. It follows from the above claim that $\cup_{i=1}^{p} Y_i$ is an independent set with size at least $\sum_{i=1}^{p} |X_i|/k = |X|/k$, as desired. □

For a triangle-free graph G, the proof is easier since X itself is independent, so there is no need to introduce the equivalence relation.

Lemma 2.5 *Let $G = (V, E)$ be an (m, k)-colorable graph. Then for any $1 \leq \ell \leq m + 1$, we have*

$$\alpha(G) \geq \frac{1}{2k} \sum_{v \in V} \frac{1 + d_1(v) + \cdots + d_{\ell-1}(v)}{1 + d_1(v) + \cdots + d_\ell(v)}.$$

Proof. The proof goes along the same line as that of the preceding lemma, so we only give a sketch here.

Randomly label the vertices of G with a permutation of the integers $1, 2, \ldots, N$, where $N = |V|$. Let X be the set of all the vertices v such that the minimum label on the vertices in $\cup_{j=0}^{\ell} N_j(v)$ is on a vertex $\cup_{j=0}^{\ell-1} N_j(v)$. Then for certain fixed permutation of the integers from 1 to N, we have

$$|X| \geq \sum_{v \in V} \frac{1 + d_1(v) + \cdots + d_{\ell-1}(v)}{1 + d_1(v) + \cdots + d_\ell(v)}.$$

We aim to prove that there is an independent set in this X with size at least $|X|/(2k)$.

To this end, define an equivalence relation R on X such that u and v satisfy R if the minimum label on the vertices in $\cup_{j=0}^{\ell} N_j(u)$ is precisely the same as that on the vertices in $\cup_{j=0}^{\ell} N_j(v)$. Then X can be partitioned into certain equivalence classes $X_1, X_2, \ldots, X_p$ for some

integer $p \geq 1$. It can be shown that for each $1 \leq i \leq p$, there exists a vertex v_i such that

• the distance between each vertex in X_i and v_i is at most $\ell - 1$; and

• v_i is the vertex with the minimum label in $\cup_{j=0}^{\ell} N_j(v)$ for each $v \in X_i$.

Based on these v_i, we can deduce that there is no edge between X_i and X_j whenever $1 \leq i \neq j \leq p$. Now partition each X_i into subsets $X_{i,j}$, $1 \leq j \leq \ell - 1$, such that the distance between every vertex in $X_{i,j}$ and v_i is j. Since each $X_{i,j}$ contains an independent set $Y_{i,j}$ of size at least $|X_{i,j}|/k$, where $1 \leq i \leq p$ and $0 \leq j \leq \ell - 1$. Thus one of $\cup_{i=1}^{p} \cup_{odd\, j} Y_{i,j}$ and $\cup_{i=1}^{p} \cup_{even\, j} Y_{i,j}$ is an independent set with size at least $\frac{1}{2k} \cup_{i=1}^{p} \cup_{j=0}^{\ell-1} X_{i,j} = \frac{|X|}{2k}$, completing the proof. □

Proof of Theorem 2.5. Applying Lemma 2.4 and Lemma 2.5 repeatedly, we have

$$\begin{aligned} \alpha(G) \;\geq\; & \frac{1}{k(m-1)} \sum_{v \in V} \left(\frac{d_1(v)}{1 + d_1(v) + d_2(v)} \right. \\ & + \frac{1 + d_1(v) + d_2(v)}{2(1 + d_1(v) + d_2(v) + d_3(v))} \\ & \left. + \cdots + \frac{1 + d_1(v) + \cdots + d_{m-1}(v)}{2(1 + d_1(v) + \cdots + d_m(v))} \right). \end{aligned}$$

Since the arithmetic mean is no less than the geometric mean, we obtain

$$\alpha(G) \geq \frac{1}{k 2^{(m-2)/(m-1)}} \sum_{v \in V} \left(\frac{d_1(v)}{1 + d_1(v) + \cdots + d_m(v)} \right)^{1/(m-1)}.$$

By the condition that $N_i(v)$ is $k-$colorable, there is an independent set in $N_i(v)$ with size $\alpha_i(v) \geq d_i(v)/k$, and by the fact that there is no edge between $N_i(v)$ and $N_j(v)$ whenever $i - j = 0 \pmod 2$,

$$2\alpha(G) \geq 1 + \alpha_1(v) + \cdots + \alpha_m(v) \geq \frac{1}{k}[1 + d_1(v) + \cdots + d_m(v)].$$

Therefore

$$\alpha(G) \geq \frac{1}{k 2^{(m-2)/(m-1)}} \sum_{v \in V} \left(\frac{d_1(v)}{2k\alpha(G)} \right)^{1/(m-1)},$$

the desired statement follows. □

Theorem 2.5 can be used to shown $r(C_{2m+1}, K_n) \leq c\left(\frac{n^{m+1}}{\log n}\right)^{1/m}$ for large n. We give proof for the case $m = 2$, and the proof for general case is similar.

Theorem 2.6 *Let $n \geq e^{e^2}$. Then*

$$r(C_5, K_n) \leq 6\frac{n^{3/2}}{\sqrt{\log n}}.$$

Proof. Let G be a graph on $N = r(C_5, K_n) - 1$ vertices that contains no C_5 and $\alpha(G) \leq n - 1$. We consider two cases depending on the value of d, the average degree of G. Note that if G is C_5-free, then G is $(2, 3)$-colorable.

Case 1. $d > 3\sqrt{n \log n}$. By Theorem 2.5, we have $\alpha(G) \geq \sqrt{Nd/18}$. It follows that $n - 1 > \sqrt{3N\sqrt{n \log n}/18}$, implying

$$N + 1 \leq \frac{6(n-1)^2}{\sqrt{n \log n}} + 1 \leq \frac{6n^{3/2}}{\sqrt{\log n}}.$$

Case 2. $d \leq 3\sqrt{n \log n}$. Since G is C_5-free, each neighborhood of G does not contain any path of length 3, according to a theorem in Lecture 1, we have

$$n - 1 \geq N\frac{\log(3\sqrt{n \log n}/3) - 1}{3\sqrt{n \log n}} \geq N\frac{\sqrt{\log n}}{6\sqrt{n}}.$$

It follows that

$$N + 1 \leq \frac{6(n-1)\sqrt{n}}{\sqrt{\log n}} + 1 \leq \frac{6n^{3/2}}{\sqrt{\log n}},$$

completing the proof. □

2.4 Pick Vertices Randomly

A *dominating set* of a graph $G = (V, E)$ is a set $U \subseteq V$ such that $U \cup N(U) = V$, where $N(U) = \cup_{u \in U} N(u)$. The *domination number*

is the is the smallest cardinality among all dominating sets of G. Let $\beta(G)$ denote the domination number of G. Clearly

$$\alpha(G) \geq \beta(G)$$

since any maximal independent set is a dominating set. We would like to ask what G look like if G is triangle-free and $\alpha(G)$ and $\beta(G)$ are close?

Recall that we have that for any triangle-free graph G with N vertices and average degree d, then $\alpha(G) \geq Nf(d)$, where $f(x) \sim \log x/x$ as $x \to \infty$. A similar bound for domination number du to Alon (1990) is as follows, in which the involved function has the same asymptotical form. We shall write $N[v]$ for $N(v) \cup \{v\}$ and $N[X]$ for $X \cup N(X)$.

Theorem 2.7 *Let G be a graph with N vertices and minimum degree $\delta \geq 1$. Then*

$$\beta(G) \leq N\frac{1+\log(\delta+1)}{\delta+1}.$$

Proof. Let us pick a vertex v of G with probability p randomly and independently, where $p = \log(\delta+1)/(\delta+1)$. Let X be the random set of picked vertices and let $Y = V \setminus N[X]$. The set $X \cup Y$ is clearly a dominating set of G which can be picked with the cardinality as small as its expected number. The expected values of $|X|$ is Np. A vertex v belongs to Y if and only if neither v nor any neighbor belongs to X, so

$$E(|Y|) = \sum_{v \in V}(1-p)^{1+d(v)} \leq N(1-p)^{1+\delta}.$$

Thus

$$E(|X|+|Y|) \leq N\left[p + (1-p)^{1+\delta}\right].$$

Since

$$(1-p)^{\delta+1} \leq e^{-p(\delta+1)} = e^{-\log(\delta+1)} = \frac{1}{\delta+1},$$

then we have

$$E(|X|+|Y|) \leq N\frac{1+\log(\delta+1)}{\delta+1},$$

as desired. □

In many cases, the probabilistic method supplies effective randomized algorithms for various problems. In some cases, these algorithms can be converted into deterministic ones. The aim of derandomization is to convert probabilistic proofs of existence of combinatorial structures into efficient deterministic algorithms for their actual structures. The following proof of Alon (1990) can be viewed as a derandomization of the proof of the above theorem.

A constructive proof of Theorem 2.7. Let us run a greedy algorithm as follows. At the initial step, let $X = \{x\}$, where $\deg(x) = \Delta(G)$, and let $Y = V \setminus N[X]$, where $V = V(G)$. The set Y consists of vertices that are not dominated by X. In general step, we shall select a vertex $z \in V \setminus X$ such that z dominates the most of vertices of Y, and enlarge X by adding z into X. We claim that z dominates at least $(\delta + 1)|Y|/N$ vertices of Y. In fact, note the trivial fact that

$$\sum_{y \in Y} |N[y]| \geq (\delta + 1)|Y|,$$

in which the sum counts the vertices in $V \setminus X$ only. On average, each vertex in $V \setminus X$ is counted

$$\frac{1}{N - |X|} \sum_{y \in Y} |N[y]| > \frac{(\delta + 1)|Y|}{N}$$

times by these subsets $N[y]$ in the sum, so some vertex $z \in V \setminus X$ appears at least $(\delta + 1)|Y|/N$ times, proving the claim.

We iteratively select a vertex in $V \setminus X$ that dominates the most of the vertices in Y. After each step of selection, the ratio of remaining vertices is at most $1 - (\delta+1)/N$. Hence after $N \log(\delta+1)/(\delta+1)$ steps, the number of remaining vertices is at most

$$N \left(1 - \frac{\delta + 1}{N}\right)^{N \log(\delta+1)/(\delta+1)} < Ne^{-\log(\delta+1)} = \frac{N}{\delta + 1}.$$

The selected vertices and these remaining vertices together form a dominating set of size at most $N(1 + \log(\delta + 1))/(\delta + 1)$. □

Chapter 3

The Lovász Local Lemma

László Lovász (born March 9, 1948, recipient of the 1999 Wolf Prize) is a Hungarian-American mathematician, best known for his work in combinatorics, who served as president of the International Mathematical Union between January 1, 2007 and December 31, 2010

In probability theory, if a large number of events are all independent of one another and each has probability less than 1, then there is a positive (possibly small) probability that none of the events will occur. The Lovász local lemma (a weaker version was proved in 1975 by Lovász and Erdős) allows one to relax the independence condition slightly: As long as the events are "mostly" independent from one another and they are not individually too likely, then there will still be a positive probability that none of them occurs. It is most commonly used in the probabilistic method, in particular to give existence proofs. Differing from the "almost all" argument, we are concerned with the existence of the event of small positive probability. There are several different versions of the lemma. The simplest and most frequently used is the symmetric version, and a complicated version is the general form, which can be used to improve some results in previous chapters or simplify the proofs. The last section is a brief description of a constrained random graph process, which makes the event appears with a larger probability in the constrained process.

3.1 The local lemma

Let $A_1, A_2, \ldots, A_n$ be the events in a probability space $(\Omega, \mathcal{F}, \Pr)$. In combinatorial application such as a coloring of edges of K_N , any A_i is a "bad" event. We wish that no "bad" event happens, that is to say, we wish to show

$$\Pr(\cap \overline{A}_i) > 0 \tag{3.1}$$

so there is a point (coloring) which is good. In proofs in Chapter 4 for lower bound of $r(n, n)$, A_S is the event that S is monochromatic for an n-set S. The event A_S is "bad". Then $\cap \overline{A}_S$ is the event in which none of n-sets is monochromatic. $\Pr(\cap \overline{A}_S) > 0$ means that there must be a coloring in which no "bad" thing happens. That is, there is an edge coloring of K_N, which determines a graph G such that G contains no K_n and whose complement contains no K_n either.

It is a trivial fact that if each $\Pr(A_i)$ is small such that

$$\sum_{i=1}^{n} \Pr(A_i) < 1,$$

then

$$\Pr(\cap \overline{A}_i) = 1 - \Pr(\cup A_i) \geq 1 - \sum \Pr(A_i) > 0.$$

On the other hand, if $A_1, \ldots, A_n$ are mutually independent events, i.e., any A_i is independent of any Boolean function of other A_j, and $\Pr(A_i) = x_i < 1$ for $1 \leq i \leq n$ then

$$\Pr(\cap \overline{A}_i) = \prod_{i=1}^{n} (1 - x_i) > 0.$$

with no further restriction on probabilities $\Pr(A_i)$.

The Lovász Local Lemma may be understood in term of taking advantage of *partial independence* of the events $A_1, A_2, \ldots, A_n$ so that (3.1) can be ensured with far weaker bounds on the probabilities $\Pr(A_i)$ than that are needed for $\sum \Pr(A_i) < 1$.

The argument to follow uses *conditional probability*. Recall that for events A and B with $\Pr(B) > 0$, the conditional probability $\Pr(A|B)$ is given by

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

If A and B are independent then $\Pr(A|B) = \Pr(A)$. By admitting that an event of zero probability is independent of any other event, then two events A and B are independent if and only if $\Pr(A \cap B) = \Pr(A)\Pr(B)$.

Let us introduce a graph to describe the dependency of events. A graph D on vertices $[n] = \{1, 2, \ldots, n\}$ (the set of indices for the events A_i) is called *dependency graph* for events $A_1, A_2, \ldots, A_n$ if for every i, the event A_i is mutually independent of all A_j with $ij \notin E(D)$ and $j \neq i$. That is, A_i is independent of any Boolean function of these events in $\{A_j : j \notin N[i]\}$. This graph must contain edges between the pairs of dependent events, and it contains such edges only in most applications so the term dependency graph is after. The original local lemma is as follows.

Theorem 3.1 *If each of the events of $A_1, A_2, \ldots, A_n$ has probability p or less, each vertex in the dependence graph has degree at most $d \geq 1$, and if*

$$4dp \leq 1,$$

then $\Pr(\cap_{i=1}^{n} \overline{A}_i) > 0$.

The following form of the Lovász Local Lemma is called its general form, see Spencer (1977).

Theorem 3.2 *Let $A_1, A_2, \ldots, A_n$ be the events in a probability space $(\Omega, \mathcal{F}, \Pr)$. Suppose that there exist real numbers $x_1, x_2, \ldots, x_n$ such that $0 < x_i < 1$ and for $i = 1, 2, \ldots, n$,*

$$\Pr(A_i) \leq x_i \Pi_{j:\, ij \in E(D)} (1 - x_j).$$

Then $\Pr(\cap_{i=1}^{n} \overline{A}_i) \geq \Pi_{i=1}^{n} (1 - x_i) > 0$.

We remark that if i is an isolated vertex in D, that is, the event A_i is mutually independent of all other events, then the neighborhood of vertex i in D is empty, and $\Pi_{j \in \emptyset}(1 - x_j) = 1$.

Proof. The desired result follows directly from the following claim.

Claim. For $S \subset [n]$, set

$$\mathcal{C}_S = \cap_{j \in S} \overline{A}_j.$$

(For $S = \emptyset$, we take $\mathcal{C}_S$ to be Ω). If $i \notin S$, then

$$\Pr(A_i|\mathcal{C}_S) \leq x_i.$$

Proof of the claim. The proof is by induction on $|S|$. If $|S| = 0$ the desired result is immediate since the hypothesis of the local lemma yields

$$\Pr(A_i|\mathcal{C}_S) = \Pr(A_i|\Omega) = \Pr(A_i) \leq x_i \Pi_{j:\, ij \in E(D)}(1 - x_j) \leq x_i.$$

Now assume that $|S| \geq 1$ and form a partition $S = (S_1, S_2)$, where

$$S_1 = \{j \in S : ij \in E(D)\} \quad \text{and} \quad S_2 = S \setminus S_1.$$

Let us write $\Pr(A_i|\mathcal{C}_S)$ as

$$\frac{\Pr(A_i \cap \mathcal{C}_S)}{\Pr(\mathcal{C}_S)} = \frac{\Pr(A_i \cap \mathcal{C}_{S_1} \cap \mathcal{C}_{S_2})}{\Pr(\mathcal{C}_{S_1} \cap \mathcal{C}_{S_2})} = \frac{\Pr(A_i \cap \mathcal{C}_{S_1}|\mathcal{C}_{S_2})}{\Pr(\mathcal{C}_{S_1}|\mathcal{C}_{S_2})},$$

and bound the numerator and denominator separately. First, since A_i and $\mathcal{C}_{S_2}$ are independent,

$$\Pr(A_i \cap \mathcal{C}_{S_1}|\mathcal{C}_{S_2}) \leq \Pr(A_i|\mathcal{C}_{S_2}) = \Pr(A_i) \leq x_i \Pi_{j \in S_1}(1 - x_j).$$

To bound the denominator, we use the induction hypothesis. If $|S_1| = 0$, then

$$\Pr(\mathcal{C}_{S_1}|\mathcal{C}_{S_2}) = \Pr(\Omega|\mathcal{C}_{S_2}) = 1$$

and the claim follows. Otherwise, suppose $S_1 = \{j_1, j_2, \ldots, j_r\}$, where $r \geq 1$. Let $\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_r$ be the events defined recursively by

$$\mathcal{D}_0 = \mathcal{C}_{S_2} = \cap_{j \in S_2} \overline{A}_j$$

and for $k = 1, 2, \ldots, r$,

$$\mathcal{D}_k = \mathcal{D}_{k-1} \cap \overline{A}_{j_k} = \left(\cap_{j \in S_2} \overline{A}_j\right) \cap \left(\cap_{t=1}^{k} \overline{A}_{j_t}\right).$$

They start with $\mathcal{D}_0 = \mathcal{C}_{S_2}$ and end with $\mathcal{D}_r = \mathcal{C}_S$. Note that for each $k = 0, 1, \ldots, r-1$, the event $\mathcal{D}_k$ has a form of $\mathcal{C}_T$ for some set $T \subseteq S$

with $|T| < |S|$ and the fact that $\Pr(\overline{A}_j|\mathcal{D}_k) = 1 - \Pr(A_j|\mathcal{D}_k)$. Using the induction hypothesis on $\mathcal{C}_T$ repeatedly, we have

$$\begin{aligned}\Pr(\mathcal{C}_{S_1}|\mathcal{C}_{S_2}) &= \frac{\Pr(\mathcal{C}_S)}{\Pr(\mathcal{D}_0)} = \frac{\Pr(\mathcal{D}_r)}{\Pr(\mathcal{D}_0)} = \frac{\Pr(\mathcal{D}_r)}{\Pr(\mathcal{D}_{r-1})} \cdots \frac{\Pr(\mathcal{D}_1)}{\Pr(\mathcal{D}_0)} \\ &= \Pr(\overline{A}_{j_r}|\mathcal{D}_{r-1}) \cdots \Pr(\overline{A}_{j_1}|\mathcal{D}_0) \\ &= (1 - \Pr(A_{j_r}|\mathcal{D}_{r-1})) \cdots (1 - \Pr(A_{j_1}|\mathcal{D}_0)) \\ &\geq \Pi_{j \in S_1}(1 - x_j).\end{aligned}$$

Combining this with what proved we have established the claim.

Note that $\cap_{i=k+1}^{n} \overline{A}_i$ has a form of $\mathcal{C}_S$ with $k \notin S$. In view of the claim just established,

$$\begin{aligned}\Pr(\cap_{i=1}^{n} \overline{A}_i) &= \Pr(\overline{A}_1 | \cap_{i=2}^{n} \overline{A}_i) \Pr(\cap_{i=2}^{n} \overline{A}_i) \\ &= \Pr(\overline{A}_1 | \cap_{i=2}^{n} \overline{A}_i) \Pr(\overline{A}_2 | \cap_{i=3}^{n} \overline{A}_i) \cdots \Pr(\overline{A}_n|\Omega) \\ &= (1 - \Pr(A_1 | \cap_{i=2}^{n} \overline{A}_i)) \cdots (1 - \Pr(A_n|\Omega)) \\ &\geq \Pi_{i=1}^{n}(1 - x_i).\end{aligned}$$

This completes the proof of the local lemma. □

Let us call the following form of the local lemma as *symmetric form.*

Theorem 3.3 *If each of the events of $A_1, \ldots, A_n$ has probability p or less, each vertex in the dependence graph has degree at most d, and if*

$$e(d+1)p \leq 1,$$

where e is natural logarithm base, then $\Pr(\cap_{i=1}^{n} \overline{A}_i) > 0$.

Proof. By taking $x_i = 1/(d+1)$ for $1 = 1, 2, \ldots, n$, we shall show

$$\Pr(A_i) \leq x_i \Pi_{j:\, ij \in E(D)}(1 - x_j).$$

Since for any i the right side is at least

$$\frac{1}{d+1}\left(1 - \frac{1}{d+1}\right)^d > \frac{1}{e(d+1)} \geq p,$$

where the first inequality holds from the fact $\left(1+\frac{1}{k}\right)^k < e$. □

Note that the original condition $4dp \leq 1$ can be implied by Theorem 3.3 as $4dp \geq e(d+1)p$ for $d \geq 3$, and if $d = 1, 2$, then $\frac{1}{d+1}\left(1-\frac{1}{d+1}\right)^d \geq p$.

We also need the following form of local lemma due to Spencer who used it to obtain his lower bound for non-diagonal classic Ramsey numbers. This form is slightly more convenient for some applications.

Corollary 3.1 *Let $A_1, A_2, \ldots, A_n$ be events in a probability space. If there exist numbers $y_1, y_2, \ldots, y_n$ such that for each i, $0 < y_i \Pr(A_i) < 1$, and*

$$\log y_i \geq -\sum_{j:\, ij \in E(D)} \log(1 - y_j \Pr(A_j)),$$

then $\Pr(\cap \overline{A_i}) > 0$.

Proof. We may suppose that for each i the probability $\Pr(A_i)$ is positive. Let x_i be as in the general form of the local lemma and set $y_i = x_i/\Pr(A_i)$ for $i = 1, 2, \ldots, n$. The hypothesis of the local lemma

$$\Pr(A_i) \leq x_i \Pi_{j:\, ij \in E(D)}(1 - x_j)$$

will take the form

$$y_i \geq \Pi_{j:\, ij \in E(D)} \frac{1}{1 - y_j \Pr(A_j)}.$$

The assertion holds from taking logarithms on both sides of the above inequality. □

Let us have an example to explain that for the local lemma, the dependency graph D may contain more edges other than these connecting pairs of dependent events.

Let $\{1, 2, 3\}$ be the vertex set of a K_3, and let the probability space Ω consists of all 2-coloring of the vertices, in which each vertex is assigned to color red or blue with probability $1/2$ randomly and independently. Then $|\Omega| = 8$. For $i < j$, let A_{ij} be the event that the edge $\{i, j\}$ is monochromatic. Clearly $\Pr(A_{12}) = \Pr(A_{13}) = \Pr(A_{23}) = 1/2$. Also events A_{12}, A_{13}, A_{23} are pairwise independent as

$$\Pr(A_{12}A_{13}) = \Pr(A_{12}A_{23}) = \Pr(A_{13}A_{23}) = \frac{1}{4}.$$

If we mistakenly use the local lemma by letting $E(D) = \emptyset$, then we would set $x_{ij} = 1/2$ with $\Pr(A_{ij}) \leq x_{ij}\Pi_{\emptyset}(1 - x_{k\ell})$. Thus we had a wrong conclusion that $\chi(K_3)$ were at most two by $\Pr(\cap\overline{A_{ij}}) > 0$.

Erdős and Spencer (1991) pointed that the dependency graph D can be replaced by a graph F on $[n]$ if F satisfies that for each i and each $S \subseteq [n] \setminus N_F[i]$,

$$\Pr\left(A_i \Big| \cap_{j\in S} \overline{A}_j\right) \leq x_i \Pi_{j:\, ij\in E(F)}(1 - x_j).$$

This condition contains conditional probabilities. To avoid to compute these probabilities in applications and to have a slightly stronger form, we shall specify their idea further. Let us have the following definitions first from Erdős and Spencer (1991), and Lu and Székely (2007).

Let A and B be events. Then B is said to be *positive* or *negative* to A if $\Pr(A|B) \geq \Pr(A)$ or $\Pr(A|B) \leq \Pr(A)$, respectively. When the inequality holds strictly, then B is said to be *strictly positive* or *strictly negative* to A, respectively. For a set S of indices of events, we write

$$\mathcal{C}_S = \cap_{j\in S}\overline{A}_j,$$

in which $\mathcal{C}_S$ is admitted as Ω if $S = \emptyset$. A graph F on $[n]$ (the set of indices for the events) is called *negative* of events $A_1, A_2, \ldots, A_n$ if for every i and any set $S \subseteq [n] \setminus N[i]$, the event $\mathcal{C}_S$ is negative to A_i, where $N[i]$ is the closed neighborhood of i in F.

We call F as *negative graph* as it is the "real" negative graph in most applications, which contains edges ij such that A_i and A_j are *strictly negative* each other.

Note that a dependency graph is a negative graph, but the latter may contain less edges, and the dependency graph in the local lemma can be replaced by any negative graph.

It is easy to verify the local lemma holds if we replace the dependency graph by a negative graph.

Property 3.1 *The local lemma holds when the dependency graph is replaced by a negative graph of the events.*

Since negative graphs are bipartite for most applications, the local lemma with negative graph will be easier to apply.

Let us remark that the dependency of events can be described by a directed graph instead of a graph. A directed graph D on vertices $[n] = \{1, 2, \ldots, n\}$ is called *directed dependency graph* for events $A_1, A_2, \ldots, A_n$ if each event A_i is is mutually independent of the events in $\{A_j : j \notin N^+[i]\}$, where $N^+[i]$ is the closed out-neighborhood of i. Then the condition to guarantee $\Pr(\cap\overline{A}_i) > 0$ is that there exist $0 < x_1, x_2, \ldots, x_n < 1$ such that

$$\Pr(A_i) \leq x_i \Pi_{j:\,(i,j)\in E(D)}(1 - x_j),$$

where (i, j) is the arc from i to j in directed dependency graph D for the events.

Similarly, the negative graph in the local lemma can be replaced by a directed negative graph.

However, no matter using negative graph or directed dependency graph in the local lemma, the idea is to reduce the edges in the dependency graph.

3.2 Applications of the local lemma

Local Lemma is one of important mathematical contributions of Lovász, and it has a lot applications in recent years in many fields. The local lemma was invented to prove a result on coloring for hypergraphs. Call a coloring of vertices of a hypergraph $\mathcal{H}$ to be *proper*, if no edge is monochromatic, and call $\mathcal{H}$ as k-colorable if there is a proper k-coloring for its vertices. Using the original condition $4dp \leq 1$, Erdős and Lovász (1975) proved that an r-uniform hypergraph $\mathcal{H}$ is 2-colorable if each edge of $\mathcal{H}$ intersects at most 2^{r-3} other edges. As the first application of the local lemma, this result becomes a specific problem in derandomization of the local lemma. See, e.g., Beck (1991). We now improve this result slightly.

Theorem 3.4 *Let $\mathcal{H} = (V, \mathcal{E})$ be an r-uniform hypergraph. If each edge of $\mathcal{H}$ intersects at most $e^{-1}2^r - 2$ other edges, the $\mathcal{H}$ is 2-colorable.*

Proof. Set $V = \{1, 2, \ldots, n\}$ and $\mathcal{E} = \{e_1, e_2, \ldots, e_m\}$. Let the probability space consist of all 2-colorings of V, in which each vertex is

colored by red and blue with probability 1/2 randomly and independently. Let A_i and B_i be the event that e_i is monochromatically red and blue, respectively. Then $\Pr(A_i) = \Pr(B_j) = 1/2^r$. Clearly events A_i and B_j are strictly negative each other if and only if $e_i \cap e_j \neq \emptyset$. Let us connect such pairs of A_i and B_j, and we thus obtain a negative graph F of events $A_1, \ldots, A_m, B_1, \ldots, B_m$, which is a bipartite graph. By assumption, each vertex of F has degree at most $d \leq e^{-1}2^r - 1$, and thus $ep(d+1) \leq 1$. So $\Pr\left((\cap_i \overline{A}_i) \cap (\cap_j \overline{B}_j)\right) > 0$, and there is a proper vertex coloring of $\mathcal{H}$ in two colors. □

Before the above-mentioned result of Erdős and Lovász, a similar result had appeared shown by basic probabilistic method as follows.

Theorem 3.5 (Erdős-Selfridge) *Let $\mathcal{H} = (V, \mathcal{E})$ be an r-uniform hypergraph. If $|\mathcal{E}| < 2^{r-1}$, the $\mathcal{H}$ is 2-colorable.*

Proof. The basic probabilistic method gives

$$\Pr(\cup A_e) \leq \sum \Pr(A_e) = \frac{|\mathcal{E}|}{2^{r-1}} < 1,$$

where A_e be the event that the edge e is monochromatic in the space defined as that in the proof of the last theorem. □

Let us return to our main purpose. The following theorem of Spencer (1975) improves Theorem 4.4 in Chapter 4 with another factor $\sqrt{2}$. This is negligible when viewed in the light of the gap between the upper bound and lower bounds, but we do what we can. The progress on this difficult problem has been slow.

Theorem 3.6 *As $n \to \infty$,*

$$r(n, n) \geq (1 - o(1))\frac{\sqrt{2}}{e} n 2^{n/2}.$$

Proof. Consider the random graph space $\mathcal{G}(N, 1/2)$ or consider to color each edge of K_N with probability 1/2, randomly and independently. Let S be a subset of size n of $V(K_N)$ and let A_S signify the event "S is monochromatic", S ranging over the n-subsets. Define a graph D with vertex set consisting all such S and connect vertices S and T in

D if and only if $|S \cap T| \geq 2$. Then A_S is mutually independent of all A_T with T not adjacent to S, since the A_T give information only about edges outside of S. Hence D is a dependency graph. We apply the local lemma with

$$p = \Pr(A_S) = 2^{1-\binom{n}{2}}.$$

And for any S, its degree d in D can be bounded as

$$d = |\{T : |S \cap T| \geq 2\} < \binom{n}{2}\binom{N}{n-2}.$$

If $ep(d+1) < 1$ then $\Pr(\cap \overline{A}_S) > 0$ thus $r(n,n) > N$. So we want

$$e\binom{n}{2}\binom{N}{n-2}2^{1-\binom{n}{2}} < 1.$$

As we did it before, the left hand side is less than

$$\frac{en^2}{2}\left(\frac{eN}{n-2}\right)^{n-2}\frac{2}{2^{n(n-1)/2}} = \frac{en^2}{2}\left(\frac{n}{n-2}\right)^{n-2}\left(\frac{eN}{\sqrt{2}n2^{n/2}}\right)^{n-2}.$$

For any $\epsilon > 0$, if we take $N = \lceil (1-\epsilon)\frac{\sqrt{2}}{e}n2^{n/2}\rceil$, then the above tends to zero. □

The first application of the general form of the local lemma was made by Spencer (1977) who gave a lower bound

$$r(m,n) \geq c\left(\frac{n}{\log n}\right)^{(m+1)/2},$$

which improves that obtained in Chapter 4. Erdős, Faudree, Rousseau and Schelp(1987), and Krivelevich(1995) generalized Spencer's lower bound from K_m to a fixed graph F; Li and Zang did it to $r(F, G_n)$, where the order of G_n is n and $e(G_n) = n^{2-o(1)}$. Dong, Li and Lin (2009) did it further. Set

$$\rho(F) = \frac{e(F)-1}{v(F)-2},$$

where $v(F)$ and $e(F)$ are the order and the size of F, respectively.

Recall the automorphism group of a graph G on n vertices defined in Chapter 4, denoted by $\mathcal{A}(G)$, for which $|\mathcal{A}(G)| \leq n!$.

Theorem 3.7 *Let F be a fixed graph with $v(F) \geq 3$, and let G_n be a graph of order n with average degree $d_n \to \infty$. Then for all large n*

$$r(F, G_n) \geq c\Big(\frac{d_n}{\log d_n}\Big)^{\rho(F)},$$

where $c = c(F) > 0$ is a constant.

Proof. Let $m = v(F)$ and $\rho = \rho(F)$. Clearly we may assume that $\rho > 1$ and d_n is sufficiently large. Color the edges of K_N by red and blue randomly and independently, in which each edge is colored red with probability p and blue with probability $q = 1 - p$. For each subgraph S of K_N that is isomorphic to F, let A_S be the event that S spans a red F. For each subgraph T of K_N that is isomorphic to G_n, let B_T be the event that T spans a blue G_n. Then $\Pr(A_S) = p^{e(F)}$ and $\Pr(B_T) = q^{e(G_n)}$. Since F has m vertices, there are $(N)_m/|\mathcal{A}(F)|$ events of form A_S. Similarly, there are $(N)_n/|\mathcal{A}(G_n)|$ events of form B_T. Obviously, a pair of distinct events are strictly negative each other if and only if they are of different types and the corresponding subgraphs have edges in common. Any pair of events of the same type are positive each other and a pair of events of different types that do not have common edges are independent. Hence, each A event is strictly negative to at most $(N)_n/|\mathcal{A}(G_n)| < N^n$ of the B events; each B event is strictly negative to at most $e(G_n)(N-2)_{m-2} < e(G_n)N^{m-2}$ of the A events. By connecting the events of different types that have common edges, we have a negative graph for these events.

We aim to prove that there exist positive numbers a and b satisfying Spencer's form of Local Lemma, namely, $ap^{e(F)} < 1$ and $bq^{e(G_n)} < 1$ hold with $y_i = a$ for each A event and $y_j = b$ for each B event. Specifically,

$$\log a \geq -N^n \log(1 - bq^{e(G_n)}), \tag{3.2}$$
$$\log b \geq -e(G_n)N^{m-2} \log(1 - ap^{e(F)}). \tag{3.3}$$

If such a and b are available, then $r(F, G_n) > N$. To this end, set $a = 2$,

$$p = \frac{6\rho \log d_n}{d_n},\ b = \exp\big(\rho\, n \log d_n\big),\ N = c\Big(\frac{d_n}{\log d_n}\Big)^{\rho},$$

where $c = c(F)$ is a constant to be chosen. Using the basic inequality $q = 1 - p < e^{-p}$ for $p > 0$, we have

$$\begin{aligned} N^n b q^{e(G_n)} &\le N^n b e^{-pe(G_n)} = \exp\Big\{n \log N + \log b - p\frac{nd_n}{2}\Big\} \\ &\le \exp\Big\{ - \rho\, n \log d_n\Big\} \to 0. \end{aligned}$$

So $bq^{e(G_n)} \to 0$ thus $\log(1-x) \sim -x$ for $x = bq^{e(G_n)}$, and the right-hand side of (3.2) tend to zero, and thus (3.2) holds for all large n.

Note that the right-hand side of (3.3) is asymptotically

$$e(G_n)N^{m-2}ap^{e(F)} = (6\rho)^{e(F)}c^{m-2}\Big(n \log d_n\Big).$$

So (3.3) holds if we choose c such that

$$\rho > (6\rho)^{e(F)}c^{m-2},$$

and the proof is completed. □

Corollary 3.2 *For fixed $m \ge 3$ as $n \to \infty$,*

$$c_m \left(\frac{n}{\log n}\right)^{(m+1)/2} \le r(m,n) \le (1+o(1))\frac{n^{m-1}}{(\log n)^{m-2}},$$

where $c_m = c(m) > 0$ is constant.

The lower bound on classical Ramsey numbers will be improved in the next section. However, the local lemma is still a generally powerful tool for applications and the proof is comparatively simple, particularly, for the problem involving system without specific structure.

The following result improves the lower bound obtained by deletion method as $r(C_m, K_n) \ge c(n/\log n)^{m/(m-1)}$.

Corollary 3.3 *Let integer $m \ge 3$ be fixed. Then there exists constant $c = c(m) > 0$ so that*

$$r(C_m, K_n) \ge c\left(\frac{n}{\log n}\right)^{(m-1)/(m-2)}$$

for all large n.

Corollary 3.4 *For any fixed integers $k \geq m \geq 2$, there exist constant $c = c(m,k) > 0$ so that*

$$r(K_{m,k}, K_n) \geq c\left(\frac{n}{\log n}\right)^{(mk-1)/(m+k-2)}$$

for all large n.

We will prove that $r(K_{m,k}, K_n) \leq c\left(\frac{n}{\log n}\right)^m$ later. Note that the exponent $(mk-1)/(m+k-2)$ in the lower bound can be arbitrarily close to the exponent m in the upper bound for fixed large k.

We have seen that the probabilistic method has a lot applications with much better results than that by elementary combinatorial method. However, we shall see some other methods, such as the algebraic methods, have much success for some topics in Ramsey theory in the next several chapters. Let us conclude this section with two jokes given by Spencer (1994) to say that for many topics, unlike that for Turán's bound for independence number shown in Chapter 4, the probabilistic method cannot provide "exact" results often. The problem that Spencer joked is serious. In order to have verisimilitude, we write the joked results as usual in "academic language" but without indices.

The proof following results is due to Joker, who used the basic probabilistic method.

Theorem (Joker) *Let S and T be nonempty sets. If $|T| > \binom{|S|}{2}$, then there exists an injection $f: S \to T$.*

Proof. Consider the probability space consisting of all maps from S to T, in which each map appears equiprobably and independently. For any unordered pair of points x and y of S, let $A_{xy} = A_{yx}$ signify the event $f(x) = f(y)$. Since for fixed pair x and y

$$|\{f : S \to T : f(x) = f(y)\}| = |T|^{|S|-1},$$

we have $\Pr(A_{xy}) = 1/|T|$ and

$$\Pr\left(\cup_{\{x,y\}\subseteq S} A_{xy}\right) \leq \sum_{\{x,y\}\subseteq S} \frac{1}{|T|} = \frac{1}{|T|}\binom{|S|}{2} < 1,$$

which implies that $\Pr(\cap_{\{x,y\}\subseteq S}\overline{A}_{xy}) > 0$ and the desired injection exists. □

Later Joker amused himself by improving the above result by using the local lemma. The new result is tight up to a multiplicative constant.

Theorem (Joker) *Let S and T be nonempty sets. If $|T| \geq 2e|S|$, then there exists an injection $f : S \to T$.*

Proof. The same as that for Joke 1 but apply the local lemma. In the dependence graph, the vertex A_{xy} is adjacent to $A_{xy'}$ with $y' \in S \setminus \{y\}$ and $A_{x'y}$ with $x' \in S \setminus \{x\}$. Let $d = 2(|S| - 1)$, then the independence graph is d regular, in which the event A_{xy} is mutually independent to all non-neighbors. As $p = 1/|T|$, the condition ensures $e(d+1)p < 1$, so the symmetric form of the Local Lemma gives that $\Pr(\cap_{\{x,y\}\subseteq S}\overline{A}_{xy}) > 0$, implying the existence of the desired injection. □

3.3 Triangle-free process ⋆

It is often difficult to show the existence of small events. The local lemma is a tool for such proof that improved most lower bounds from basic probabilistic method. The key for the proof of the local lemma itself is conditional probability. A revolutionary idea for finding the small events is also "conditional". If we know a certain condition in which the event is likely to appear, then the probability for event is large under the condition. In other word, we try to switch a small event to be a large one conditionally. However, we may encounter the difficulty to finger the conditional probability out.

Obtaining the right order of magnitude of $r(m, n)$ even $r(3, n)$ was certainly a challenge in decades. A celebrated result of Kim (1995) showed that of $r(3, n)$ is $n^2/\log n$, which was obtained again by Bohman (2009). They used different analysis on the same random graph process, called the triangle-free process. For general constrained graph process, see, e.g., Ruciński and Wormald (1992), Erdős, Suen and Winkler (1995), Bollobás and Riordan (2000), and Osthus and Taraz (2001).

The triangle-free process can be described as follows. We begin with the empty graph, denoted by G_0, on N vertices. At step i we form the graph G_i by adding a new edge to G_{i-1} chosen uniformly at random

from the collection of pairs of vertices that neither appear as edges in G_{i-1} nor form triangles when added as edges to G_{i-1}. The process terminates at a maximal triangle-free graph G_M, for which the random variable M is the number of edges of G_M. Note that a maximal triangle-free graph is connected and the number of edges in a triangle-free graph of order N is at most $N^2/4$ (see Chapter 8), we have

$$N - 1 \leq M \leq \frac{N^2}{4}.$$

However, Bohman (2009) proved that almost surely (a.s.)

$$c_1 N^{3/2}\sqrt{\log N} \leq M \leq c_2 N^{3/2}\sqrt{\log N}.$$

From a result in Chapter 3, we have that the independence numbers of such graphs are at least $\Omega(\sqrt{N \log N})$. Remarkably, Kim and Bohman show that a.s. independence numbers of such graphs are at most $O(\sqrt{N \log N})$, which implies that $r(3, n) \geq \Omega(n^2/\log n)$.

Theorem 3.8 *For some constant $c > 0$,*

$$r(3, n) \geq \frac{cn^2}{\log n}.$$

Let us talk a bit more on the process employed by Bohman. For a set V, let $V^{(2)}$ be the set of all pairs u, v of V, which is the edge set of complete graph on V. The vertex set of our complete graph of order N is on $[N] = \{1, 2, \ldots, N\}$. In the evolution of the triangle-free process, we shall track the some random sets. Recall that G_i is the graph given by the first i edges selected by the process. The graph G_i partitions $[N]^{(2)}$ into three parts: E_i, O_i and C_i. The set E_i is simply the edge set of G_i. A pair of $[N]^{(2)}$ is open, and in the set O_i, if it can still be added as an edge without violating the triangle-free condition. A pair of $[N]^{(2)}$ is closed, and in the set C_i, if it is neither an edge in the graph nor open; that is, a pair $e = \{u, v\}$ is in C_i if $e \notin E_i \cup O_i$ and there exists a vertex w such that $\{u, w\}, \{v, w\} \in E_i$. Note that e_{i+1} is chosen uniformly at random from O_i. That is to say, each edge of O_i has the same probability $1/|O_i|$ to be chosen as e_{i+1}. We do not

express this as $\Pr(e_{i+1} \in O_i) = 1/|O_i|$ as only edges in random set O_i are available.

The proofs of Kim and Bohman are hard and tedious and thus are omitted. With more complicated analysis on K_4-free process, Bohman (2009) also improved the known lower bound of $r(4, n)$, and generally, Bohman and Keevash (2010) improved the lower bound of $r(m, n)$ obtained from the local lemma by a factor $(\log n)^{1/(m-2)}$ as

$$r(m, n) \geq c\Big(\frac{n}{\log n}\Big)^{(m+1)/2}(\log n)^{1/(m-2)}.$$

References

J. Beck, An algorithmic approach to the Lovász local lemma, *Random Structures Algorithms*, **2** (1991), 343-365.

T. Bohman, The triangle-free process, *Adv. Math.*, **221** (2009), 1653-1677.

T. Bohman and P. Keevash, The early evolution of H-free process, *Invent. Math.*, **181** (2010), 291-336.

B. Bollobás, O. Riordan , Constrained graph processes, *Electron. J. Combin.* **7** (2000) R18.

L. Dong, Y. Li and Q. Lin, Ramsey numbers involving graphs with large degrees, *Appl. Math. Lett.*, **22** (2009), 1577-1580.

P. Erdős, R. Faudree, C. Rousseau, and R. Schelp, A Ramsey problem of Harary on graphs with prescribed size, *Discrete Math.*, **67** (1987), 227-233.

P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraph and some related questions, in: *Infinite and Finite Sets (to Paul Erdős on His 60th Birthday 1973) II, A. Hajnal, R. Rado and V. Sos Eds.*, Colloq. Math. Soc. Janos Bolyai, North-Holland, Amsterdam/London, 1975.

P. Erdős and J. Spencer, Lopsided Lovász lemma and Latin transversals, *Discrete Appl. Math.*, 30 (1991), 151-154.

P. Erdős, S. Suen, P. Winkler, On the size of a random maximal graph, *Random Structures Algorithms* **6** (1995) 309-318.

J. Kim, The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$, *Random Structures Algorithms* **7** (1995), 173-207.

M. Krivelevich, Bounding Ramsey numbers through large deviation inequalities, *Random Structures Algorithms*, **7** (1995), 145-155.

Y. Li and W. Zang, Ramsey numbers involving large dense graphs and bipartite Turán numbers, *J. Combin. Theory Ser. B*, **87** (2003), 280-288.

L. Lu and L. Székely, Using Lovász local lemma in the space of random injections, Electron. J. Combin. 14 (2007), no. 1, 63

D. Osthus, A. Taraz, Random maximal H-free graphs, *Random Structures Algorithms*, **18** (2001), 61-82.

A. Ruciński, N. Wormald, Random graph processes with degree restrictions, *Combin. Probab. Comput.* **1** (1992) 169-180.

J. Spencer, Ramsey's theorem-a new lower bound, *J. Combin. Theory Ser. A*, **18** (1975), 108-115.

J. Spencer, Asymptotic lower bound for Ramsey functions, *Discrete Math.*, **20** (1977), 69-76.

J. Spencer, *Ten Lectures on the Probabilistic Method, 2nd Edition*, SIAM, Philadelphia, 1994.

Chapter 4

Concentration

4.1 The Chernoff's Inequality

The probability space we consider in graph Ramsey theory has only finite many possible outcomes, and the random variable is often non-negative. Let X be a random variable, the expected value of X is defined to be $E(X) = \sum_i a_i \Pr(X = a_i)$, where the summation is taken over all values a_i that X can take.

Theorem 4.1 (Markov's Inequality) *Let $a > 0$ and let X be a non-negative random variable. Then*

$$\Pr(X \geq a) \leq \frac{E(X)}{a}.$$

Proof. Suppose that $\{a_i\}$ is the set of all values that X takes. Then

$$\begin{aligned} E(X) &= \sum_i a_i \Pr(X = a_i) \\ &\geq \sum_{a_i \geq a} a_i \Pr(X = a_i) \\ &\geq a \sum_{a_i \geq a} \Pr(X = a_i) = a \Pr(X \geq a), \end{aligned}$$

as required. □

Corollary 4.1 *If a random variable X only takes nonnegative integer values and $E(X) < 1$, then $\Pr(X \geq 1) < 1$ hence $\Pr(X = 0) > 0$.*

This is exactly what we used to obtain lower bounds of Ramsey numbers in the last chapter.

For a positive integer k, the *kth moment* of a real-valued random variable X is defined to be $E(X^k)$, and so the first moment is simply the expected value. Denote by $\mu = E(X)$, and define the *variance* of X as $E((X - \mu)^2)$, which is denoted by σ^2. Call

$$\sigma = \sqrt{E((X - \mu)^2)}$$

as the *standard deviation* of X. A basic equality is as follows.

$$\sigma^2 = E(X^2) - \mu^2.$$

Theorem 4.2 (Chebyshev's Inequality) *Let X be a random variable and let a be a positive number. Then*

$$\Pr(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}.$$

Proof. By Markov's inequality, for any $a > 0$,

$$\begin{aligned} \sigma^2 &= E((X - \mu)^2)) \\ &\geq a^2 \Pr((X - \mu)^2 \geq a^2) \\ &= a^2 \Pr(|X - \mu| \geq a). \end{aligned}$$

It follows by the required statement. □

In importance, the second moment $E(X^2)$ is second to the first moment $E(X)$.

Lemma 4.1 (Second Moment Method) *If X is a random variable, then*

$$\Pr(X = 0) \leq \frac{\sigma^2}{\mu^2} = \frac{E(X^2) - \mu^2}{\mu^2},$$

where $\mu = E(X)$. In particular, $\Pr(X = 0) \to 0$ if $E(X^2)/\mu^2 \to 1$.

The proof follows from Chebyshev's Inequality and the trivial fact that $\Pr(X = 0) \leq \Pr(|X - \mu| \geq \mu)$ immediately. Intuitively, if σ grows more slowly than μ grows, than $\Pr(X = 0) \to 0$ since σ "pulls" X close to μ thus far away from zero.

The Chebshev's inequality is in fact the Markov's inequality on random variable $|X - \mu|$. However, Chebshev's inequality states the probability of a random variable X apart from $E(X)$ is bounded. When this is the case, we say that X is *concentrated*. A concentration bounds is used to show that a random variable is very close to its expected value with high probability, so it behaves approximately as one may "expect" it to be. When S_n is the sum of n independent variables, each variable equals to 1 with probability p and -1 with probability $1 - p$, respectively, the bound can be sharper. Such random variables are bounded in Chernoff's inequality. Most of the results in this chapter may be found in, or immediately derived from, the seminal paper of Chernoff (1952) while our proofs are self-contained. A set of random variables $X_1, X_2, \ldots$ are said to be mutually independent means each X_i is independent of any Boolean expression formed from other $(X_j)'s$. In any form of Chernoff bounds, we assume that

Assumption A : *On the independence of variables in Chernoff bound* Let $X_1, X_2, \ldots$ be mutually independent variables and they have the same binomial distribution. Set

$$S_n = \sum_{i=1}^{n} X_i.$$

All concentration bounds in the remaining part of this section are Chernoff bounds of different forms, which estimate the probability of

$$\Pr(S_n \geq n(\mu + \delta)),$$

where $\mu = E(X_i)$. The symmetric bound on $\Pr(S_n \leq n(\mu - \delta))$ can be obtained similarly.

Theorem 4.3 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$$

for $i = 1, 2, \ldots$. Then for any $\delta > 0$,

$$\Pr(S_n \geq n\delta) < \exp\{-n\delta^2/2\},$$

and for any $a > 0$,

$$\Pr(S_n \geq a) < \exp\{-a^2/(2n)\}.$$

Proof. Let $\lambda > 0$ be arbitrary. Then

$$E(e^{\lambda X_i}) = \frac{e^\lambda + e^{-\lambda}}{2}.$$

Note that

$$\begin{aligned} E(e^{\lambda S_n}) &= E(e^{\lambda X_1})E(e^{\lambda X_2}) \ldots E(e^{\lambda X_n}) \\ &= \left(\frac{e^\lambda + e^{-\lambda}}{2}\right)^n = \left(\sum_{j=0}^{\infty} \frac{\lambda^{2j}}{(2j)!}\right)^n \\ &< \left(\sum_{j=0}^{\infty} \frac{1}{j!}\left(\frac{\lambda^2}{2}\right)^j\right)^n = e^{n\lambda^2/2}, \end{aligned}$$

where we use the fact that $(2j)! \geq 2^j j!$ for all $j \geq 0$ with strict inequality when $j \geq 2$. Now by Markov's inequality,

$$\begin{aligned} \Pr(S_n \geq n\delta) &= \Pr(e^{\lambda S_n} \geq e^{\lambda n\delta}) \\ &\leq \frac{E(e^{\lambda S_n})}{e^{\lambda n\delta}} \\ &< \exp\{n(\lambda^2/2 - \lambda\delta)\}, \end{aligned}$$

for all $\lambda > 0$. Setting $\lambda = \delta$, we obtain the desired result. □

For large n, the central limit theorem implies that S_n is approximately normal with zero mean and standard deviation $\sqrt{n}$. For any fixed u,

$$\lim_{n\to\infty} \Pr(S_n \geq u\sqrt{n}) \int_u^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt < e^{-u^2/2}.$$

However, the Chernoff bound holds for all positive n and a.

Since X_i is often an indicator variable of some random event, so X_i takes 1 when the event appears and 0 otherwise. The following form of Chernoff bound may be used in more cases.

Theorem 4.4 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = \Pr(X_i = 0) = \frac{1}{2}$$

for $i = 1, 2, \ldots$. *Then for any* $\delta > 0$,

$$\Pr(S_n \geq n(1+\delta)/2) < \exp\{-n\delta^2/2\}.$$

Namely

$$\Pr(S_n \geq n(1/2+\delta) < \exp\{-2n\delta^2\}.$$

Proof. Set $Y_i = 2X_i - 1$ and $T_n = \sum_{i=1}^n Y_i = 2S_n - n$. Then

$$\Pr(Y_i = 1) = \Pr(Y_i = -1) = \frac{1}{2},$$

and $\{Y_i\}$ satisfies Assumption A. Note that $T_n \geq n\delta$ if and only if $S_n \geq n(1+\delta)/2$. Applying Theorem 4.3 to $\{Y_i\}$ and T_n, we have

$$\Pr(S_n \geq n(1+\delta)/2) = \Pr(T_n \geq n\delta) < \exp\{-n\delta^2/2\}$$

as claimed. □

Under Assumption A, suppose

$$\Pr(X_i = 1) = p, \quad \text{and} \quad \Pr(X_i = 0) = 1 - p$$

for $i = 1, 2, \ldots$. Then we say that the sum $S_n = \sum_{i=1}^n X_i$ has binomial distribution, denoted by $B(n, p)$. Involved in Theorem 4.4 is special binomial distribution $B(n, 1/2)$. For general case , the calculation is slightly more complicated, but the technique is the same. As usual, denote by q for $1 - p$.

Theorem 4.5 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = p \quad \textit{and} \quad \Pr(X_i = 0) = q$$

for $i = 1, 2, \ldots$. *Then there exists* $\delta_0 = \delta_0(p) > 0$ *so that if* $0 < \delta < \delta_0$, *then*

$$\Pr(S_n \geq n(p+\delta)) < \exp\{-n\delta^2/(3pq)\}.$$

Proof. Denote by a for $p+\delta$. By the same argument as used before,

$$\begin{aligned}\Pr(S_n \geq na) &= \Pr(e^{\lambda S_n} \geq e^{\lambda na}) \\ &\leq \frac{1}{e^{\lambda na}} E(e^{\lambda S_n}) \\ &= \frac{1}{e^{\lambda na}} (pe^{\lambda} + q)^n \\ &= (pe^{\lambda(1-a)} + qe^{-\lambda a})^n\end{aligned}$$

for all $\lambda > 0$. Let $c = 1 - a = q - \delta > 0$, then $a + c = 1$. By taking $\lambda = \log(aq/cp)$, we have

$$\begin{aligned}\min_{\lambda>0}(pe^{\lambda c} + qe^{-\lambda a}) &= e^{-\lambda a}(pe^{\lambda} + q) \\ &= \left(\frac{cp}{aq}\right)^a \frac{q}{c} \left(\frac{p}{a}\right)^a \left(\frac{q}{c}\right)^c.\end{aligned}$$

Setting $0 < \delta < 1 - p$, and expanding in powers of δ, with the fact that

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + O(x^4),$$

we find

$$\begin{aligned}\log\left(\frac{p}{a}\right)^a &= (p+\delta)\log\left(1 - \frac{\delta}{p+\delta}\right) \\ &= -\delta - \frac{\delta^2}{2(p+\delta)} - \frac{\delta^3}{3(p+\delta)^2} + o(\delta^3),\end{aligned}$$

and

$$\begin{aligned}\log\left(\frac{q}{c}\right)^c &= (q-\delta)\log\left(1 + \frac{\delta}{q-\delta}\right) \\ &= \delta - \frac{\delta^2}{2(q-\delta)} + \frac{\delta^3}{3(q-\delta)^2} + o(\delta^3).\end{aligned}$$

Adding them by terms, the first sum vanishes, and the second is

$$\begin{aligned}\frac{-\delta^2}{2}\left(\frac{1}{p+\delta} + \frac{1}{q-\delta}\right) &= \frac{-\delta^2}{2}\left(\frac{1}{p(1+\delta/p)} + \frac{1}{q(1-\delta/q)}\right) \\ &= \frac{-\delta^2}{2}\left(\frac{1}{pq} - \frac{(q^2-p^2)\delta}{p^2q^2} + o(\delta)\right) \\ &= \frac{-\delta^2}{2pq} + \frac{(q-p)\delta^3}{2p^2q^2} + o(\delta^3),\end{aligned}$$

and the third is

$$\begin{aligned}\frac{\delta^3}{3}\left(\frac{1}{(q-\delta)^2}-\frac{1}{(p+\delta)^2}\right) &= \frac{\delta^3}{3}\left(\frac{1}{q^2}-\frac{1}{p^2}+o(1)\right)\\ &= \frac{-(q-p)\delta^3}{3p^2q^2}+o(\delta^3).\end{aligned}$$

We have for small $\delta > 0$

$$\log\left[\left(\frac{p}{a}\right)^a\left(\frac{q}{c}\right)^c\right] = \frac{-\delta^2}{2pq}+\frac{(q-p)\delta^3}{6p^2q^2}+o(\delta^3) < \frac{-\delta^2}{3pq}.$$

Thus

$$\Pr(S_n \geq n(p+\delta)) < \exp\{-n\delta^2/(3pq)\},$$

completing the proof. □

From above proof for $p > q$ and Theorem 4.4 for $p = q = 1/2$, we see that if $p \geq 1/2$, the bound can be slightly better as

$$\Pr(S_n > n(p+\delta)) < \exp\{-n\delta^2/(2pq)\}.$$

We now write out a symmetric form for Theorem 4.5, and omit those for Theorem 4.3 and Theorem 4.4.

Theorem 4.6 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = p \quad \textit{and} \quad \Pr(X_i = 0) = q$$

for $i = 1, 2, \ldots$. *Then there exists* $\delta_0 = \delta_0(p) > 0$ *so that if* $0 < \delta < \delta_0$, *then*

$$\Pr(S_n \leq n(p-\delta)) < \exp\{-n\delta^2/(3pq)\}.$$

Therefore

$$\Pr(|S_n - np| > n\delta)) < 2\exp\{-n\delta^2/(3pq)\}.$$

□

From the above proof, we have

$$\begin{aligned}\Pr(S_n \geq na) &\leq \left(\left(\frac{p}{a}\right)^a \left(\frac{q}{c}\right)^c\right)^n \\ &= \exp\left\{n\left(a\log\frac{p}{a} + (1-a)\log\frac{q}{1-a}\right)\right\},\end{aligned}$$

where $a = p + \delta$ and $c = 1 - a$. Set $k = na$, then $k > np$ and

$$\Pr(S_n \geq k) \leq \exp\left\{n\left((k/n)\log\frac{p}{k/n} + (1-k/n)\log\frac{q}{1-k/n}\right)\right\}.$$

Let $H(x)$ signify the entropy function

$$H(x) = x\log\frac{p}{x} + (1-x)\log\frac{q}{1-x}, \quad 0 < x < 1,$$

then

$$\Pr(S_n \geq k) \leq \exp\{nH(k/n)\},$$

which is valid also for $k = np$ since $H(p) = 0$. The following form of Chernoff's inequality was used by Beck (1983).

Theorem 4.7 *Under Assumption A, suppose*

$$\Pr(X_i = 1) = p \quad \textit{and} \quad \Pr(X_i = 0) = q$$

for $i = 1, 2, \ldots$. *If* $k \geq np$, *then*

$$\Pr(S_n \geq k) \leq \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}.$$

Consequently,

$$\Pr(S_n \geq k) \leq \left(\frac{npe}{k}\right)^k.$$

Proof. The right hand side of the first inequality is just $\exp\{nH(k/n)\}$. For the second inequality, simply note that

$$\left(\frac{nq}{n-k}\right)^{n-k} \leq \left(\frac{n}{n-k}\right)^{n-k} = \left(1 + \frac{k}{n-k}\right)^{n-k} < e^k.$$

Thus the required result follows. □

4.2 Applications of Chernoff's Bounds

Let us first see that a. a. graphs are nearly regular.

Theorem 4.8 *Let $0 < p < 1$ and $\epsilon > 0$ be fixed. Then almost all graphs G in $\mathcal{G}(n,p)$ satisfy that*

$$|\deg(v) - (n-1)p| \leq \epsilon(n-1)p$$

for each vertex v.

Proof. Let G be a random graph in $\mathcal{G}(n,p)$ and let v ba a fixed vertex of G. Then $\deg(v)$ has binomial distribution $B(n-1,p)$. From Chernoff's Theorems, we have

$$\begin{aligned} \Pr(|\deg(v) - (n-1)p| > \epsilon(n-1)p) &< 2\exp(-(n-1)\epsilon^2/(3pq)) \\ &\sim 2\exp(-n\epsilon^2/(3pq)). \end{aligned}$$

Hence we bound the probability that there is at least one vertex v such that $|\deg(v) - (n-1)p| > (n-1)p)$ by $(2+o(1))n\exp(-n\epsilon/(3pq))$, which tends to zero as $n \to \infty$. □

The condition that fixed p can be weakened as $p = (\log n/n)\omega(n)$ with $\omega(n) \to \infty$, see Alon and Spencer (1992).

Let us enjoy an application of Chernoff bound that is of Erdős style, which disproved a conjecture with almost all graphs.

A suspended path in graph G is a path $(x_0, x_1, \ldots, x_k)$ in which $x_1, \ldots, x_{k-1}$ have degree two in G. A graph H is a *subdivision* of G if H is obtained from G replacing edges of G by suspended paths, that is to say, H is obtained by adding vertices on the edges of G.

A often used measure for sparseness of graphs is K_r-freeness as we have met previously. However, there are K_3-free graphs whose chromatic number can be arbitrarily large, see Mycielski's construction (1955) in the exercises. A more general measure for sparseness is to forbid subdivision. Hajós conjectured that every graph G with $\chi(G) \geq r$ contains a subdivision of K_r as a subgraph. This conjecture is trivial for $r = 2, 3$, and it is confirmed by Dirac (1952) for $r = 4$, and it is open for $r = 5, 6$. Catlin (1979) disproved the conjecture for $r \geq 7$ by

a constructive proof, but the disproof of Erdős and Fajtlowicz (1981) was more powerful. Let $\gamma(G)$ denote the largest r such that G contains a subdivision of K_r as a subgraph. Hajós conjecture is equivalent to that $\gamma(G) \geq \chi(G)$.

Theorem 4.9 *Almost all graphs $G_p \in \mathcal{G}(n, 1/2)$ satisfy*

$$\chi(G) \geq \frac{n}{2\log_2 n}, \quad \textit{and} \quad \gamma(G) \leq \sqrt{6n}.$$

Proof. Set $k = \lfloor 2\log_2 n \rfloor$. Since

$$\Pr(\alpha(G) \geq k) \leq \binom{n}{k} 2^{-\binom{k}{2}} < \left(\frac{e\sqrt{2}n}{k2^{k/2}}\right)^k \to 0,$$

and the fact

$$\alpha(G)\chi(G) \geq n$$

for any graph G, the first statement follows immediately. Set $r = \lceil\sqrt{6n}\rceil$. Then $n \leq r^2/6$. There are

$$\binom{n}{r} \leq \left(\frac{en}{r}\right)^r \leq \left(\frac{er}{6}\right)^r$$

potential K_r subdivisions, one for each $r-$element subset of $V(G)$. If we fix such a subset X, then we notice that since each subdivided edge has to use a distinct vertex of $V(G) \setminus X$. There are $\binom{r}{2}$ suspended pathes in a subdivision, and at most $n - r$ of them are of length two or more, which are "really" subdivided edges. So the number of edges in subgraph induced by X is at least

$$\binom{r}{2} - (n - r) \geq \binom{r}{2} + r - \frac{r^2}{6} \geq \frac{2}{3}\binom{r}{2}$$

edges. But the number of edges in subgraph induced by X, denoted by $e(X)$, has binomial distribution $B(N, 1/2)$, where $N = \binom{r}{2}$. From Chernoff bound in the last section,

$$\Pr(e(X) \geq N(1+\delta)/2) \leq \exp\{-N\delta^2/2\},$$

by taking $\delta = 1/3$ hence $\frac{2}{3}\binom{r}{2} = \binom{r}{2}(1+\delta)/2$, we have

$$\Pr\left(e(X) \geq \frac{2}{3}\binom{r}{2}\right) \leq \exp\{-N\delta^2/2\} = \exp\left\{-\frac{1}{18}\binom{r}{2}\right\}.$$

Thus we bound the probability that our random graph G contains a subdivision of K_r as follows.

$$\begin{aligned}\Pr(\gamma(G) \geq r) &\leq \sum_X \Pr\left(e(X) \geq \frac{2}{3}\binom{r}{2}\right) \\ &\leq \binom{n}{r}\exp\left\{-\frac{1}{18}\binom{r}{2}\right\} \\ &= \left(\frac{er\exp\{-(r-1)/36\}}{6}\right)^r,\end{aligned}$$

which tends to zero as $n \to \infty$. □

Since for almost all G in $\mathcal{G}(n, 1/2)$,

$$\chi(G) - \gamma(G) \geq \frac{n}{2\log_2 n} - \sqrt{6n} \to \infty$$

as $n \to \infty$, so Hajós conjecture failed badly, and almost all graphs in $\mathcal{G}(n, 1/2)$ are counterexamples. Further more, the gap between the truth and the conjecture is big.

The following is an application of Chernoff's bounds for Ramsey number $r(K_{m,n}, K_n)$.

Theorem 4.10 *Let integer $m \geq 2$ be fixed. Then there exists a constant $c = c(m) > 0$ such that*

$$r(K_{m,n}, K_n) \geq c\frac{n^{m+1}}{(\log n)^m}.$$

Proof. The lower bound is obtained through a simple application of Chernoff bound (Theorem 4.7). Let

$$N = \left\lfloor \frac{n^{m+1}}{3(2m\log n)^m} \right\rfloor,$$

and let $G(N,p)$ be a random graph of order N and edge probability $p = (2m\log n)/n$. The probability that m chosen vertices in $G(N,p)$ are connecting with another fixed vertex is p^m. So the probability that they have at least n common neighbors is $\Pr(S \geq n)$, where S has the binomial distribution $B(N-m, p^m)$. Then $n > Np^m$ and Theorem 4.7 yields

$$\begin{aligned}\Pr(K_{m,n} \subseteq G(N,p)) &\leq \binom{N}{m}\left(\frac{(N-m)p^m e}{n}\right)^n \\ &< \frac{N^m}{m!}\left(\frac{Np^m e}{n}\right)^n < c_1\frac{n^{m(m+1)}}{(\log n)^{m^2}}\left(\frac{e}{3}\right)^n,\end{aligned}$$

where $c_1 = c_1(m) > 0$ is a constant. Hence $\Pr(K_{m,n} \subseteq G(N,p)) \to 0$. At the time, by standard estimates that $\binom{N}{n} \leq (Ne/n)^n$ and $1 - p < e^{-p}$, we obtain a bound of the probability that $G(N,p)$ has an independent set of size at least n as follows

$$\begin{aligned}\Pr(\alpha(G(N,p)) \geq n) &\leq \binom{N}{n}(1-p)^{n(n-1)/2} \\ &\leq \left(\frac{Ne}{n}e^{-p(n-1)/2}\right)^n \leq \left(\frac{c_2}{3(2m\log n)^m}\right)^n,\end{aligned}$$

where $c_2 = c_2(m) > 0$ is a constant, so $\Pr(\alpha(G(N,p)) \geq n) \to 0$. Hence the probability that $G(N,p)$ contains neither $K_{m,n}$ as a subgraph nor an independent set of size n is positive (in fact, close to 1). Thus $r(K_{m,n}, K_n) > N$. □.

Using an upper bound for Turán number of $K_{m,n}$ and the main result in Chapter 2, we can show that the lower bound in above theorem is the right order of $r(K_{m,n}, K_n)$.

4.3 Martingales on Random Graphs ⋆

Most parameters of a random graph are concentrated around their expectations. To describe such phenomena, martingale is a powerful tool, which may liberate us from drudgery computations.

Let X and Y be random variables on a probability space Ω. Given $Y = y$ with $\Pr(Y = y) > 0$, we define a conditional expectation $E(X|Y = y)$ as

$$E(X|Y = y) = \sum_x x \Pr(X = x|Y = y),$$

which is a number depending on y. As Y is random, we have a new random variable $E(X|Y)$. For an element $s \in \Omega$, if $Y(s) = y$, then $E(X|Y)$ takes value $E(X|Y = y)$ at s.

Lemma 4.2 $E[E(X|Y)] = E[X]$.

Proof. From the definition, we have

$$\begin{aligned} E[E(X|Y)] &= \sum_y E[X|Y = y] \Pr(Y = y) \\ &= \sum_y \left(\sum_x x \Pr[X = x|Y = y] \right) \Pr(Y = y) \\ &= \sum_x x \left(\sum_y \Pr[X = x|Y = y] \Pr(Y = y) \right) \\ &= \sum_x x \Pr(X = x) = E(X) \end{aligned}$$

as asserted. □

A *martingale* is a sequence $X_0, X_1, \cdots, X_m$ of random variables so that for $0 \leq i < m$,

$$E(X_{i+1}|X_i) = X_i;$$

namely, $E(X_{i+1}|X_i = x) = x$ for any given $X_i = x$.

Imagine one walks on a line randomly, at each step he moves one unit to the left or right with probability p, or stands still with probability $1 - 2p$. Let X_i be the position of i step. This is a martingale as the expected position after $i + 1$ steps equals the actual position after i steps.

Let us look some martingales used in graph theory. The first is called the *edge exposure martingale* on chromatic numbers, in which we reveal G_p *one edge-slot at a time*. Let the random graph space

$\mathcal{G}(n,p)$ be the underlying probability space. Set $m = \binom{n}{2}$, and label the potential edges on vertex set $[n]$ by $e_1, e_2, \cdots, e_m$ in any manner. We define $X_0(H), X_1(H), \cdots, X_m(H)$ for a given graph H on vertex set $[n]$, which are random variables if H is a random graph in $\mathcal{G}(n,p)$. Let $X_0(H) = E(\chi(G_p))$. For general i,

$$X_i(H) = E[\chi(G_p) | e_j \in E(G_p) \text{ iff } e_j \in E(H), 1 \le j \le i].$$

In other words, $X_i(H)$ is the expected value of $E[\chi(G_p)]$ under the condition that the set of the first i edges of G_p equals that of H while the remaining edges are not seen and considered to be random. Note that X_0 is a constant $E(\chi(G_p))$ and $X_m = \chi(H)$.

Figure 1 shows why this is a martingale on the random space $\mathcal{G}(3, 0.5)$. Of cause, we can consider some other graph parameters.

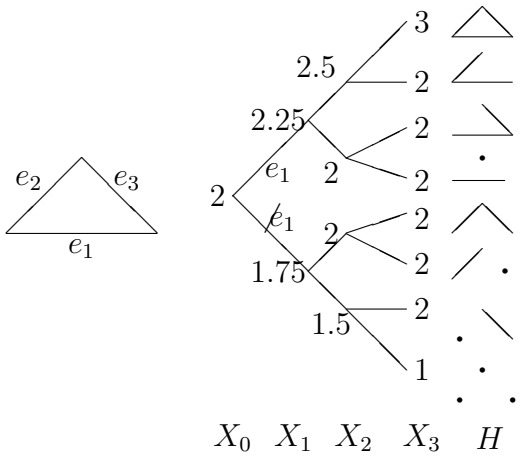

Fig. 5.1 An edge exposure martingale

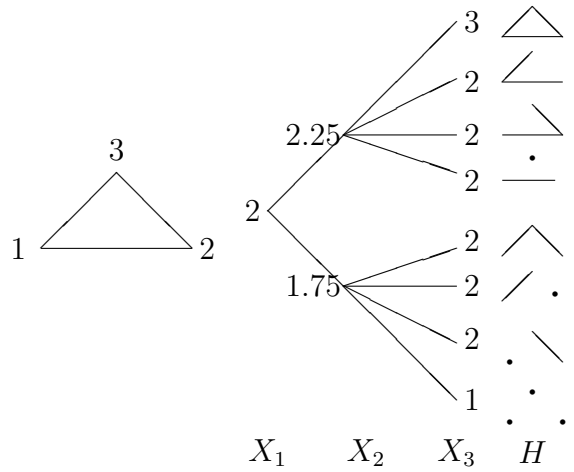

Fig. 5.2 A vertex exposure martingale

The second is called the *vertex exposure martingale* on chromatic numbers, in which we reveal G_p *one vertex-slot at a time*. Let the random graph space $\mathcal{G}(n,p)$ be the underlying probability space. We define $X_1 = E(\chi(G_p))$ and

$$X_i(H) = E[\chi(G_p)|E_i(G_p) = E_i(H)],$$

where $E_i(H)$ is the edge set induced by vertex set $\{1, \cdots, i\}$. In other words, $X_i(H)$ is the expected value of $E[\chi(G_p)]$ under the condition that the set of the edges of G_p induced by the first i vertices equals that of H while the remaining edges are not seen and considered to be random. Note that X_1 is a constant $E(\chi(G_p))$ and $X_n = \chi(H)$. Note that the vertex exposure martingale is a subsequence of the edge exposure martingale.

In Fig. 5.1, The probability space is $\mathcal{G}(3, 0.5)$, so $X_0 = E(\chi(G_p)) = 2$, and $X_1(H) = 2.75$ if $e_1 \in E(H)$, and $X_1(H) = 1.75$ otherwise. Thus $E(X_1|X_0) = 2 = X_0$. The random variables X_2 and X_3 take 4 values and 8 values, respectively, and $E(X_{i+1}|X_i) = X_i$.

Lemma 4.3 *Let Y be a (discrete) random variable such that $E(Y) = 0$ and $|Y| \leq 1$. Then $E(e^{tY}) \leq (e^t + e^{-t})/2$ for all $t \geq 0$.*

Proof. For fixed $t \geq 0$, set

$$h(y) = \frac{e^t + e^{-t}}{2} + \frac{e^t - e^{-t}}{2} y, \quad -1 \leq y \leq 1.$$

Note that the function $f(y) = e^{ty}$ is convex, and $h(y)$ is a line through the point $(-1, f(-1))$ and $(1, f(1))$ as $f(-1) = h(-1)$ and $f(1) = h(1)$, hence $e^{ty} \leq h(y)$, and

$$E(e^{tY}) \leq E(h(Y)) = \frac{e^t + e^{-t}}{2}$$

as $E(Y) = 0$, and thus the assertion follows. □

Theorem 4.11 (Azuma's Inequality) *Let $X_0, X_1, \cdots, X_m$ be a martingale with*

$$|X_{i+1} - X_i| \leq 1$$

for all $0 \leq i < m$, *and let* $\lambda > 0$. *Then*

$$\Pr[X_m - X_0 \geq \lambda\sqrt{m}] < e^{-\lambda^2/2},$$

and

$$\Pr[X_m - X_0 \leq -\lambda\sqrt{m}] < e^{-\lambda^2/2}.$$

Proof. We may assume that $X_0 = 0$ by translation. Set $Y_i = X_i - X_{i-1}$, then $|Y_i| \leq 1$ and $E(Y_i|X_{i-1}) = 0$. Then Lemma 4.3 yields that

$$E(e^{tY_i}|X_{i-1}) \leq \frac{e^t + e^{-t}}{2} \leq e^{t^2/2}$$

for any $t > 0$, where the last inequality has been proved in the first section of this chapter. Hence by Lemma 4.2, we have

$$\begin{aligned}
E(e^{tX_m}) &= E\left[e^{tX_{m-1}}e^{tY_m}\right] \\
&= E\left[E\left(e^{tX_{m-1}}e^{tY_m}|X_{m-1}\right)\right] \\
&= \sum_x E\left(e^{tX_{m-1}}e^{tY_m}|X_{m-1} = x\right)\Pr(X_{m-1} = x) \\
&= \sum_x e^{tx}E\left(e^{tY_m}|X_{m-1} = x\right)\Pr(X_{m-1} = x) \\
&\leq e^{t^2/2}\sum_x e^{tx}\Pr(X_{m-1} = x) \\
&= e^{t^2/2}E(e^{tX_{m-1}}).
\end{aligned}$$

This and the induction gave $E(e^{tX_m}) \leq e^{mt^2/2}$. Using Markov's Inequality, we obtain

$$\begin{aligned}
\Pr(X_m \geq \lambda\sqrt{m}) &= \Pr(e^{tX_m} \geq e^{t\lambda\sqrt{m}}) \\
&\leq \frac{E(e^{tX_m})}{e^{t\lambda\sqrt{m}}} \leq \frac{e^{mt^2/2}}{e^{t\lambda\sqrt{m}}}.
\end{aligned}$$

The assertion follows by letting $t = \lambda/\sqrt{m}$. □

4.4 Parameters of Random Graphs

We are ready to discuss some parameters of random graph G_p for fixed p. It is easy to see some parameters are concentrated around their expectations. The following result was due Shamir and Spencer (1987).

Theorem 4.12 *Let n and p be arbitrary and let $G_p \in \mathcal{G}(n,p)$. Then*

$$\Pr\left(|\chi(G_p) - E(\chi(G_p))| > \lambda\sqrt{n-1}\right) < 2e^{-\lambda^2/2}.$$

Proof. Consider the vertex exposure martingale $X_1, \cdots, X_n$ on $\mathcal{G}(n,p)$ with the parameter $\chi(G)$. A single vertex can always be given a new color so Azuma's Inequality can apply. □

Similarly, we have

$$\Pr\left(|\omega(G_p) - E(\omega(G_p))| > \lambda\sqrt{n-1}\right) < 2e^{-\lambda^2/2},$$

and

$$\Pr\left(|e(G_p) - E(e(G_p))| > \lambda\sqrt{m}\right) < 2e^{-\lambda^2/2},$$

where $m = \binom{n}{2}$. However, the proofs give no clue that what are these expectations.

Lemma 4.4 *Let $0 < p < 1$, $a = 1/p$ and $\epsilon > 0$ be fixed, and let $f(x) = \binom{n}{x}p^{\binom{x}{2}}$ for $0 \le x \le n$. Define an integer k such that*

$$f(k-1) > 1 \ge f(k).$$

Then as $n \to \infty$,

$$\lceil \omega_n - \epsilon \rceil \le k \le \lfloor \omega_n + \epsilon \rfloor + 1,$$

where

$$\omega_n = 2\log_a n - 2\log_a \log_a n + 2\log_a(e/2) + 1,$$

and $f(k-4) > c\left(\frac{n}{\log_a n}\right)^3 = n^{3-o(1)}$, where $c > 0$ is a constant.

Proof. It is easy to know that $k \to \infty$ and $k = o(\sqrt{n})$, thus by Stirling's formula, we have

$$f(k) = \binom{n}{k} p^{\binom{k}{2}} \sim \frac{n^k}{k!} p^{k(k-1)/2} \sim \frac{1}{\sqrt{2\pi k}} \left(\frac{en}{k} p^{(k-1)/2} \right)^k.$$

So if $\delta > 0$ fixed, for all large n,

$$\frac{en}{k} p^{(k-1)/2} \leq 1 + \delta$$

as $f(k) \leq 1$. This is equivalent to that

$$k \geq 2\log_a n - 2\log_a k + 2\log_a e + 1 - 2\log_a(1+\delta).$$

Let us set $k \sim 2\log_a n$ first. Then the difference between the right hand side in the above inequality and ω_n is

$$2\log_a \frac{2\log_a n}{k} - 2\log_a(1+\delta) \to -2\log_a(1+\delta),$$

so $k - \omega_n \geq -2\log_a(1+\delta) + o(1) \geq -\epsilon$ if we take δ small enough. Hence $k \geq \omega_n - \epsilon$.

Similarly, from

$$f(k-1) \sim \frac{1}{\sqrt{2\pi(k-1)}} \left(\frac{en}{k-1} p^{(k-2)/2} \right)^{k-1},$$

we have $\frac{en}{k-1} p^{(k-2)/2} \geq 1$, which gives

$$k \leq 2\log_a n - 2\log_a(k-1) + 2\log_a e + 2.$$

Furthermore, by taking $k \sim 2\log_a n$ first, we obtain $k \leq \omega_n + 1 + o(1) \leq \omega_n + \epsilon + 1$, the desired upper bound for k follows.

Finally, note that

$$f(k-2) > \frac{f(k-2)}{f(k-1)} = \frac{k-1}{n-k+2} a^{k-2} \sim p^2 \frac{k}{n} a^k > \frac{c\,n}{\log n},$$

the assertion for $f(k-4)$ follows immediately. □

Lemma 4.5 *For fixed $0 < p < 1$, $a = 1/p$ and $\epsilon > 0$, almost all graphs $G_p \in \mathcal{G}(n,p)$ satisfy*

$$\omega(G_p) < \lfloor \omega_n + \epsilon \rfloor < 2\log_a n,$$

where ω_n is defined in Lemma 4.4.

Proof. Let X_r be the number of r-cliques, where r is referred as an integer. Then

$$E(X_r) = f(r) = \binom{n}{r} p^{\binom{r}{2}} \leq \frac{n^r}{r!} p^{r(r-1)/2} < \frac{1}{\sqrt{2\pi r}} \left(\frac{en}{r} p^{(r-1)/2} \right)^r.$$

We shall find some $r = r(n) \to \infty$ such that $E(X_r) \to 0$. This is certainly true if $enp^{(r-1)/2}/r \leq 1$ (hence $r \to \infty$). The same argument in the proof of Lemma 4.4 applies that if $r = \lceil \omega_n + \epsilon \rceil$, then $E(X_r) \to 0$, thus $\Pr[\omega(G_p) \geq r] \to 0$ and $\Pr[\omega(G_p) \leq \lfloor \omega_n + \epsilon \rfloor] \to 1$. □

Remark. The above result can be stated as

$$\Pr\left(\omega(G_p) \leq \lceil \omega_n + \epsilon \rceil - 1\right) \to 1.$$

Matula (1970, 1972, 1976) was the first to notice that for fixed values of p almost all $G_p \in \mathcal{G}(n,p)$ have clique numbers concentrated on (at most) two values,

$$\lfloor \omega_n - \epsilon \rfloor \leq \omega(G_p) \leq \lfloor \omega_n + \epsilon \rfloor.$$

Results asserting this phenomenon were proved by Grimmett and McDiarmid (1975); and these were further strengthened by Bollobás and Erdős (1976).

In order to reduce the difficulty of the proof and preserve the typical flavor, we slightly weaken the above lower bound $\lfloor \omega_n - \epsilon \rfloor$ by having its asymptotical form a little bit later. Let us discuss the chromatic numbers first. A technical lemma is as follows.

Lemma 4.6 *Let k be the integer defined in Lemma 4.4 and let $\ell = k - 4$. Let $Y = Y(G)$ be the maximum size of a family of edge-disjoint cliques of size ℓ in $G \in \mathcal{G}(n,p)$. Then*

$$E(Y) \geq \frac{c\,n^2}{\ell^4},$$

where $c > 0$ is a constant.

Proof. Let $\mathcal{L}$ denote the family of ℓ-cliques of G. Then by Lemma 4.4, we have

$$\mu = E(|\mathcal{L}|) = f(\ell) = \binom{n}{\ell} p^{\binom{\ell}{2}} \geq c_1 \left(\frac{n}{\ell}\right)^3.$$

Let W denote the number of unordered pairs $\{A, B\}$ of ℓ-cliques of G with $A \sim B$, where $A \sim B$ signifies that $2 \leq |A \cap B| < \ell$. Let

$$\Delta = \sum_{A \sim B} \Pr(AB),$$

where the sum is taken over all ordered pairs $\{A, B\}$. Then $E(W) = \Delta/2$ and

$$\begin{aligned} \Delta &= \binom{n}{\ell} \sum_{i=2}^{\ell-1} \binom{\ell}{i} \binom{n-\ell}{\ell-i} p^{2\binom{\ell}{2} - \binom{i}{2}} \\ &= \mu \sum_{i=2}^{\ell-1} \binom{\ell}{i} \binom{n-\ell}{\ell-i} p^{\binom{\ell}{2} - \binom{i}{2}} = \mu \sum_{i=2}^{\ell-1} R_i. \end{aligned}$$

Setting $a = 1/p$, we have

$$\frac{R_{i+1}}{R_i} = \frac{(\ell-i)^2}{(i+1)(n-2\ell+i+1)} a^i.$$

If i is small, say bounded, then this ratio is $O((\log_a n)^2/n)$, and if i is large, say $\ell - i = O(1)$, then the ratio is at least $\sqrt{n}$. It is increasing on i, so

$$\Delta = \mu \sum_{i=2}^{\ell-1} R_i \leq 2\mu(R_2 + R_{r-1}).$$

Here

$$\begin{aligned} R_2 &= \binom{\ell}{2} \binom{n-\ell}{\ell-2} p^{\binom{\ell}{2}-1} \\ &= \frac{\ell^2(\ell-1)^2}{2p(n-\ell+2)(n-\ell+1)} \mu \leq \frac{\ell^4}{2pn^2} \mu, \end{aligned}$$

and

$$R_{\ell-1} = \ell(n-\ell) p^{\binom{\ell}{2} - \binom{\ell-1}{2}} \leq n\ell p^{\ell-1},$$

thus

$$\Delta = 2\mu\left(\frac{\ell^4}{2pn^2}\mu + n\ell p^{\ell-1}\right) \le C\,\frac{\mu^2\ell^4}{n^2}.$$

Let $\mathcal{C}$ be a random subfamily of $\mathcal{L}$ defined by setting for each $A \in \mathcal{L}$,

$$\Pr[A \in \mathcal{C}] = p_1,$$

where $0 < p_1 < 1$ will be determined. Then $E(|\mathcal{C}|) = \mu p_1$. Let W' be the number of unordered pairs $\{A, B\}$ of ℓ-cliques in $\mathcal{C}$ with $A \sim B$. Then

$$E(W') = E(W)p_1^2 = \frac{\Delta p_1^2}{2}.$$

Delete from $\mathcal{C}$ one set from each such pair $\{A, B\}$. This yields a set $\mathcal{C}^*$ of edge-disjoint ℓ-cliques of G and

$$E(Y) \ge E(|\mathcal{C}^*|) \ge E(|\mathcal{C}|) - E(W') = \mu p_1 - \frac{\Delta p_1^2}{2}.$$

By choosing $p_1 = \frac{\mu}{\Delta} < 1$, we have

$$E(Y) \ge \frac{\mu^2}{2\Delta} \ge \frac{c\,n^2}{\ell^4}$$

as asserted. □

Theorem 4.13 (Bollobás) *Let $0 < p < 1$, $a = 1/p$ be fixed, and let $m = \lceil n/\log_a^2 n\rceil$. Then for almost all graphs $G_p \in \mathcal{G}(n,p)$, each induced subgraph of order m of G_p has a clique of size at least $r = 2\log_a n - 7\log_a\log_a n$.*

Proof. Let S be an m-set of vertices. We shall bound the probability that S induces no r-clique by $e^{-m^{1+\delta}}$ for all large n (hence all large m), where $\delta > 0$ is a constant. So the probability that there exists an m-set with no r-clique is at most

$$\binom{n}{m}e^{-m^{1+\delta}} < \left(\frac{en}{m}\right)^m e^{-m^{1+\delta}} = \exp\left(m\log_e\frac{en}{m} - m^{1+\delta}\right),$$

which goes to zero, and the assertion follows.

Let X be the maximum number of pairwise edge-disjoint r-cliques sets in this graph (induced by S), where *edge-disjoint* means they share at most one vertex. We shall show that $X \geq 1$ holds almost surely. To do this, we invoke Azuma's Inequality. Consider the edge exposure martingale for X that results from revealing G one-edge slot at a time. We have $X_0 = E(X)$ and $X_{\binom{m}{2}} = X$. Clearly the Lipschitz condition $|X_{i+1} - X_i| \leq 1$ is satisfied, so Azuma's Lemma gives

$$\begin{aligned}
\Pr(X = 0) &\leq \Pr[X - E(X) \leq -E(X)] \\
&= \Pr\left[X - E(X) \leq -\lambda \binom{m}{2}^{1/2}\right] \leq e^{-\lambda^2/2} \\
&= \exp\left(-\frac{E^2(X)}{m(m-1)}\right),
\end{aligned}$$

where $\lambda = E(X)/\binom{m}{2}^{1/2}$. Hence it suffices to find $\delta > 0$ such that $E^2(X) \geq m^{3+\delta}$ for all large n.

Now, let t_0 be the integer such that $f(t_0 - 1) > 1 \geq f(t_0)$, where $f(x) = \binom{m}{x} p^{\binom{x}{2}}$, and let $t = t_0 - 4$. Then by Lemma 4.4, we have

$$t \geq 2\log_a m - 2\log_a \log_a m - 3 > 2\log_a n - 7\log_a \log_a n,$$

so $t > r$. Let T be the maximum number of edge-disjoint cliques of size t, Then $E(X) \geq E(T)$ and $E(T) \geq cm^2/t^4$ by Lemma 4.6, hence

$$E(X) \geq \frac{cm^2}{t^4} \sim \frac{cn^2}{16(\log_a n)^8},$$

implying that $E^2(X) \geq n^{4-o(1)} \geq n^{3+\delta}$ for any $1 > \delta > 0$ if n is large, which completes the proof. □

Theorem 4.14 (Bollobás, 1988) *Let $0 < p < 1$, and $b = 1/q = 1/(1-p)$. Then for any fixed $\epsilon > 0$, almost all graphs $G_p \in \mathcal{G}(n,p)$ satisfy*

$$\frac{n}{2\log_b n} \leq \chi(G_p) \leq (1+\epsilon)\frac{n}{2\log_b n}.$$

Proof. The lower bound holds because almost all G_p satisfy $\alpha(G_p) \leq 2\log_b n$ and $\chi(G)\alpha(G) \geq n$. The upper bound follows from the above theorem, which is applied for independent sets instead of cliques, because we can almost always select independent set of size $2\log_b n - 7\log_b \log_b n$ until we have only $n/\log_b^2 n < (\epsilon/2)n/(2\log_b n)$ vertices left. We first use at most

$$\frac{n}{2\log_b n - 7\log_b \log_b n} < \left(1 + \frac{\epsilon}{2}\right) \frac{n}{2\log_b n}$$

colors, and then we can complete the coloring by using distinct new colors on each of the remaining vertices. □

Let us remark that Achlioptas and Naor recently obtained a result on sparser random graphs as follows. Given $d > 0$, let k_d be the smallest integer k such that $d < 2k\log k$. Then $\chi(G_p)$ for almost all $G_p \in \mathcal{G}(n, d/n)$ is either k_d or $k_d + 1$. This result improves an earlier result of Łuczak (1991) by specifying the form of k_d.

Theorem 4.15 *Let $0 < p < 1$ and $\epsilon > 0$ be fixed. Then almost all graphs $G_p \in \mathcal{G}(n,p)$ satisfy*

$$(1-\epsilon)2\log_b n \leq \alpha(G_p) < 2\log_b n.$$

Proof. The upper bound is the complement of that in Lemma 4.5. The lower bound follows from Theorem 4.14 and the fact that $\alpha(G) \geq n/\chi(G)$. □

Theorem 4.16 *Let $0 < p < 1$ and $\epsilon > 0$ be fixed. Then almost all graphs $G_p \in \mathcal{G}(n,p)$ satisfy*

$$(1-\epsilon)2\log_a n \leq \omega(G_p) < 2\log_a n.$$

Proof. This is complement of Theorem 4.15. □

For some graph parameter $f(G)$, we have seen that there is a function $g(n)$ such that almost all graphs G_p in $\mathcal{G}(n,p)$ satisfy that

$$(1-\epsilon)g(n) \leq f(G_p) \leq (1+\epsilon)g(n),$$

hence $f(G)$ concentrate in a small range. We shall call the function $g(n)$ as a *threshold* for the parameter f. We will discuss the threshold for probability $p = p(n)$ instead of fixed p, and will consider some other graph parameters in the next chapter.

Chapter 5

Quasi-random graphs

Random graphs have been proven to be one of most important tools in modern graph theory. Their tremendous triumph raises the following general question: what are the essential properties and how can we tell when a given graph behaves like a random graph G_p in $\mathcal{G}(n,p)$? Here a typical property of random graphs is that almost all G_p satisfy. This leads us to a concept of *quasi-random graphs*. It was Thomason (1987) who introduced the notation of jumbled graphs to measure the similarity between the edge distribution of quasi-random graphs and random graphs. Quasi-random graphs are also called pseudo-graphs. A cornerstone contribution of Chung, Graham and Wilson (1989) showed that many properties of different nature are equivalent to the notation of quasi-random graphs. For a survey on quasi-random graphs, see Krivelevich and Sudakov (2006). This chapter focuses on quasi-random graphs. In recent years, there are some quasi-random families of graphs appearing, which are all constructed by finite fields. Their algebraic parameters are easier to compute, some of which are related to characters of finite fields and thus the third section is devoted to the topics. The last section is application for quasi-random graphs in Ramsey theory.

5.1 Properties of dense graphs

Speaking formally, a quasi-random G of order n is a graph that behaves like a random graph $G(n,p)$ with $p = e(G)/\binom{n}{2}$. For $0 < p < 1 \le \alpha$, a graph G is called (p,α)-*jumbled* if each induced subgraph H on h vertices of G satisfies that

$$\left|e(H) - p\binom{h}{2}\right| \le \alpha h.$$

Equivalently, G is (p,α)-jumbled if the average degree $d(H)$ of each induced subgraph H of G satisfies that

$$|d(H) - p(h-1)| \le 2\alpha.$$

The following result of Thomason (1987) contains a simple local condition of a graph of being jumbled.

Theorem 5.1 *Let G be a graph of order n with $\delta(G) \ge pn$. If any pair of vertices has at most $p^2n + \ell$ common neighbors, where $\ell > 0$, then G is $(p, \sqrt{(p+\ell)n}/2)$-jumbled.*

Proof. Let H be an induced subgraph of G of order h with $d(H) = d$, where $h < n$. Write $V(G) = \{v_1, v_2, \ldots, v_n\}$ and $V(H) = \{v_1, v_2, \ldots, v_h\}$, say. Let d_i be the number of neighbors of v_i in H for $1 \le i \le n$. Then $\sum_{i=1}^{h} d_i = hd$ and

$$\sum_{j=h+1}^{n} d_j \ge \sum_{i=1}^{h}(pn - d_i) = h(pn - d).$$

Since any pair of vertices are covered by at most $p^2n + \ell$ vertices, and at most that in H particularly, we have

$$\sum_{i=1}^{n}\binom{d_i}{2} \le \binom{h}{2}(p^2n + \ell).$$

The above and the convexity of the function $\binom{x}{2}$ imply that

$$h\binom{d}{2} + (n-h)\binom{h(pn-d)/(n-h)}{2} \le \binom{h}{2}(p^2n + \ell).$$

Equivalently,

$$(d-ph)^2 \leq \frac{n-h}{n}\Big[(h-1)\ell + p(1-p)n\Big],$$

which gives that

$$|d-p(h-1)| \leq \sqrt{(p+\ell)n}$$

as claimed. Finally, note that the same inequality holds for $h=n$. □

For given graphs G and H, let $N^*_G(H)$ be the number of labeled occurrences of H as an induced subgraph of G, which is the number of adjacency-preserving injections from $V(H)$ to $V(G)$ whose image is the set of vertices of an induced copy of H of G. Namely, these injections are both adjacency-preserving and non-adjacency-preserving. Let $N_G(H)$ be the number of labeled copies of H as a (not necessarily induced) subgraph of G. Then

$$N_G(H) = \sum_{H'} N^*_G(H'),$$

where H' ranges over all graphs on $V(H)$ obtained from H by adding a set of edges. For example, if $G = H = C_t$, then $N^*_G(H) = N_G(H) = 2t$, and if $G = K_n$ and $n \geq t \geq 4$, then $N^*_G(C_t) = 0$ and $N_G(C_t) = N^*_G(K_t) = (n)_t$. If $G = K_{n/2,n/2}$ and n is even, then $N_G(C_4) = 2\Big(\frac{n}{2}(\frac{n}{2}-1)\Big)^2 \sim 2\cdot\Big(\frac{n}{2}\Big)^4$ for large n.

Let G be a (p,α)-jumbled graph of order n, where $\alpha = \alpha_n = o(n)$ as $n \to \infty$. Then, as shown by Thomason, for fixed p and fixed graph H of order h

$$N^*_G(H) \sim p^{e(H)}(1-p)^{\binom{h}{2}-e(H)}n^h.$$

Let x and y be vertices of G. Denote by $s(x,y)$ the number of vertices of G adjacent to x and y the same way: either to both or none. Let λ_i be eigenvalues of G with $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. Let $\lambda = \lambda(G) = |\lambda_2|$. For two (not necessarily disjoint) subsets B and C, let $e(B,C)$ denote the number of edges from B to C, in which each edge in $B \cap C$ is counted twice. If $B \cap C = \emptyset$, then $e(B,C)$ is simply the number of edges between B and C.

The *quasi-random graph* defined by Chung, Graham and Wilson is in fact a family of simple graphs, which satisfy any (hence all) of

equivalent properties in the following theorem. It is remarkable that these properties ignore "small" local structures. The expressions of the properties are related to the edge density p, here $p = 1/2$.

Theorem 5.2 *Let $\{G\}$ be a sequence of graphs, where $G = G_n$ is a graph of order n. Then the following properties are equivalent:*

$P_1(h)$**:** *For any fixed $h \geq 4$ and graph H of order h, $N^*_G(H) \sim (\frac{1}{2})^{\binom{h}{2}} n^h$.*

$P_2(t)$**:** $e(G) \sim \frac{n^2}{4}$ *and* $N_G(C_t) \leq (\frac{n}{2})^t + o(n^t)$ *for any even* $t \geq 4$.

P_3**:** $e(G) \geq \frac{n^2}{4} + o(n^2)$, $\lambda_1 \sim \frac{n}{2}$ *and* $\lambda_2 = o(n)$.

P_4**:** *For each* $U \subseteq V(G)$, $e(U) = \frac{1}{2}\binom{|U|}{2} + o(n^2)$.

P_5**:** *For each* $U \subseteq V(G)$ *with* $|U| = \lfloor \frac{n}{2} \rfloor$, $e(U) \sim \frac{n^2}{16}$.

P_6**:** $\sum_{x,y} \left| s(x,y) - \frac{n}{2} \right| = o(n^3)$.

P_7**:** $\sum_{x,y} \left| |N(x) \cap N(y)| - \frac{n}{4} \right| = o(n^3)$.

Proof.* The steps of proof of Chung, Graham and Wilso are $P_1(h{+}1) \Rightarrow P_1(h)$ and

$$P_1(2h) \Rightarrow P_2(2t) \Rightarrow P_2(4) \Rightarrow P_3 \Rightarrow P_4 \Longleftrightarrow P_5 \Rightarrow P_6 \Rightarrow \ P_1(2h),$$

so that all but P_7 are proven to be equivalent. They then add P_7 to the equivalent chain by proving that

$$P_2(t) \Rightarrow P_7 \Rightarrow P_6.$$

Here, we omit some steps but keep most of them and preserve the typical flavor.

Fact 1. $P_1(h+1) \Rightarrow P_1(h)$, and $P_1(3)$ implies the property

$$P_0 : \ \sum_v \left| \deg(v) - \frac{n}{2} \right| = o(n^2).$$

Let us remark that P_0 is equivalent to that

$$P_0' : \ \text{All but } o(n) \text{ vertices of } G \text{ have degree } (1 + o(n))\tfrac{n}{2}$$

by Cauchy-Schwarz inequality, and P_0 implies that

$$e(G) \sim \frac{n^2}{4}.$$

Assume that $P_1(h+1)$ holds. Let H be a graph of order h. There are 2^h ways to extend it to a graph H' of order $h+1$, and each copy of H is contained in $n-h$ subgraphs H' of order $h+1$. By $P_1(h+1)$, we have

$$N_G^*(H') \sim n^{h+1}2^{-\binom{h+1}{2}},$$

thus

$$N_G^*(H) \sim n^{h+1}2^{-\binom{h+1}{2}}\frac{2^h}{n-h} \sim n^h 2^{-\binom{h}{2}},$$

which is $P_1(h)$. Suppose that $\{G\}$ satisfies $P_1(3)$. Let H_i be the graph of order 3 and i edges, $1 \le i \le 3$. By counting how often each edge can contribute to the various $N_G^*(H_i)$, we have

$$(n-2)\sum_v \deg(v) = N_G^*(H_1) + 2N_G^*(H_2) + N_G^*(H_3) \sim \frac{n^3}{2},$$

thus $\sum_v \deg(v) \sim \frac{n^2}{2}$ and $e(G) \sim \frac{n^2}{4}$. Also

$$\sum_v \deg(v)(\deg(v)-1) = N_G^*(H_2) + N_G^*(H_3) \sim \frac{n^3}{4},$$

implying that $\sum_v \deg^2(v) \sim \frac{n^3}{4}$. Then, by Cauchy-Schwarz,

$$\begin{aligned}\sum_v \Big|\deg(v) - \frac{n}{2}\Big| &\le \sqrt{n}\Big(\sum_v \Big|\deg(v) - \frac{n}{2}\Big|^2\Big)^{1/2} \\ &= \sqrt{n}\Big(\sum_v \deg^2(v) - n\sum_v \deg(v) + \frac{n^3}{4}\Big)^{1/2},\end{aligned}$$

which is $o(n^2)$.

Fact 2. $P_1(2t) \Rightarrow P_2(2t)$ $(t \ge 2)$. Fact 1 has proved that $e(G) \sim \frac{n^2}{4}$. We then show that

$$N_G(C_{2t}) = \sum_{H'} N_G^*(C_{2t}) \le (1+o(1))\Big(\frac{n}{2}\Big)^{2t}.$$

As H' ranges over all graphs on $V(H)$ obtained from H by adding to it a set of edges, the number of such sets is $2^{\binom{2t}{2}-2t}$. This and $P_1(2t)$ imply $P_2(2t)$.

Fact 3. $P_2(2t) \Rightarrow P_2(4) \Rightarrow P_3$. There is nothing to prove for the first implication and we prove the second. Let A be the adjacency matrix of G and d the average degree of G. We first claim that

$$\lambda_1 \geq d.$$

Let us verify that for any unit vector X, $\lambda_1 \geq X^t AX$. Let Λ be the diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$ and P a normal orthogonal matrix such that $PAP^t = \Lambda$. Then PX is a unit vector, and

$$\lambda_1 = \lambda_1 (PX)^t \cdot (PX) \geq (PX)^t \Lambda (PX) = X^t (P^t \Lambda P) X = X^t AX.$$

By taking $X = \frac{1}{\sqrt{n}} J$, where $J = (1, 1, \ldots, 1)^t$, we obtain that

$$\lambda_1 \geq \frac{1}{n} J^t AJ = \frac{1}{n} \sum_v \deg(v) = d$$

as claimed. This and $e(G) \sim \frac{n^2}{2}$ imply $\lambda_1 \geq \frac{n}{2} + o(n)$. Next, consider the trace of A^4. Clearly,

$$tr(A^4) = \sum_{i=1}^{n} \lambda_i^4 \geq \lambda_1^4 \geq (1 + o(1)) \frac{n^4}{16}.$$

On the other hand, as this trace is precisely the number of labeled and closed walks of length 4 in G, i.e., the number of sequences $v_0, v_1, v_2, v_3, v_4 = v_0$ such that $v_i v_{i+1}$ is an edge. This number is $N_G(C_4)$ plus the number of such sequences in which $v_2 = v_0$, and plus the number of such sequences in which $v_2 \neq v_0$. Thus

$$\sum_{i=1}^{n} \lambda_i^4 = N_G(C_4) + o(n^4) \sim \left(\frac{n}{2}\right)^4.$$

It follows that $tr(A^4) \sim \frac{n^4}{16}$, thus $\lambda_1 \sim \frac{n}{2}$ and $\sum_{i=2}^{n} \lambda_i^4 = o(n^4)$ hence $\lambda_2 = o(n)$ as desired.

Fact 4. $P_3 \Rightarrow P_4$. To simply the proof, we suppose that G is regular. Then the Fact 4 follows from Corollary 5.2 in the next section.

Fact 5. $P_4 \Longleftrightarrow P_5$. The implication $P_4 \Rightarrow P_5$ is immediate, so we show $P_5 \Rightarrow P_4$. By ignoring one vertex possibly, we assume that n is even so that $n/2$ is an integer. Suppose that for any subset S with $|S| = n/2$, $\left|e(S) - \frac{n^2}{16}\right| < \epsilon n^2$, where $\epsilon > 0$ is fixed. We shall show that for any subset T,

$$\left|e(T) - \frac{1}{2}\binom{t}{2}\right| < 20\epsilon n^2,$$

where $t = |T|$. Let us consider two cases, .

Case 1. $t = |T| \geq n/2$. By averaging over all $S \subseteq T$ with $|S| = n/2$, we have

$$e(T) = \frac{1}{\binom{t-2}{n/2-2}} \sum \left\{e(S) : \ S \subseteq T, |S| = n/2\right\}$$

as each edge is counted exactly $\binom{t-2}{n/2-2}$ times. Thus

$$e(T) \leq \frac{\binom{t}{n/2}}{\binom{t-2}{n/2-2}}\left(\frac{n^2}{16} + \epsilon n^2\right) \leq \binom{t}{2}\left(\frac{1}{2} + 9\epsilon\right).$$

Similarly,

$$e(T) \geq \binom{t}{2}\left(\frac{1}{2} - 9\epsilon\right).$$

Case 2. $t = |T| < n/2$. We shall show that the assumption

$$e(T) \geq \frac{1}{2}\binom{t}{2} + 20\epsilon n^2$$

leads to a contradiction. Set $\bar{T} = V \setminus T$. Then $|\bar{T}| = n - t > n/2$ and by Case 1, we have

$$\binom{n-t}{2}\left(\frac{1}{2} - 9\epsilon\right) < e(\bar{T}) < \binom{n-t}{2}\left(\frac{1}{2} + 9\epsilon\right).$$

Consider the average value A of $e(T \cup T')$, where T' ranges over all subsets of $\bar{T}$ with $|T'| = n/2 - t$ so that $|T \cup T'| = n/2$, so

$$A = \binom{n-t}{n/2-t}^{-1} \sum_{T'} \left\{e(T \cup T') : \ T' \subseteq \bar{T}, |T'| = n/2 - t\right\}.$$

Counting how much different edges contribute to the sum, we know that the sum equals to

$$e(T)\binom{n-t}{n/2-t} + e(\bar{T})\binom{n-t-2}{n/2-t-2} + e(T,\bar{T})\binom{n-t-1}{n/2-t-1}.$$

From the fact that $e(T,\bar{T}) = e(G) - e(T) - e(\bar{T})$, we obtain that

$$A = \frac{n/2}{n-t}e(T) - \frac{(n/2-t)n/2}{(n-t)(n-t-1)}e(\bar{T}) + \frac{n/2-t}{n-t}e(G),$$

which satisfies

$$\begin{aligned} A \geq\ & \frac{n/2}{n-t}\left\{\frac{1}{2}\binom{t}{2} + 20\epsilon n^2\right\} - \frac{(n/2-t)n/2}{(n-t)(n-t-1)}\binom{n-t}{2}\left(\frac{1}{2}+9\epsilon\right) \\ & + \frac{n/2-t}{n-t}\binom{n}{2}\left(\frac{1}{2}-9\epsilon\right) > \frac{n^2}{16} + \epsilon n^2. \end{aligned}$$

Similarly, the assumption

$$e(T) < \frac{1}{2}\binom{t}{2} - 20\epsilon n^2,$$

leads a contradiction to the property P_5, too. □

A property is called *a quasi-random property* for $p = 1/2$ if it is equivalent to any property in Theorem 5.2. It is surprised that $P_2(4)$, which seems to be weaker, is a quasi-random property for $p = 1/2$.

Theorem 5.3 *The property*

$$P_2(4):\ e(G) \sim \frac{n^2}{4} \quad \textit{and} \quad N_G(C_4) \leq \left(\frac{n}{2}\right)^4 + o(n^4)$$

is a quasi-random property for $p = 1/2$.

Proof. See Fact 3 in the proof of the last theorem. □

Some other properties can be added to the list, one of which is in the next theorem.

Theorem 5.4 *The property*

$$P_8: \text{ For all } U, V \subseteq V(G),\ e(U,V) = \frac{1}{2}|U||V| + o(n^2)$$

is a quasi-random property for $p = 1/2$.

Proof. Let us prove the result by $P_4 \Longleftrightarrow P_8$. It suffices to show that $P_4 \Rightarrow P_8$. Suppose that P_4 holds. If U and V are disjoint, then

$$\begin{aligned} e(U,V) &= e(U \cup V) - e(U) - e(V) \\ &= \frac{1}{4}(u+v)^2 - \frac{1}{4}u^2 - \frac{1}{4}v^2 + o(n^2) \\ &= \frac{1}{2}uv + o(n^2), \end{aligned}$$

where $u = |U|$ and $v = |V|$. In case U and V are not disjoint, write $|U \cap V| = x$, from P_4 and what we just proved, we know that $e(U,V)$ equals to

$$\begin{aligned} & e(U \setminus V, V \setminus U) + e(U \cap V, U \setminus V) + e(U \cap V, V \setminus U) + 2e(U \cap V) \\ &= \frac{1}{2}(u-x)(v-x) + \frac{1}{2}x(u-x) + \frac{1}{2}x(v-x) + o(n^2) \\ &= \frac{1}{2}uv + o(n^2), \end{aligned}$$

which is P_8. □

The following theorem is for general edge density p. However, $0 < p < 1$ is fixed.

Theorem 5.5 *Let* $\{G\}$ *be a sequence of graphs, where* $G = G_n$ *is a graph of order* n. *Let* $0 < p < 1$ *be fixed. Then the following properties are equivalent:*

$P_1(h)$**:** *For any fixed* $h \geq 4$ *and graph* H *of order* h,

$$N_G^*(H) \sim p^{e(H)}(1-p)^{\binom{h}{2}-e(H)}n^h.$$

$P_2(t)$**:** $e(G) \sim \frac{pn^2}{2}$ *and* $N_G(C_t) \leq (pn)^t + o(n^t)$ *for any even* $t \geq 4$.

P_3: $e(G) \geq \frac{pn^2}{2} + o(n^2)$, $\lambda_1 \sim pn$ *and* $\lambda_2 = o(\lambda_1)$.

P_4: *For each* $U \subseteq V(G)$, $e(U) = p\binom{|U|}{2} + o(n^2)$.

P_5: *For each* $U \subseteq V(G)$ *with* $|U| = \lfloor \frac{n}{2} \rfloor$, $e(U) \sim \frac{p}{8}n^2$.

P_6: $\sum_{x,y} \left| s(x,y) - (p^2 + (1-p)^2)n \right| = o(n^3)$.

P_7: $\sum_{x,y} \left| |N(x) \cap N(y)| - p^2 n \right| = o(n^3)$.

5.2 Paley Graphs

Let q be a prime power. An element $a \in F(q)$ is called to be *quadratic* if $a = b^2$ for some $b \in F(q)$. A quadratic element of $Z_p = F(p)$ is usually called a *quadratic residue* (mod p) to signify the modulo operations.

Let us define a function $\chi(x)$ on $F(q)$ as

$$\chi(x) = x^{(q-1)/2}.$$

This function is usually called the *quadratic residue character* of $F(q)$.

Lemma 5.1 *Let q be an odd prime power. Then $\chi(x) \in \{-1, 0, 1\}$. If $x \neq 0$, then $\chi(x) = 1$ if and only if x is quadratic, namely,*

$$\chi(x) = \begin{cases} 1 & x \text{ is quadratic, } x \neq 0, \\ 0 & x = 0, \\ -1 & x \text{ is not quadratic.} \end{cases}$$

Furthermore, half of elements of $F^*(q)$ are quadratic, and half of that are non-quadratic.

Proof. Let x be an element of $F^*(q)$. Then $\chi(x) = \pm 1$ as

$$(\chi(x) - 1)(\chi(x) + 1) = \chi^2(x) - 1 = x^{q-1} - 1 = 0.$$

Let ν be a primitive element of $F(q)$. Then

$$F^*(q) = \{\nu, \nu^2, \ldots, \nu^{q-2}, \nu^{q-1} = 1\}.$$

As ν is primitive, it not quadratic and $\chi(\nu) \neq 1$ hence the set of non-zero quadratic elements is

$$S_0 = \{\nu^2, \nu^4, \ldots, \nu^{q-1} = 1\},$$

and the set of non-quadratic elements is

$$S_1 = \{\nu, \nu^3, \ldots, \nu^{q-2}\}.$$

Using the facts that $\chi(\nu) = -1$ and $\chi(\nu^k) = \chi^k(\nu)$, we have $\chi(x) = 1$ if and only if $x \in S_0$, as claimed. □

Note that, asymptotically, there are half primes $p \leq n$ of the form of $p \equiv 1 \,(\mathrm{mod}\, 4)$ and half of the form of $p \equiv 3 \,(\mathrm{mod}\, 4)$.

The *Paley graph* P_q is defined as follows. Let $q \equiv 1 \,(\mathrm{mod}\, 4)$ be a prime power. The vertex set of P_q is $F(q)$, and distinct vertices x and y are adjacent if and only if

$$\chi(x - y) = (x - y)^{(q-1)/2} = 1.$$

So x and y are adjacent if and only if $x - y$ is non-zero quadratic. Note that $\chi(x - y) = \chi(y - x)$ as $\chi(-1) = 1$ as $q \equiv 1 \,(\mathrm{mod}\, 4)$.

Let A be an additive group and let S be an inverse-closed subset of A^*. A graph, called the *Cayley graph* with respect to S, is defined as follows: its vertex set is A and u and v are adjacent if $u - v \in S$. Clearly, the Paley graph is the Cayley graph with respective to the subset of non-zero quadratic elements. As an example, it is easy to verify that P_5 is C_5, which is the Ramsey graph for $r(3, 3)$.

A graph G of order n is said to be a *strongly regular graph* with parameters n, d, λ, μ, denoted by $srg(n, d, \lambda, \mu)$ in short, if it is d-regular, and any pair of vertices have λ common neighbors if they are adjacent, and μ common neighbors otherwise. For example, C_5 is an $srg(5, 2, 0, 1)$. Strongly regular graphs were introduced by Bose (1963). It is easy to see that the complement of an srg is also an srg.

Proposition 5.1 *Let G be an $srg(n, d, \lambda, \mu)$. Then its complement is also an $srg(n, d_1, \lambda_1, \mu_1)$, where*

$$\begin{aligned} d_1 &= n - d - 1, \\ \lambda_1 &= n - 2d + \mu - 2, \\ \mu_1 &= n - 2d + \lambda. \end{aligned}$$

Proof. The value of d_1 can be determined by $d+d_1 = n-1$. Let u and v be distinct vertices of G. If they are non-adjacent, then $|N(u) \cup N(v)| = 2d-\mu$. The remaining $n-2d+\mu-2$ vertices are the common neighbor of u and v in $\overline{G}$, giving λ_1 as claimed. If u and v are adjacent, then $\{u, v\} \subseteq N(u) \cup N(v)$ and $|N(u) \cup N(v)| = 2d - \lambda$. The remaining $n - 2d + \lambda$ vertices are common neighbors of u and v in $\overline{G}$, yielding μ_1 as claimed. □

For vertex disjoint graphs G and H, let $G \cup H$ be the graph on vertex set $V(G) \cup V(H)$ edge set $E(G) \cup E(H)$, which is called the union of G and H. Let mG be the union of m copies of G. the union mK_k is an $srg(mk, k-1, k-2, 0)$. On the other hand, if G is an $srg(n, k, \lambda, 0)$, then G is a union of complete graphs with the same order. We sometimes exclude complete and empty graphs as an srg to avoid to define μ and λ, respectively. A relation among the parameters is as follows.

Proposition 5.2 *Let G be an $srg(n, d, \lambda, \mu)$. Then*

$$d(d - \lambda - 1) = \mu(n - d - 1).$$

Proof. Let v be a vertex and let $M(v)$ be the set of non-neighbors of v. Consider the partition $V(G) = \{v\} \cup N(v) \cup M(v)$. By the definition, $N(v)$ contains d vertices, and $M(v)$ contains $n - d - 1$ vertices. Each vertex of $N(v)$ is adjacent to λ vertices in $N(v)$, and hence $d - \lambda - 1$ vertices in $M(v)$. Each vertex in $M(v)$ is adjacent to μ vertices in $N(v)$. Counting the edges between $N(v)$ and $M(v)$ in two ways, the required equality follows. □

A graph G is called vertex-transitive if for any two vertices a and b of G, there is an automorphism mapping a to b, it is called edge-transitive if for any two edges ab and uv of G, there is an automorphism mapping $\{a, b\}$ to $\{u, v\}$.

Theorem 5.6 *Let $q \equiv 1 \,(mod\ 4)$ be a prime power. Then the Paley graph P_q is an*

$$srg\Big(q, \frac{q-1}{2}, \frac{q-5}{4}, \frac{q-1}{4}\Big).$$

Furthermore, it is self-complementary, vertex-transitive and edge-transitive.

Proof. Lemma 5.1 implies that P_q is $(q-1)/2$-regular. Note that $\sum_x \chi(x) = 0$, and the number of common neighbors of two vertices a and b is

$$\begin{aligned} & \sum_{x \neq a,b} \frac{1+\chi(x-a)}{2} \cdot \frac{1+\chi(x-b)}{2} \\ = \quad & \frac{q-2}{4} - \frac{\chi(a-b)}{2} + \frac{1}{4} \sum_{x \neq a,b} \chi(x-a)\chi(x-b). \end{aligned}$$

Using the fact that for $x \neq b$, $\chi(x-b) = \chi((x-b)^{-1})$ and the multiplicity of χ, we can write the last term as

$$\frac{1}{4} \sum_{x \neq a,b} \chi\Big(\frac{x-a}{x-b}\Big) = \frac{1}{4} \sum_{x \neq 0,1} \chi(x) = \frac{-1}{4}.$$

Thus the number of common neighbors of a and b is $(q-3)/4 - \chi(a-b)/2$, which is $(q-5)/4$ if a and b are adjacent and $(q-1)/4$ otherwise.

Fix $a \in F^*(q)$ with $\chi(a) = -1$, define a map ϕ_0 as

$$\phi_0 : V(P_q) \to V(P_q),\ \phi_0(x) = ax.$$

Then the map ϕ_0 is an automorphism between P_q and $\overline{P_q}$. Hence P_q is self-complementary.

It is easy to verify that the map $\phi_1(x) = a+b-x$ is an automorphism mapping a to b, and the map $\phi_2(x) = \frac{u-v}{a-b}(x-b)+v$ is an automorphism mapping an edge ab to an edge uv. □.

Then P_q is $(q-1)/2$-regular, and the distinct eigenvalues of P_q are $(q-1)/2$, $(\sqrt{q}-1)/2$ and $-(\sqrt{q}-1)/2$. Therefore,

$$e(P_q) = \frac{q(q-1)}{4} \sim \frac{q^2}{4},\ \lambda_1 = \frac{q-1}{2} \sim \frac{q}{2},\ \lambda = \frac{\sqrt{q}-1}{2} = o(q).$$

Thus P_q satisfies quasi-random property P_3 hence all other quasi-random properties with $p = 1/2$.

5.3 Graph with small second eigenvalue

The last section was devoted to the quasi-random graphs of fixed edge density. Let us now switch to the case of density $p = p(n) = o(1)$, which is more important for some applications.

In applications, we shall allow the graphs to be semi-simple, that is, each vertex is attached at most a loop. When $p \to 0$, the situation is significantly more complicated as revealed by Chung and Graham (2002). The first remarked fact is that the properties defined for quasi-random graphs with fixed edge density may be not equivalent anymore. Let E_q^o be the Erdő-Rényi graph of order $n = q^2 + q + 1$. The graph is $(q+1)$-regular, in which $q+1$ vertices have loops (each of such vertices has one). So the edge density $p \sim \frac{1}{\sqrt{n}}$. We have found in Chapter 9 that $\lambda_1 = q + 1 \sim pn$, and $\lambda = \sqrt{q} = o(d)$. So the property P_3 holds. However,

$$p^4(1-p)^2n^4 \sim n^2,$$

and thus the property $P_1(4)$ does not hold as E_q^o does not contain C_4.

Recall that the quasi-random property P_3, the magnitude of $\lambda = \lambda(G)$ is a measure of quasi-randomness. As called by Alon, a graph G is an (n, d, λ)-graph if G is d-regular with n vertices and

$$\lambda = \lambda(G) = \max\{|\lambda_i| : \ 2 \le i \le n\},$$

where $\lambda_1 = d$, and $\lambda_2, \ldots, \lambda_n$ are all eigenvalues of G. Here he connected quasi-randomness to the eigenvalue gap. For sparse graphs with $p = o(1)$, Chung and Graham (2002) found some equivalent properties under certain conditions. One of the properties is that $\lambda_1 \sim pn$ and $\lambda = o(\lambda_1)$.

We shall have more results on (n, d, λ)-graphs, which are due to Alon et.al, particularly Alon and Spencer (2008). For two (not necessarily disjoint) subsets B and C, we have defined $e(B, C)$ as the number of ordered pairs (u, v) with $u \in B$ and $v \in C$. If G is simple, then $e(B, C)$ is the same as defined in the last section, i.e., it counts each edge from $B \setminus C$ to $C \setminus B$ once, and each edge in $B \cap C$ twice. When G is semi-simple, it also counts each loop in $B \cap C$ once. For disjoint subsets B and C in a random graph, $e(B, C)$ is expected to be $\frac{d}{n}|B||C|$, which is close to the right-hand side of the inequality in the following theorem if λ is much smaller than d.

Theorem 5.7 *Let $G = (V, E)$ be a semi-simple (n, d, λ)-graph. Then for each partition of V into disjoint subsets B and C,*

$$e(B, C) \geq \frac{(d - \lambda)|B||C|}{n}$$

Proof. Let A be the adjacency matrix of G and I the identity matrix of order n. Observe that for any real vector x of dimension n (as a real valued function on V), we have

$$((dI - A)x, x) = \sum_{u \in V} \Big(dx_u^2 - \sum_{v:\ uv \in E} x_v x_u \Big)$$

$$= \ d \sum_{u \in V} x_u^2 - 2 \sum_{uv \in E} x_v x_u = \sum_{uv \in E} \Big(x_u - x_v \Big)^2.$$

Set $b = |B|$ and $c = |C| = n - b$. Define a vector $x = (x_v)$ by

$$x_v = \begin{cases} -c & v \in B, \\ b & v \in C. \end{cases}$$

Note that $dI - A$ and A have the same eigenvectors, and that the eigenvalues of $dI - A$ are precisely $d - \mu$ as μ ranges over all eigenvalues of A. Also, d is the largest eigenvalue of A corresponding to the eigenvector $J = (1, 1, \ldots, 1)^t$ and $(x, J) = 0$. Hence x is orthogonal to the eigenvector of the smallest eigenvalue of $dI - A$.

Since $dI - A$ is a symmetric matrix, its eigenvectors are orthogonal each other and form a basis of the n-dimensional space and x is a linear combination of these eigenvectors other than that of $J/\sqrt{n}$. This together with the fact that $d - \lambda$ is the second smallest eigenvalue of $dI - A$, we have

$$((dI - A)x, x) \geq (d - \lambda)(x, x) = (d - \lambda)(bc^2 + cb^2) = (d - \lambda)bcn.$$

However, as B and C form a partition of V,

$$\sum_{uv \in E} (x_u - x_v)^2 = e(B, C)(b + c)^2 = e(B, C)n^2,$$

implying the desired inequality. □

The next theorem bounds some kind of variance. In a random d-regular graph, we expect that a vertex v has $\frac{d}{n}|B|$ neighbors in B. The theorem shows that if λ is small, then $|N_B(v)|$ is not too far from the expectation for most vertices v, where $N_B(v) = N(v) \cap B$.

Theorem 5.8 *Let $G = (V, E)$ be a semi-simple (n, d, λ) graph. Then for each $B \subseteq V$,*

$$\sum_{v \in V} \Big(|N_B(v)| - \frac{d}{n}|B| \Big)^2 \leq \lambda^2 \frac{|B|(n - |B|)}{n}.$$

Proof. Let A be the adjacency matrix of G. Define a vector $f : V \to R$ by

$$f_u = \begin{cases} 1 - \frac{b}{n} & u \in B, \\ -\frac{b}{n} & u \notin B, \end{cases}$$

where $b = |B|$. Then $\sum_u f_u = 0$, and f is orthogonal to the eigenvector $J = (1, 1, \ldots, 1)^t$ of the largest eigenvalue d of A. Thus f is a linear combination of eigenvectors other than J, and

$$(Af, Af) \leq \lambda^2 (f, f) = \lambda^2 \frac{b(n - b)}{n}.$$

Let A_v be the row of A corresponding to vertex v. Then the coordinate $(Af)_v$ of Af at v is

$$A_v f = \Big(1 - \frac{b}{n}\Big)|N_B(v)| - \frac{b}{n}(d - |N_B(v)|) = |N_B(v)| - \frac{db}{n},$$

and thus

$$(Af, Af) = \sum_v \Big(|N_B(v)| - \frac{db}{n} \Big)^2,$$

the desired inequality follows. □

Corollary 5.1 *Let $G = (V, E)$ be a semi-simple (n, d, λ)-graph. Then for every two subsets B and C of G, we have*

$$\Big| e(B, C) - \frac{d}{n}|B||C| \Big| \leq \lambda \sqrt{|B||C|}.$$

Proof. Set $b = |B|$ and $c = |C|$. Note that

$$\begin{aligned} \Big| e(B, C) - \frac{dbc}{n} \Big| &= \Big| \sum_{v \in C} \Big(|N_B(v)| - \frac{db}{n} \Big) \Big| \leq \sum_{v \in C} \Big| |N_B(v)| - \frac{db}{n} \Big| \\ &\leq \sqrt{c} \Big[\sum_{v \in C} \Big(|N_B(v)| - \frac{db}{n} \Big)^2 \Big]^{1/2}, \end{aligned}$$

where the Cauchy-Schwarz inequality is used. From Theorem 5.8, we have

$$\begin{aligned}\Big|e(B,C)-\frac{dbc}{n}\Big| &\le \sqrt{c}\Big[\sum_{v\in V}\Big(|N_B(v)|-\frac{db}{n}\Big)^2\Big]^{1/2}\\ &\le \lambda\sqrt{c}\sqrt{b(1-\frac{b}{n})}\le\lambda\sqrt{bc}\end{aligned}$$

as desired. □

Let $e(B)$ and $\ell(B)$ be the number of edges and loops in B, respectively. Then

$$e(B,B)=2\,e(B)+\ell(B).$$

Note that $\ell(B)\le|B|$ if G is semi-simple.

Corollary 5.2 *Let $G=(V,E)$ be a semi-simple (n,d,λ) graph, and let B be a subset of G. Then*

$$\Big|e(B)-\frac{d}{2n}|B|^2\Big|\le\frac{\lambda+1}{2}|B|.$$

Remark. By setting $e(B)=0$, we have $\alpha(G)\le\frac{\lambda+1}{d}n$, which is slightly weaker than a similar bound obtained in Chapter 9.

For an (n,d,λ)-graph $G=(V,E)$ and $B\subseteq V$, define $\overline{B}$ as the set of vertices u so that the proportion of $N(u)$ in B, which is $|N_B(u)|/|B|$, is at most half of that in V. Then $|B||\overline{B}|$ is at most $\Theta(n^2/d)$ if $\lambda=\Theta(\sqrt{d})$.

Corollary 5.3 *Let $G=(V,E)$ be a semi-simple (n,d,λ)-graph and $B\subseteq V$. Define*

$$\overline{B}=\{u\in V:\ |N_B(u)|\le\frac{d}{2n}|B|\},$$

where $N_B(u)=N(u)\cap B$. Then

$$|B||\overline{B}|\le\Big(\frac{2\lambda n}{d}\Big)^2.$$

Consequently, $|B\cap\overline{B}|\le\frac{2\lambda n}{d}$.

Proof. From Theorem 5.8, we have

$$\sum_{v\in V}\Big(|N_B(v)| - \frac{d}{n}|B|\Big)^2 \le \lambda^2\frac{|B|(n-|B|)}{n} \le \lambda^2|B|.$$

Each $v \in \overline{B}$ contributes to the left-hand side more than $\Big(\frac{d|B|}{2n}\Big)^2$, thus

$$|\overline{B}|\Big(\frac{d|B|}{2n}\Big)^2 \le \lambda^2|B|,$$

implying what as claimed.

For an (n, d, λ)-graph, the spectral gap between d and λ is a measure for its quasi-random property. The smaller the value of λ compared to d, the closer is edge distribution to the ideal uniform distribution. How small can be λ?

Theorem 5.9 *Let G be an (n, d, λ)-graph and let $\epsilon > 0$. If $d \le (1-\epsilon)n$, then*

$$\lambda \ge \sqrt{\epsilon d}.$$

Proof. Let A be the adjacency matrix of G. Then

$$\begin{aligned} nd &= 2e(G) = tr(A^2) = \sum_{i=1}^{n}\lambda_i^2 \\ &\le d^2 + (n-1)\lambda^2 \le (1-\epsilon)nd + n\lambda^2, \end{aligned}$$

which follows by what claimed.

On this estimate, we can say, not precisely, that an (n, d, λ)-graph with $\lambda = \Omega(\sqrt{d})$ has good quasi-randomness. Recall a result in Chapter 2, if G is an $srg(n, d, \mu_1, \mu_2)$ with $n \ge 3$. Then, except $\lambda_1 = d$, the other eigenvalues are solutions of the equation

$$\lambda^2 + (\mu_2 - \mu_1)\lambda + (\mu_2 - d) = 0.$$

Thus when $\mu_1 - \mu_2$ is small compared to d, which implies that λ is close to $\sqrt{d}$, G has good quasi-randomness.

5.4 Erdős-Rényi graphs

The starting point of a problem involving complete bipartite graph is usually at $C_4 = K_{2,2}$. We begin with a construction of a graph of Erdős-Rényi (1962), which contains no C_4.

Let $F = F(q)$ be the Galois field with q elements. Define an equivalence relation $\equiv$ on $(F^3)^* = F^3 \setminus \{(0,0,0)\}$ by letting $(a_1, a_2, a_3) \equiv (b_1, b_2, b_3)$ if there is $\lambda \in F^* = F \setminus \{0\}$ such that $(a_1, a_2, a_3) = \lambda(b_1, b_2, b_3)$. Let $\langle a_1, a_2, a_3 \rangle$ denote the equivalence class containing (a_1, a_2, a_3), and let V be the set of all equivalence classes.

Define a graph E_q on vertex set V by letting distinct vertices $\langle a_1, a_2, a_3 \rangle$ and $\langle x_1, x_2, x_3 \rangle$ be adjacent if and only if

$$a_1x_1 + a_2x_2 + a_3x_3 = 0.$$

This definition is clearly compatible, i.e., it does not depend on the choice of representative elements of the equivalence classes. It is trivial to see that

$$|V| = \frac{q^3 - 1}{q - 1} = q^2 + q + 1.$$

For a vertex $A = \langle a_1, a_2, a_3 \rangle$, since $a_1x_1 + a_2x_2 + a_3x_3 = 0$ has $q^2 - 1$ solutions forming $q + 1$ vertices,

$$\deg(A) = \begin{cases} q & \text{if } a_1^2 + a_2^2 + a_3^2 = 0, \\ q + 1 & \text{otherwise.} \end{cases}$$

We now come to the point to see the most important fact on E_q what contains no C_4.

Theorem 5.10 *The graph E_q contains no C_4.*

Proof. Let $\langle a_1, a_2, a_3 \rangle$ and $\langle b_1, b_2, b_3 \rangle$ be distinct vertices. Then the vectors (a_1, a_2, a_3) and (b_1, b_2, b_3) are linearly independent. Consider the equation system

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 &= 0 \\ b_1x_1 + b_2x_2 + b_3x_3 &= 0, \end{cases}$$

which has exactly $q - 1$ solutions forming only one vertex. The desired assertion follows. □

Let $n = q^2 + q + 1$ and let $e(E_q)$ be the number of edges of E_q. Then, as $q \to \infty$,

$$ex(n; C_4) \geq e(E_q) \geq (1 - o(1))\frac{1}{2}n^{3/2}.$$

Let us associate the graph E_q with a more general construction, which is a $(q+1)$-uniform hypergraph $(X, \mathcal{L})$ called projective plane. However, the members in $\mathcal{L}$ are called lines, and the order of such a plane does not mean the cardinality of X. We will need projective planes in the next section.

A *projective plane* of order q consists of a set X of q^2+q+1 elements called *points*, and a family $\mathcal{L}$ of subsets of X called *lines*, having the following properties:

(P1) Every line has $q + 1$ points.

(P2) Any pair of distinct points lie on a unique line.

The only possible projective plane of order $q = 1$ is a triangle. The unique projective plane of order $q = 2$ is the famous *Fano plane*. It contains 7 points, 7 lines, in which each line has 3 points, see Fig 9.1

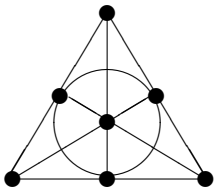

Fig. 9.1. The Fano plane

Additional properties of projective planes are as follows.

Corollary 5.4 *A projective plane of order q has the properties as follows.*

(P3) Any point lies on $q+1$ lines.
(P4) There are q^2+q+1 lines.
(P5) Any two lines meet in a unique point.

Proof. Fix a point $x \in X$. There are $q(q+1)$ points different from x; each line through x contains q further points, and there are no other overlaps between these lines (apart from x). So $q(q+1)$ points of $X\setminus\{x\}$ are partitioned equally into parts by these lines. Therefore there must be $q+1$ lines through x.

To show (P4), let us count in two ways the pairs (x, L) with $x \in L$, we obtain $|\mathcal{L}|(q+1) = (q^2+q+1)(q+1)$, so $|\mathcal{L}| = q^2+q+1$.

Finally, let L_1 and L_2 be distinct lines, and let x be a point of L_1. Then the $q+1$ points of L_2 are joined to x by different lines. Since there are only $q+1$ lines through x, they all meet L_2 in a point. In particular, L_1 meets L_2 in a point. □

A nice property of projective planes is their duality. Let $(X, \mathcal{L})$ be a projective plane of order q, and let $M = (m_{x,L})$ be its incidence matrix, in which the rows and columns correspond to points and lines. Each row and column of M has exactly $q+1$ $1'$s, and any two rows and any two columns share exactly one 1. The transpose of M leaves the matrix unchanged.

Return to the graph E_q, whose vertex set is V consisting of q^2+q+1 points (equivalence classes in $(F_q^3)^*$). Let $\langle a_1, a_2, a_3\rangle$ be a point of V. Define a line $L(a_1, a_2, a_3)$ to be the set of all points $\langle x_1, x_2, x_3\rangle$ in V (not vectors (x_1, x_2, x_3) in $(F_q^3)^*$) for which

$$a_1x_1 + a_2x_2 + a_3x_3 = 0.$$

It is easy to see that the definition for lines is compatible, and each line contains exactly $q+1$ points. Note that some lines $L(a_1, a_2, a_3)$ contain point $\langle a_1, a_2, a_3\rangle$ and some do not. Any pair of distinct points $\langle x_1, x_2, x_3\rangle$ and $\langle y_1, y_2, y_3\rangle$ lie on a unique line $L(a_1, a_2, a_3)$ with

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 &= 0, \\ a_1y_1 + a_2y_2 + a_3y_3 &= 0. \end{cases}$$

Therefore, we obtain a projective plane $(V, \mathcal{L})$, where $\mathcal{L}$ consists of all lines defined. This projective plane is usually denoted by $PG(2, q)$.

Some authors use $PG(2,q)$ to signify the Erdős-Rényi graph or a bipartite graph, whose bipartition are points and lines, in which a point is adjacent to a line if and only if the point is contained in the line.

No projective plane of order non-prime power is known to exist, and it is conjectured that there is none. It is known that there is no projective plane of order 6, 10 or 14. It is not known whether there is a projective plane of order 12.

We hope to have an exact expression of $e(E_q)$.

Lemma 5.2 *Let $q = p^m$ with prime p and odd m. Then there are precisely $q^2 - 1$ non-zero solutions (x_1, x_2, x_3) of the equation*

$$x_1^2 + x_2^2 + x_3^2 = 0$$

in $F(q)$, and hence precisely $q + 1$ vertices in E_q of degree q and q^2 vertices of degree $q + 1$.

Proof. Label the vertex set of E_q as

$$V(E_q) = \{A, B, \cdots, X, \cdots, Y, \cdots, Z\}$$

in some order. We write $X \perp Y$ if and only if

$$x_1y_1 + x_2y_2 + x_3y_3 = 0,$$

where $X = \langle x_1, x_2, x_3\rangle$ and $Y = \langle y_1, y_2, y_3\rangle$. Let $n = q^2 + q + 1$ and define an $n \times n$ real matrix $M = (m_{ij})$ by

$$m_{ij} = \begin{cases} 1 & \text{if } X \perp Y, \\ 0 & \text{otherwise,} \end{cases}$$

where X and Y represent the ith vertex and the jth vertex, and X and Y are associated with the ith row and jth row in M, respectively. Note the diagonal elements of M are different from that in the adjacency matrix of E_q. We have $m_{ii} = 1$ if $X \perp X$, that is, X lies on the conic $x_1^2 + x_2^2 + x_3^2 = 0$. All that remains to show is that

$$tr(M) = q + 1,$$

where $tr(M) = \sum_{i=1}^{n} m_{ii}$ is the trace of M. We know that the trace equals the sum of eigenvalues.

Fact 1. Any row of M contains precisely $q+1$ ones hence $q+1$ is an eigenvalue of M.

This is because $ML = (q+1)L$, where $L = (1, 1, \ldots, 1)^T$.

Fact 2. For $i \neq j$, there is exactly one column with 1 in both the ith row and the jth row. Namely, $M_i \cdot M_j = 1$, where M_i and M_j are the ith row and the jth row of M, respectively.

Suppose that X and Y represent (the vertices) the ith row and the jth row, respectively. Then there is a unique (vertex) row, say the kth row, corresponding to the solution (w_1, w_2, w_3) of the equation system

$$\begin{cases} x_1w_1 + x_2w_2 + x_3w_3 &= 0, \\ y_1w_1 + y_2w_2 + y_3w_3 &= 0. \end{cases}$$

That is to say, $m_{ik} = m_{jk} = 1$. Note that M is symmetric, so we see that only in the kth column, the elements in both the ith row and the jth row are 1.

Using these two facts and the symmetry of M, we have

$$M^2 = \begin{pmatrix} q+1 & 1 & \cdots & 1 & 1 \\ 1 & q+1 & \cdots & 1 & 1 \\ \vdots & & & & \vdots \\ 1 & 1 & \cdots & 1 & q+1 \end{pmatrix} = qI + J,$$

where I is the identity matrix and J is the all-ones-matrix. It is easy to see J has the eigenvalues $n = q^2 + q + 1$ (of multiplicity 1) and 0 (of multiplicity $n-1 = q^2+q$). This follows by that M^2 has the eigenvalues $q+n = (q+1)^2$ (of multiplicity 1) and q (of multiplicity $n-1 = q^2+q$).

Since M is symmetric and hence it is diagonalizable into

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

where $\lambda_1, \ldots, \lambda_n$ are eigenvalues of M. This implies that M^2 is diagonalizable into

$$\begin{pmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_n^2 \end{pmatrix},$$

implying that $\lambda_1^2 = (q+1)^2$ and $\lambda_2^2 = \ldots = \lambda_n^2 = q$, in which the eigenvalues may be relabelled. So $\lambda_1 = q+1$ as $q+1$ is an eigenvalue of M, and $\lambda_i = \pm\sqrt{q}$ for $i = 2, \ldots, n$. Let s and t be the numbers of eigenvalues of M equal to $\sqrt{q}$ and $-\sqrt{q}$, respectively. Then $s+t = n-1$ and

$$tr(M) = (q+1) + (s-t)\sqrt{q}.$$

Since the trace is an integer and $q = p^m$ with m odd, we must have $s = t$ hence $tr(M) = q+1$, completing the proof. □

Let us remark that in fact for any prime power q, the assertion in above lemma holds.

Theorem 5.11 *The order of the graph E_q is q^2+q+1 and*

$$e(E_q) = \frac{1}{2}q(q+1)^2$$

for prime power q. □

Remak. If we define a loop for a vertex A when $a_1^2 + a_2^2 + a_3^2 = 0$, and a graph E_q^o from E_q by attaching a loop for a vertex $\langle a_1, a_2, a_3 \rangle$, then E_q^o is $(q+1)$-regular. The proof of Lemma 5.2 in fact gives that the distinct eigenvalues of E_q^o are $q+1$, $\sqrt{q}$ and $-\sqrt{q}$. So $\lambda \sim \sqrt{d}$. These eigenvalues are close to that of E_q.

Lemma 5.3 *Let A and B be real symmetric matrices of order n. Let the eigenvalues $\lambda_i(A)$, $\lambda_i(B)$ and $\lambda_i(A+B)$ of A, B and $A+B$, respectively, be labelled in non-increasing order. Then, for each $1 \leq i \leq n$, we have*

$$\lambda_i(A) + \lambda_1(B) \geq \lambda_i(A+B) \geq \lambda_i(A) + \lambda_n(B).$$

Therefore, the eigenvalues λ_i of E_q satisfy $q \le \lambda_1 \le q+1$ and the other positive λ_i has $\sqrt{q} \le \lambda_i \le \sqrt{q}+1$, and negative ones have $-\sqrt{q}+1 \le \lambda_i \le -\sqrt{q}+1$.

Note that E_q does not have a fixed positive density. However, it has a good quasi-randomness as $\lambda_1 = O(\sqrt{d})$.

5.5 Applications of characters ⋆

We shall find the spectrum of $G_{q,t}$ defined in Chapter 9. Let us define the characters of a finite field $F(q)$, which are group homomorphisms from $F(q)$ or $F^*(q)$ to

$$S^1 = \{z : |z| = 1\} = \{e^{i\theta} : 0 \le \theta < 2\pi\},$$

respectively, where S^1 is a multiplicative group of complex numbers. An *additive character* of $F(q)$ is a function $\psi : F(q) \to S^1$ such that for any $x, y \in F(q)$,

$$\psi(x+y) = \psi(x)\psi(y).$$

Clearly $\psi(0) = 1$ and $\psi(-x) = \overline{\psi(x)}$. The trivial function ψ_0 with $\psi_0(x) \equiv 1$ is also called the *principle additive character* of $F(q)$.

A *multiplicative character* of $F(q)$ is a function $\chi : F^*(q) \to S^1$ such that for any $x, y \in F^*(q)$,

$$\chi(xy) = \chi(x)\chi(y).$$

Clearly $\chi(1) = 1$ and $\chi(x^{-1}) = \overline{\chi(x)}$. The trivial function χ_0 with $\chi_0(x) \equiv 1$ is also called the *principal multiplicative character* of $F(q)$. It is often to extend the domain of a multiplicative character χ to all of $F(q)$ by $\chi(0) = 0$ if $\chi \neq \chi_0$, and $\chi_0(0) = 1$. The character χ with $\chi(x) = x^{(q-1)/2}$ is often called *quadratic residue character*.

In the following proofs, we shall not distinguish the elements of $F(p)$ from the integers of $\{0, 1, \ldots, p-1\}$.

Lemma 5.4 *The numbers of additive characters and multiplicative characters of $F(q)$ are q and $q-1$, respectively.*

Proof. Let us begin with the multiplicative group $F^*(q)$, which is a cyclic group of order $q-1$ with $F^*(q) = \{1, \mu, \ldots, \mu^{q-2}\}$, where μ is a primitive element of $F(q)$. Each multiplicative character χ of $F(q)$ is uniquely determined by $\chi(\mu)$. From $1 = \chi(\mu^{q-1}) = \chi(\mu)^{q-1}$, we have that $\chi(\mu) = \zeta_{q-1}^k$ for some $0 \leq k \leq q-2$, where $\zeta_{q-1} = e^{2\pi i/(q-1)}$. If we use χ_1 to signify the multiplicative character of $F(q)$ with $\chi_1(\mu) = \zeta_{q-1}$, then the set of all multiplicative characters of $F(q)$ is $\{\chi_1^k : 0 \leq k \leq q-2\}$, in which χ_1^0 is the trivial character χ_0. Thus $F(q)$ has $q-1$ multiplicative characters, forming a group isomorphic to $F^*(q)$.

Let $q = p^m$ and let $\zeta_p = e^{2\pi i/p}$. For each $a = (a_1, a_2, \ldots, a_m) \in F^m(p)$, set

$$\psi_a : F(q) \to S^1, \psi_a(x) = \zeta_p^{a_1x_1 + a_2x_2 + \cdots + a_mx_m},$$

where $x = (x_1, x_2, \ldots, x_m)$ is the unique expression of x as a vector of $F^m(p)$. Then ψ_a is an additive character of $F(q)$. For $a \neq a'$, we show that $\psi_a \neq \psi_{a'}$. It suffices to show that ψ_a is not the trivial character for $a \neq 0$. Since for $a \neq 0$ there is some k such that $1 \leq a_k \leq p-1$, so for $e_k = (0, \ldots, 0, 1, 0, \ldots, 0)$, the unit vector with 1 at the kth coordinate, $\psi_a(e_k) = \zeta_p^{a_k} \neq 1$. Thus the group of additive characters of $F(q)$ contains at least hence exactly q elements. □

As usual, the function $\delta(x, y)$ is the Kronecker's symbol defined as

$$\delta(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise.} \end{cases}$$

The proof of Lemma 5.4 implies the following result.

Lemma 5.5 *Let χ be a multiplicative character of $F(q)$. Then*

$$\sum_{t \in F(q)} \chi(t) = q\,\delta(\chi, \chi_0).$$

Let us define the Gaussian sum as

$$G(\chi, \psi) = \sum_{x \in F(q)} \chi(x)\psi(x).$$

Theorem 5.12 *Let χ be a multiplicative character of $F(q)$ and ψ be an additive character of $F(q)$. Then $G(\chi_0, \psi_0) = q$, and $G(\chi, \psi) = 0$ if exactly one of χ and ψ is trivial. Furthermore*

$$|G(\chi, \psi)| = \sqrt{q}$$

if none of χ and ψ is trivial.

Proof. The first two equalities are easy, and we shall verify the last. In order to simplify the proof, we prove it for the case that q is a prime p, which is the most important special case. The proof for general q will be given later, which is more involved and the readers are encouraged to skip.

Let ψ and χ be additive character and multiplicative character of $F(p)$, none of which is trivial. From the proof of Lemma 5.4, we have $\psi(x) = \zeta_p^{ax}$, where $\zeta_p = e^{2\pi i/p}$ and $a \neq 0$. Let $g_a(\chi) = \sum_{x \in F(p)} \chi(x)\zeta_p^{ax}$, which is $G(\chi, \psi)$ on $F(p)$.

We shall verify that $\overline{g_a(\chi)} = \chi(a)\overline{g_1(\chi)}$. This follows from that

$$\begin{aligned} g_a(\chi) &= \sum_{x \in F(p)} \chi(x)\zeta_p^{ax} = \sum_{y \in F(p)} \chi(a^{-1}y)\zeta_p^{y} \\ &= \chi(a^{-1}) \sum_{y \in F(p)} \chi(y)\zeta_p^{y} = \overline{\chi(a)} g_1(\chi), \end{aligned}$$

and

$$|g_a(\chi)|^2 = g_a(\chi)\overline{g_a(\chi)} = |\chi(a)|^2 |g_1(\chi)|^2 = |g_1(\chi)|^2.$$

That is to say, $|g_a(\chi)|^2$ have the same value for any $a \neq 0$. On the other hand, for any $a \in F(p)$

$$g_a(\chi)\overline{g_a(\chi)} = \sum_{x \in F(p)} \chi(x)\zeta_p^{ax} \sum_{y \in F(p)} \overline{\chi(y)}\zeta_p^{-ay} = \sum_{x,\, y \in F(p)} \chi(x)\overline{\chi(y)}\zeta_p^{a(x-y)}.$$

It is easy to see that $\sum_{a \in F(p)} \zeta_p^{a(x-y)} = p\,\delta(x, y)$ as $a(x-y)$ ranges all of $F(p)$ for $x - y \neq 0$, which and the fact that $\chi(0) = 0$ as $\chi \neq \chi_0$ imply that

$$\sum_{a \in F(p)} g_a(\chi)\overline{g_a(\chi)} = \sum_{x,\, y \in F(p)} \chi(x)\overline{\chi(y)}\delta(x, y)p = (p-1)p.$$

Since $g_0(\chi) = 0$ as $\chi \neq \chi_0$, we obtain that $(p-1)|g_1(\chi)|^2 = (p-1)p$ hence $|g_a(\chi)| = |g_1(\chi)| = \sqrt{p}$. □

The proof of Theorem 5.12 for general $q = p^m$

The forms of additive characters in the proof of Lemma 5.4 are simple, but we shall express them in the other way for proving Theorem 5.12 in general case.

For $\alpha \in F(q) = F(p^m)$, define the *trace* of α to be

$$tr(\alpha) = \alpha + \alpha^p + \alpha^{p^2} + \cdots + \alpha^{p^{m-1}}.$$

Lemma 5.6 *If $\alpha, \beta \in F(q)$ and $a \in F(p)$, then*
(a) $tr(\alpha) \in F(p)$.
(b) $tr(\alpha + \beta) = tr(\alpha) + tr(\beta)$.
(c) $tr(a\alpha) = a\, tr(\alpha)$.
(d) For fixed $\alpha \neq 0$, $tr(\alpha x)$ maps $F(q)$ onto $F(p)$.

Proof. The properties (a),(b) and (c) follow from the facts that $tr^p(\alpha) = tr(\alpha)$, $(\alpha+\beta)^p = \alpha^p+\beta^p$, $\alpha^q = \alpha$ and $a^p = a$. To show the property (d), consider the fact that the polynomial $tr(\alpha x)$ has at most p^{m-1} roots and αx ranges all p^m elements of $F(q)$, there is $x \in F(q) = F(p^m)$ such that $tr(\alpha x) = c \neq 0$, where $c \in F(p)$. If $b \in F(p)$, the using property (c) we see that $tr((b/c)\alpha x) = (b/c)tr(\alpha x) = b$. Thus the trace $tr(\alpha x)$ is onto. □

For fixed $\alpha \in F(q)$, we now define $\psi_\alpha : F(q) \to S^1$ by

$$\psi_\alpha(x) = \zeta_p^{tr(\alpha x)},$$

where $\zeta_p = e^{2\pi i/p}$. Note that ψ_0 is the trivial additive character of $F(q)$. For the case $q = p$, $\psi_\alpha(x) = \zeta_p^{\alpha x}$ is exactly what we have used.

Lemma 5.7 *The function ψ_α has the following properties.*
(a) $\psi_\alpha(x + y) = \psi_\alpha(x)\psi_\alpha(y)$ for any $x, y \in F(q)$.
(b) If $\alpha \neq 0$, then there is $x \in F(q)$ such that $\psi_\alpha(x) \neq 1$.
(c) If $\alpha \neq 0$, then $\sum_{x \in F(q)} \psi_\alpha(x) = 0$; if $x \neq 0$, then $\sum_{\alpha \in F(q)} \psi_\alpha(x) = 0$.

Proof. The property (a) follows from that $tr(\alpha(x+y)) = tr(\alpha x) + tr(\alpha y)$. The property (b) follows from the fact that $tr(\alpha x)$ is onto as $\alpha \neq 0$, so there $x \in F(q)$ such that $tr(\alpha x) = 1$. Then $\psi_\alpha(x) = \zeta_p \neq 1$. As $\psi_\alpha(x) = \psi_x(\alpha)$, we shall only verify the first equality in the property (c). Let $S = \sum_{x \in F(q)} \psi_\alpha(x)$. Choose y such that $\psi_\alpha(y) \neq 1$, thus

$$\psi_\alpha(y)S = \sum_{x \in F(q)} \psi_\alpha(x)\psi_\alpha(y) = \sum_{x \in F(q)} \psi_\alpha(x+y) = S,$$

thus $S = 0$. □

Lemma 5.8 *For any fixed $\alpha \in F(q)$, the function ψ_α is an additive character of $F(q)$, and any additive character of $F(q)$ is of such form. Furthermore, for any $x, y \in F(q)$,*

$$\sum_{\alpha \in F(q)} \psi_\alpha(x-y) = q\,\delta(x,y).$$

Proof. The first assertion follows the property (a) in Lemma 5.7. For the second, we shall verify that the number of such functions is q. It suffices to show that if $\alpha \neq \beta$, the functions ψ_α and ψ_β are distinct. If $\psi_\alpha(x) = \psi_\beta(x)$ for any $x \in F(q)$, then

$$\zeta_p^{tr((\alpha-\beta)x)} = \psi_{\alpha-\beta}(x) = 1$$

for any $x \in F(q)$, implying that $\alpha = \beta$ from the property (b) in Lemma 5.7.

Since

$$\sum_{\alpha \in F(q)} \psi_\alpha(x-y) = \sum_{\alpha \in F(q)} \zeta_p^{tr(\alpha(x-y))},$$

which is q for $x = y$. If $x \neq y$, equality follows from the fact that $\alpha(x-y)$ ranges over all of $F(q)$ and the property (c) in Lemma 5.7. □

We now write Gaussian sum in the form

$$G(\chi, \psi_\alpha) = \sum_{x \in F(q)} \chi(x)\psi_\alpha(x).$$

We shall prove that if $\chi \neq \chi_0$ and $\alpha \neq 0$, then

$$|G(\chi, \psi_\alpha)| = \sqrt{q}.$$

Proof. The proof is an analogy of that for the case $q = p$. For any $\alpha \neq 0$, we first verify that $\overline{G(\chi, \psi_\alpha)} = \chi(\alpha)\overline{G(\chi, \psi_1)}$. This is because that

$$\begin{aligned} G(\chi, \psi_\alpha) &= \sum_{x \in F(q)} \chi(x)\zeta_p^{tr(\alpha x)} = \sum_{y \in F(q)} \chi(\alpha^{-1}y)\zeta_p^{tr(y)} \\ &= \chi(\alpha^{-1}) \sum_{y \in F(q)} \chi(y)\zeta_p^{tr(y)} = \overline{\chi(\alpha)}G(\chi, \psi_1). \end{aligned}$$

Therefore, we have $|G(\chi, \psi_\alpha)|^2 = |G(\chi, \psi_1)|^2$ for $\alpha \neq 0$. On the other hand, for any $\alpha \in F(q)$,

$$\begin{aligned} G(\chi, \psi_\alpha)\overline{G(\chi, \psi_\alpha)} &= \sum_{x \in F(q)} \chi(x)\zeta_p^{tr(\alpha x))} \sum_{y \in F(q)} \overline{\chi(y)}\zeta_p^{-tr(\alpha y))} \\ &= \sum_{x,\, y \in F(q)} \chi(x)\overline{\chi(y)}\zeta_p^{tr(\alpha(x-y))}. \end{aligned}$$

Since $\chi(0) = 0$ as $\chi \neq \chi_0$,

$$\sum_{\alpha \in F(q)} G(\chi, \psi_\alpha)\overline{G(\chi, \psi_\alpha)} = \sum_{x,\, y} \chi(x)\overline{\chi(y)}\delta(x, y)q = q(q-1).$$

Observing that $G(\chi, \psi_0) = 0$ as $\chi \neq \chi_0$, we have

$$\sum_{\alpha \in F(q)} G(\chi, \psi_\alpha)\overline{G(\chi, \psi_\alpha)} = \sum_{\alpha \in F^*(q)} |G(\chi, \psi_1)|^2 = (q-1)|G(\chi, \psi_1)|^2,$$

yielding that $(q-1)|G(\chi, \psi_1)|^2 = q(q-1)$ hence $|G(\chi, \psi_\alpha)| = |G(\chi, \psi_1)| = \sqrt{q}$. The proof for general case of Theorem 5.12 is completed. □

The order of a multiplicative character χ is the smallest positive integer d such that $\chi^d = \chi_0$. A more sophisticated result on character sum is the Weil's theorem as follows. Let χ be the multiplicative character of $F_q = F(q)$ of order $d > 1$ and $f(x)$ a polynomial over F_q. If $f(x)$ has precisely s distinct zeros and it is not the form $c(g(x))^d$, where $c \in F_q$ and $g(x) \in F_q[x]$, then

$$\left| \sum_{x \in F(q)} \chi(f(x)) \right| \leq (s-1)\sqrt{q}. \tag{5.1}$$

In particular, the inequality holds when χ is the square residue character and $f(x)$ is not the form $cg^2(x)$, where $c \in F_q$ and $g(x) \in F_q[x]$. Similarly, for an additive character $\psi \neq \psi_0$, if $g(x)$ is a polynomial of degree $n < q$ with g.c.d. $(n, q) = 1$, then $\left|\sum_{x \in F(q)} \psi(g(x))\right| \leq (n-1)\sqrt{q}$.

The prepared results are enough for introducing the following result of Szabó (2003) on spectrum of $G^o_{q,t}$, which is constructed in Chapter 9. Note that $G^o_{q,t}$ is $q^{t-1}(q-1)$-regular, in which each vertex $(A, a) \in F(q^2) \times F^*(q)$ with $N(2A) = a^2$ has a loop.

Theorem 5.13 *Let $t \geq 2$ be an integer and q be an odd prime power. The spectrum of $G^o_{q,t}$ is as follows.*

eig.	$q^{t-1}-1$	$q^{(t-1)/2}$	1	0	-1	$-q^{(t-1)/2}$
mul.	1	$\frac{(q^{t-1}-1)(q-2)}{2}$	$\frac{q^{t-1}-1}{2}$	$q-2$	$\frac{q^{t-1}-1}{2}$	$\frac{(q^{t-1}-1)(q-2)}{2}$

Proof. Let M be the adjacency matrix of $G^o_{q,t}$. Let ψ be an additive character of $F(q^{t-1})$ and χ be a multiplicative character of $F(q)$. Let $V(\psi, \chi)$ be the column vector whose coordinates are labeled by vertices of $G^o_{q,t}$, whose entry at (X, x) is $\psi(X)\chi(x)$. Then the entry of the column vector $MV(\psi, \chi)$ at (A, a) is

$$\begin{aligned}
\sum_{\substack{(B,b)\in F(q^{t-1})\times F^*(q)\\ N(A+B)=ab}} \psi(B)\chi(b) &= \sum_{B\in F(q^{t-1})\setminus\{-A\}} \psi(B)\chi\left(\frac{N(A+B)}{a}\right)\\
&= \sum_{C\in F^*(q^{t-1})} \psi(C-A)\chi\left(\frac{N(C)}{a}\right)\\
&= \left(\sum_{C\in F^*(q^{t-1})} \psi(C)\chi(N(C))\right)\overline{\psi(A)}\,\overline{\chi(a)}.
\end{aligned}$$

Setting

$$\lambda = \lambda(\psi, \chi) = \sum_{C\in F^*(q^{t-1})} \psi(C)\chi(N(C)),$$

then $\overline{\lambda(\chi, \psi)} = \lambda(\overline{\psi}, \overline{\chi})$,

$$MV(\psi, \chi) = \lambda V(\overline{\psi},\ \overline{\chi}), \tag{5.2}$$

and $MV(\overline{\psi},\ \overline{\chi}) = \overline{\lambda}V(\psi,\chi)$. Thus we have

$$M^2V(\psi,\chi) = \lambda\,\overline{\lambda}V(\psi,\chi) = |\lambda|^2V(\psi,\chi).$$

Hence $V(\psi,\chi)$ is an eigenvector of M^2 with corresponding eigenvalue $|\lambda(\psi,\chi)|^2$.

Observe that in the multiplicative group consisting of additive characters, the inverse ψ^{-1} of ψ is $\overline{\psi}$, and the similar statement holds also for the multiplicative group consisting of multiplicative characters. We claim that the eigenvectors of the form $V(\psi,\chi)$ are pairwise orthogonal. Let $(\psi',\chi') \neq (\psi,\chi)$, and let $\psi'' = \psi'\psi^{-1} = \psi'\overline{\psi}$ and $\chi'' = \chi'\chi^{-1} = \chi'\overline{\chi}$. Then $(\psi'',\chi'') \neq (\psi_0,\chi_0)$, where ψ_0 and χ_0 are trivial additive character and trivial multiplicative character, respectively. The inner product of complex vectors $V(\psi',\chi')$ and $V(\psi,\chi)$ is

$$\begin{aligned} V^T(\psi',\chi')\overline{V(\psi,\chi)} &= V^T(\psi',\chi')V(\overline{\psi},\overline{\chi}) \\ &= \sum_{(X,x)\in F(q^{t-1})\times F^*(q)} \psi''(X)\chi''(x) \\ &= \sum_{X\in F(q^{t-1})} \psi''(X) \sum_{x\in F^*(q)} \chi''(x) = 0, \end{aligned}$$

as one of sum in the last row is 0. The number of vectors of form $V(\psi,\chi)$ is equal to the order of $G^o_{q,t}$ by Lemma 5.4, hence all eigenvalues of M^2 are of form $|\lambda(\psi,\chi)|^2$. Therefore, any eigenvalue of M is of form

$$\pm|\lambda(\psi,\chi)| = \pm\left|\sum_{C\in F^*(q^{t-1})} \psi(C)\chi(N(C))\right|.$$

When $\psi = \psi_0$ and $\chi = \chi_0$, the corresponding eigenvalue is $q^{t-1}-1$, which can be obtained from Perron-Frobenius Theorem with multiplicity 1.

Let μ be a primitive element of $F(q^{t-1})$, and let

$$A_k = \{\mu^{k+j(q-1)} : 0 \le j \le \ell - 1\},$$

where $\ell = (q^{t-1}-1)/(q-1)$. Then $A_0, A_1, \ldots, A_{q-2}$ form a partition of $F^*(q)$ with $|A_k| = \ell$. It is easy to see $N(x) = N(y)$ if x and y are in the same A_k. Therefore, when $\psi = \psi_0$ and $\chi \neq \chi_0$, as

$$|\lambda(\psi_0,\chi)| = \left|\sum_{C\in F^*(q^{t-1})} \chi(N(C))\right| = \left|\ell \sum_{c\in F^*(q)} \chi(c)\right| = 0,$$

thus 0 is an eigenvalue of M with multiplicity $q-2$ which is the number of multiplicative characters of $F(q)$ except χ_0.

When $\psi \neq \psi_0$ and $\chi = \chi_0$,

$$\lambda(\psi, \chi_0) = \sum_{C \in F^*(q^{t-1})} \psi(C) = -\psi(0) = -1.$$

So 1 is an eigenvalue of M^2 with multiplicity $q^{t-1}-1$, then ± 1 are eigenvalues of M with the sum of the multiplicities being $q^{t-1}-1$. Let $W(\psi) = V(\psi, \chi_0) + V(\overline{\psi}, \chi_0)$. It follows from (5.2) that for any $\psi \neq \psi_0$, $MW(\psi) = -W(\psi)$. For any $\psi, \psi' \neq \psi_0, \psi \neq \overline{\psi'}$, it is easy to see that the complex vectors $W(\psi)$ and $W(\psi')$ are orthogonal. So -1 is an eigenvalue of M with multiplicity at least $(q^{t-1}-1)/2$. Similarly, by considering $V(\psi, \chi_0) - V(\overline{\psi}, \chi_0)$, we know that 1 is an eigenvalue of M with multiplicity at least $(q^{t-1}-1)/2$, hence each multiplicity is exactly $(q^{t-1}-1)/2$.

When $\psi \neq \psi_0$ and $\chi \neq \chi_0$, observing that χN is a non-trivial multiplicative character of $F(q^{t-1})$, by Theorem 5.12 on Gaussian sum,

$$|\lambda| = \left| \sum_{C \in F^*(q^{t-1})} \psi(C)\chi(N(C)) \right| = q^{(t-1)/2}.$$

Let S and T be the multiplicities of the eigenvalues $q^{(t-1)/2}$ and $-q^{(t-1)/2}$ of M, respectively. As $(q^{t-1}-1)(q-2)$ is the number of vectors of form $V(\psi, \chi)$ with $\psi \neq \psi_0$ and $\chi \neq \chi_0$,

$$S + T = (q^{t-1}-1)(q-2).$$

By the definition, the graph $G^o_{q,t}$ has a loop at each vertex (A, a) if and only if $N(2A) = a^2$. Since exactly $(q-1)/2$ elements of $F^*(q)$ are squares and the equation $N(X) = y$ has $(q^{t-1}-1)/(q-1)$ solutions in X for each fixed $y \in F^*(q)$, there are $(q^{t-1}-1)/2$ elements $A \in F(q^{t-1})$ with $N(2A)$ being a non-zero square. Once $N(2A)$ is a non-zero square, there are two distinct elements $a, -a \in F^*(q)$ with $N(2A) = a^2 = (-a)^2$. Thus $G^o_{q,t}$ contains $q^{t-1}-1$ loops, which is the trace of M. Hence

$$q^{t-1} - 1 = tr(M) = \sum_{j=1}^{q^{t-1}(q-1)} \lambda_j$$

$$= \quad q^{t-1} - 1 + \frac{q^{t-1}-1}{2} - \frac{q^{t-1}-1}{2} + q^{(t-1)/2}(S-T),$$

implying that $S = T = (q^{t-1}-1)(q-2)/2$. □

The above theorem has the following corollary, which and the lower bound in Chapter 9 imply $\alpha(G_{q,t}) = \alpha(G^o_{q,t}) = \Theta(q^{(t+1)/2})$.

Corollary 5.5 *Let $t \geq 2$ be an integer and q be an odd prime power. Then*

$$\alpha(G_{q,t}) \leq \frac{(q^t - q^{t-1})(q^{(t-1)/2}+1)}{q^{t-1}+q^{(t-1)/2}-1} \sim q^{(t+1)/2} \sim n^{(t+1)/(2t)},$$

where $n = q^{t-1}(q-1)$ is the order of $G_{q,t}$.

Let us conclude this section with an algebraic construction that almost matches the probabilistic bound $r_k(K_{m,n}) \geq k^m n - n^{1/2+o(1)}$ in Chapter 5.

Theorem 5.14 *Let positive integers k and m be fixed. Then*

$$r_k(K_{m,n}) \geq k^m n - n^{0.525}.$$

for large n.

Proof. As the assertion is trivial for $k = 1$, we assume that $k \geq 2$. Let $p \equiv 1 \pmod{2k}$ be a prime and F_p the finite field of p elements. Let μ be a primitive element of F_p. Define a logarithmic-like function $\log_\mu(x) : F_p^* \to Z_{p-1} = \{0, 1, \ldots, p-2\}$ as

$$\log_\mu(x) = \ell \text{ if } x = \mu^\ell,\ 0 \leq \ell \leq p-2.$$

For every j with $0 \leq j \leq k-1$, define a graph H_j on vertex set F_p, in which x and y is adjacent in H_j if and only if

$$\log_\mu(x-y) \equiv j \ (\mathrm{mod}\, k).$$

As $p \equiv 1 \pmod{2k}$ and $(-1)^2 = 1$, we have $-1 = \mu^{(p-1)/2}$, and thus $\log_\mu(x-y) \equiv \log_\mu(y-x) \pmod{k}$, so the definition is compatible. In case $k = 2$, the graph H_0 is the Paley graph.

Lemma 5.9 *Let $k \geq 2$ be an integer and $p \equiv 1 \pmod{2k}$ be a prime. Let H_j, $0 \leq j \leq k-1$, be the graph defined with respect to a primitive element of μ of F_p. Then these H_j are pairwise isomorphic.*

Proof. We shall verify that each H_j is isomorphic to H_0. Define a bijection ϕ on $F(p)$ as $\phi(z) = \mu^j z$. Then $\{x, y\}$ is an edge of H_0 if and only if $x - y = \mu^\ell$ for some $\ell \equiv 0 \pmod{k}$. As $\phi(x) - \phi(y) = \mu^{j+\ell}$, thus $\{x, y\}$ is an edge of H_0 if and only if $\{\phi(x), \phi(y)\}$ is an edge of H_j. Thus H_j is isomorphic to H_0. For any vertex x, its neighborhood in H_0 is

$$\{x + \mu^k, x + \mu^{2k}, \ldots, x + \mu^{p-1}\},$$

so the degree of x in H_0 is $((p-1)/k)$. This proves the lemma. □

Let $\zeta_k = e^{2\pi i/k}$. It is easy to see the following identity holds

$$(x - \zeta_k)(x - \zeta_k^2) \cdots (x - \zeta_k^{k-1}) = 1 + x + x^2 + \cdots + x^{k-1}. \tag{5.3}$$

Define a function χ on F_p^* as

$$\chi(x) = \zeta_k^\ell, \quad \text{where } \ell \equiv \log_\mu x \pmod{k}.$$

Extend χ to all of F_p by $\chi(0) = 0$. Then χ is a multiplicative character of F_p of order k.

Let $U \subseteq F_p$ be a subset of vertices of the graph H_0 with $|U| = m$. Denote by $J(U)$ for $\cap_{u \in U} N(u)$. If $|J(U)| < n$ for any such U, then $r_k(K_{m,n}) > p$ from Lemma 5.9. For a fixed U, define a function $f(x)$ on $x \in F_p$ as

$$f(x) = \prod_{u \in U} \prod_{j=1}^{k-1} (\chi(x-u) - \zeta_k^j) = \prod_{u \in U} \sum_{j=0}^{k-1} \chi^j(x-u),$$

where we use the identity (5.3). For $x \notin U$, if $x \notin J(U)$, then $f(x) = 0$ as $\chi(x-u) = \zeta_k^j$ for some j with $1 \leq j \leq k-1$. If $x \in J(U)$, then $f(x) = k^m$ as $\log_\mu(x-u) \equiv 0 \pmod{k}$ hence $\chi(x-u) = 1$. Therefore, we have

$$\sum_{x \notin U} f(x) = k^m |J(U)|.$$

Set $U = \{u_1, u_2, \ldots, u_m\}$. Note that χ is multiplicative thus

$$\begin{aligned} f(x) &= \prod_{t=1}^{m} \Big(1 + \chi(x-u_t) + \cdots + \chi^{k-1}(x-u_t)\Big) \\ &= \sum_{0\le j_1,\ldots,j_m\le k-1} \chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big) \\ &= 1 + \sum_{\substack{0\le j_1,\ldots,j_m\le k-1 \\ j_1+\ldots+j_m\ge 1}} \chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big). \end{aligned}$$

Applying the Weil's theorem for the the polynomial $(x-u_1)^{j_1}\cdots(x-u_m)^{j_m}$ with $j_1+\cdots+j_m \ge 1$, which is not the form $ch^k(x)$ with $c \in F_p$ and $h(x) \in F_p[x]$ as $j_1, \ldots, j_m < k$, from (5.1), we have

$$\Big|\sum_{x\in F_p} \chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big)\Big| \le (m-1)\sqrt{p}.$$

Hence we obtain that

$$\begin{aligned} \Big|p - \sum_{x\in F_p} f(x)\Big| &= \Big|\sum_{x\in F_p} \sum_{\substack{0\le j_1,\ldots,i_m\le k-1 \\ j_1+\ldots+j_m\ge 1}} \chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big)\Big| \\ &= \Big|\sum_{\substack{0\le j_1,\ldots,j_m\le k-1 \\ j_1+\ldots+j_m\ge 1}} \sum_{x\in F_p} \chi\Big((x-u_1)^{j_1}\cdots(x-u_m)^{j_m}\Big)\Big| \\ &\le \sum_{\substack{0\le j_1,\ldots,j_m\le k-1 \\ j_1+\ldots+j_m\ge 1}} (m-1)\sqrt{p}. \end{aligned}$$

It is well-known that the number of solutions of nonnegative integers $(j_1, \ldots, j_m)$ of the equation

$$j_1 + j_2 + \cdots + j_m = s$$

is $\binom{s+m-1}{s}$ for a fixed integer s. Omitting the constraint that $j_1, \ldots, j_m \le k-1$, we obtain that

$$\Big|p - \sum_{x\in F_p} f(x)\Big| \le \sum_{s=1}^{m(k-1)} \binom{s+m-1}{s}(m-1)\sqrt{p} = A\sqrt{p},$$

where $A = A(k,m)$ is independent of p. Note that $|f(x)| \le k^m$, thus $|\sum_{x\in U} f(x)| \le mk^m$ and

$$\begin{aligned}\Big|p - k^m|J(U)|\Big| &= \Big|p - \sum_{x\notin U} f(x)\Big| \le \Big|p - \sum_{x\in F_p} f(x)\Big| + \Big|\sum_{x\in U} f(x)\Big| \\ &\le A\sqrt{p} + mk^m \le (A+1)\sqrt{p}\end{aligned}$$

for large p, which implies that

$$k^m|J(U)| \le p + (A+1)\sqrt{p}.$$

It is known that there are asymptotically $N/(\phi(2k)\log N)$ primes p in the form $p \equiv 1 \pmod{2k}$ between 1 and N, where $\phi(2k)$ is the number of integers from 1 to $2k$ that are relatively prime to $2k$. Let $p \equiv 1 \pmod{2k}$ be a prime between $k^m n - n^{0.525}$ and $k^m n - n^{0.525}/2$. The existence of such prime for large n is ensured by results for estimating the difference between consecutive primes, see Baker, Harman and Pintz (2001). The constant 0.525 is in the process of improvement to $0.5 + o(1)$ implied by the famous Riemann hypothesis. By choosing such p, we have

$$|J(U)| \le n - \frac{n^{0.525}}{2} + (A+1)\sqrt{k^m n} < n,$$

for large n. Thus H_0 contains no $K_{m,n}$, implying

$$r_k(K_{m,n}) > p \ge k^m n - n^{0.525}$$

as each H_i is isomorphic to H_0. □

The largest difference between consecutive primes is conjectured as $p^{1/2+o(1)}$. If so, we have $r_k(K_{m,n}) \ge k^m n - n^{1/2+o(1)}$, which is the same as that in Chapter 5.

5.6 Some multi-color Ramsey numbers

For $H_1 = H_2 = \cdots = H_k = H$, let us write

$$r_{k+1}(H; H_{k+1}) = r(H_1, \ldots, H_k, H_{k+1})$$

Alon and Rödl (2005) gave sharp bounds for multi-color Ramsey numbers in form of $r_{k+1}(H; K_n)$ with $k \geq 2$, where H is a (some kind) bipartite graph or K_3. Their main idea is to estimate the number of independent sets of given size in a quasi-random graph G, which contains no H, and to consider random shifts of G. The number of shifts is k. The bigger k, the tighter is the bound. When $k = 1$, which means no shift actually, the method is ineffective (for bounding $r(H, K_n)$). It is interesting that G is Turán's graph when H is bipartite. Recall that the graphs constructed in Chapter 9 are regular, which contains loops but each vertex has at most one loop. We call such graphs to be semi-simple.

Theorem 5.15 *Let $G = (V, E)$ be a semi-simple (N, d, λ)-graph, and let $n_0 = \frac{2N \log N}{d}$. Then for any $n \geq n_0$, the number M of independent sets of size n in G satisfies that*

$$M \leq \left(\frac{edn}{2\lambda n_0}\right)^{n_0} \left(\frac{2e\lambda N}{dn}\right)^{n}.$$

Proof. Consider the number of ways to choose an ordered set $v_1, v_2, \ldots, v_n$ of n vertices which form an independent set. Starting with $B_0 = V$, we choose v_1 arbitrarily. Define

$$B_i = V \setminus \cup_{j=1}^{i} N[v_j].$$

Then B_i is the set of vertices by deleting $\{v_1, v_2, \ldots, v_i\}$ and their neighbors, where $v_1, \ldots, v_i$ have been chosen. Clearly $\{B_i\}$ is decreasing in the sense $B_i \supseteq B_{i+1}$, and v_{i+1} has to lie in B_i. Define

$$\overline{B}_i = \left\{u \in V : \ |N_{B_i}(u)| \leq \frac{d}{2N}|B_i|\right\},$$

where $N_{B_i}(u) = N(u) \cap B_i$. If the next chosen vertex v_{i+1} from B_i is not in $\overline{B}_i$, then B_{i+1} is obtained by deleting v_{i+1} and at least $\frac{d}{2N}|B_i|$ vertices from B_i and so

$$|B_{i+1}| < \big(1 - \frac{d}{2N}\big)|B_i|.$$

Hence throughout the process there cannot be more than n_0 choices like that, since otherwise the corresponding set of non-neighbors will be empty before the process terminates from

$$\left(1-\frac{d}{2N}\right)^{n_0} = \left(1-\frac{d}{2N}\right)^{2N\log N/d} < \frac{1}{N}.$$

It follows that with at most n_0 possible exceptions, each vertex v_{i+1} has to lie in $B_i \cap \overline{B_i}$. By Corollary 5.3, we have

$$|B_i \cap \overline{B_i}| \leq \frac{2\lambda N}{d}.$$

Therefore, the total number of choices for the ordered set $v_1, v_2, \ldots, v_n$ is at most

$$\binom{n}{n_0} N^{n_0} \left(\frac{2\lambda N}{d}\right)^{n-n_0} \leq \left(\frac{edn}{2\lambda n_0}\right)^{n_0} \left(\frac{2\lambda N}{d}\right)^{n}.$$

Indeed, there are $\binom{n}{n_0}$ possibilities to choose a set of indices covering all indices i for which the vertex v_{i+1} has not been chosen in $B_i \cap \overline{B_i}$. Then there are at most N ways to choose each such vertex v_i, and at most $\frac{2\lambda N}{d}$ ways to choose each vertex v_{j+1} for each other index j.

Now, dividing the above bound by $n!$, we obtain an upper bound for the number of unordered independent sets of size n as claimed. □

Theorem 5.16 *Let G be a graph of order N that contains no H, and let M be the number of independent sets of size n in G. If*

$$M^k < \binom{N}{n}^{k-1},$$

then $r_{k+1}(H; K_n) > N$.

Proof. For each i, $1 \leq i \leq k$, let G_i be a random copy of G on the same vertex set V, that is, a graph obtained from G by mapping its vertices to those of V according to a random one to one mapping. The probability that a fixed set of n vertices of V will be an independent set in each G_i is

$$\left(\frac{M}{\binom{N}{n}}\right)^k < \frac{1}{\binom{N}{n}},$$

implying that with a positive probability there is no such independent set.

Color each edge of K_N on $V(G)$ by minimum i if the edge belongs to G_i. Otherwise, color the edge by $k+1$. Then, there is no monochromatic H in first k colors and no K_n in the last color $k+1$, so $r_{k+1}(H, K_n) > N$. □

Let us recall an upper bound in Chapter 8 that for any fixed integers $s \geq t \geq 2$,

$$r_{k+1}(K_{t,s}; K_n) \leq c\left(\frac{n}{\log n}\right)^t,$$

where $c = c(k, t, s) > 0$ is a constant.

Theorem 5.17 *The Ramsey number $r_{k+1}(C_4; K_n)$ satisfies the following:*

(1) For any fixed $k \geq 3$, $r_{k+1}(C_4; K_n) = \Theta\Big(\frac{n}{\log n}\Big)^2$.

(2) There are positive constants c_1 and c_2 such that

$$c_1\Big(\frac{n \log\log n}{(\log n)^2}\Big)^2 \leq r(C_4, C_4, K_n) \leq c_2\Big(\frac{n}{\log n}\Big)^2.$$

Proof. It suffices to prove the lower bounds. Consider the Erdős-Rényi graph E_q^o of order $N = q^2 + q + 1$ and let M be the number of independent sets of size n. By Lemma 5.16, we shall show that $M^k < \binom{N}{n}^{k-1}$. The graph E_q^o is d-regular, where $d = q + 1$. Let $n_0 = \frac{2N \log N}{d}$. Then for large q,

$$4q \log q < n_0 < 4(q+1) \log q.$$

From Theorem 5.15, if $n \geq n_0$, which can be guaranteed by taking $c > 4$, we have

$$M \leq \left(\frac{edn}{2\lambda n_0}\right)^{n_0} \left(\frac{2e\lambda N}{dn}\right)^n,$$

where $\lambda = \sqrt{q}$.

(1) For $k \geq 3$, it suffices to show that $r(C_4, C_4, C_4, K_n) = \Theta\Big(\frac{n}{\log n}\Big)^2$ as $r_{k+1}(C_4; K_n) \geq r(C_4, C_4, C_4, K_n)$. Set $n = cq \log q$, where c is a large

constant to be chosen. We shall show that

$$M^{3/n} < \binom{N}{n}^{2/n}. \tag{5.4}$$

Substituting d, λ, n_0, n, N by values in terms of q, we have

$$M^{3/n} \le \Big(\frac{ce(q+1)}{8\sqrt{q}}\Big)^{\frac{12}{c}(1+1/q)}\Big(\frac{2e\sqrt{q}}{c\log q}\Big)^3 \sim c_1\frac{q^{12/c+3/2}}{(\log q)^3},$$

where c_1 is a positive constant, and

$$\binom{N}{n}^{2/n} \sim \Big(\frac{eN}{n}\Big)^2 \sim \Big(\frac{eq}{c\log q}\Big)^2.$$

Thus the inequality (5.4) holds if $12/c+3/2 \le 2$, which is $c \ge 24$. Then we have $n > n_0$ and $N \sim q^2 \sim n^2/(c\log n)^2$ as $q \to \infty$, completing the proof for $k \ge 4$.

(2) For $r(C_4, C_4, K_n)$, set $n = cq\log^2 q/\log\log q$, where c is a positive constant to be chosen. We shall show that $M^{2/n} < \binom{N}{n}^{1/n}$. Note that for some constant $c_i > 0$,

$$\begin{aligned} M^{2/n} &\le c_1\Big(\frac{\sqrt{q}\log q}{\log\log q}\Big)^{\frac{8\log\log q}{c\log q}}\Big(\frac{\sqrt{q}\log\log q}{\log^2 q}\Big)^2 \\ &\le c_2 q^{1+\frac{4\log\log q}{c\log q}}\Big(\frac{\log\log q}{\log^2 q}\Big)^2 = c_2\frac{q(\log\log q)^2}{(\log q)^{4-4/c}}, \end{aligned}$$

and

$$\binom{N}{n}^{1/n} \ge c_3\frac{eN}{n} \ge c_4\frac{q\log\log q}{\log q}.$$

We are done by taking $c > 4/3$ so that $4 - 4/c > 1$. □

Note that we have found the spectrum of the projective norm graph $G_{q,t}$ in the last section and it contains no $K_{t,s}$ for $s \ge (t-1)! + 1$. Similar argument gives the following result.

Theorem 5.18 *For any fixed $t \ge 2$ and $s \ge (t-1)! + 1$, the Ramsey number $r_{k+1}(K_{t,s}; K_n)$ satisfies the following:*

(1) For any $k \geq 3$,

$$r_{k+1}(K_{t,s}; K_n) = \Theta\Big(\frac{n}{\log n}\Big)^t.$$

(2) There are positive constants c_1 and c_2 such that

$$c_1\Big(\frac{n \log\log n}{(\log n)^2}\Big)^t \leq r(K_{t,s}, K_{t,s}, K_n) \leq c_2\Big(\frac{n}{\log n}\Big)^t.$$

Alon and Rödl (2005) also solved a conjecture of Erdő and Sós conjectured that

$$\lim_n \frac{r(K_3, K_3, K_n)}{r(K_3, K_n)} = \infty.$$

Recall that in Chapter 2 we define how to "blow up" G with H, where each vertex v of G is replaced by a copy of H, denoted by H_v, in which any pair of vertices from distinct H_u and H_v are adjacent if and only if u and v are adjacent. If H is an r-independent set, we call it an r-blow up of G.

Lemma 5.10 *There is some constant $c = c_k > 0$ such that*

$$r_{k+1}(K_3; K_n) \geq \frac{c\, n^{k+1}}{(\log n)^{2k}}$$

for all large n.

Proof. Let $N = c_1 s^2/\log s < r(K_3, K_{s+1})$, where $c_1 > 0$ is a fixed constant. Then there is a graph F of order N with no K_3 and $\alpha(F) \leq s$. Let G be the r-blow up of F and M the number of independent sets of size n in G, where $r = r(s)$ will be chosen. Then

$$M \leq \frac{\binom{N}{s}(rs)^n}{n!} < \Big(\frac{eN}{s}\Big)^s\Big(\frac{ers}{n}\Big)^n.$$

The first inequality follows from the facts that each independent set A of size n can be partitioned into at most s blocks (blown vertices); there are at most $\binom{N}{s}$ ways to choose these blocks, and each vertex of A is one of the rs vertices in a block.

Since G has rN vertices and it contains no K_3, by Theorem 5.4, we have $r_{k+1}(K_3; K_n) > rN$ if

$$M^k < \binom{rN}{n}^{k-1}.$$

We now take $r = s^{k-1}(\log s)^{2-k}$ and $n = cs\log s$, where $c > 0$ is a constant to be chosen. Then

$$M^{k/n} < \Big(\frac{c_1 es}{\log s}\Big)^{k/(c\log s)}\Big(\frac{es^{k-1}}{c(\log s)^{k-1}}\Big)^k \le \frac{c_2}{c^k}\Big(\frac{s}{\log s}\Big)^{k(k-1)},$$

where c_2 and henceforth c_3 and c_4 are positive constants that is independent of c, and

$$\binom{rN}{n}^{(k-1)/n} > c_3\Big(\frac{erN}{n}\Big)^{k-1} \ge \frac{c_4}{c^{k-1}}\Big(\frac{s}{\log s}\Big)^{k(k-1)}.$$

Thus the condition is satisfied if we take c large such that $c_2/c^k < c_4/c^{k-1}$, and hence

$$r_{k+1}(K_3; K_n) > rN = \frac{c_1 s^{k+1}}{(\log s)^{k-1}} = \Theta\Big(\frac{n^{k+1}}{(\log n)^{2k}}\Big),$$

completing the proof. □

Theorem 5.19 *There are constants $c_i = c_i(k) > 0$ such that*

$$\frac{c_1\, n^{k+1}}{(\log n)^{2k}} \le r_{k+1}(K_3; K_n) \le \frac{c_2\, n^{k+1}}{(\log n)^k}$$

for all large n.

Proof. It remains to show the upper bound, which holds for $k = 0$ and $k = 1$. We next prove the result by induction on $k \ge 2$. Assuming that the result holds on $k-1$, we prove it for k. Let $N = r_{k+1}(K_3; K_n) - 1$. There is an edge-coloring of K_N by colors $1, 2, \ldots, k+1$ with no monochromatic K_3 in any of the first k colors, and no monochromatic

K_n in the color $k+1$. Consider the graph T consisting of all edges of the first k colors. Then $D = \Delta(T)$ satisfies that

$$D \le k(r_k(K_3; K_n) - 1) < kr_k(K_3, K_n).$$

The neighborhood $N(v)$ of v in T is $\cup_{i=1}^{k} N_i(v)$, where $N_i(v)$ is the set of neighbors of v that are connected to v by an edge in the color i, $1 \le i \le k$. For a vertex u in $N(v)$, we consider the neighbor of u in the subgraph induced by $N(v)$ in T. Suppose $u \in N_1(v)$, say. Such neighbors are these in

$$N(u) \cap N(v) = \cup_{j=1}^{k} \Big(\cup_{i=1}^{k} \Big(N_i(u) \cap N_j(v) \Big) \Big).$$

First of all, $N_1(u) \cap N_1(v) = \emptyset$ since there is no monochromatic triangle in the color 1. For $2 \le i \le k$, $N_i(u) \cap N_1(v)$ contains no edge in the colors i and 1, thus $|N_i(u) \cap N_1(v)| \le r_{k-1}(K_3, K_n) - 1$. Hence

$$\Big| \cup_{i=1}^{k} \Big(N_i(u) \cap N_1(v) \Big) \Big| < (k-1) r_{k-1}(K_3, K_n).$$

Similarly, for $2 \le j \le k$,

$$\Big| \cup_{i=1}^{k} \Big(N_i(u) \cap N_j(v) \Big) \Big| < (k-1) r_{k-1}(K_3, K_n).$$

Thus the maximum degree of subgraph induced by $N(v)$ in T is at less than $m = k^2 r_{k-1}(K_3, K_n)$. By the main result in Chapter 3, we have

$$n \ge \alpha(T) \ge N \frac{\log(D/m) - 1}{D}.$$

Then, using the induction hypothesis for D and Lemma 5.10 for m, we have the desired upper bound. □

References

N. Alon and V. Rödl, Sharp bounds for some multicolor Ramsey numbers, *Combinatoria*, **25** (2005), 125-141.

N. Alon and J. Spencer, *The Probabilistic Method, 3rd ed.*, Wiley-Interscience, New York, 2008.

R. Baker, G. Harman and J. Pintz, The difference between consecutive primes, II, *Proc. Lond. Math. Soc.*, **83** (2001), 532-562.

F. R. Chung, R. Graham, Sparse quasi-random graphs, *Combinatoria*, **22** (2002), 217-244.

F. R. Chung, R. Graham and R. Wilson, Quasi-random graphs, *Combinatoria*, **9** (1989), 345-362.

M. Krivelevich and B. Sudakov, Pseudo-random graphs, *Bolyai Soc. Math. Stud.*, **15** (2006), 199-262.

J. Seidel, A survey of two-graphs, in: Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), vol I, Atti dei Convegni Lincei, No. 17, Accad. Naz. Lincei, Rome, 1976, 481C 511.

T. Szabó, On the spectrum of projective norm-graphs, *Inform. Process. Lett.*, **86** (2) (2003), 71-74.

A. Thomason, Pseudo-random graphs, in: *Proceedings of Random Graphs, Poznań 1985, M. Karoński, Eds.*, *Ann. Discrete Math.* **33** (1987), 307-331.

Chapter 6

Real-world Networks

Complex systems from various fields, such as physics, biology, or sociology, can be systematically analyzed using their network representation. A network (also known as a graph) is composed of vertices (or nodes) and edges, where vertices represent the constitutes in the system and edges represent the relationships between these constitutes.

We shall introduce some basic concepts for real-world networks in this chapter.

Note that in some papers on graphs produced by a random process, "typical" graphs (instead of random graphs) are chosen to present the graphs in the random graph space.

6.1 Data and empirical research

Big data is a term for large or complex data. The term often refers simply to that traditional data processing applications are inadequate, and seldom to a particular size of data set. Most big data come from real-world networks, and analysis of such data sets can find new correlations to spot business trends, prevent diseases, combat crime and so on.

Empirical research is research using empirical evidence, where the empirical evidence is the record of direct observations or experiences in form of data and big data. Through quantifying the data, a researcher can answer empirical questions from real-world. In particular,

we are interested in the empirical research and the data from real-world networks.

It is usual that researchers aim the common case, so they often describe the behavior by ignoring some case that are not significant for the research. This is similar to the case in random graph, we describe an event by saying "almost all" to signify that the probability of event goes to 1. It is often that we are more concerned with the average of parameters since they are concentrated at the average in most cases. Average of some parameters may be more meaningful than the extremal value of them.

However, this is not always the case. When investigating a social networks, the nodes of large degrees, called "hubs" such as internet celebrities, attracted much attention as these nodes are very important for the structure of the networks.

Collecting data that needed is an challenge before analysis. For example, the data in Barabási and Albert (2009) came from a software designed to collect the links in World Wide Web pointing from one page to another. The data in Backstrom and Kleinberg (2014) and Ugander, Backstrom, Marlow and Kleinberg(2012) came from Facebook Inc. directly as several of co-authors are employees of the company.

6.2 Six degrees of separation

Six degrees of separation is the theory that any pair of persons is six or fewer steps away in the world as a network connected by friendship, which means the maximum distance of nodes in the network is at most six in language of graph theory. However, as claimed before, "any pair" for social networks for sociology may mean most pairs.

The term *small world* became famous since a paper of S. Milgram (1967) who was an American psychologist. Some seminal works have been conducted before Milgram took up the experiments reported as the small world problem, and the experiment is called "the small-world experiment", in which Milgram and other researchers examined the average path length for social networks of people in the United States. The research suggested that human society is a small-world-type network, and the experiments are often associated with the phrase "six

degrees of separation", although Milgram did not use this term himself.

Milgram's experiment developed out of a desire to learn more about the probability that two randomly selected people would know each other. This is one way of looking at the small world problem.

Though the experiment went through several variations, Milgram typically chose individuals in cities of Omaha, Nebraska, and Wichita, Kansas, to be the starting points and Boston, Massachusetts, to be the end point of a chain of correspondence. These cities were selected because they were thought to represent a great distance in US, both socially and geographically.

Information packets (a letter,a roster and postcards) were initially sent to randomly selected individuals in Omaha or Wichita. In the more likely case that the person did not personally know the target, then the person was to think of a friend or relative he knew personally who was more likely to know the target. He was then directed to sign his name on a roster in the information packet and forward the packet to that person. When and if the package eventually reached the contact person in Boston, the researchers could examine the roster to count the number of times it had been forwarded from person to person.

However, a significant problem was that often people refused to pass the letter forward, and thus the chain never reached its destination. In one case, only 64 of the 296 letters eventually did reach the target contact. Among these chains, the average path length fell around five and a half or six. Hence, the researchers concluded that people in US are separated by about six people on average.

Smaller communities, such as mathematicians and actors, have been found to be densely connected by chains of personal or professional associations. Mathematicians have created the Erdős number to describe their distance from Paul Erdős based on shared publications. A similar exercise has been carried out for the actor Kevin Bacon and other actors who appeared in movies together with him.

In 2001, D. Watts attempted to recreate Milgram's experiment on the Internet, using an e-mail message as the "package" that needed to be delivered, with $48,000$ senders and 19 targets (in 157 countries). Watts found that the average number of intermediaries was around six, reported in Watts (1998). Today, the phrase "six degrees of sep-

aration" is often used as a synonym for the idea of the "small world" phenomenon.

Watts and Strogatz (1998) showed that the average path length between two nodes in a random network is equal to $\log N / \log K$, where N is number of nodes and K is degree of acquaintances per node. Thus, assuming 10% of population of US is too young to participate and $N = 300,000,000$ (90% of the US population) and $K = 30$, the Degrees of Separation 5.7. If $N = 6,000,000,000$ (90% of the World population) and $K = 30$, then Degrees of Separation 6.6.

However, the convenient way of communication in a social network will make the average distance smaller and smaller. Facebook's data team released data in online papers described that amongst all Facebook users at the time of research, the average distances of friendship links were 5.28 in 2008, 4.74 in 2011 and 3.57 in February 2016 (this year). The world changes from "six degrees of separation" to "four degrees of separation".

6.3 Clustering coefficient

An important measure of network topology, called *clustering coefficient*, assesses the triangular pattern as well as the connectivity in a vertex's neighborhood: a vertex has a high clustering coefficient if its neighbors tend to be directly connected with each other. The clustering coefficient c_v of a vertex v can be calculated as

$$c_v = \begin{cases} 0, & \text{if } d_v = 0 \\ \frac{e_v}{\binom{d_v}{2}}, & \text{if } d_v \geq 2. \end{cases}$$

For $d_v = 1$, it is a convention to define $c_v \in [0, 1]$ depending on the situation. Thus $0 \leq c_v \leq 1$. The clustering coefficient c_v for $d_v \geq 2$ is the ratio of number of triangles and all possible triangles that share vertex v.

Let G^k be a graph obtained from G by adding new edges connecting vertices of distance at most k in G. It is to see if $n \geq 8$, then $c_v = 1/2$ for each v in circular lattice C_n^2.

For a graph G of order N (i.e., G contains N vertices) and minimum degree $\delta(G) \geq 2$, its *average clustering coefficient* is defined as

$$\overline{c}(G) = \frac{1}{N}\sum_{v\in V} c_v = \frac{2}{N}\sum_{v\in V}\frac{e_v}{d_v(d_v-1)}.$$

Average clustering coefficient explains the clustering (triangulation) within a network by averaging the clustering coefficients of all its nodes. The idea of clustering coefficient is proposed (especially in the analysis of social networks) to measure the local connectivity or "cliqueness" of a social network. If a network has a high average clustering coefficient and a small average distance, it is often called a "small-world" network.

Let us label the vertices of G of order N as $v_1, v_2, \ldots, v_N$. Recall that $A = (a_{ij})_{N\times N}$ is the adjacency matrix of G, where

$$a_{ij} = \begin{cases} 1, & \text{if } v_iv_j \in E, \\ 0, & \text{otherwise.} \end{cases}$$

We also call the eigenvalues of A as eigenvalues of G. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ be eigenvalues of G in a non-increasing order. Set

$$\lambda = \lambda(G) = \max\{|\lambda_i| :\ 2 \leq i \leq N\}.$$

As called by Alon, a graph G is an (N, d, λ)-graph if G is d-regular with N vertices and $\lambda = \lambda(G)$. Note that a d-regular connected graph satisfies that $\lambda_1 = d$. For an (N, d, λ)-graph, the spectral gap between d and λ is a measure for its quasi-random property. The smaller the value of λ compared to d, the closer is the edge distribution to the ideal uniform distribution (i.e., it becomes a random graph). We may say, not precisely, that an (N, d, λ)-graph with $\lambda = O(\sqrt{d})$ has good quasi-randomness. Generally, this is a weak condition as most random graphs are such graphs.

Theorem 6.1 *Let G be an (N, d, λ)-graph that is connected. If $\lambda = O(\sqrt{d})$ as $d \to \infty$, then*

$$\overline{c}(G) \sim \frac{d}{N}.$$

Proof. Let A be adjacency matrix of G. Note that A is symmetric, and thus it is diagonalizable. Let $\lambda_1, \lambda_2, \ldots, \lambda_N$ be the eigenvalues of A. Then the eigenvalues of A^k are $\lambda_1^k, \lambda_2^k, \ldots, \lambda_N^k$. Note that the (i, j) element of A^k is the number of walks from vertex v_i to vertex v_j, and a closed walk of length 3 is exactly a triangle. Thus the ith diagonal element of A^3 is $2e_{v_i}$, and

$$\overline{c}(G) = \frac{2}{N} \sum_v \frac{e_v}{d_v(d_v - 1)} = \frac{1}{Nd(d-1)} \sum_{i=1}^{N} \lambda_i^3 = \frac{1}{Nd(d-1)} \left(d^3 + \sum_{i=2}^{N} \lambda_i^3 \right),$$

where we used the fact that $\lambda_1 = d$ as G is d-regular and connected. The assumption $\lambda = O(\sqrt{d})$ implies that

$$\frac{\left| \sum_{i=2}^{N} \lambda_i^3 \right|}{Nd(d-1)} \leq \frac{N\lambda^3}{Nd(d-1)} = \frac{O(d^{3/2})}{d^2} \to 0.$$

Thus

$$\overline{c}(G) \sim \frac{d^2}{N(d-1)} \sim \frac{d}{N}$$

for large d. □

6.4 Small-world networks

The small-world phenomenon is typical for random graphs that have small maximum distances. A definition for small-world network describes it as a network, in which the typical distance L between two randomly chosen nodes grows proportionally to $\log N$, where N is the numbers of nodes in the network. Namely,

$$E(L) = \Theta(\log N),$$

which grows slowly as $N \to \infty$ and the average distance of nodes is small.

A certain category of small-world networks were identified as a class of random graphs by D. Watts and S. Strogatz in (1998). They measured that in fact many real-world networks have a small average distance, but also a clustering coefficient significantly higher than expected

by random chance. They noted that graphs could be classified according to two independent structural features, namely the clustering coefficient, and average distance. Purely random graphs, built according to the Erdős-Rényi (ER) model, exhibit a small average distance (typically as $\Theta(\log N)$) along with a small clustering coefficient (typically d/N where d is the expected value of degrees).

Many biological,technological and social networks lie between completely regular and completely random. Typically, these networks have many vertices that are sparse in sense that the average degrees are much less than the number of vertices.

Watts and Strogatz modelled the small-world networks by starting at a graph C_n^k with

$$n \gg k \gg \log n \gg 1,$$

where $k \gg \log n$ guarantees that a random graph will be connected. Then, they choose vertices in order and edges adjacent to the chosen vertices and reconnect these edges to vertices chosen uniformly at random. In the process, the average clustering coefficient decreases slowly and the average distance decreases rapidly and thus they obtained a network between regular ring lattice C_n^k and completely random network. The obtained network has a large average clustering coefficient and a small average distance, which is called small-world network.

6.5 Power law and scale-free networks

Let X be a discrete random variables taking positive integers. If

$$\Pr(X = k) = \frac{c}{k^\gamma},$$

where $c, \gamma > 0$ are constants, then X is said to have a *power law* distribution. This distribution is also call Pareto distribution as economist Pareto originally used it to describe the allocation of wealth that a larger portion of the wealth of society is owned by a smaller percentage of the people (so called 80-20 rule). Contrast to exponential distribution that decreases rapidly, power law is also called heavy-tailed distribution.

Let $P(k)$ be the fraction of nodes in the network G that have degree k, namely

$$P(k) = \frac{|\{v : \deg(v) = k\}|}{N},$$

where N is number of nodes in G. If $P(k)$ is equal (or close) to power law, then the network G is said to be *scale-free*. For such networks, it is typical $2 < \gamma \leq 3$.

The networks of citations between scientific papers are interesting. In 1965, D. Price (1965) found that vertices of degree k in such networks had a heavy-tailed distribution following a power law. He did not however use the term "scale-free network", which was not coined until some decades later. In 1976, Price also proposed a mechanism to explain the occurrence of power laws in citation networks, which he called *cumulative advantage* but which is today more commonly known under the name *preferential attachment*.

In 1999, A. Barabási and colleagues coined the term *scale-free network* when they found some nodes had a much bigger degrees than the that "expected" in random network, and they were surprised and used the term "scale-free", which now is used to describe the class of networks that exhibit a power-law degree distribution.

Albert L. Barabási is a physicist, best known for his work in the research of network theory, and Réka Albert, the co-author of paper (2009), is a professor of physics and biology.

In their earlier study (1999), Albert, Jeong and Barabási found that the World Wide Web is not a random network, but the number of links per node, often called the *degree distribution*, follows a power law. Subsequently, researchers found that not only the WWW, but many other networks, follow the same distribution. These different datasets together indicated that we are dealing with a potentially universal behavior, which might have a common explanation.

Barabási and Albert (2009) proposed a generative mechanism to explain the appearance of power-law distributions, which they called "preferential attachment" and which is essentially the same as that proposed by Price. Analytic solutions for this mechanism (also similar to the solution of Price) were presented earlier by Dorogovtsev, Mendes and Samukhin (2002). Finally, it was rigorously proved by mathematicians Bollobás, Riordan, Spencer and Tusnády (2001).

To explain this phenomenon, Barabási and Albert (2009) suggested the following random graph process as a model, called BA model.

Consider a random graph process in which vertices are added to the graph one at a time and joined to a fixed number of earlier vertices, selected with probabilities proportional to their degrees. Let $v_1, v_2, \ldots$ be a sequence of vertices. Assume that $m_0 \geq 2$ is the number of vertices to start at the process, and let $d(v_i)$ be the degree for the early vertex v_i in the existing graph.

They described the process to start with a small number m_0 of vertices, at every time step we add a new vertex with $m \leq m_0$ edges that link the new vertex to m different vertices already present in the system. If the new vertex is v_{t+1}, then the probability that v_{t+1} is adjacent to v_i is proportional to

$$\frac{d(v_i)}{\sum_{j=1}^{t} d(v_j)}.$$

The above probability signifies the new vertex to incorporate preferential attachment. Note that, to be clear to start, there should exist at least one edge in the first m_0 vertices. A question is if we connect each early vertex and new vertex randomly by above probability, then the excepted number of new edges is one.

The research in Barabási and Albert (2009) is empirical, and the proof is heuristic. The process defined in Bollobás et. al. (2001) preserves the idea of preferential attachment, and the description is much more complex, and power law has been shown for degrees at most $N^{1/15}$ with $\gamma = 3$.

On a theoretical level, some other abstract definitions of scale-free have been proposed. For example, Li et. al. (2005) offered a potentially more precise "scale-free metric". Let $G = (V, E)$ be a simple graph and $s(G) = \sum_{uv \in E} d(u)d(v)$ and $S(G) = s(G)/s_{\max}$, where $s_{\max}$ as the maximum value of $s(H)$ among simple graphs H on same vertex set V with degree distribution identical to G. The notation $S(G)$ gives a metric between 0 and 1, where a G with small $S(G)$ is "scale-rich", and G with $S(G)$ close to 1 is "scale-free". Note that $s(G)$ is maximized when high-degree nodes are connected to other high-degree nodes and $S(G)$ captures the notion of self-similarity implied in the name "scale-free".

Some properties are often listed as the characteristics of scale-free networks, which are as follows.

- Power-law degree distribution;
- Generated by certain random processes with preferential attachment;
- Highly connected hubs that hold the network together with the "robust yet fragile" feature of error tolerance, which is robust when attacked by removing some nodes randomly and fragile by removing some hubs deliberately;
- Generic in the sense of being preserved under random degree-preserving rewiring;
- Self-similar;
- Universal in the sense of not depending on domain-specific details.

6.6 Network Structure

As pointed by Newman (2003), the research on networks may provide new insight into the study of complex systems. Networks have many notable properties, such as the small-world property, the scale-free property, the community structure property, and the links between two objects usually display diversity.

By collecting data from mobile phones, Fagle, Pentland and Lazer (2009) found that the data have the potential to provide insight into the relational dynamics of individuals, and allow the prediction of individual-level outcomes such as job satisfaction.

The concept of contagion has expanded from its original grounding in epidemic disease to describe many processes that spread across networks such as fads, political opinions, the adoption of new technologies, and financial decisions, see, e.g. R. Pastor-Satorras and A. Vespignani (2001) and M. Newman, D. Watts and S. Strogatz (2002).

In traditional models of social contagion, the probability that an individual is affected by the contagion grows by monotonically with the

size of neighborhood. By analyzing the growth of Facebook, Ugander, Backstrom, Marlow and Kleinberg (2012) find that the probability of contagion is tightly controlled by the number of connected components in an individual neighborhood, rather than by the actual size of the neighborhood.

A crucial task in the analysis of on-line social-networking systems is to identify important people liked by strong social ties. Drawing data from e-mail, Kossinets and Watts has developed a method of analyzing and estimating tie strength in on-line domains (2006), in which the key structure is *embeddedness*–the number $|N(u) \cap N(v)|$ of mutual friends of two people u and v, a quantity that typically increases with tie strength.

The embeddedness is not necessarily to be the most appropriate for characterizing particular types of strength ties. Backstrom and Kleinberg (2014) proposed a networks-based characterization for intimate relationships, those involving spouses or romantic partners. Using data from a large sample of Facebook users, they try to recognize these people with high accuracy. They found that embeddedness is in fact a comparatively weak means of characterizing romantic relations, and that an alternate network measure that they term *dispersion* is significantly more effective. Roughly, a link between two people has high dispersion when their mutual friends are not well connected to one another. Their research has an important contingent nature: given that a user has declared a relationship partner, they want to understand how effectively they can find partner.

Note that the links to a person's relationship partner or other closest friends may have lower embeddedness, but they often involve mutual neighbors from several foci, reflecting the fact that the social orbits of these close friends are not bounded within any one focus–consider, for example, a husband who knows several of his wife's co-workers, family members, and former classmates, even though these people being to different foci and do not know each other. Thus, Backstrom and Kleinberg proposed some definition as follows.

For a network $G = (V, E)$ and a pair of nodes u and v, denote by $C_{uv} = N(u) \cap N(v)$, the set of mutual friends of u and v, and $c_{uv} = |C_{uv}|$. Let $d(s, t, G)$ be the graph-theoretic distance between u and v in G. For

distinct s and t in C_{uv}, define

$$d_{uv}(s,t) = \begin{cases} 1, & d(s,t,C_{uv}) \geq 3, \\ 0, & d(s,t,C_{uv} \leq 2. \end{cases}$$

Then, define the *absolute dispersion* of u and v as

$$disp(u,v) = \sum_{s,t \in C_{uv},\, s \neq t} d_{uv}(s,t).$$

Note that $disp(u,v)$ depends on both of C_{uv} and $d(s,t,C_{uv})$, and define

$$norm(u,v) = \frac{disp(u,v)}{c_{uv}},$$

which is called *normalized dispersion*. Predicting u's partner to be the individual v with

$$norm(u,v) = \max\{norm(u,x) : x \in V\}$$

gives the correct answer in 48.0% of all instances.

There are two ways to strengthen normalized dispersion that lead to increased performance. The first is to rank pair of u and v by a function of the form

$$\frac{(disp(u,v)+b)^{\alpha}}{(c_{uv}+c)}.$$

Searching over choices α, b and c leads to maximum performance of 50.5% at

$$\alpha = 0.61, \quad b = 0, \quad c = 5.$$

The second way is by applying the idea of dispersion recursively. For a fixed node u, define first $x_v =$ for all neighbors v of u. Then, iteratively update each x_v to be

$$\frac{\sum_{w \in C_{uv}} x_w^2 + 2\sum_{s,t \in C_{uv}} d_{uv}(s,t) x_s x_t}{c_{uv}} \to x_v.$$

Note that after the first iteration, $x_v = 1 + 2 \cdot norm(u,v)$, and hence ranking nodes by x_v after the first iteration is equivalently to ranking nodes by $norm(u,v)$. Backstrom and Kleinberg found that the highest performance ranking nodes by values of x_v after the third iteration, call such x_v as *recursive dispersion*. The performance by embeddedness and recursive dispersion for romantic relationships is 24.7% and 50.6%, respectively; and that for (married) spouses is 32.1% and 60.7%, respectively.

6.7 References

R. Albert, H. Jeong and A.L. Barabási, Internet-diameter of the World-Wide Web, *Nature*, 401 (6749) (1999),130-131.

L. Backstrom and J. Kleinberg, Romantic partnerships and the dispersion of socialties: A network analysis of relation status on Facebook, Proc. 17th ACM conference on computer supported cooperative work and social computing, 2014.

A. Barabási and R. Albert, Emergence of scaling in random networks, *Science*, 286 (5439) (1999), 509-512.

B. Bollobás, O. Riordan, J. Spencer and G. Tusnády, The degree sequence of a scale-free random graph process, *Random Struct. Algor.*, 18 (2001), 279-290.

S. Dorogovtsev, J. Mendes, Evolution of networks, *Advances in Physics*, 51 (4) (2002),1079.

N. Fagle, A. Pentland and D. Lazer, Infering friendship network structure by using mobile phone data, *Proc. Natl. Acad. Sci. USA*, 106 (36) (2009), 15274-15278.

G. Kossinets and D. Watts, Empirical analysis of an evolving social network, *Science*, 311 (2006), 88-90.

L. Li, D. Alderson, J. Doyle and W. Willinger, Towards a theory of Scale-free graphs: Definitions, properties and implications, *Internet Math.*, 2 (4) (2005), 431-523.

S. Milgram, The small world problems, *Psychology Today*, 2 (1967), 60-67.

M. Newman, The structure and function of complete networks, *SIAM Review*, 45 (2003), 167-256.

M. Newman, D. Watts and S. Strogatz, Random graph model for social networks, *Proc. Natl. Acad. Sci. USA*, 99 (Suppl 1) (2002), 2566-2572.

R. Pastor-Satorras and A. Vespignani, Epidemic spreading in scale-free networks, *Phys. Rew. Lett.*, 86 (2001), 3200-3203.

D. Price, Networks of scientific papers, *Science*, 149 (3683) (1965), 510-515.

J. Ugander, L. Backstrom, C. Marlow and J. Kleinberg, Structural diversity in social contagion, *Proc. Natl. Acad. Sci. USA*, 109 (16) (2012), 5962-5966.

D. Watts and S. Strogatz, Collective dynamics of 'small-world' networks, *Nature*, 393 (6684) (1998), 440-442.