Extremal and Probabilistic Combinatorics

Publicações Matemáticas

Extremal and Probabilistic Combinatorics

Robert Morris
IMPA

Roberto Imbuzeiro Oliveira
IMPA

impa

28º Colóquio Brasileiro de Matemática

Impresso no Brasil / Printed in Brazil

Capa: Noni Geiger / Sérgio R. Vaz

28º Colóquio Brasileiro de Matemática

- Cadeias de Markov e Teoria do Potencial - Johel Beltrán
- Cálculo e Estimação de Invariantes Geométricos: Uma Introdução às Geometrias Euclidiana e Afim - M. Andrade e T. Lewiner
- De Newton a Boltzmann: o Teorema de Lanford - Sérgio B. Volchan
- Extremal and Probabilistic Combinatorics - Robert Morris e Roberto Imbuzeiro Oliveira
- Fluxos Estrela - Alexander Arbieto, Bruno Santiago e Tatiana Sodero
- Geometria Aritmética em Retas e Cônicas - Rodrigo Gondim
- Hydrodynamical Methods in Last Passage Percolation Models - E. A. Cator e L. P. R. Pimentel
- Introduction to Optimal Transport: Theory and Applications - Nicola Gigli
- Introduction to Stochastic Variational Analysis - Roger J-B Wets
- Introdução à Aproximação Numérica de Equações Diferenciais Parciais Via o Método de Elementos Finitos - Juan Galvis e Henrique Versieux
- Matrizes Especiais em Matemática Numérica - Licio Hernanes Bezerra
- Mecânica Quântica para Matemáticos em Formação - Bárbara Amaral, Alexandre Tavares Baraviera e Marcelo O. Terra Cunha
- Multiple Integrals and Modular Differential Equations - Hossein Movasati
- Nonlinear Equations - Gregorio Malajovich
- Partially Hyperbolic Dynamics - Federico Rodriguez Hertz, Jana Rodriguez Hertz e Raúl Ures
- Processos Aleatórios com Comprimento Variável - A. Toom, A. Ramos, A. Rocha e A. Simas
- Um Primeiro Contato com Bases de Gröbner - Marcelo Escudeiro Hernandes

ISBN: 978-85-244-319-4

Distribuição: IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
E-mail: ddic@impa.br
http://www.impa.br

Contents

Preface

"I count a lot of things that there's no need to count," Cameron said. "Just because that's the way I am. But I count all the things that need to be counted."

Richard Brautigan

These lecture notes give a very brief introduction to Extremal and Probabilistic Combinatorics. *Extremal Combinatorics*, as one might guess, is about extreme behaviours. For instance, what is the largest number of edges in a graph with no triangles? How 'small' must a set of integers be if it does not contain any arithmetic progressions of length 3? *Probabilistic Combinatorics* deals with the behaviour of random discrete structures, and how they can be harnessed to understand deterministic problems. It might seem odd that Probability, which usually deals with typical phenomena, is somehow related to extremal behaviour, which would seem exceptional, but we will see that the field is filled with tantalizing examples of this.

Combinatorics is unique among current research areas in Mathematics in that its prerequisites are minimal. Many important problems in the field can be understood by newcomers, and a few have actually been *solved* by them! This peculiarity has allowed us to present both classical results and recent ideas, while keeping most of the text accessible to students with very little mathematical background. Chapter 5 is the main exception to this rule, but one may still be able to get a sense of what is going on there without delving too deeply into abstract Mathematics.

Of course, saying that Combinatorics is (mostly) elementary does not imply that it is easy. Many people get hooked to Combinatorics precisely because its problems can be so challenging. Our readers should be prepared to take up the challenge and attempt to solve all of the exercises in the text, as well as to fill in missing steps and proofs that are 'left to the reader'. No-one should be discouraged if a few of those problems turn out to be unwieldy. Unsuccessful attempts may give one a better idea about what is not obvious about a given problem and will make the study of other people's ideas more profitable. Bollobas' book [2] may be especially useful as a source of complete proofs and pointers to the literature.

Although we shall only be able to scratch the surface of the modern study of Combinatorics, we hope that the reader will come away with a flavour of the subject, and will be inspired to explore some of the 'further reading' recommended at the end of each chapter. The area is overflowing with beautiful and accessible problems, and, in Bollobás' words: "just as a bear-cub can acquire life-skills through play, so the reader can learn skills for a mathematical life simply by solving, or trying to solve such problems." If you keep your "brain open" then, with a little luck, you will be able to say (as Gauss did), "Like a sudden flash of lightning, the riddle happened to be solved!"

Roberto Imbuzeiro and Rob Morris

Rio de Janeiro, May 2011

Chapter 1

Preliminaries

In Geometry (which is the only science that it hath pleased God hitherto to bestow on mankind), men begin at settling the significations of their words; which ... they call Definitions.

Thomas Hobbes, Leviathan

1.1 Numbers and sets

In these notes, $\mathbb{N} = \{1, 2, \dots\}$ denotes the set of positive integers, and $[n] = \{1, \dots, n\}$ denotes the set of all positive integers from 1 to n.

Given a set S, $|S|$ is the number of elements of S (which may be infinite). For $k \in \mathbb{N}$, $\binom{S}{k}$ is the set whose elements are all k-element subsets of S. Thus for any finite S with $|S| = n$, and any $0 \leqslant k \leqslant n$, the number of elements in $\binom{S}{k}$ is the binomial coefficient:

$$\left|\binom{S}{k}\right| = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

The classical lower bound $k! \geqslant (k/e)^k$, valid for all $k \in \mathbb{N}$, implies the following useful upper bound for the binomial coefficients, which is not far from the (easy) lower bound:

$$\left(\frac{n}{k}\right)^k \leqslant \binom{n}{k} \leqslant \left(\frac{en}{k}\right)^k.$$

The set $\binom{[n]}{2}$ will be special for us due to its connection with the edge set of a graph (see below). We will think of $\binom{S}{2}$ as the set of *unordered pairs* of elements in S and will often write su or us for an element $\{u, s\} \in S$.

1.2 Asymptotics

Given sequences $\{a_n\}, \{b_n\}$ of positive numbers, we write $a_n = O(b_n)$ if there exists a constant $C > 0$ such that $|a_n| \leqslant C\, b_n$ for all large enough n. We write $a_n = o(b_n)$ or $a_n \ll b_n$ if $a_n/b_n \to 0$ as $n \to \infty$.

1.3 Graphs

A *graph* G is pair (V, E) where V is a set of elements, called *vertices*, and $E \subset \binom{V}{2}$ is a set of *edges*. We shall write $|G|$ (instead of $|V|$) for the number of vertices in G and $e(G)$ (instead of $|E|$) for the number of edges in G. Given $v, w \in V$, we say that v and w are *adjacent* in G, and write $v \sim w$, if $vw \in E$. G is said to be complete if $v \sim w$ for all distinct $v, w \in V$, or equivalently, if $E = \binom{V}{2}$. The complete graph with vertex set $[n]$ will be denoted by K_n.

A *path* in G is a sequence of distinct vertices $(v_0, v_1, \ldots, v_k)$ with $v_{i-1} \sim v_i$ for each $i \in [k]$. If $v_k \sim v_0$ as well, we have a cycle. G is *connected* if for any two distinct vertices $v, w \in V$ there exists a path $(v_0, v_1, \ldots, v_k)$ with $v_0 = v$ and $v_k = w$. A connected graph with no cycles is called a *tree*. A graph is *bipartite* if it contains no cycles of odd length, which means that there exists some $A \subset V$ such that there are no edges $\{v, w\} \subset A$, and no edges $\{v, w\} \subset V \setminus A$.

A *subgraph* G' of G is a graph $G' = (V', E')$ with $V' \subset V$ and $E' \subset E \cap \binom{V'}{2}$. If $E' = E \cap \binom{V'}{2}$, G' is said to be *induced* by V' and we write $G' = G[V']$. For disjoint $A, B \subset V(G)$, we write $G[A, B]$ for the *induced* bipartite graph whose vertex set is $A \cup B$ which contains all edges $ab \in E$ with $a \in A$, $b \in B$.

A subset $S \subset V(G)$ of size k such that $G[S]$ is a complete graph is called a *k-clique*. A set $S \subset V$ such that $G[S]$ has no edges is called an *independent set*.

1.4 Probability

In these notes we only consider finite probability spaces. A finite probability space is defined by a non-empty finite set Ω called the *state space* and a function $\mathbb{P} : \Omega \to [0, 1]$ that assigns a probability $\mathbb{P}(\omega)$ to each $\omega \in \Omega$. The probabilities are assumed to satisfy:

$$\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1.$$

Subsets $E \subset \Omega$ are called *events*, and we write $\mathbb{P}(E) = \sum_{\omega \in E} \mathbb{P}(\omega)$ (by definition, this sum is 0 if E is empty). The most basic case of the Inclusion-Exclusion principle shows that, for all $E_1, E_2 \subset \Omega$:

$$\mathbb{P}(E_1 \cup E_2) \; = \; \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2) \; \leqslant \; \mathbb{P}(E_1) + \mathbb{P}(E_2).$$

A mapping $X : \Omega \to \mathbb{R}$ is called a *random variable*. The *expectation* of X is given by:

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \, \mathbb{P}(\omega).$$

The *variance* of X is $\mathrm{Var}(X) := \mathbb{E}\big[(X - \mathbb{E}(X))^2\big]$, which is also equal to $\mathbb{E}(X^2) - \mathbb{E}(X)^2$. This implies in particular that $\mathbb{E}(X)^2 \leqslant \mathbb{E}(X^2)$ for any random variable.

We shall often write expressions such as $\{X > t\}$ and $\{X = x\}$ to denote the events $\{\omega \in \Omega \,:\, X(\omega) > t\}$ and $\{\omega \in \Omega \,:\, X(\omega) = x\}$ (respectively). In the text we will usually not bother to define the specific probability space we are using; rather, we shall focus on events and random variables.

Two final definitions: A sequence of events $E_1, \ldots, E_m \subset \Omega$ is said to be *independent* if, for every $\varnothing \neq A \subset [m]$,

$$\mathbb{P}\bigg(\bigcap_{i \in A} E_i \bigg) \; = \; \prod_{i \in A} \mathbb{P}(E_i).$$

Notice that the events remain independent if some of them are replaced by their complements in Ω. Finally, a sequence $X_1, \ldots, X_m$ of random variables is independent if, for all $x_1, \ldots, x_m \in \mathbb{R}$, the events $\{X_1 = x_1\}, \ldots, \{X_m = x_m\}$ are independent. In this case,

$$\mathrm{Var}\left(\sum_{i=1}^{m} X_i\right) = \sum_{i=1}^{m} \mathrm{Var}(X_i),$$

i.e., the variance of a sum of independent random variables equals the sum of the variances.

Chapter 2

Ramsey Theory

My brain is open
Paul Erdős

Ramsey Theory is the study of finding order in chaos. We shall begin with some motivation from Number Theory.

2.1 Number Theory

The problem of solving equations in the integers is one of the oldest in mathematics. For example, Pell's equation

$$x^2 - ny^2 = \pm 1$$

was studied by Brahmagupta in the 7th century, and Fermat's Last Theorem, that

$$x^n + y^n = z^n$$

has no solutions if $n \geqslant 3$, was open for more than 300 years, before being proven by Andrew Wiles in 1995. Hilbert's 10th problem was to determine whether or not all such 'Diophantine' equations are solvable; the answer (no) was given by Matiyasevich in 1970.

We shall be interested in a different, but related problem: that of solving equations (or, more generally, systems of equations) in *subsets* of the integers. For example, consider the following question:

Question 2.1.1. *What is the maximum 'density' of a set $A \subset \mathbb{N}$ which contains no solution to the equation $x + y = z$?*

The odds have density 1/2, and are sum-free. To see that we can't do any better, consider the sets $A_n = A \cap [n]$ and $n - A_n := \{n - a : a \in A_n\}$, where $[n] = \{1, \ldots, n\}$. If $n \in A$ then these sets are disjoint subsets of $[n]$, and both have size $|A_n|$. Thus, $|A_n| \leqslant n/2$, as required.

Not all such questions are so easy, however. We shall solve the following problem, known as Roth's Theorem, later in the course.

Question 2.1.2. *What is the maximum 'density' of a set $A \subset \mathbb{N}$ which contains no solution to the equation $x + y = 2z$?*

2.2 Ramsey's Theorem

In 1916, whilst studying Fermat's Last Theorem in the group $\mathbb{Z}_p$, Issai Schur came across (and solved) the following, slightly different question:

Question 2.2.1. *Suppose we colour the positive integers with r colours. Must there exist a monochromatic solution of the equation $x + y = z$?*

Here and in what follows, a "colouring" of a non-empty set S with $r \in \mathbb{N}$ colours is just a colourful way of denoting a map $c : S \to [r]$. The number of $\{1, \ldots, r\}$ are usually the colours, and $s \in S$ is painted with colour $i \in [r]$ if $c(s) = i$. A monochromatic set is a subset $A \subset S$ whose elements are all painted with the same colour. Note that it should be much easier to find a monochromatic solution than to find a solution in a given set (i.e., colour class); however, if r is large then each colour class may be quite sparse.

In order to answer Schur's question, we shall skip forward in time to 1930, and jump from Germany to England. There we find Frank Ramsey, a precocious young logician, who is best remembered for the following beautiful theorem. He considered it only a minor lemma; as is often the case in Combinatorics, however, the lemma turned out to be more important than the theorem it was used to prove.

Ramsey's Theorem. *Let $r \geqslant 1$. Every colouring $c : \binom{\mathbb{N}}{2} \to [r]$ of the pairs in $\mathbb{N}$ contains an infinite monochromatic subset.*

In order to prove Theorem 2.2, we shall only need the (infinite) pigeonhole principle:

"If an infinite number of letters lie in a finite number of pigeonholes, then some pigeonhole must contain an infinite number of letters."

Given a colouring $c : \binom{\mathbb{N}}{2} \to [r]$, a colour i and a vertex $v \in \mathbb{N}$, let $N_i(v) = \{w : c(vw) = i\}$ denote the colour i neighbourhood of v.

Proof of Theorem 2.2. We first construct a sequence $(x_1, x_2, \ldots)$ of vertices such that the colour of the edge $x_i x_j$ is determined by $\min\{i, j\}$. Indeed, let $x_1 = 1$ and $X_1 = \mathbb{N}$, and observe that (by the pigeonhole principle) $N_{c(1)}(x_1)$ is infinite for some $c(1) \in [r]$. Let X_2 be that infinite set (if more than one is infinite then choose one arbitrarily).

Now, given $(x_1, \ldots, x_{n-1})$ and an infinite set X_n, choose $x_n \in X_n$ arbitrarily, note that $N_{c(n)}(x_n) \cap X_n$ is infinite for some $c(n) \in [r]$, and let X_{n+1} be that infinite set. Since the sets X_n are nested, it follows that $c(x_i x_j) = c(i)$ for every $j > i$.

Finally, consider the sequence $(c(1), c(2), \ldots)$, and observe that it contains an infinite monochromatic subsequence. Let $(a(1), a(2), \ldots)$ be the indices of this subsequence. Then $\{x_{a(1)}, x_{a(2)}, \ldots\}$ is an infinite monochromatic subset of $\mathbb{N}$. □

We remark that Ramsey actually proved the following, more general theorem.

Ramsey's Theorem for k-sets. *Every r-colouring of the k-subsets of $\mathbb{N}$ contains an infinite monochromatic set.*

(Here, a set A is monochromatic if the colouring c is constant on subsets of A.)

2.3 Schur's Theorem

We can now answer Schur's question. Let us call (x, y, z) a *Schur triple* if $x + y = z$.

Theorem 2.3.1 (Schur, 1916). *Every colouring $c : \mathbb{N} \to [r]$ contains a monochromatic Schur triple.*

Proof. Let $c : \mathbb{N} \to [r]$ be given, and consider the colouring $c' : \binom{\mathbb{N}}{2} \to [r]$ defined by

$$c'\big(\{a, b\}\big) \;:=\; c\big(|a - b|\big).$$

By Ramsey's Theorem, there exists a monochromatic triangle, $\{x, y, z\}$, with $x < y < z$, say. But then $c(y - x) = c(z - x) = c(z - y)$, and so these form a monochromatic Schur triple, as required. □

We can also deduce the following theorem, which was Schur's motivation for proving the result above.

Corollary 2.3.2 (Schur, 1916)**.** *For every $n \in \mathbb{N}$, the equation*

$$x^n + y^n = z^n \pmod p$$

has a non-trivial solution for every sufficiently large prime p.

Proof. Let $n \in \mathbb{N}$, consider the subgroup $H_n = \big\{x^n : x \in \mathbb{Z}_p^*\big\} \leqslant \mathbb{Z}_p^*$, and partition $\mathbb{Z}_p^*$ into cosets,

$$\mathbb{Z}_p^* \;=\; a_1 H_n \cup \ldots \cup a_r H_n.$$

We claim that $r \leqslant n$. Indeed, since $x \mapsto x^n$ is a homomorphism, and $|H_n|$ is the size of the image, it follows that $r = |\{x : x^n = 1\}|$, and $x^n = 1$ has at most n roots.

Thus we have an r-colouring of $\{1, \ldots, p-1\}$, so, by Theorem 2.3.1, if p is sufficiently large then there exists a monochromatic Schur triple. That is, there exist integers x, y and z, none of which is divisible by p, such that

$$a_i x^n + a_i y^n = a_i z^n \pmod p$$

for some $i \in [r]$. But a_i is invertible, so $x^n + y^n = z^n \pmod p$, as required. □

2.4 Finite Ramsey's Theorem

Question 2.4.1. *Suppose there are six people at a party. Is it true that there are either three people all of whom know each other, or three people none of whom know each other?*

Answer: Yes!

Proof. Suppose without loss of generality (by symmetry) that Alex knows at least three people, Beta, Carla and Dudu. If any two of these know each other then they form a 'good' triple with Alex, otherwise they form a good triple themselves. □

More generally, define $R(k)$, the k^{th} Ramsey number, to be the smallest value of n such that, given any red-blue colouring of the edges of the complete graph on n vertices, there exists a monochromatic complete subgraph on k vertices. It follows from Theorem 2.2 that $R(k)$ exists for every $k \in \mathbb{N}$. The example above shows that $R(3) = 6$.

Theorem 2.4.2 (Erdős and Szekeres, 1935; Erdős, 1947)**.**

$$(\sqrt{2})^k \ll R(k) \ll 4^k.$$

For the upper bound, we shall generalize the theorem (in an appropriate way), and use induction. Indeed, for every $k, \ell \in \mathbb{N}$, let $R(k,\ell)$ denote the smallest n such that any two-colouring of $[n]$ contains either a red K_k or a blue K_ℓ. Clearly $R(k) = R(k,k)$.

Proof of the upper bound. We claim that

$$R(k,\ell) \leqslant R(k-1,\ell) + R(k,\ell-1)$$

for every $k, \ell \in \mathbb{N}$. Indeed, let $n \geqslant R(k-1,\ell) + R(k,\ell-1)$, and choose a vertex $v \in [n]$. By the (finite) pigeonhole principle, it has either at least $R(k-1,\ell)$ red neighbours, or at least $R(k,\ell-1)$ blue neighbours. By symmetry assume the former. By the definition of $R(k-1,\ell)$, the induced colouring on the red neighbours of v contain either a red K_{k-1} or a blue K_ℓ. Adding v (if necessary), we are done.

It follows from the claim (by induction) that

$$R(k,\ell) \leqslant \binom{k+\ell}{k}$$

for every $k, \ell \in \mathbb{N}$. This is because the binomial coefficients satisfy:

$$\binom{k+\ell}{k} = \binom{k+\ell-1}{k-1} + \binom{k+\ell-1}{k}$$

for every $k, \ell \in \mathbb{N}$ and $\binom{k+\ell}{k} = R(k,\ell)$ for $k = 1$ or $\ell = 1$ (exercise!). So the upper bound follows. □

It turns out to be surprisingly difficult to construct colourings with no large complete monochromatic subgraphs. However, in 1947 Erdős gave the following shockingly simple proof of an exponential lower bound on $R(k)$. Despite its simplicity, this has been one of the most influential proofs in the development of Combinatorics.

Proof of the lower bound. Let $c : \binom{[n]}{2} \to \{0,1\}$ be a random 2-colouring of K_n. To be precise, let the probability that $c(ij)$ is red be $1/2$ for every edge $ij \in E(K_n)$, and choose the colours of edges independently.

Let X be the number of monochromatic cliques in K_n under c. The expected value of X is the total number of k-cliques in K_n, which is $\binom{n}{k}$, times the probability that a given clique is monochromatic. This gives:

$$\binom{n}{k}\left(\frac{1}{2}\right)^{\binom{k}{2}-1} \leqslant 2\left(\frac{en}{k}\left(\frac{1}{\sqrt{2}}\right)^{k-1}\right)^{k} \ll 1$$

if $n < \frac{1}{e\sqrt{2}}k2^{k/2}$.Since the expected value of random variable X (defined on colourings) is less than 1, then there must exist a colouring in which $X = 0$. Hence there exists a colouring with no monochromatic k-clique, and we are done. $\square$

It is surprisingly difficult to construct large colourings with no large monochromatic clique. The first constructive super-polynomial bound on the Ramsey numbers was given by Frankl in 1977. The best known construction is due to Frankl and Wilson, and uses

$$\exp\left(\left(\frac{1}{4}-\varepsilon\right)\frac{(\log k)^2}{\log\log k}\right)$$

vertices. We shall see their proof in Chapter 5.

2.5 Van der Waerden's Theorem

Let us next turn our attention to the second equation mentioned earlier, $x + y = 2z$, and notice that the solutions of this equation are arithmetic progressions of length three. Since it is difficult to determine the maximum density of a set with no such triple (see Chapter 6), we shall first answer the 'Ramsey' version.

Question 2.5.1. *If we colour the integers* $\mathbb{N}$ *with two colours, must there exist a monochromatic arithmetic progression of length three?*

Answer: Yes! In fact there must exist one in $\{1, \ldots, 9\}$.

The following theorem generalizes this simple fact to arithmetic progressions of length k.

Theorem 2.5.2 (Van der Waerden, 1927)**.** *Every two-colouring of* $\mathbb{N}$ *contains arbitrarily long monochromatic arithmetic progressions.*

We would like to prove Van der Waerden's Theorem by induction, but in its present form this seems to be difficult.

Key idea: Strengthen the induction hypothesis! Prove that the same result holds for an *arbitrary* number of colours.

When using induction, it is often easier to prove a stronger result than a weaker one!

Proof of Van der Waerden's Theorem. For each pair $k, r \in \mathbb{N}$, define $W(r,k)$ to be the smallest integer n such that, if we r-colour the set $[n]$ then there must exist a monochromatic arithmetic progression of length k. We claim that $W(r,k)$ exists for every $k, r \in \mathbb{N}$, and will prove it by double induction. Note that the result is trivial if $k \leqslant 2$ (for all r), so assume that $W(r, k-1)$ exists, for every $r \in \mathbb{N}$. We shall denote the arithmetic progression $\{a, a+d, a+2d, \ldots, a+(k-1)d\}$ by $AP(a,d,k)$: it has common difference d and length k.

We make the following important definition: Let $A_1, \ldots, A_t$ be arithmetic progressions of length ℓ, with $A_j = AP(a_j, d_j, \ell)$. We say that $A_1, \ldots, A_t$ are *focused* at z if

$$a_i + \ell d_i \;=\; a_j + \ell d_j \;=\; z$$

for every $i, j \in [t]$, and we say they are *colour-focused* if moreover they are all monochromatic and of different colours.

Let $k, r \in \mathbb{N}$ be arbitrary. We make the following claim:

Claim: For all $s \leqslant r$, there exists n such that whenever $[n]$ is r-coloured, there exists either a monochromatic arithmetic progression of length k, or s colour-focused arithmetic progressions of length $k-1$.

Proof of Claim. We shall prove the claim by induction on s. When $s = 1$ the claim follows by the main induction hypothesis: we may take $n = W(r, k-1)$. So let $s > 1$, and suppose n is suitable for $s-1$; we will show that $N = 2n \cdot W(r^{2n}, k-1)$ is suitable for s.

Indeed, partition $[N]$ into blocks of length $[2n]$, and observe that each contains either a monochromatic arithmetic progression of length k (in which case we are done), or $s-1$ colour-focused arithmetic progressions of length $k-1$, *together with their focus*, which moreover has a colour different from each of them.

Next, note that the r-colouring of $[N]$ induces an r^{2n}-colouring of the blocks, and that therefore, by the definition of N, there exists an arithmetic progression $\{B(x), B(x+y), \ldots, B(x+(k-2)y)\}$ of blocks, such that these blocks are coloured identically.

Finally, let $A_j = AP(a_j, d_j, k-1)$ for $1 \leqslant j \leqslant s-1$ be the $s-1$ colour-focused APs in $B(x)$, let z be their focus, and observe that the following s APs of length $k-1$ are colour-focused at $z + 2yn(k-1)$:

$$A'_j \;:=\; AP(a_j, d_j + 2yn, k-1),$$

for $1 \leqslant j \leqslant s-1$, and $AP(z, 2yn, k-1)$. This completes the induction step, and hence proves the claim. □

The theorem follows from the claim by setting $s = r$. □

Because of the double induction, the proof above gives *very* bad bounds on $W(r,k)$. For example, even for 3-APs it gives only

$$W(r,3) \;\leqslant\; k^{k^{\cdot^{\cdot^{\cdot^{k}}}}},$$

where the 'tower' has height $k-1$.

More generally, let $f_1(x) = 2x$, and

$$f_{n+1}(x) \;=\; f_n^{(x)}(1) \;=\; f_n(f_n(\ldots(1)\ldots)$$

for all $n \geqslant 1$. (This is the Ackermann or Grzegorczyk hierarchy.) Thus $f_2(x) = 2^x$, $f_3(x)$ is a tower of 2s of height x, and so on. (Exercise: write down as many values of $f_4(x)$ as possible.)

Our proof of van der Waerden's Theorem gives a bound on $W(2,k)$ which grows faster than f_n for every integer n! Shelah vastly improved this bound, proving that

$$W(r,k) \leqslant f_4(r+k).$$

A further (and similarly huge) improvement was given by Gowers, who showed that $W(r,k)$ is at most a tower of 2s of height 7.

Van der Waerden's Theorem and Schur's Theorem played an important (historic) role in the development of several areas of mathematics, including Hypergraph Regularity, Additive Combinatorics, and Ergodic Theory. The most important step was the following theorem of Szemerédi, proved in 1975, which resolved a conjecture of Erdős and Turán from 1936. The upper density of a set $A \subset \mathbb{N}$ is

$$\overline{d}(A) \;=\; \limsup_{n\to\infty} \frac{|A \cap [n]|}{n}.$$

Szemerédi's Theorem. *Any subset of* $\mathbb{N}$ *with positive upper density contains arbitrarily long arithmetic progressions.*

Furstenburg (in 1977) and Gowers (in 2001) gave important alternative proofs of this result. The following theorem was proved recently by Ben Green and Terence Tao.

Theorem 2.5.3 (Green and Tao, 2008)**.** *There exist arbitrarily long arithmetic progressions in the primes.*

In fact they proved the theorem for any subset of the primes of positive (relative) density.

2.6 An unprovable theorem

The following theorem may be proved using the Axiom of Choice.

Theorem 2.6.1 (The Strengthened Ramsey Theorem)**.** *Given* $m,k,r \in \mathbb{N}$*, there exists* $N \in \mathbb{N}$ *such that the following holds. If we* r*-colour the* k*-subsets of* $[n]$*, then there exists a monochromatic set* $Y \subset S$ *with* $|Y| \geqslant m$*, and with* $|Y| \geqslant \min\{y : y \in Y\}$*.*

Problem 2.6.2. *Deduce Theorem 2.6.1 from the infinite version of Ramsey's Theorem. [Hint: use a compactness argument.]*

But not without!

Theorem 2.6.3 (Paris and Harrington, 1977)**.** *The Strengthened Ramsey Theorem cannot be proved in Peano Arithmetic.*

2.7 Recommended Further Reading

R. Graham, B.L. Rothschild and J. Spencer, Ramsey Theory (2nd edition), Wiley (1990).

Imre Leader, Part III Lecture Notes, available at
http://www.dpmms.cam.ac.uk/∼par31/notes/ramsey.pdf

2.8 Exercises

We are usually convinced more easily by reasons we have found ourselves than by those which occurred to others.
Blaise Pascal

1. Show that $R(3,4) = 9$ and $R(4) = 18$. [Hint: consider the graph on $\{1,\dots,17\}$, where ij is an edge iff $i-j$ is a square (mod 17).]

2. Prove that any sequence of integers has a monotone subsequence.

3. Colour $\mathbb{N}$ with two colours. Must there exist a monochromatic infinite arithmetic progression?

4. Let $C(s)$ be the smallest n such that every connected graph on n vertices has, as an induced subgraph, either a complete graph K_s, a star $K(1,s)$ or a path P_s of length s. Show that $C(s) \leqslant R(s)^s$.

5. Prove Ramsey's Theorem for k-sets. [Hint: Use induction on k.]

6. Prove that any sequence of integers has an infinite subsequence which is either convex or concave.

7. Let A be a set of n points in $\mathbb{R}^2$, such that no three points of A are co-linear. Prove that, if n is sufficiently large, then A contains k points forming a convex k-gon.

8. Show that there is an infinite set S of positive integers such that the sum of any two distinct elements of S has an even number of distinct prime factors.

9. Does there exist a 2-colouring of the infinite subsets of $\mathbb{N}$ with no infinite monochromatic subset?

Chapter 3

Extremal Graph Theory

Theorems are fun, especially when you are the prover; but then the pleasure fades. What keeps us going are the unsolved problems.
Carl Pomerance

In the previous chapter we considered, for subsets of the integers, and for colourings, questions of the following form:

> "Suppose a certain sub-structure (e.g., monochromatic clique, Schur triple) is forbidden. How large can our base set be?"

In this chapter, we extend this line of inquiry to graphs.

3.1 Turán's Theorem

Consider a group of people, and suppose that there do not exist three people in the group who are all friends with each other. How many pairs of friends can there be in the group? In other words:

Question 3.1.1. *What is the maximum number of edges in a triangle-free graph on n vertices?*

A complete bipartite graph (with part sizes as equal as possible) has $\lfloor \frac{n}{2} \rfloor \lceil \frac{n}{2} \rceil \approx \frac{n^2}{4}$ edges and no triangle.

Theorem 3.1.2 (Mantel, 1907)**.** *If G is a triangle-free graph on n vertices, then*

$$e(G) \leqslant \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil.$$

Proof. By induction on n. Consider an edge $uv \in G$; since G is triangle-free, it follows that there at most $n-1$ edges incident with u or v. Since $G - \{u, v\}$ is triangle-free, it follows (by induction) that

$$e\big(G - \{u, v\}\big) \leqslant \left(\left\lfloor \frac{n}{2} \right\rfloor - 1\right)\left(\left\lceil \frac{n}{2} \right\rceil - 1\right) = \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil - n + 1,$$

and so the result follows. □

In 1941, Turán generalized the result above in the following way. Write K_r for the complete graph on r vertices, and write $T_{r,n}$ for the r-partite graph on n vertices with the maximum number of edges; this graph is called the Turán graph. (Exercise: Show that, for r fixed, $\left(1 - \frac{1}{r}\right)\binom{n}{2} - O(r) \leqslant e(T_{r,n}) \leqslant \left(1 - \frac{1}{r}\right)\binom{n}{2}$.)

Turán's Theorem. *If G is a K_{r+1}-free graph on n vertices, then*

$$e(G) \leqslant e(T_{r,n}) = \left(1 - \frac{1}{r}\right)\binom{n}{2} - O(r).$$

Moreover, equality holds if and only if $G = T_{r,n}$.

Problem 3.1.3. *Extend the proof of Mantel's Theorem, above, to prove Turán's Theorem. [Hint: Remove a copy of K_r, and observe that $e(T_{r,n}) - e(T_{r,n-r}) = n(r-1) - \binom{r}{2}$.]*

The following extension of Turán's Theorem is due to Erdős.

Theorem 3.1.4 (Erdős, 1970)**.** *Let G be a K_{r+1}-free graph on n vertices. Then there exists an r-partite graph H on the same vertex set, with $d_H(v) \geqslant d_G(v)$ for every $v \in V(G)$.*

Proof. Let G be a K_{r+1}-free graph, and let $w \in V(G)$ be a vertex of maximum degree in G. 'Zykov symmetrization' is the following operation: for every vertex $v \in V(G) \setminus N(w)$, we remove the edges incident with v, and add an edge between v and each neighbour of w. Observe that, for every vertex $u \in V(G)$, the degree of u has not decreased (since the degree of w was maximal), and the graph is still K_{r+1}-free.

We use induction on r, and the operation above. Indeed, let $w \in V(G)$ be a vertex of maximal degree in G, and note that the graph $G_1 = G[N(w)]$ is K_r-free. Thus by the induction hypothesis, there exists an $(r-1)$-partite graph H_1 on $N(w)$ with $d_{H_1}(v) \geqslant d_{G_1}(v)$ for every $v \in N(w)$. Let H be the graph obtained from G by performing Zykov symmetrization at the vertex w, and replacing G_1 by H_1. Since in each step the degrees did not decrease, H is the required r-partite graph. $\square$

What can we say about the graph G if it is K_{r+1}-free and contains almost as many edges as the Turán graph $T_{r,n}$?

Theorem 3.1.5 (Erdős-Simonovits Stability Theorem, 1966). *For every $\varepsilon > 0$ there exists a $\delta > 0$ such that the following holds. If G is a K_{r+1}-free graph on n vertices, and*

$$e(G) \;\geqslant\; e(T_{r,n}) \;-\; \delta n^2,$$

then G may be transformed into the Turán graph by adding or removing at most εn^2 edges.

Problem 3.1.6. *Prove the Erdős-Simonovits Stability Theorem for triangles.*

3.2 Bipartite graphs

In Turán's Theorem, the extremal examples all contain large complete bipartite graphs. What happens then if we forbid a given small bipartite graph? Erdős first studied this question in order to answer the following problem in combinatorial number theory. Sequences of integers with pairwise different sums are known as Sidon sequences; Erdős studied the following multiplicative version of Sidon's problem.

Question 3.2.1. *Let $A = \{a_1, \dots, a_t\} \subset [n]$ be such that $a_i a_j \neq a_k a_\ell$ unless $\{i,j\} = \{k,\ell\}$. What is the maximum possible size of A?*

An obvious lower bound: $\pi(n)$, the number of primes in $[n]$. Is this close to the truth?

First some notation. Given a graph H, we shall write

$$\mathrm{ex}(n,H) \;:=\; \max\big\{e(G) : |G| = n \text{ and } G \text{ is } H\text{-free}\big\}.$$

Thus Turán's Theorem can be restated as follows:

$$\mathrm{ex}(n,K_{r+1}) = e(T_{r,n}) \approx \left(1 - \frac{1}{r}\right)\binom{n}{2}.$$

Erdős proved that for the 4-cycle the extremal number is much smaller.

Theorem 3.2.2 (Erdős, 1938)**.**

$$\mathrm{ex}\big(n,C_4\big) = O\big(n^{3/2}\big).$$

Proof. We count 'cherries'. Indeed, consider triples $(x,\{y,z\})$ of (distinct) vertices in G such that $xy, xz \in E(G)$. The number of such triples is

$$\sum_{v \in V(G)} \binom{d(v)}{2} \;\geqslant\; \frac{n}{2}\left(\frac{2e(G)}{n} - 1\right)^2,$$

by convexity, and since $\sum d(v) = 2e(G)$ (exercise). Now simply observe that if there exist two such triples on the same pair $\{y,z\}$, then the graph contains a C_4. Hence the number of such triples in a C_4-free graph is at most $\binom{n}{2}$. So

$$\frac{n}{2}\left(\frac{2e(G)}{n} - 1\right)^2 \;\leqslant\; \binom{n}{2},$$

and hence $e(G) = O(n^{3/2})$, as required. ☐

Erdős used this result to answer Question 3.2.1.

Corollary 3.2.3 (Erdős, 1938)**.** *Let $A \subset [n]$ be a multiplicative Sidon set. Then*

$$|A| \;\leqslant\; \pi(n) + O(n^{3/4}).$$

Erdős later said that, "Being struck by a curious blindness and lack of imagination, I did not at that time extend the problem from C_4 to other graphs, and thus missed founding an interesting and fruitful new branch of graph theory."

The following natural extension of Theorem 3.2.2 is therefore instead due to Kővári, Sós and Turán. For each $s,t \in \mathbb{N}$, let $K(s,t)$ denote the complete bipartite graph with part sizes s and t.

Theorem 3.2.4 (Kővári, T. Sós and Turán, 1954). *If $s \leqslant t$, then*

$$\mathrm{ex}\big(K(s,t)\big) \;=\; O(n^{2-1/s}).$$

Problem 3.2.5. *Generalise the proof of Theorem 3.2.2 above to prove the Kővári-Sós-Turán Theorem. [Hint: count copies of $K(1,s)$.]*

How sharp are these bounds?

Theorem 3.2.6 (Klein, Brown, Kollár-Rónyai-Szabó). *If $s \leqslant t$ are integers, with either $s \in \{2,3\}$, or $t > (s-1)!$, then*

$$\mathrm{ex}\big(K(s,t)\big) \;=\; \Theta\left(n^{2-1/s}\right).$$

Problem 3.2.7 (+). *Show that* $\dfrac{\mathrm{ex}(n,K(4,4))}{\mathrm{ex}(n,K(3,3))} \to \infty$.

It follows from Theorem 3.2.4 that $\mathrm{ex}(n,H) = o(n^2)$ for every bipartite graph H. Can we prove more precise results for specific bipartite graphs, like trees or cycles?

Conjecture 3.2.8 (Erdős and Sós, 1963). *If $e(G) > \frac{1}{2}(k-2)|G|$, then G contains every tree on k vertices.*

Theorem 3.2.9 (Ajtai, Komlós, Simonovits and Szemerédi, 2011). *The Erdős-Sós conjecture is true for sufficiently large values of k .*

Given a family of graphs $\mathcal{L}$, write $\mathrm{ex}(n,\mathcal{L})$ for the maximum number of edges in a graph on n vertices which is L-free for every graph $L \in \mathcal{L}$.

Theorem 3.2.10 (Erdős, 1959). *If $\mathcal{L}$ is a finite family of graphs which does not contain a forest, then there exists $k \in \mathbb{N}$ such that*

$$\mathrm{ex}(n,\mathcal{L}) \;\geqslant\; \mathrm{ex}(n,\{C_3,C_4,\ldots,C_k\}) \;\geqslant\; n^{1+c}$$

for any $c < 1/k$, and all sufficiently large n.

Sketch of proof. Take a random bipartite graph G with n^{1+c} edges, for some $c < 1/k$; we claim that, with high probability, G has very few short cycles. Indeed, let $p = e(G)/\binom{n}{2}$; then the expected number of cycles of length 2ℓ is at most

$$(2\ell)!\binom{n}{2\ell}p^{2\ell} \approx n^{2c\ell} = o(n),$$

if $2\ell \leqslant k$, since $c < 1/k$. Thus with high probability there are $o(n)$ cycles of length at most k. Removing one edge from each of these gives the result. □

The following upper bound of Bondy and Simonovits is over 35 years old, and has only been improved by a constant factor.

Theorem 3.2.11 (Bondy and Simonovits, 1974)**.**

$$\mathrm{ex}(n, C_{2k}) = O(n^{1+1/k}).$$

Despite this, the Bondy-Simonovits Theorem is only known to be sharp for C_4, C_6 and C_{10}. For all other values of k, the correct order of magnitude is unknown.

3.3 The Erdős-Stone Theorem

The following result is sometimes called as the fundamental theorem of graph theory. It determines asymptotically the extremal number of an arbitrary graph H. In order to state the theorem, we need to define the following fundamental parameter.

Definition. *The chromatic number $\chi(G)$ of a graph G is the minimum number of colours in a proper colouring of G. That is,*

$$\chi(G) = \min\big\{r : \exists\, c : V(G) \to [r] \text{ such that } c(u) \neq c(v) \text{ for every } uv \in E(G)\big\}.$$

Note that a graph G is bipartite if and only if $\chi(G) \leqslant 2$, and more generally is r-partite if and only if $\chi(G) \leqslant r$.

The Erdős-Stone Theorem (Erdős and Stone, 1946)**.** *Let H be an arbitrary graph. Then*

$$\mathrm{ex}(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right)\binom{n}{2}.$$

Godsil

Note that when H is bipartite this says only that $\mathrm{ex}(n,H) = o(n^2)$, which we already knew from the Kővári-Sós-Turán Theorem (in fact we know something a bit stronger, that $\mathrm{ex}(n,H) \leqslant n^{2-\varepsilon}$ for some $\varepsilon = \varepsilon(H)$).

Proof. For simplicity, we shall consider only the case $\chi(H) = 3$; the proof in the general case is essentially the same. Let $\varepsilon > 0$, let $t = |H|$, and let G be a graph on n vertices with

$$e(G) \geqslant \left(\frac{1}{2} + \varepsilon\right)\binom{n}{2}.$$

We shall show that, if n is sufficiently large, then G contains a copy of $K_3(t)$, the complete 3-partite graph with t vertices in each part, and hence that it contains a copy of H.

We begin by removing all vertices of small degree. To be precise, we define a sequence of graphs $G \supset G_{n-1} \supset \ldots \supset G_N$, where $|G_k| = k$ and G_{k-1} is obtained from G_k by deleting a vertex with degree at most $(1/2 + \varepsilon/2)\,k$, if such a vertex exists. Observe (by counting edges) that this process must stop at some G_N with $N \geqslant \varepsilon n$, and that every vertex in G_N has degree at least $(1/2 + \varepsilon/2)\,N$.

Now, by the Kővári-Sós-Turán Theorem, there exists a complete bipartite graph $F = K(q,q)$ in G_N, where $q = 4t/\varepsilon$. Suppose that G_N is $K_3(t)$-free; we claim that, if N is sufficiently large, then there are at most

$$\big(q + t\big)N$$

edges between $V(F)$ and $X = V(G_N) \setminus F$. Indeed, let Y denote the set of vertices in X with at least $q + t$ neighbours in F. Then, since each copy of $K(t,t)$ in F can be covered by at most $t - 1$ vertices of X, we have

$$\sum_{v \in Y} \binom{q}{t} \;\leqslant\; \sum_{v \in Y} \max_{a \geqslant t} \binom{a}{t}\binom{d(v) - a}{t} \;\leqslant\; t\binom{q}{t}^2,$$

which implies that $|Y| \leqslant t\binom{q}{t}$. Hence the number of edges between $V(F)$ and X is at most $(q - t - 1)N + qt\binom{q}{t}$, as required.

Finally, recall that every vertex in F has degree at least $\left(\frac{1}{2}+\frac{\varepsilon}{2}\right)N$. Thus the number of edges between $V(F)$ and X is at least

$$2q\left(\frac{1}{2}+\frac{\varepsilon}{2}\right)N \,-\, (2q)^2 \;\geqslant\; \left(q+t+\frac{\varepsilon}{4}\right)N,$$

which is a contradiction. □

3.4 Counting H-free graphs

We end this section by briefly describing an important modern branch of extremal graph theory: the problem of counting the *number* of graphs on n vertices which avoid a given graph H. The following conjecture of Erdős is the central open problem in the area.

Conjecture 3.4.1 (Erdős, 1970s)**.** *Given a graph H, let $\mathcal{P}(n,H)$ denote the collection of graphs on n vertices which do not contain H. Suppose H contains a cycle. Then*

$$|\mathcal{P}(n,H)| \;=\; 2^{(1+o(1))\mathrm{ex}(n,H)}.$$

In the case $\chi(H) \geqslant 3$, this conjecture was proved by Erdős, Frankl and Rödl in 1986. However, it is open for *every single* bipartite graph. Even in a weaker form (with $1+o(1)$ replaced by some large constant), results have been proved in only a few special cases:

- Kleitman and Winston (1982): $|\mathcal{P}(n,C_4)| = 2^{O(n^{3/2})}$,
- Kleitman and Wilson (1996):

$$|\mathcal{P}(n,C_6)| = 2^{O(n^{4/3})} \quad \text{and} \quad |\mathcal{P}(n,C_8)| = 2^{O(n^{5/4})},$$

- Kohayakawa, Kreuter and Steger (1998):

$$\left|\mathcal{P}\big(n,\{C_4,C_6,\ldots,C_{2k}\}\big)\right| = 2^{O(n^{1+1/k})}.$$

- Balogh and Samotij (2010): $\left|\mathcal{P}\big(n,K(s,t)\big)\right| = 2^{O(n^{2-1/s})}$.

3.5 Recommended Further Reading

The standard reference for the area is Bollobás [2].

3.6 Exercises

1. Determine $\mathrm{ex}(n, P_4)$, where P_4 denotes the path of length four. How about for a path of length k?

2. Show that Conjecture 3.4.1 is false for $H = K(1,t)$ (this graph is called a 'star'). Is it true for paths?

3. Let G be a graph on n vertices with $e(T_{r-1,n}) + 1$ edges. By Turán's Theorem, G contains a copy of K_r. Show that G also has a copy of $K_{r+1} - e$, the complete graph minus an edge.

4. Prove the following weaker form of the Erdős-Sós Conjecture: If $e(G) > (k-2)|G|$ then G contains every tree on k vertices.

5. Construct a C_4-free graph on n vertices with $\Theta\big(n^{3/2}\big)$ edges. [Hint: try a projective plane.]

6. Let $g(n)$ be the largest number of edges in a graph G on n vertices with the following property: there is a 2-colouring of the edges of G with no monochromatic triangle.

   (a) Show that $g(n)/\binom{n}{2}$ converges.

   (b) Find c such that $g(n)/\binom{n}{2} \to c$ as $n \to \infty$.

Chapter 4

The Random Graph

Creativity is the ability to introduce
order into the randomness of nature
Eric Hoffer

Most real-world systems are too complex for us to study directly: the weather, the stock market, and social interactions are subject to too many unpredictable variables and influences. However, many of these influences have the *appearance* of randomness; by modelling them as random variables, we can sometimes still say something useful about the whole system.

In this chapter we shall study a particularly simple and famous such random model: the Erdős-Rényi random graph. In order to motivate the model, we shall begin with an application.

4.1 High girth and chromatic number

Question 4.1.1. *Does there exist a triangle-free graph with chromatic number at least k?*

The answer is yes, but it is non-trivial to construct such graphs. (The first examples were given by Zykov in 1949.) We begin this section by giving a short proof of a much stronger result.

The *girth* $g(G)$ of G is the length of the shortest cycle in G. The following theorem was one of the earliest (and most shocking) applications of the probabilistic method.

Theorem 4.1.2 (Erdős, 1959). *There exist graphs whose girth and chromatic number are both arbitrarily large.*

Proof. We take a 'random' graph, and perform a small amount of 'surgery', c.f. the proof of Theorem 3.2.10. Let $k \in \mathbb{N}$ be arbitrary; we shall show that there exists a graph G with $\chi(G) \geqslant k$ and $g(G) \geqslant k$.

As it turns out, proving $\chi(G) \geqslant k$ directly is often hard, so we will deal with the so-called *independence number* $\alpha(G)$ instead. Recall from Chapter 1 that an independent set is a subset of vertices containing no edges. By definition, $\alpha(G)$ is the size of the largest independent set in G.

To see the connection between $\alpha(G)$ and the chromatic number, suppose G is painted with $\chi(G)$ colours and no adjacent vertices have the same colour. Given $i \in [\chi(G)]$, the set of vertices S_i with colour i is independent. Since the sets $S_1, \ldots, S_{\chi(G)}$ partition G,

$$|G| \;=\; \sum_{i=1}^{\chi(G)} |S_i| \;\leqslant\; \chi(G) \max_i |S_i| \;\leqslant\; \chi(G)\, \alpha(G).$$

In particular, if $\alpha(G) \leqslant |G|/k$, then $\chi(G) \geqslant k$.

Now let $\varepsilon > 0$ be sufficiently small, let $p = n^{-(1-\varepsilon)}$, and let G be a random graph on n vertices, in which each edge is chosen independently with probability p. That is to say, we assume that we have a probability space Ω and independent random variables I_{vw} for each $vw \in \binom{[n]}{2}$, such that

$$\mathbb{P}\big(I_{vw} = 1\big) \;=\; 1 - \mathbb{P}\big(I_{vw} = 0\big) \;=\; p,$$

and we let G be the graph with vertex set $[n]$ and edge set $\big\{vw \in \binom{[n]}{2} \,:\, I_{vw} = 1\big\}$.

We claim that, with high probability, G has at most $n/2$ cycles of length at most k, and contains no independent set of size $n/2k$. The result will follow, since we may remove a *vertex* of each cycle, leaving a graph on $n/2$ vertices with girth at least k, and with no independent set of size $n/2k$, and hence chromatic number at least k.

To show that G has the desired properties, we simply count the expected number of bad events, and use **Markov's inequality**: that for any random variable which takes non-negative values, and any $a > 0$,

$$\mathbb{P}(X > a) \leqslant \frac{\mathbb{E}(X)}{a}.$$

Problem 4.1.3. *Prove Markov's inequality.*

Indeed, the expected number of cycles of length $\ell \leqslant k$ in G is at most

$$n^\ell p^\ell \leqslant n^{\varepsilon \ell} \ll n,$$

and the expected number of independent sets of size $n/2k$ is at most

$$2^n p^{n^2/8k^2} \ll 1.$$

Thus,

$$\frac{\mathbb{E}\big[|\{C_\ell \subset G : \ell \leqslant k\}|\big]}{n/2} + \mathbb{E}\Big[\big|\{S : |S| \geqslant n/2k, e(X) = 0\}\big|\Big] \to 0$$

as $n \to \infty$. Hence, by Markov, G has the desired properties with high probability, and so we are done. $\square$

The key point in the previous proof was the following: that we could choose *whichever* probability distribution (on the set of all graphs on n vertices) happened to be suitable for our needs. This gives us a great deal of flexibility in constructing proofs. In this section we shall study the simple, but surprisingly useful family of distributions that appeared in the above proof. It was first introduced by Gilbert, and independently by Erdős and Rényi, in 1959. This distribution is usually called simply 'the random graph'.

Definition. *The (Erdős-Renyi) random graph, $G_{n,p}$, is the graph on n vertices obtained by choosing each edge independently at random with probability p. That is to say, we assume that we have a probability space Ω and independent random variables I_{vw} for each $vw \in \binom{[n]}{2}$, such that*

$$\mathbb{P}(I_{vw} = 1) = 1 - \mathbb{P}(I_{vw} = 0 = p),$$

and we let G be the graph with vertex set $[n]$ and edge set $\{vw \in \binom{[n]}{2} : I_{vw} = 1\}$.

Note that $G_{n,p}$ is not actually a graph; it is a probability distribution on the graphs on n vertices. However, it will be convenient to treat it as if it were a graph.

4.2 Subgraphs of the random graph

We shall ask questions of the following form: Let $\mathcal{P}$ be a *property* of graphs (i.e., a collection of graphs closed under isomorphism, e.g., being connected, being planar, containing a triangle). For which values of p does a typical member of $G_{n,p}$ have the property $\mathcal{P}$? We say an event E (e.g., the event that $G_{n.p} \in \mathcal{P}$) holds *with high probability* (whp) if $\mathbb{P}(E) \to 1$ as $n \to \infty$. We will usually let p depend on n when computing such limits.

Question 4.2.1. *For which p is $G_{n,p}$ likely to contain a triangle?*

Note that the property 'containing a triangle' is monotone, i.e., if it holds for a graph H, then it holds for all graphs containing H. The following result says that moreover there is a *threshold* for this property, and the threshold is $1/n$.

Theorem 4.2.2.

$$\mathbb{P}\big(G_{n,p} \text{ contains a triangle}\big) \;\to\; \begin{cases} 0 & \text{if} \quad p \ll 1/n \\ 1 & \text{if} \quad p \gg 1/n \end{cases}$$

Proof. We shall prove the two parts of the theorem using the *1st and 2nd moment methods*, respectively. We have already seen an example of the 1st moment method; this simply refers to the use of Markov's inequality. Thus

$$\begin{aligned} \mathbb{P}\big(G_{n,p} \text{ contains a triangle}\big) &\leqslant \mathbb{E}\big(\text{number of triangles in } G_{n,p}\big) \\ &\leqslant \binom{n}{3} p^3 \;\ll\; 1 \end{aligned}$$

if $pn \ll 1$.

The 2nd moment method refers to the use of **Chebychev's inequality**: Let X be a random variable with mean μ and variance σ^2, and let $\lambda > 0$. Then

$$\mathbb{P}\big(|X - \mu| \geqslant \lambda\sigma\big) \;\leqslant\; \frac{1}{\lambda^2}.$$

Problem 4.2.3. *Prove Chebychev's inequality. [Hint: use Markov.]*

It follows easily from Chebychev that

$$\mathbb{P}\big(X = 0\big) \leqslant \frac{\mathrm{Var}(X)}{\mathbb{E}(X)^2}.$$

Thus, in order to prove the second part of the theorem, we need to bound the variance of the random variable

$$X = X(G) := \text{the number of triangles in } G.$$

To do so, we recall that $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$, and observe that we are summing over ordered pairs (u, v) of (potential) triangles, and that pairs which are independent of one another appear in both terms, and hence cancel out. To be precise,

$$\begin{aligned} \mathbb{E}(X^2) - \mathbb{E}(X)^2 &= \mathbb{E}\left(\sum_{(u,v)} \mathbf{1}[u]\mathbf{1}[v] \right) - \left(\sum_u \mathbb{P}(u) \right)^2 \\ &= \sum_{(u,v)} \Big(\mathbb{P}(u \wedge v) - \mathbb{P}(u)\mathbb{P}(v) \Big), \end{aligned}$$

where $\mathbf{1}$ denotes the indicator function, and u and v denote the events that specific triangles occur in $G_{n,p}$. Hence

$$\mathrm{Var}(X) \leqslant \sum_{|u \cap v| = 2} \mathbb{P}(u \wedge v) + \sum_{|u \cap v| = 3} \mathbb{P}(u \wedge v) \leqslant n^4 p^5 + n^3 p^3,$$

and thus, by Chebychev,

$$\mathbb{P}\big(G_{n,p} \text{ contains no triangles}\big) \leqslant \frac{\mathrm{Var}(X)}{\mathbb{E}(X)^2} \leqslant \frac{n^4 p^5 + n^3 p^3}{n^6 p^6} \ll 1$$

if $pn \gg 1$, as required. $\square$

This method can easily be extended to find the threshold for any complete graph K_r.

Theorem 4.2.4.

$$\mathbb{P}\big(G_{n,p} \text{ contains a copy of } K_r\big) \to \begin{cases} 0 & \text{if} \quad p \ll n^{-2/(r-1)} \\ 1 & \text{if} \quad p \gg n^{-2/(r-1)} \end{cases}$$

Problem 4.2.5. *Prove Theorem 4.2.4. Deduce that, for every function $p = p(n)$ and every n, there exists $k = k(n)$ such that the size of the largest clique in $G_{n,p}$ is, with high probability, either k or $k+1$.*

One might be tempted to conjecture that, for any graph H, the threshold for the property 'containing a copy of H' is given by

$$\mathbb{E}\big(X_H(G_{n,p})\big) \approx 1,$$

where $X_H(G)$ denotes the number of copies of H in G. However, for many graphs H the method above does not work, since the variance can be too large; moreover, the conjecture is false!

Theorem 4.2.6. *Let H denote the graph on five vertices consisting of a copy of K_4 plus an edge. Then*

$$\mathbb{P}\big(G_{n,p} \text{ contains a copy of } H\big) \to \begin{cases} 0 & \text{if} \quad p \ll n^{-2/3} \\ 1 & \text{if} \quad p \gg n^{-2/3} \end{cases}$$

Note that, for the H defined in the theorem, $\mathbb{E}\big(X_H(G_{n,p})\big) \approx n^5 p^7$, so our simple-minded prediction of the threshold was $n^{-5/7}$.

Proof. We shall apply Theorem 4.2.4 and a simple, but extremely useful technique, known as 'sprinkling'. The key observation is that the main 'obstruction' to the event $H \subset G_{n,p}$ is the event that $G_{n,p}$ contains a copy of K_4.

If $p \ll n^{-2/3}$ then the result follows immediately from Theorem 4.2.4, since if $K_4 \not\subset G$ then $H \not\subset G$. So assume that $p \gg n^{-2/3}$, and consider the graph G formed by the union of two (independent) copies of $G_{n,p/2}$. Note that $G \sim G_{n,p'}$ for $p' = p - (p/2)^2 < p$, so

$$\mathbb{P}(H \subset G) \leqslant \mathbb{P}(H \subset G_{n,p}).$$

We shall show that G contains a copy of H with high probability. Indeed, let G_1 and G_2 be the two copies of $G_{n,p/2}$, and observe that G_1 contains a copy of K_4 with high probability, by Theorem 4.2.4. Moreover, the probability that G_2 contains an edge with exactly one endpoint in this copy of K_4 is $1 - (1-p)^{4(n-4)} \approx 1 - e^{-pn} \to 1$ as $n \to \infty$. Hence $G = G_1 \cup G_2$ contains a copy of H with high probability, as required. $\square$

The reader might be wondering why we needed to 'sprinkle' the extra edges of G_2 in order to create our copy of H; why not use the edges of G_1? The reason is that, by choosing a copy of K_4, we have changed the distribution on the remaining edges: by choosing that particular copy, we have made every other edge slightly less likely to be present. Using new edges allows us to preserve independence.

Problem 4.2.7. *Come up with a conjecture regarding the threshold for the appearance of an arbitrary graph H in $G_{n,p}$. Can you prove your conjecture?*

4.3 The Janson inequalities

In the previous section we were able to prove various results using only the concepts of expected value and variance, together with the inequalities of Markov and Chebychev. In this section we shall consider two powerful theorems, which allow us to prove much stronger results when answering questions of the following type:

Question 4.3.1. *Let $X_1, \ldots, X_k$ be a collection of 'bad' events which have some weak / limited dependence on one another. What is the probability that none of the bad events occurs?*

In particular, we shall reconsider the question: What is the probability that $G_{n,p}$ contains no triangles? We shall prove the following theorem.

Theorem 4.3.2. *Let $0 < c \in \mathbb{R}$ be fixed and let $p = \dfrac{c}{n}$. Then*

$$\limsup_{n\to\infty} \mathbb{P}\big(G_{n,p} \text{ contains a triangle}\big) \;=\; 1 - e^{-c^3/6}.$$

The upper bound will follow from the FKG inequality; the lower bound from Janson's inequality.

4.3.1 The FKG inequality

We begin with the upper bound. Given a collection $\mathcal{A}$ of graphs on n vertices, we say that $\mathcal{A}$ is *monotone increasing* if $G \in \mathcal{A}$ and

$E(G) \subset E(H)$ implies that $H \in \mathcal{A}$, and that it is *monotone decreasing* if $G \in \mathcal{A}$ and $E(H) \subset E(G)$ implies that $H \in \mathcal{A}$. For example, the event "G contains a triangle" is monotone increasing, and the event "G is planar" is monotone decreasing.

The following theorem was first proved by Harris in 1960. Its generalization to a wider family of measures, proved by Fortuin, Kasteleyn and Ginibre in 1971, is a fundamental tool in Percolation Theory. For this reason it is commonly referred to as the FKG inequality.

The FKG Inequality. *Let $\mathcal{A}$ and $\mathcal{B}$ be monotone increasing, and let $\mathcal{C}$ be monotone decreasing. Then*

$$\mathbb{P}\big((G_{n,p} \in \mathcal{A}) \wedge (G_{n,p} \in \mathcal{B})\big) \geqslant \mathbb{P}\big(G_{n,p} \in \mathcal{A}\big)\mathbb{P}\big(G_{n,p} \in \mathcal{B}\big),$$

and

$$\mathbb{P}\big((G_{n,p} \in \mathcal{A}) \wedge (G_{n,p} \in \mathcal{C})\big) \leqslant \mathbb{P}\big(G_{n,p} \in \mathcal{A}\big)\mathbb{P}\big(G_{n,p} \in \mathcal{C}\big).$$

Problem 4.3.3. *Prove the FKG inequality.*

Note that, by symmetry (or by inclusion-exclusion), the first inequality also holds of $\mathcal{A}$ and $\mathcal{B}$ are both monotone decreasing.

We can now prove an upper bound on the probability that there is a triangle in $G_{n,p}$.

Proof of the upper bound in Theorem 4.3.2. For each triple $X \subset [n]$, let B_X denote the event that these vertices form a triangle in $G_{n,p}$, i.e., if $X = \{x, y, z\}$, the event that $\{xy, xz, yz\} \subset E(G_{n,p})$. The events $\overline{B_X}$ are all monotone decreasing, and have probability $(1-p)^3$. Hence, by the FKG inequality,

$$\begin{aligned}\mathbb{P}\big(\triangle \not\subset G_{n,p}\big) &= \mathbb{P}\Bigg(\bigwedge_{X \in \binom{[n]}{3}} \overline{B_X}\Bigg) \geqslant \prod_{X \in \binom{[n]}{3}} \mathbb{P}\left(\overline{B_X}\right) \\ &= \big(1 - p^3\big)^{\binom{n}{3}} \to e^{-c^3/6}\end{aligned}$$

as $n \to \infty$, as required. □

4.3.2 Janson's inequality

In this section we shall prove that, under certain circumstances, the bound given by the FKG inequality is close to being sharp.

Let $A_1, \ldots, A_m$ be subsets of $[N]$, and let the elements of a set $Y \subset [N]$ be chosen independently at random, with $\mathbb{P}(j \in Y) = p$ for each j. (In our application, we shall take $N = \binom{n}{2}$.) Let

$$B_j \ := \ \big\{ A_j \subset Y \big\},$$

and observe that these events are monotone increasing, and that if A_i and A_j are disjoint then B_i and B_j are independent.

Define $\mu = \sum_j \mathbb{P}(B_j)$ denote the expected number of bad events, and let

$$\Delta \ = \ \sum_{i \sim j} \mathbb{P}\big(B_i \wedge B_j\big),$$

where the sum is over ordered pairs (i,j) such that A_i and A_j intersect.

The Janson Inequalities (Janson, 1987). *Let $A_1, \ldots, A_m \subset [n]$ and let $B_1, \ldots, B_m$ be the events defined above. Then*

$$\mathbb{P}\bigg(\bigwedge_{j=1}^m \overline{B_j} \bigg) \ \leqslant \ e^{-\mu + \Delta/2}.$$

Moreover, if $\Delta \geqslant \mu$, then

$$\mathbb{P}\bigg(\bigwedge_{j=1}^m \overline{B_j} \bigg) \ \leqslant \ e^{-\mu^2/2\Delta}.$$

The FKG inequality implies that

$$\mathbb{P}\bigg(\bigwedge_{j=1}^m \overline{B_j} \bigg) \ \geqslant \ \prod_j \big(1 - \mathbb{P}(B_j)\big) \ \approx \ e^{-\mu},$$

if $\mathbb{P}(B_j) \ll 1$ for each j, so the Janson inequalities are close to being sharp. Moreover, when they apply, they are much more powerful than Chebychev's inequality. Indeed, let $X = \sum_j \mathbf{1}[B_j]$, and note that $\mathrm{Var}(X) \leqslant \mu + \Delta$. Thus if $1 \ll \mu \ll \Delta \ll \mu^2$ then, setting $\gamma = \mu^2/\Delta$, we obtain bounds on $\mathbb{P}(X = 0)$ of roughly $1/\gamma$ from Chebychev, and $e^{-\gamma}$ from Janson.

The original proofs of Janson are based on estimates of the Laplace transform of an appropriate random variable. The proof we give is due to Boppana and Spencer.

Proof of the 1st Janson inequality. First note that

$$\mathbb{P}\bigg(\bigwedge_{j=1}^{m} \overline{B_j} \bigg) \;=\; \prod_{j=1}^{m} \mathbb{P}\bigg(B_j \,\bigg|\, \bigwedge_{i=1}^{j-1} \overline{B_i} \bigg).$$

Using this identity, and the inequality $1 - x \leqslant e^{-x}$, it suffices to show that, for each $j \in [m]$,

$$\prod_{j=1}^{m} \mathbb{P}\bigg(B_j \,\bigg|\, \bigwedge_{i=1}^{j-1} \overline{B_i} \bigg) \;\geqslant\; \mathbb{P}(B_j) \;-\; \sum_{i \sim j,\, i<j} \mathbb{P}(B_i \wedge B_j).$$

To prove this, we shall break up the event $\bigwedge_{i=1}^{j-1} \overline{B_i}$ into two parts, E and F. Let E denote the event $\bigwedge_{i \in I} \overline{B_i}$, where

$$I \;=\; \big\{ i \in [j-1] \,:\, i \sim j \big\}$$

(i.e., the set of indices i such that $A_i \cap A_j$ is non-empty), and let F denote the event $\bigwedge_{i \in J} \overline{B_i}$, where $J = [j-1] \setminus I$. Note that the event B_j is independent of the event E, and that the events E and F are both monotone decreasing. Therefore, by the FKG inequality,

$$\begin{aligned}
\mathbb{P}\big(B_j \,\big|\, E \cap F\big) \;&\geqslant\; \mathbb{P}\big(B_j \wedge E \,\big|\, F\big) \;=\; \mathbb{P}\big(B_j \,|\, F\big)\mathbb{P}\big(E \,\big|\, B_j \wedge F\big) \\
&\geqslant\; \mathbb{P}\big(B_j\big)\mathbb{P}\big(E \,|\, B_j\big) \;\geqslant\; \mathbb{P}(B_j)\bigg(1 \;-\; \sum_{i \in I} \mathbb{P}\big(B_i \,|\, B_j\big) \bigg) \\
&=\; \mathbb{P}(B_j) \;-\; \sum_{i \sim j,\, i<j} \mathbb{P}(B_i \wedge B_j),
\end{aligned}$$

as required. □

The proof of the second inequality uses a little bit of magic.

Proof of the 2nd Janson inequality. We deduce the desired bound by applying the 1st Janson inequality 2^m times: we use it to bound the probability of $\bigwedge_{j \in S} \overline{B_j}$ for every subset $S \subset [m]$. Indeed, we have

$$\log\left(\mathbb{P}\bigg(\bigwedge_{j \in S} \overline{B_j}\bigg)\right) \leqslant -\sum_{j \in S} \mathbb{P}(B_j) + \frac{1}{2} \sum_{i,j \in S,\, i \sim j} \mathbb{P}(B_i \wedge B_j)$$

for any set $S \subset [m]$.

Now the magic: we choose a set S *randomly*! To be precise, choose the elements of S independently, each with probability q. It follows that

$$\mathbb{E}\left[\log\left(\mathbb{P}\bigg(\bigwedge_{j \in S} \overline{B_j}\bigg)\right)\right] \leqslant -q\mu + q^2 \frac{\Delta}{2},$$

so, for every $q \in [0,1]$, there must exist a set S_q such that

$$\log\left(\mathbb{P}\bigg(\bigwedge_{j \in S} \overline{B_j}\bigg)\right) \leqslant -q\mu + q^2 \frac{\Delta}{2}.$$

Finally, we choose $q = \dfrac{\mu}{\Delta}$ to minimize the right-hand side, and obtain

$$\mathbb{P}\bigg(\bigwedge_{j=1}^{m} \overline{B_j}\bigg) \leqslant \mathbb{P}\bigg(\bigwedge_{j \in S_q} \overline{B_j}\bigg) \leqslant e^{-\mu^2/2\Delta},$$

as required. □

We can now complete the proof of Theorem 4.3.2

Proof of the lower bound in Theorem 4.3.2. Once again, for each triple $X = \{x, y, z\} \subset [n]$ we write B_X for the event that these vertices form a triangle in $G_{n,p}$. Thus we have a set of $\binom{n}{3}$ events B_X of the form $A_X \subset E(G_{n,p}) \subset [N]$, where $N = \binom{n}{2}$.

Observe that $\mu = \sum_X \mathbb{P}(B_X) = p^3 \binom{n}{3}$, and that $\Delta = \sum_{X \sim Y} \mathbb{P}(B_X \wedge B_Y) = O(n^4 p^5) \ll 1$. Hence by the 1st Janson inequality,

$$\mathbb{P}\big(\triangle \not\subset G_{n,p}\big) \leqslant e^{-\mu + \Delta/2} \to e^{-c^3/6}$$

as $n \to \infty$, as required. □

4.4 Connectedness

In the previous section we showed that some 'local' events – that $G_{n,p}$ contains a copy of a given small subgraph – have thresholds of the following form: if $\mathbb{P}(H \subset G_{n,p}) \geqslant \varepsilon$ as $n \to \infty$, then $\mathbb{P}(H \subset G_{n,Cp}) \geqslant 1 - \varepsilon$, as $n \to \infty$, for some large constant C. In this section we shall investigate whether a similar phenomenon occurs for the following 'global' property:

Question 4.4.1. *For which p is $G_{n,p}$ likely to be connected?*

Note that the property 'being connected' is again monotone – adding edges cannot disconnect a graph. The following theorem says that this property undergoes a much more sudden transition than those considered in the previous section.

Theorem 4.4.2. *For every $\varepsilon > 0$,*

$$\mathbb{P}\big(G_{n,p} \text{ is connected}\big) \;\to\; \begin{cases} 0 & \text{if} \quad p \leqslant \dfrac{(1-\varepsilon)\log n}{n} \\[2ex] 1 & \text{if} \quad p \geqslant \dfrac{(1+\varepsilon)\log n}{n} \end{cases}$$

as $n \to \infty$.

Proof. We shall use the 1st and 2nd moment methods to prove the second and first parts, respectively. The key point is to choose the right thing to count! It turns out that, for the random graph, the obstruction to being connected is avoiding having isolated vertices.

Indeed, let $X = X(G)$ denote the number of isolated vertices in G, and let $p = \frac{(1-\varepsilon)\log n}{n}$. Then

$$\mathbb{E}\big(X(G_{n,p})\big) \;=\; n(1-p)^{n-1} \;>\; ne^{-pn-p^2 n} \;\gg\; n^{\varepsilon/2}$$

for sufficiently large n. Moreover, the variance of $X(G_{n,p})$ is bounded by

$$\begin{aligned} \mathrm{Var}\big(X(G_{n,p})\big) \;&\leqslant\; \sum_{(u,v)} \Big(\mathbb{P}(u \wedge v) - \mathbb{P}(u)\mathbb{P}(v)\Big) \\ &\leqslant\; \mathbb{E}(X) \,+\, n^2\Big((1-p)^{2n-3} - (1-p)^{2n-2}\Big) \\ &=\; \mathbb{E}(X) \,+\, \frac{p}{1-p}\mathbb{E}(X)^2. \end{aligned}$$

Hence $\text{Var}(X) \ll \mathbb{E}(X)^2$, and so, by Chebychev,

$$\mathbb{P}\big(G_{n,p} \text{ is connected}\big) \;\to\; 0$$

as $n \to \infty$.

To prove the second part, we consider the random variable $Y_k(G)$, which counts the number of components of size k in G. We have

$$\mathbb{E}\big(Y_k(G_{n,p})\big) \;\leqslant\; \binom{n}{k}(1-p)^{k(n-k)} \;\leqslant\; \left(\frac{en}{k}e^{-(1-\varepsilon)pn}\right)^k$$

if $k \leqslant \varepsilon n$, and similarly $\mathbb{E}\big(Y_k(G_{n,p})\big) \leqslant 2^n e^{-\varepsilon n^2/2}$ if $k \geqslant \varepsilon n$.

Let $p = \frac{(1+3\varepsilon)\log n}{n}$, and note that if G is not connected, there must exist a component of size at most $n/2$. Hence, by Markov,

$$\begin{aligned}\mathbb{P}\big(G_{n,p} \text{ is not connected}\big) \;&\leqslant\; \sum_{k=1}^{n/2} \mathbb{E}\big(Y_k(G_{n,p})\big)\\ &\leqslant\; \sum_{k=1}^{\varepsilon n} \left(\frac{en}{k}e^{-(1-\varepsilon)pn}\right)^k \;+\; \sum_{k=\varepsilon n}^{n/2} 2^n e^{-\varepsilon n^2/2}\\ &\leqslant\; \sum_{k=1}^{\infty} \left(\frac{e}{k}n^{-\varepsilon}\right)^k \;+\; e^{-\varepsilon n^2/3} \;\to\; 0\end{aligned}$$

as $n \to \infty$, as required. $\square$

In the proof above we used the inequality $1 - x > e^{-x-x^2}$, which holds for sufficiently small $x > 0$. We also used the usual bound $\binom{n}{k} \leqslant \left(\frac{en}{k}\right)^k$ on the binomial coefficient.

4.5 The giant component

The most famous property of the random graph was also one of the first to be discovered, by Erdős and Renyi in 1960. Despite this, it is still the object of extensive ongoing research. It deals with the following question:

Question 4.5.1. *How big is the largest connected component of $G_{n,p}$?*

Let $C_1(G)$ denote the largest component of a graph G, i.e., the connected set with the largest number of vertices. Most of the following theorem was proved by Erdős and Renyi in their original paper on $G_{n,p}$; the final part was added much later by Bollobás, who also determined the size of the 'window' where the phase transition occurs.

Theorem 4.5.2 (Erdős and Renyi 1960, Bollobás 1984)**.** *Let $c > 0$ be a constant, and let $p = c/n$. Then*

$$C_1\big(G_{n,p}\big) = \begin{cases} O\big(\log n\big) & \text{if} \quad c < 1 \\ \Theta\big(n^{2/3}\big) & \text{if} \quad c = 1 \\ \Theta(n) & \text{if} \quad c > 1 \end{cases}$$

with high probability as $n \to \infty$. Moreover, if $c \neq 1$ then the size of the second largest component is $O(\log n)$ with high probability.

We call the unique component of size $\Theta(n)$ in the case $c > 1$ the *giant component.*

Problem 4.5.3. *By counting the expected number of paths of length k, show that if $c < 1$ then $C_1\big(G_{n,p}\big) = O(\log n)$ with high probability.*

Problem 4.5.4. *Prove that if $c > 1$ and $p = c/n$, then $C_1\big(G_{n,p}\big) = \Theta(n)$ with high probability.*

4.6 Recommended Further Reading

Bollobás' seminal book helped popularize the field and is still a useful reference.

B. Bollobás, Random Graphs (2nd edition), Cambridge University Press (2001).

More recent results are covered in Alon and Spencer [1] and:

S. Janson, T. Łuczak and A. Rucinski, Random Graphs, Wiley Interscience (2000).

4.7 Exercises

The clique number $\omega(G)$ of a graph G is the size of the largest clique in G. We shall write whp to mean with high probability as $n \to \infty$.

1. Show that $\omega(G_{n,1/2}) \in \{k, k+1\}$ whp, for some $k = k(n)$.

2. In a tournament, each pair out of n teams plays each other once, and there are no draws. For $n \geqslant n_0(k)$ large, must there exist a set A of k teams such that no team beat everyone in A?

3. (*a*) Show that $\omega(G_{n,1/2}) = \big(2 + o(1)\big) \log n$ whp.

   (*b*) Deduce that $\chi(G_{n,1/2}) \gtrsim \dfrac{n}{2 \log n}$ whp.

   (*c*) Show that, whp, *every* subset of $V(G_{n,1/2})$ of size $n/(\log n)^2$ contains an independent set of size $k - 4$. [Hint: use Janson.]

   (*d*) Deduce that $\chi(G_{n,1/2}) \sim \dfrac{n}{2 \log n}$ whp.

4. Find the threshold function for the property "G is bipartite". Is it sharp?

5. Prove the following 'asymptotic' version of Mantel's Theorem on $G_{n,p}$ (with p constant):

Mantel's Theorem on $G_{n,p}$. *With high probability, every subgraph $H \subset E(G_{n,p})$ with*

$$e(H) \geqslant \left(\frac{1}{2} + \varepsilon\right) e(G_{n,p})$$

contains a triangle.

Can you prove it when $p = p(n) \to 0$? Show it is false when $p \ll 1/\sqrt{n}$.

Chapter 5

Topological and Algebraic Methods

Everyone else would climb a peak by looking for a path somewhere in the mountain. Nash would climb another mountain altogether, and from that distant peak would shine a searchlight back onto the first peak.

Donald Newman

In earlier chapters, we have seen how ideas from Probability can sometimes be used to solve problems in Combinatorics. We shall now discuss some surprising applications from two other areas of mathematics: Topology and Linear Algebra. In particular, we shall see a famous application of the Borsuk-Ulam Theorem, due to Lovász, and the stunning disproof of Borsuk's Conjecture, by Kahn and Kalai.

5.1 The Borsuk-Ulam Theorem

A Ham Sandwich is a family of three (disjoint) measureable subsets of $\mathbb{R}^3$, which we shall call B (the bread), H (the ham) and X (the cheese). The sandwich is 'fairly divided in two' if each of B, H and X is partitioned into two equal-size pieces.

Question 5.1.1. *Given a Ham Sandwich, must there exist a plane which fairly divides the sandwich in two?*

This question was first posed by Steinhaus in 1938, and was solved by Banach. The proof uses the following famous theorem, which was conjectured by Ulam and proved by Borsuk in 1933. Let $S^n \subset \mathbb{R}^{n+1}$ denote the n-dimensional sphere.

The Borsuk-Ulam Theorem. *For any continuous function $f : S^n \to \mathbb{R}^n$, there exists a point $x \in S^n$ such that $f(x) = f(-x)$.*

To get a feel for this theorem, let's use it to prove the Ham Sandwich Theorem, i.e., that the answer to Question 5.1.1 is yes.

Proof of the Ham Sandwich Theorem. Given a Ham Sandwich, we define a continuous function $f : S^2 \to \mathbb{R}^2$ as follows. For each $y \in S^2$, let $P(y)$ denote the plane perpendicular to y which splits the bread into two equal pieces. Now define

$$f(x) \;:=\; \big(a(y), b(y)\big),$$

where $a(y)$ is the amount of ham on the 'positive' side of $P(y)$, and $b(y)$ is defined analogously for the cheese.

Since f is continuous, by the Borsuk-Ulam Theorem there exists $y \in S^2$ such that $f(y) = f(-y)$. But $P(y) = P(-y)$, so this means that there is an equal amount of bread, ham and cheese on both sides of $P(y)$, as required. $\square$

In the next section we shall give a more combinatorial application of the Borsuk-Ulam Theorem. In order to do so, it will be useful to restate the theorem as follows.

Theorem 5.1.2 (Lyusternik and Shnirelman, 1930)**.** *Let $A_0, \ldots, A_n$ be a cover of S^n, such that each A_j is either open or closed. Then there is a set A_j which contains two antipodal points, x and $-x$.*

Proof. Suppose first that the sets A_j are all closed, and define a continuous function $f : S^n \to \mathbb{R}^n$ by

$$f(x) \;:=\; \big(d(x, A_1), \ldots, d(x, A_n)\big),$$

where $d(x, A_j) = \inf_{y \in A_j} d(x, y)$ denotes the ℓ_2-distance (say) from x to A_j. By the Borsuk-Ulam Theorem, we have $f(x) = f(-x)$ for some $x \in S^n$. Thus, if $d(x, A_j) = 0$ for some $j \in [n]$, then $x, -x \in A_j$, since the A_j are closed. But if not, then $x, -x \in A_0$, since the A_j form a cover, so we are done.

Next suppose that the sets A_j are all open. By compactness, there exist closed sets $B_j \subset A_j$ which also cover S^n. (To see this, for each $x \in A_j$ choose an open set $N_x \ni x$ such that $\overline{N}_x \subset A_j$, and take a finite sub-cover.) Thus the result follows by the first case.

Finally, suppose $F_1, \ldots, F_k$ are closed and the rest are open, and replace each closed set with its ε-neighbourhood. Applying the previous case for a sequence $\varepsilon(j) \to 0$, and passing to a subsequence, we find that (wlog) there exist $x_j, -x_j \in (F_1)_{\varepsilon(j)}$ for each $j \in \mathbb{N}$. Taking a convergent subsequence, and recalling that F_1 is closed, the theorem follows. $\square$

In fact, it turns out that the Theorem 5.1.2 is equivalent to the Borsuk-Ulam Theorem; we leave the proof of the reverse implication to the reader.

5.2 The Kneser graph

At the start of Chapter 4, we remarked that it is non-trivial to construct triangle-free graphs with high chromatic number. The following family of graphs have this property, and were first studied by Martin Kneser in 1955.

Definition. *The Kneser graph,* $\text{Kn}(n, k)$*, has vertex set* $\binom{[n]}{k}$*, the* k*-subset of* $[n]$*, and edge set* $E = \big\{\{S, T\} : S \text{ and } T \text{ are disjoint}\big\}$.

Note that if $k \geqslant n/3$, then $\text{Kn}(n, k)$ is triangle-free, and has independent sets of size $\binom{n}{k-1} \geqslant |V|/3$ (take every k-set containing element 1, say). Nevertheless, Kneser conjectured that the chromatic number of $\text{Kn}(n, k)$ is $n - 2k + 2$, and hence grows with n if $k \approx n/3$, say. Lovász's proof of this conjecture, in 1978, was the first application of topological methods in Combinatorics.

Theorem 5.2.1 (Lovász, 1978)**.** *If* $n \geqslant 2k - 1$*, then*

$$\chi\big(\text{Kn}(n, k)\big) \;=\; n - 2k + 2.$$

Proof. To colour $\mathrm{Kn}(n,k)$ with $n-2k-2$ colours, simply give colour $2k-1$ to every subset of $[2k-1]$, and colour every other k-set with its maximal element. Then every two sets with the same colour intersect, so this a proper colouring.

To prove that $d = n-2k+1$ colours do not suffice, let $X \subset S^d$ be a set of n points in general position; that is, every hyperplane through the origin contains at most d points. Let $\mathcal{A}_j$ denote the collection of vertices of $\mathrm{Kn}(n,k)$ (i.e., k-subsets of X) of colour j, for each $j \in [d]$.

Now, define a cover of S^d as follows: for each $j \in [d]$ and $x \in S^d$, let $x \in U_j$ if there is a k-set of colour j in the open hemisphere whose pole is x. Let $U_0 = S^d \backslash (U_1 \cup \ldots \cup U_d)$, and observe that each U_j $(j \geqslant 1)$ is open, and hence U_0 is closed.

By the Borsuk-Ulam Theorem (applied as in Theorem 5.1.2), there exist $x \in S^d$ and $0 \leqslant j \leqslant d$ such that $x, -x \in U_j$. If $j = 0$ then $x, -x \notin U_j$ for every $j \geqslant 1$, and so the hemispheres with poles x and $-x$ each contain at most $k-1$ elements of X. But then the equator must contain at least $n - 2(k-1) = d+1$ points of X, which is a contradiction (since they are in general position).

So $x, -x \in U_j$ for some $j \geqslant 1$, which means that there are two k-sets of colour j lying in opposite hemispheres. But these k-sets are disjoint, so this is again a contradiction, which proves the theorem. $\square$

The proof above is somewhat simpler that Lovász's original proof, and is due to Joshua Greene, who was an undergraduate student at the time. For many other combinatorial applications of the Borsuk-Ulam Theorem, see the book by Matoušek [4].

5.3 Linear Algebra

We next turn to applications from a different, but equally fundamental area of mathematics: Linear Algebra. The basic idea is straightforward: we translate a combinatorial problem into one about (for example) vector spaces, or polynomials; then we apply a general result about such objects. We shall give only a taste of the plethora of beautiful combinatorial results which can be proven in this way; many of these have no known purely combinatorial proof.

We begin with a simple question.

Question 5.3.1 (Oddtown)**.** *Suppose $\mathcal{A} \subset \mathcal{P}(n)$ is a collection of clubs, each of which has an odd number of members, and any two of which share an even number of members. How large can $|\mathcal{A}|$ be?*

Answer: $|\mathcal{A}| \leqslant n$.

Proof. Let $\mathcal{A} \subset \mathcal{P}(n)$ be a set-system such that $|A|$ is odd for every $A \in \mathcal{A}$ and $|A \cap B|$ is even for every $A, B \in \mathcal{A}$ with $A \neq B$. We shall map the family $\mathcal{A}$ into the vector space $(\mathbb{F}_2)^n$ in the simplest way possible: A maps to its *characteristic vector*.

To be precise, for each $A \in \mathcal{P}(n)$, let $\chi_{\mathbf{A}} \in (\mathbb{F}_2)^n$ denote the vector such that $(\chi_{\mathbf{A}})_j = 1$ if and only if $j \in A$. Then the condition "$|A|$ is odd" translates to "$\|\chi_{\mathbf{A}}\| \neq 0$", and the condition "$|A \cap B|$ is even" translates to "$\chi_{\mathbf{A}} \cdot \chi_{\mathbf{B}} = 0$". (Note that we are working over the field with two elements!)

We claim that the vectors $\Xi = \{\chi_{\mathbf{A}} : A \in \mathcal{A}\}$ are independent. Indeed, suppose that

$$\sum_{A \in \mathcal{A}} \lambda_A \chi_{\mathbf{A}} \;=\; 0$$

for some collection $\lambda_A \in \mathbb{R}$. Then, taking the scalar product with some set $B \in \mathcal{A}$, we obtain $\lambda_B = 0$. But B was arbitrary, so the collection Ξ is indeed linearly independent.

But the dimension of $(\mathbb{F}_2)^n$ is n, which means that there do not exist $n+1$ linearly independent vectors, and so we are done.

Finally, note that this bound is best possible, since each resident of Oddtown may set up a club consisting only of himself. □

Problem 5.3.2 (+)**.** *Find a purely combinatorial proof that Oddtown can have at most n clubs.*

5.4 The Frankl-Wilson inequalities

Given $L \subset [n]$, say a set-system $\mathcal{A} \subset \mathcal{P}(n)$ (that is, a collection of subsets of $[n]$) is *L-intersecting* if $|A \cap B| \in L$ for every $A, B \in \mathcal{A}$ with $A \neq B$. In this section we shall consider the following question.

Question 5.4.1. *How large can an L-intersecting set-system be?*

In the following few pages we shall answer this question in a 'modular' setting (i.e., in the group $\mathbb{Z}_p$), and give a beautiful application, due to Kahn and Kalai. Two more applications can be found in the exercises.

The following theorem was proved by Deza, Frankl and Singhi in 1983, but for historical reasons (see Theorem 5.4.5 below) it is usually referred to as the Modular Frankl-Wilson inequality.

The Modular Frankl-Wilson Inequality. *Let p be prime, and let $L \subset \{0, \ldots, p-1\}$ with $|L| = s$. Let $\mathcal{A}$ be a family of subsets of $[n]$ such that $|A| \notin L \pmod p$ for every $A \in \mathcal{A}$, but $|A \cap B| \in L \pmod p$ for any pair of distinct sets $A, B \in \mathcal{A}$. Then*

$$|\mathcal{A}| \leqslant \sum_{j=0}^{s} \binom{n}{j}.$$

Note that Theorem 5.4, applied in the case $p = 2$ and $L = \{0\}$, says that Oddtown has at most $n + 1$ clubs.

The proof of Theorem 5.4 is similar to the Oddtown proof described above, except that instead of considering the characteristic vectors of the sets in $\mathcal{A}$, we shall define a family of polynomials, one for each set $A \in \mathcal{A}$, which are linearly independent in some vector space. We shall use the following easy lemma to show that the polynomials are linearly independent.

Lemma 5.4.2. *Let $\mathbb{F}$ be a field and $v_1, \ldots, v_m \in \mathbb{F}^n$. Suppose there exist functions $u_1, \ldots, u_m : \mathbb{F}^n \to \mathbb{F}$ such that $u_j(v_i) = 0$ for all $1 \leqslant i < j \leqslant m$ and $u_i(v_i) \neq 0$ for all $i \in [m]$. Then $v_1, \ldots, v_m$ are linearly independent.*

Proof. Suppose $\sum_i \lambda_i v_i = 0$, and apply u_j. It follows that $\lambda_j = 0$ for all $j \in [m]$. $\square$

We can now prove the Modular Frankl-Wilson inequality.

Proof of Theorem 5.4. For each $A \in \mathcal{A}$, we define a polynomial $f_A : (\mathbb{F}_p)^n \mapsto \mathbb{F}_p$ as follows:

$$f_A(\mathbf{x}) := \prod_{j=1}^{s} \big(\mathbf{x} \cdot \chi_{\mathbf{A}} - \ell_j\big).$$

Observe that, for each $A, B \in \mathcal{A}$ with $A \neq B$,

$$f_A(\chi_{\mathbf{A}}) \;=\; \prod_{j=1}^{s} \big(|A| - \ell_j\big) \;\neq\; 0,$$

and that

$$f_A(\chi_{\mathbf{B}}) \;=\; \prod_{j=1}^{s} \big(|A \cap B| - \ell_j\big) \;=\; 0.$$

Hence, by Lemma 5.4.2, it follows that the polynomials $\{f_A : A \in \mathcal{A}\}$ are linearly independent. (Exercise: write down a linear functional corresponding to the vector $\chi_{\mathbf{A}}$.)

Let F denote the space of polynomials of degree t most s, and note that $f_A \in F$ for all $A \in \mathcal{A}$. Hence

$$|\mathcal{A}| \;\leqslant\; \dim(F) \;\leqslant\; \sum_{j=0}^{s} \binom{n+j-1}{j},$$

since there are exactly $\binom{n+j-1}{j}$ non-negative integer solutions to the equation $d_1 + \dots d_n \;=\; j$, and the dimension of F is equal to the number of monomials $x_1^{d_1} \dots x_n^{d_n}$ with $d_1 + \cdots + d_n \leqslant s$.

This bound is close to, but not quite as good as the one we want. The following cute trick resolves the situation:

Replace each polynomial f_A by the polynomial $\overline{f}_A$ obtained by reducing, for each monomial, the exponent of each variable to one. (For example, if $f(x, y, z) \;=\; x^2y + y^3z^4 + xyz$, then $\overline{f}(x, y, z) \;=\; xy + yz + xyz$.) The crucial (but trivial) observation is that f and $\overline{f}$ take the same values on $\{0, 1\}^n$, and hence that these 'reduced' polynomials still satisfy the properties we want, i.e., $\overline{f}_A(\chi_{\mathbf{A}}) \neq 0$ and $\overline{f}_A(\chi_{\mathbf{B}}) = 0$ for every $A \neq B \in \mathcal{A}$.

Hence the polynomials $\{\overline{f}_A : A \in \mathcal{A}\}$ are linearly independent, by Lemma 5.4.2, and so, writing F' for the space of *multilinear* polynomials of degree at most s, we obtain

$$|\mathcal{A}| \;\leqslant\; \dim(F') \;\leqslant\; \sum_{j=0}^{s} \binom{n}{j},$$

as required. $\square$

Sometimes non-modular results can be proved using the MFWI.

Theorem 5.4.3 (The Weak Fisher inequality). *Let $0 \leqslant \ell \leqslant n$, and suppose that $\mathcal{A} \subset \mathcal{P}(n)$ is such that $|A \cap B| = \ell$ for every $A, B \in \mathcal{A}$ with $A \neq B$. Then $|\mathcal{A}| \leqslant n + 1$.*

Finally, we note two other related theorems, which can be proved in the same way (see that exercises). Recall that $\mathcal{A}$ is k-uniform if $|A| = k$ for every $A \in \mathcal{A}$, and is L-intersecting if $|A \cap B| \in L$ for every $A, B \in \mathcal{A}$ with $A \neq B$.

Theorem 5.4.4 (Ray-Chaudhuri and Wilson, 1975). *Let $L \subset \mathbb{N}$ with $|L| = s$, and let $\mathcal{A} \subset \mathcal{P}(n)$ be k-uniform and L-intersecting. Then*

$$|\mathcal{A}| \leqslant \binom{n}{s}.$$

Theorem 5.4.5 (Frankl and Wilson, 1981). *Let $L \subset \mathbb{N}$ with $|L| = s$, and let $\mathcal{A} \subset \mathcal{P}(n)$ be L-intersecting. Then*

$$|\mathcal{A}| \leqslant \sum_{j=0}^{s} \binom{n}{j}.$$

5.5 Borsuk's Conjecture

In 1933, the same year as his famous proof of the Borsuk-Ulam Theorem, Borsuk asked the following question. Suppose $X \subset \mathbb{R}^n$, into how many pieces must we divide X such that each piece has smaller diameter than X?

Definition. *For each $n \in \mathbb{N}$, let* $\mathrm{Bor}(n)$ *denote the smallest integer such that every bounded set in $\mathbb{R}^n$ can be partitioned into* $\mathrm{Bor}(n)$ *sets of smaller diameter.*

The regular n-dimensional simplex of diameter 1, shows that $\mathrm{Bor}(n) \geqslant n + 1$, since each vertex must be in a different set. Borsuk conjectured that this bound is sharp.

Borsuk's Conjecture. $\mathrm{Bor}(n) = n + 1$ *for every $n \in \mathbb{N}$.*

The conjecture was open for 60 years, before the following spectacular disproof was discovered by Kahn and Kalai.

Theorem 5.5.1 (Kahn and Kalai, 1993)**.** $\mathrm{Bor}(m) > 1.1^{\sqrt{m}}$ *for all sufficiently large* $m \in \mathbb{N}$*. In particular, Borsuk's Conjecture is false in all sufficiently high dimensions.*

Proof. The plan is as follows: we shall construct an injection Φ from a set system $\mathcal{A}$ to a sphere in m dimensions. We shall show that, if $\mathcal{B} \subset \mathcal{A}$ is such that $\Phi(\mathcal{B})$ has smaller diameter than $\Phi(\mathcal{A})$, then there is a restriction on the possible intersection sizes of sets in $\mathcal{B}$. This will allow us to bound $|\mathcal{B}|$ from above (using the MFWI), and hence bound $\mathrm{Bor}(m)$ from below.

With foresight, choose $n \in \mathbb{N}$ so that $m/2 \leqslant \binom{n}{2} + n \leqslant m$, and choose a prime p such that $n/4 \leqslant p = n/4 + o(n)$ (this is possible by known results on the distribution of primes). Set $\mathcal{A} = \binom{[n]}{2p-1}$; we shall show that if $\mathcal{B} \subset \mathcal{A}$ corresponds to set of sub-maximal diameter, then $|\mathcal{B}| \leqslant 2\binom{n}{p-1}$.

We begin by choosing an orthonormal basis

$$\big\{e_i : i \in [n]\big\} \,\cup\, \big\{f_{ij} : 1 \leqslant i < j \leqslant n\big\}$$

of $\mathbb{R}^k$, where $k = \binom{n}{2} + n$. Define a function $\Phi : \mathbb{R}^n \to \mathbb{R}^k$ by

$$\Phi(x_1, \ldots, x_n) \;:=\; \sum_{i<j} x_i x_j f_{ij} \;+\; \alpha \sum_i x_i e_i,$$

where $\alpha > 0$ will be chosen later, and note that Φ is injective. A simple calculation (exercise) shows that

$$\Phi(\mathbf{x}) \cdot \Phi(\mathbf{y}) \;=\; \frac{1}{2}(\mathbf{x} \cdot \mathbf{y})^2 \;+\; \alpha^2\, \mathbf{x} \cdot \mathbf{y} \;-\; \frac{n}{2},$$

if $\mathbf{x} \in \{-1, 1\}^n$. Since $\mathbf{x} \cdot \mathbf{x} = n$ for such $\mathbf{x}$, it follows that

$$\|\Phi(\mathbf{x})\|^2 \;=\; \frac{n^2}{2} \;+\; \alpha^2\, n \;-\; \frac{n}{2}$$

for every $\mathbf{x} \in \{-1, 1\}^n$.

Let us identify the family $\mathcal{A}$ with the set $\big\{\mathbf{v_A} : A \in \mathcal{A}\big\} \subset \{-1, 1\}^n$, where $(\mathbf{v_A})_j = 1$ if $j \in A$, and $(\mathbf{v_A})_j = -1$ otherwise. By the observations above, $\Phi(\mathcal{A})$ is a subset of a sphere centred at the origin. It follows that two points of $\Phi(\mathcal{A})$ are at distance $\mathrm{diam}(\Phi(\mathcal{A}))$ if and only if their scalar product is equal to $\min\big\{\Phi(\mathbf{x}) \cdot \Phi(\mathbf{y}) : \mathbf{x}, \mathbf{y} \in \mathcal{A}\big\}$.

The minimum of the quadratic $(\mathbf{x} \cdot \mathbf{y})^2/2 + \alpha^2(\mathbf{x} \cdot \mathbf{y}) - n/2$ is obtained at $\mathbf{x} \cdot \mathbf{y} = -\alpha^2$. We would like to choose α so that this minimum is attained by $\Phi(\mathcal{A})$, and so that the minimum corresponds to a particular intersection size mod p. Note that

$$\mathbf{v_A} \cdot \mathbf{v_B} \;=\; n - 2\big(|A| + |B| - 2|A \cap B|\big),$$

and set $\alpha = \sqrt{4p - n}$. Then, for each $A, B \in \mathcal{A} = \binom{[n]}{2p-1}$, we have

$$|A \cap B| \;=\; p - 1 \;\Leftrightarrow\; \mathbf{v_A} \cdot \mathbf{v_B} \;=\; n - 4p \;=\; -\alpha^2.$$

The diameter of $\Phi(\mathcal{A})$ is thus realised by all pairs $\{A, B\} \subset \mathcal{A}$ such that $|A \cap B| = p - 1$. Therefore, if $\mathcal{B} \subset \mathcal{A}$ is such that $\Phi(\mathcal{B})$ has a smaller diameter than $\Phi(\mathcal{A})$, then $|A \cap B| \neq p - 1 \pmod p$ for every $A, B \in \mathcal{B}$ with $A \neq B$.

Note also that, since $\mathcal{A} = \binom{[n]}{2p-1}$, we have $|A| = p - 1 \pmod p$ for every $A \in \mathcal{B}$. Hence, by the MFWI applied with $L = [0, p-2]$, we obtain

$$|\mathcal{B}| \;\leqslant\; \sum_{j=0}^{p-1} \binom{n}{j} \;\leqslant\; 2\binom{n}{p-1}.$$

But then, using Stirling's formula (exercise),

$$\mathrm{Bor}(m) \;\geqslant\; \mathrm{Bor}(k) \;\geqslant\; \frac{|\mathcal{A}|}{2\binom{n}{p-1}} \;=\; \frac{\binom{n}{2p-1}}{2\binom{n}{p-1}} \;\geqslant\; \big(1.1 + o(1)\big)^{\sqrt{m}},$$

as required. $\square$

Remark. *The best known bounds on* $\mathrm{Bor}(n)$ *are only*

$$(1.225)^{\sqrt{n}} \;<\; \mathrm{Bor}(n) \;<\; (1.3)^n.$$

The lower bound follows from a slight modification of the argument above; the upper bound was proved by Schramm in 1988. Borsuk's Conjecture is true in two or three dimensions, and in various other special cases (e.g., for bodies with a smooth surface). The smallest n for which the conjecture is known to fail is 298.

5.6 Recommended further reading

For many more applications of Borsuk-Ulam, see Matoušek [4]. Much of the material from this section is discussed in:

Y. Kohayakawa and C. Moreira, Tópicos em Combinatória Contemporânea, 23o. Colóquio Brasileiro de Matemática (2001)

and also

L. Babai and P. Frankl, Linear Algebra Methods in Combinatorics With Applications to Geometry and Computer Science (1992).

The articles by Alon, Godsil (with an appendix by Lovász) and Björner in the following book are also very useful:

R. Graham, M. Grötschel and L. Lovász, Handbook on Combinatorics, MIT Press (1996).

Finally, we highly recommend the collection of problems:

B. Bollobás, The Art of Mathematics: Coffee Time in Memphis, Cambridge University Press (2006),

which covers everything in this chapter, and much much more.

5.7 Exercises

1. Use the Borsuk-Ulam Theorem to solve the Necklace Problem.

The Necklace Problem. *Suppose a necklace has n kinds of beads, and an even number of each kind. How many cuts are required in order to divide the beads equally between two thieves?*

2. Prove the Ray-Chaudhuri-Wilson Theorem (Theorem 5.4.4).

3. Prove the Frankl-Wilson Theorem (Theorem 5.4.5).

The chromatic number of $\mathbb{R}^n$:

Let $\chi(\mathbb{R}^n)$ denote the chromatic number of the (infinite) graph with vertex set $\mathbb{R}^n$, and edge set $\big\{xy : \|x - y\| = 1\big\}$.

4. (a) Show that $4 \leqslant \chi(\mathbb{R}^2) \leqslant 7$.

   (b) Show that $\chi(\mathbb{R}^n) \leqslant C^n$ for some $C > 0$.

   (c) Use the MFWI to show that $\chi(\mathbb{R}^n) \geqslant (1+\varepsilon)^n$ for some $\varepsilon > 0$.

5. Give a constructive super-polynomial lower bound on $R(k)$.

Hints

2. Consider the polynomials $f_A(\mathbf{x}) = \prod_i \Big(\sum_{j \in A} x_j - \ell_i \Big)$.

3. Consider the polynomials $f_A(\mathbf{x}) = \prod_{L \ni \ell < |A|} \Big(\langle \chi_{\mathbf{A}}, \mathbf{x} \rangle - \ell \Big)$.

4. (c) Choose a prime p with $4p < n < 8p$, let $k = 2p - 1$ and

   $Y = \left\{ \chi_{\mathbf{A}} \,:\, A \in \binom{[n]}{k} \right\}$, and let the forbidden distance be $\sqrt{2p}$.

5. Let $N = \binom{n}{p^2 - 1}$, and apply the MFWI to the colouring

$$c(AB) \text{ is red} \quad \Leftrightarrow \quad |A \cap B| = -1 \pmod{p}.$$

Chapter 6

Szemerédi's Regularity Lemma

To ask the right question is harder than to answer it.
Georg Cantor

In Chapter 2 we saw that in any finite colouring of the positive integers, some colour class must contain arbitrarily long arithmetic progressions. In 1936, Erdős and Turán conjectured that the colouring here is just a distraction; the same conclusion should hold for the 'largest' colour class.

6.1 The Erdős-Turán Conjecture

Recall that the upper density of a set $A \subset \mathbb{N}$ is

$$\overline{d}(A) \;=\; \limsup_{n \to \infty} \frac{|A \cap [n]|}{n}.$$

Conjecture 6.1.1 (Erdős and Turán, 1936)**.** *If $\overline{d}(A) > 0$, then A contains arbitrarily long arithmetic progressions.*

Roth proved the case $k = 3$ of the conjecture in 1953, and Szemerédi proved the case $k = 4$ in 1969, before finally resolving the question in 1975. As mentioned in Section 6.3.4, Furstenburg (in 1977) and Gowers (in 2001) provided important alternative proofs of Szemerédi's Theorem. (According to Terry Tao, there are now at least 16 different proofs known!)

Although the conjecture of Erdős and Turán may look like little more than a curiosity, it is perhaps the greatest triumph of the 'Hungarian school' of mathematics, in which the aim is to pose (and solve) beautiful and difficult problems, and allow deep theories and connections between areas to present themselves. For more on the modern (and very active) study of Additive Combinatorics, which grew out of the proofs of Roth, Szemerédi, Furstenburg and Gowers, see the excellent recent book by Tao and Vu [5].

In this section we shall prove Roth's Theorem (the case $k = 3$ of Szemerédi's Theorem), using the Regularity Lemma introduced by Szemerédi. We shall then see how one may use this powerful tool to prove several other beautiful results. The lemma has proven so useful, it is little exaggeration to say that whenever you see a graph, the the first thing you should think to do is to take its Szemerédi partition.

6.2 The Regularity Lemma

Szemerédi's Regularity Lemma, perhaps the most powerful tool in Graph Theory, may be thought of as an answer to the following, fairly vague question.

Question 6.2.1. *How well can an arbitrary graph be approximated by a collection of random graphs?*

Szemerédi's answer to this question was as follows:

> "Given any graph G, we can partition the vertex set $V(G)$ into a bounded number of pieces (at most k, say), such that for almost every pair (A, B) of parts, the induced bipartite graph $G[A, B]$ is approximated well by a random graph."

We shall quantify the statements 'almost every' and 'is approximated well by a random graph' in terms of a parameter $\varepsilon > 0$. The crucial point is that the bound k depends on ε, but not on $n = |V(G)|$.

To quantify 'almost every' is easy: it simply means all but εk^2 pairs. To make precise the intuitive idea that a bipartite graph 'looks random' is more complicated, and we shall need the following technical definition. Given subsets $X, Y \subset V(G)$, we shall write $e(X, Y)$ for the number of edges from X to Y, i.e., the size of the set $\{xy \in E(G) : x \in X, y \in Y\}$.

Definition. *Let $\varepsilon > 0$ and let $A, B \subset V(G)$ be disjoint sets of vertices in a graph G. The pair (A, B) is said to be* ε-regular *if*

$$\left| \frac{e(A,B)}{|A||B|} - \frac{e(X,Y)}{|X||Y|} \right| \leqslant \varepsilon$$

for every $X \subset A$ and $Y \subset B$ with $|X| \geqslant \varepsilon|A|$ and $|Y| \geqslant \varepsilon|B|$.

In other words, for every pair (X, Y) of sufficiently large subsets of A and B, the density of (X, Y) is about the same as that of (A, B). It is easy to see (using Chernoff's inequality, see [1]) that in the random graph $G_{n,p}$, any (sufficiently large) pair of subsets (A, B) is ε-regular with high probability.

Before stating the Regularity Lemma, let us see why the definition above is useful by proving a couple of easy consequences of ε-regularity. Indeed, if we believe that ε-regular pairs 'look like' random graphs, then we would like ε-regular graphs and random graphs to share some basic properties.

The simplest property of a random graph is that most vertices have roughly the same number of neighbours. This property is shared by ε-regular pairs.

Property 1. *Let (A, B) be an ε-regular pair of density d in a graph G. Then all but at most $2\varepsilon|A|$ vertices of A have degree between $(d - \varepsilon)|B|$ and $(d + \varepsilon)|B|$ in $G[A, B]$.*

Proof. Let $X \subset A$ be the set of low degree vertices in A, i.e.,

$$X \;:=\; \Big\{ v \in A \,:\, |N_G(v) \cap B| < (d - \varepsilon)|B| \Big\},$$

and let $Y = B$. Then the density of the pair (X, Y) is less than $d - \varepsilon$, so $|X| < \varepsilon|A|$, by the definition of ε-regularity. By symmetry, the same holds for the high degree vertices. $\square$

Another nice property of $G_{n,p}$ is that a (randomly chosen) induced subgraph of $G_{n,p}$ is also a random graph, with the same density.

Property 2. *Let (A, B) be an ε-regular pair of density d in a graph G. If $X \subset A$ and $Y \subset B$, with $|X| \geqslant \delta|A|$ and $|Y| \geqslant \delta|B|$, then (X, Y) is an $2\varepsilon/\delta$-regular pair, of density between $d - \varepsilon$ and $d + \varepsilon$.*

Proof. This follows easily from the definitions, so we leave its proof as an exercise. □

The properties above give us some idea of the motivation behind the definition of ε-regularity. However, the main reason that the definition is useful is the following 'embedding lemma'. It says that, for the purpose of finding small subgraphs in a graph G, we can treat 'dense' ε-regular pairs like complete bipartite graphs.

Given a graph H and an integer m, let $H(m)$ denote the graph obtained by 'blowing up' each vertex of H to size m, i.e., each vertex $j \in V(H)$ is replaced by a set A_j of size m. Thus $H(m)$ has vertex set $\bigcup_j A_j$, and edge set $\{uv : u \in A_i, v \in A_j \text{ for some } ij \in E(H)\}$.

Given a graph H, an integer m and $\delta > \varepsilon > 0$, let $\mathcal{G}(H, m, \varepsilon, \delta)$ denote the family of graphs G such that $V(G) = V(H(m))$, $G \subset H(m)$, and $G[A_i, A_j]$ is ε-regular and has density at least δ whenever $ij \in E(H)$.

The Embedding Lemma (simple version). *Let H be a graph, and let $\delta > 0$. There exists $\varepsilon > 0$ and $M \in \mathbb{N}$ such that if $m \geqslant M$ and $G \in \mathcal{G}(H, m, \varepsilon, \delta)$, then $H \subset G$.*

For most of our applications the simple version of the embedding lemma will be sufficient. The full version below is a bit harder to remember, but is not much harder to prove.

The Embedding Lemma. *Let $\Delta \in \mathbb{N}$, let $\delta > 0$, and let $\varepsilon_0 = \delta^{\Delta}/(\Delta + 2)$. Let R be a graph, let $m, t \in \mathbb{N}$ with $t \leqslant \varepsilon_0 m$, and let $H \subset R(t)$ with maximum degree at most Δ.*

If $\varepsilon_0 > \varepsilon > 0$ and $G \in \mathcal{G}(R, m, \varepsilon, \delta + \varepsilon)$, then G contains at least $(\varepsilon_0 m)^{|V(H)|}$ copies of H.

Proof. We shall prove the simple embedding lemma for the triangle $H = K_3$, and leave the proof of the full statement to the reader. To find a triangle in $G \in \mathcal{G}(K_3, m, \varepsilon, \delta)$, we simply pick vertices one by one. Indeed, choose $v_1 \in A_1$ arbitrarily amongst the vertices with at least $(\delta - \varepsilon)m$ neighbours in both A_2 and A_3.

Let $X = N(v) \cap A_2$ and $Y = N(v) \cap A_3$. Since $|X| \geqslant \varepsilon|A_2|$ and $|Y| \geqslant \varepsilon|A_3|$, it follows that there exists an edge between X and Y, so we are done. Alternatively (and more instructively for the general case), note that, by Property 2 above, the pair (X, Y) is ε'-regular and has density at least $\delta - \varepsilon$ (where $\varepsilon' = 2\varepsilon/(\delta - \varepsilon)$). Hence, by Property 1, all but $2\varepsilon'|X|$ vertices of X have degree at least $\delta - 2\varepsilon$ in $G[X, Y]$, and so we have found the desired triangle. $\square$

Problem 6.2.2. *Prove the full embedding lemma. [Hint: choose vertices one by one, and keep track of their common neighbourhoods.]*

Inspired by the Embedding Lemma, we shall (throughout this section) refer to a graph on n vertices as *sparse* if it has $o(n^2)$ edges, and *dense* otherwise, i.e., if it has at least δn^2 edges for some $\delta > 0$. Note that (as stated) this definition is not precise; to understand it we should think of a sequence of graphs (G_n), where $|G_n| = n$ for each $n \in \mathbb{N}$, and consider the limit $n \to \infty$. However, our precise statements will hold for (large) fixed graphs.

Having (we hope!) convinced the reader that our definition of ε-regularity is a useful one, we are ready to state the Regularity Lemma.

Theorem 6.2.3 (The Szemerédi Regularity Lemma, 1975)**.** *Let $\varepsilon > 0$, and let $m \in \mathbb{N}$. There exists a constant $M = M(m, \varepsilon)$ such that the following holds.*

For any graph G, there exists a partition $V(G) = A_0 \cup \ldots \cup A_k$ of the vertex set into $m \leqslant k \leqslant M$ parts, such that

- $|A_1| = \ldots = |A_k|$,
- $|A_0| \leqslant \varepsilon|V(G)|$,
- *all but εk^2 of the the pairs (A_i, A_j) are ε-regular.*

The Regularity Lemma can be a little difficult to fully grasp at first sight; the reader is encouraged not to worry, and to study the applications below. We remark that the proof of the lemma (see

Section 6.7) is not difficult; the genius of Szemerédi was to imagine that such a statement could be true!

We note also that although the lemma holds for *all* graphs, it is only useful for large and (fairly) dense graphs. Indeed, if $n := |V(G)| < M(m, \varepsilon)$ then the statement is vacuous (since each part has size at most one), and if $e(G) = o(n^2)$ then G is well-approximated by the empty graph. Finally, we note that the best possible function $M(m, \varepsilon)$ is known to be (approximately) a tower of height $f(\varepsilon)$, for some function $\log(1/\varepsilon) \leqslant f(\varepsilon) \leqslant (1/\varepsilon)^5$.

6.3 Applications of the Regularity Lemma

The easiest way to understand the Regularity Lemma is to use it. In this section we shall give three simple (and canonical) applications: the Erdős-Stone Theorem, which we proved in Chapter 3; the 'triangle removal lemma', which we shall use to deduce Roth's Theorem; and bounds on Ramsey-Turán numbers.

6.3.1 The Erdős-Stone Theorem

Recall the statement of Theorem 3.3.

The Erdős-Stone Theorem. *Let H be an arbitrary graph. Then*

$$\mathrm{ex}(n, H) \;=\; \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right)\binom{n}{2}.$$

We shall first give a sketch proof, and then fill in some of the details. The reader is encouraged to spend some time turning this sketch into a rigorous proof. Given a graph G, a typical application of the Regularity Lemma (SzRL) goes as follows:

1. Apply the SzRL (for some sufficiently small $\varepsilon > 0$). We obtain a partition $(A_0, \ldots, A_k)$ of the vertex set as described above.

2. Remove edges inside parts, between irregular pairs, and between sparse pairs. There are at most $O(\varepsilon n^2)$ such edges.

3. Consider the 'reduced graph' R which has vertex set $[k]$ and edge set
$$\big\{ij : \text{ the pair } (A_i, A_j) \text{ is dense and } \varepsilon\text{-regular}\big\}.$$
4. Apply a standard result from Graph Theory (e.g., Hall's Theorem, Dirac's Theorem, Turán's Theorem) to R.
5. Use the Embedding Lemma to find a copy of the desired subgraph H in G.

Now let us see how to apply this approach in the case of Erdős-Stone. The statement we are trying to prove is that, for every $\delta > 0$ there exists $N(\delta) \in \mathbb{N}$, such that if $n \geqslant N(\delta)$, $|G| = n$ and
$$e(G) \geqslant \left(1 - \frac{1}{\chi(H) - 1} + \delta\right)\binom{n}{2},$$
then $H \subset G$.

Proof of the Erdős-Stone Theorem. Let G be a graph on n vertices as described above, let $\varepsilon = \varepsilon(H, \delta) > 0$ be sufficiently small, and let $m = 1/\varepsilon$. Apply the SzRL to obtain a partition $(A_1, \ldots, A_k)$, and form the reduced graph R.

<u>Claim</u>: $e(R) > \left(1 - \dfrac{1}{\chi(H) - 1} + \dfrac{\delta}{2}\right)\dbinom{k}{2}$.

Proof of Claim. Consider the edges inside parts, between irregular pairs, and between pairs of density at most $\delta/4$; we claim that there are at most $(\delta/3)\binom{n}{2}$ such edges. Indeed, there are at most $k(n/k)^2$ edges inside parts (recall that $k \geqslant m = 1/\varepsilon$); there are at most $\varepsilon k^2(n/k)^2$ edges between irregular pairs (since there are at most εk^2 such pairs); and there are at most $\binom{k}{2}(\delta/4)(n/k)^2$ edges between 'sparse' pairs.

Thus only considering edges between dense, regular pairs, we obtain a graph $G' \subset G$ with
$$e(G') \geqslant \left(1 - \frac{1}{\chi(H) - 1} + \frac{2\delta}{3}\right)\binom{n}{2}.$$

Now, the edges of G' are all between pairs of parts which correspond to edges of R, so we have $e(G') \leqslant e(R)(n/k)^2$, and the claim follows. $\square$

Applying Turán's Theorem to the reduced graph, we deduce from the claim that R contains a complete graph on $r = \chi(H)$ vertices. Let $i(1), \ldots, i(r)$ be the labels of the parts corresponding to the vertices of this complete graph, and let $A = \bigcup_{j=1}^{r} A_{i(j)}$.

Then $G'[A] \in \mathcal{G}(K_r, n', \varepsilon, \delta/2)$ for some $n' \geqslant n/2k$, by the definition of R. Also, since H is a fixed graph with $\chi(H) = r$, we have $H \subset R(t)$ for $t = |H|$. Hence, by the Embedding Lemma, $H \subset G$ as required. $\square$

6.3.2 The Triangle Removal Lemma

The next application is a famous and important one; until recently, there was no known proof of it which avoided the Regularity Lemma.

The Triangle Removal Lemma (Ruzsa and Szemerédi, 1976)**.** *For every $\varepsilon > 0$ there exists a $\delta > 0$ such that the following holds:*

If G is a graph with at most δn^3 triangles, then all the triangles in G can be destroyed by removing εn^2 edges.

Proof. We shall use the same approach as in the proof above. Indeed, let $\varepsilon' = \varepsilon'(\varepsilon, \delta) > 0$ be sufficiently small, and apply the SzRL for ε'. Remove all edges inside parts, between irregular pairs, and between sparse pairs; there are at most εn^2 such edges.

We claim that the remaining graph, G', is triangle-free. Indeed, the only edges remaining correspond to edges of the reduced graph, R. Thus, if G' contains a triangle, it follows that R contains a triangle. But then, by the Embedding Lemma, G must contain at least $f(\varepsilon)n^3 > \delta n^3$ triangles, which is a contradiction. $\square$

6.3.3 Roth's Theorem

Finally, let's deduce Roth's Theorem from the triangle removal lemma.

Roth's Theorem (Roth, 1954)**.** *If $A \subset \mathbb{N}$ has positive upper density, then A contains an arithmetic progression of length three.*

Proof. Given A, we first form a graph G as follows. Choose $0 < \varepsilon < \overline{d}(A)$, and choose n sufficiently large, with $|A \cap [n]| > \varepsilon n$. Let $V(G) = X \cup Y \cup Z$, where X, Y and Z are disjoint copies of $[n]$, and let

$$\begin{aligned} E(G) &= \big\{\{x,y\} : x \in X,\ y \in Y,\ \text{and } y = x + a \text{ where } a \in A\big\} \\ &\cup \big\{\{y,z\} : y \in Y,\ z \in Z,\ \text{and } z = y + a \text{ where } a \in A\big\} \\ &\cup \big\{\{x,z\} : x \in X,\ z \in Z,\ \text{and } z = x + 2a \text{ where } a \in A\big\} \end{aligned}$$

The key observation is that all but n^2 of the triangles in G correspond to 3-APs in A. Indeed, if $\{x,y,z\}$ is a triangle in G, then a, b and $(a+b)/2$ are in A, where $a = y - x$ and $b = z - y$. So if A contains no 3-AP, then the only triangles in G correspond to triples (a,a,a). There are $n|A \cap [n]| \leqslant n^2$ such triangles (each is determined by two of its vertices).

Apply the triangle removal lemma to G. Since $n^2 < \delta(\varepsilon)n^3$ (if n is sufficiently large), it follows that we can destroy all triangles in G by removing at most εn^2 edges. But the triangles in G corresponding to the triples (a,a,a) are *edge-disjoint*! Hence $n|A \cap [n]| \leqslant \varepsilon n^2$, which is a contradiction. Thus A must contain a 3-AP, as claimed. □

Problem 6.3.1. *Prove Roth's Theorem in an arbitrary abelian group.*

6.3.4 Ramsey-Turán numbers

In Turán's Theorem, the extremal K_r-free graphs (i.e., the Turán graphs) have very large independent sets. What happens to the extremal number if we require in addition that our K_r-free graphs have small independence number?

Question 6.3.2. *Let G be a triangle-free graph on n vertices, and suppose that G has no independent set of size $f(n)$. If $f(n) = o(n)$, does it follow that $e(G) = o(n^2)$?*

Answer: Yes!

Proof. The maximum degree of a vertex in G is at most $f(n)$. □

Question 6.3.3. *Let G be a K_r-free graph on n vertices, for some $r \geqslant 5$, and suppose that G has no independent set of size $f(n)$. If $f(n) = o(n)$, does it follow that $e(G) = o(n^2)$?*

Answer: No!

Proof. Recall from Theorem 4.1.2 that there exist arbitrarily large graphs with girth at least four and no linear size independent set; call such a graph an Erdős graph.

Now let $H = T_{\lfloor (r-1)/2 \rfloor}(n)$, the $\lfloor (r-1)/2 \rfloor$-partite Turán graph on n vertices, and let G be the graph obtained from H by placing an Erdős graph in each part. Then G contains no K_r, no independent set of linear size, and at least $n^2/4$ edges. □

Given a 'forbidden' graph H, and a function $f : \mathbb{N} \to \mathbb{N}$, define the Ramsey-Turán number of the pair (H, f) to be

$$RT\big(n, H, f(n)\big) \;:=\; \max\big\{e(G) : |G| = n, H \not\subset G \text{ and } \alpha(G) \leqslant f(n)\big\},$$

where $\alpha(G)$ denotes the independence number of G.

From Questions 6.3.2 and 6.3.3, we know that $RT(n, K_r, o(n))$ is $o(n^2)$ if $r \leqslant 3$, and is $\Theta(n^2)$ if $r \geqslant 5$. But what about $r = 4$?

Theorem 6.3.4 (Szemerédi, 1972; Bollobás and Erdős, 1976)**.**

$$RT(n, K_4, o(n)) \;=\; \frac{n^2}{8} + o(n^2).$$

We shall prove the upper bound using the SzRL; the lower bound follows from an ingenious construction of Bollobás and Erdős, which we shall describe briefly below.

Proof of the upper bound in Theorem 6.3.4. We shall follow the usual strategy, but this time we shall need a couple of extra (fairly simple) ideas. Indeed, let G be a graph with n vertices, and apply the SzRL to G, for some sufficiently small $\varepsilon > 0$. Form the reduced graph R of ε-regular pairs with density at least δ, in the usual way.

We make the following two claims.

Claim 1: R is triangle-free.

Proof of Claim 1. Let $\{A, B, C\}$ be a triangle of parts in R, and choose $a \in A$ with $|N(a) \cap B|, |N(a) \cap C| \geqslant (\delta/2)|C|$. Since the pair $(N(a) \cap B, N(a) \cap C)$ is ε'-regular (by Property 6.2), we can choose a vertex $b \in N(a) \cap B$ such that $|N(b) \cap N(a) \cap C| \geqslant (\delta/2)^2|C|$.

But $\alpha(G) = o(n)$, so $N(a) \cap N(b) \cap C$ is not an independent set. But then G contains a copy of K_4, so we are done. $\square$

By Claim 1 and Mantel's Theorem, R has at most $k^2/4$ edges. The upper bound is thus an immediate consequence of the following claim.

Claim 2: G contains no ε-regular pair with density bigger than $1/2+\delta$.

Proof of Claim 2. Let (A, B) be an ε-regular pair with density $1/2+\delta$. Since A has no linear size independent set, we can choose a pair $x, y \in A$ such that $xy \in E(G)$, and $|N(x) \cap B|, |N(y) \cap B| \geqslant (1/2 + \delta/2)|B|$.

But then $|N(x) \cap N(y) \cap B| \geqslant \delta|B|$, and so either $\alpha(G) \geqslant \delta|B|$, or the set $N(x) \cap N(y) \cap B$ contains an edge, in which case $K_4 \subset G$. In either case, we have a contradiction. $\square$

The result follows by counting edges. There are at most

$$\frac{k^2}{4}\left(\frac{1}{2} + \delta\right)\left(\frac{n}{k}\right)^2 = \frac{n^2}{8} + o(n^2)$$

edges between dense regular pairs, and $o(n^2)$ other edges, so G has at most $n^2/8 + o(n^2)$ edges, as required. $\square$

When Szemerédi proved the upper bound in Theorem 6.3.4 in 1972, most people expected the correct answer to be $o(n^2)$. It was therefore very surprising when Bollobás and Erdős gave the following construction, which shows that Szemerédi's bounds is in fact sharp!

The Bollobás-Erdős graph. *Let $U = S^d$ be a sufficiently high-dimensional sphere, and scatter $n/2$ red points and $n/2$ blue points at random on U. Join two points of the same colour if they are at distance less than $\sqrt{2} - \varepsilon$, and join points of different colours if they are at distance greater than $2 - \varepsilon$, for some suitably chosen $\varepsilon = \varepsilon(d) > 0$.*

Problem 6.3.5. *Show that the Bollobás-Erdős graph has $n^2/8+o(n^2)$ edges, contains no copy of K_4, and has independence number $o(n)$.*

6.4 Graph Limits

Another way of viewing the Regularity Lemma is topological in nature: it allows us to compactify the space of all large dense graphs.

Definition (Convergence of graph sequences). *A sequence $G_n = ([n], E_n)$ of graphs is* convergent *if for any graph H, the density of copies of H in G_n converges to a limit as $n \to \infty$.*

Theorem 6.4.1 (Lovász and Szegedy, 2004). *Every sequence of graphs has a convergent subsequence.*

In fact, one can say something stronger.

Definition (Graphons). *A* graphon *is a symmetric measurable function $p : [0, 1] \times [0, 1] \to [0, 1]$.*

For each graphon, define a generalised Erdős-Rényi graph $G(n, p)$ on n vertices as follows: first give each vertex v a 'colour' $x_v \in [0, 1]$, chosen uniformly at random; then add each edge vw with probability $p(x_v, x_w)$, all independently.

Let $G(\infty, p)$ denote the (formal) limit of the random graphs $G(n, p)$.

Theorem 6.4.2 (Lovász and Szegedy, 2004). *Every sequence of graphs has a subsequence converging to $G(\infty, p)$ for some graphon p.*

Various other metrics have been proposed for the space of graph sequences; interestingly, although the definitions are quite different from one another, almost all have turned out to be equivalent.

A closely related topic is that of *graph testing*: roughly speaking, a property of graphs $\mathcal{P}$ is *testable* if we can (with high probability as $k \to \infty$) test whether or not a large graph G is in $\mathcal{P}$ by looking at a random subgraph of G on k vertices. (More precisely, it accepts any graph satisfying $\mathcal{P}$ with probability 1, and rejects any which is $\varepsilon = \varepsilon(k)$-far from $\mathcal{P}$ (in a given metric) with probability $1 - \varepsilon$.)

The strongest result along these lines is due to Austin and Tao, who generalized results of Alon and Shapira (for graphs) and Rödl and Schacht (for hypergraphs).

Theorem 6.4.3 (Rödl and Schacht, 2007). *Every hereditary property of hypergraphs is testable.*

This theorem has some surprising consequences: for example, it implies Szemerédi's Theorem.

Another related question asks for a fast algorithm which produces a Szemerédi partition of a graph G. This problem has been studied by various authors, and is (at least partly) so challenging because, as noted earlier, the bound on the number of parts in a Szemerédi partition is *very* large.

Motivated by this shortcoming, Frieze and Kannan introduced the following notion of 'weak regularity': given a partition $\mathcal{P}$, with given edge densities d_{ij}, let

$$e_{\mathcal{P}}(S,T) \; = \; \sum_{i,j} d_{ij} \cdot |V_i \cap S| \cdot |V_j \cap T|,$$

denote the expected number of edges of G between S and T, and say that the partition $\mathcal{P}$ is *weakly regular* if $|e(S,T) - e_{\mathcal{P}}(S,T)| \leqslant \varepsilon n^2$ for every $S, T \subset V(G)$.

Theorem 6.4.4 (Frieze and Kannan, 1999)**.** *For every $\varepsilon > 0$ and every graph G, there exists a weakly regular partition of $V(G)$ into at most $2^{2/\varepsilon^2}$ classes.*

Lovász and Szegedy showed that by iterating the Weak Regularity Lemma, one can obtain the usual Regularity Lemma as a corollary. They also proved a generalization in the setting of an arbitrary Hilbert space, and described an algorithm which constructs the weak Szemerédi partition as Voronoi cells in a metric space.

6.5 Recommended further reading

An excellent introduction to the area is provided by the survey:

J. Komlos and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, DIMACS Technical Report (1996).

For a very different perspective, see

L. Lovász and B. Szegedy, Szemerédi's Lemma for the analyst, *J. Geom. and Func. Anal.*, **17** (2007), 252–270.

6.6 Exercises

1. Use Szemerédi's Regularity Lemma to prove:

The $(6,3)$**-Theorem** (Ruzsa and Szemerédi, 1976). *Let* $\mathcal{A} \subset \mathcal{P}(n)$ *be a set system such that* $|A| = 3$ *for every* $A \in \mathcal{A}$*, and such that for each set* $B \subset [n]$ *with* $|B| = 6$*, we have*

$$\big|\{A \in \mathcal{A} \,:\, A \subset B\}\big| \leqslant 2.$$

Then $|\mathcal{A}| = o(n^2)$.

2. Let $r_3(n)$ be the size of the largest subset of $[n]$ with no 3-AP, and let $f(k,n)$ denote the maximum number of edges in a graph on n vertices which is the union of k induced matchings.

   (a) Prove that $r_3(n) \leqslant \dfrac{f(n,5n)}{n}$.

   [Hint: consider a graph with edge set $(x + a_i, x + 2a_i)$.]

   (b) Using Szemerédi's Regularity Lemma, deduce Roth's Theorem.

In the next exercise, we shall prove the following theorem of Thomassen.

Thomassen's Theorem. *If* G *is a triangle-free graph with minimum degree* $\big(\frac{1}{3} + \varepsilon\big)|G|$*, then* $\chi(G) \leqslant C$*, for some constant* $C = C(\varepsilon)$.

Given a Szemerédi partition $A_0 \cup \ldots \cup A_k$ of $V(G)$, and $d > 0$, consider the auxiliary partition $V(G) = \bigcup_{I \subset [k]} X_I$, where

$$X_I \;:=\; \Big\{ v \in V(G) : \; i \in I \Leftrightarrow |N(v) \cap A_i| \geqslant d|V_i| \Big\}.$$

3. (a) Show that if $|I| \geqslant 2k/3$, then X_I is empty.

   (b) Show that if $|I| \leqslant 2k/3$, then X_I is an independent set.

   (c) Deduce Thomassen's Theorem.

[You may assume the following strengthening of the Regularity Lemma: that the reduced graph R has minimum degree $\big(1/3 + \varepsilon/2\big)|R|$.]

4. Use the Kneser graph to show that Thomassen's Theorem is sharp.

6.7 Proof the of Regularity Lemma

The Regularity Lemma is not very hard to prove; the hard part was imagining that it could be true. In this section we shall help the reader to prove the lemma for himself.

The idea is as follows: starting with an arbitrary equipartition P, we shall repeatedly refine P, each time getting 'closer' to a Szemerédi partition. The key point is to find he right notion of the 'distance' of a partition from an ideal partition.

Given a partition P of V into parts $V_0, \ldots, V_k$, we define the index of P to be

$$\text{ind}(P) \ := \ \frac{1}{k^2} \sum_{i \neq j} \Big(d(V_i, V_j) \Big)^2,$$

where $d(X,Y)$ denotes the density of the pair (X,Y), i.e., $\frac{e(X,Y)}{|X||Y|}$.

Problem 6.7.1. *Prove that if P is not a Szemerédi partition, i.e., more than εk^2 of the pairs are irregular, then there exists a refinement Q of P such that*

$$\text{ind}(Q) \ \geqslant \ \text{ind}(P) + \delta$$

for some $\delta = \delta(\varepsilon)$.

In order to solve Problem 6.7.1, you may need to use the following 'defect' form of the Cauchy-Schwarz inequality.

Improved Cauchy-Schwarz Inequality. *Let $a_1, \ldots, a_n > 0$, and let $m \in [n]$ and $\delta \in \mathbb{R}$. Suppose that*

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

Then

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

We leave the remainder of the proof of Theorem 6.2.3 to the reader.

Problem 6.7.2. *Deduce Szemerédi's Regularity Lemma.*

Chapter 7

Dependent Random Choice

Choose well. Your choice is brief, and yet endless.
Johann Wolfgang von Goethe

In this final chapter we shall discuss a simple, yet surprisingly powerful technique, known as Dependent Random Choice, which was introduced by Gowers in his re-proof of Szemerédi's Theorem, and has since proved to have many other beautiful applications.

The technique may be summarized by the following basic lemma.

The Dependent Random Choice Lemma. *Let G be a graph with n vertices and m edges. Suppose that*

$$\frac{(2m)^t}{n^{2t-1}} - \binom{n}{s}\left(\frac{k}{n}\right)^t \geqslant a.$$

Then there exists a subset $U \subset V(G)$ of at least a vertices, such that every set of s vertices in U has at least k common neighbours.

Proof. Pick a (multi-)set $T \subset V(G)$ at random, by choosing t times a random vertex of $V(G)$, with repetition. Thus $|T| \leqslant t$ (as a set), and $\mathbb{P}(v \notin T) = (1 - 1/n)^t$. Let

$$A \; := \; N(T) \; = \; \bigcap_{j=1}^{t} N(x_t),$$

where $T = \{x_1, \ldots, x_t\}$. Then A is the set we want!

Indeed, let $X = |A|$ and let Y denote the number of s-sets in A with at most k common neighbours. The probability that a vertex v is in A, is just the probability that T is a subset of its neighbourhood. Hence, by the convexity of x^t,

$$\mathbb{E}(X) \; = \; \sum_{v \in V(G)} \left(\frac{|N(v)|}{n} \right)^t \; \geqslant \; n \left(\frac{1}{n} \sum_{v \in V(G)} \frac{|N(v)|}{n} \right)^t \; = \; \frac{(2m)^t}{n^{2t-1}}.$$

Now, suppose the pair $\{u, v\}$ has at most k common neighbours; then the probability that $\{u, v\} \subset A$ is at most $(k/n)^t$, since each element of T must lie in the common neighbourhood of u and v, and so $\mathbb{E}(Y) \leqslant \binom{n}{s}(k/n)^t$.

By linearity of expectation,

$$\mathbb{E}(X - Y) \; \geqslant \; \frac{(2m)^t}{n^{2t-1}} - \binom{n}{s} \left(\frac{k}{n} \right)^t \; \geqslant \; a,$$

and thus there must exist a choice of T such that $X - Y \geqslant a$. Now simply remove one element from each s-set in A with at most k common neighbours, to obtain U as required. $\square$

We shall give three applications of the Dependent Random Choice Lemma: a bound on extremal numbers of bipartite graphs, a bound on Ramsey-Turán numbers, and the Balog-Szemerédi-Gowers Theorem. We refer the reader to the survey [3] for a more detailed treatment of these and many other applications.

7.1 Extremal numbers of bipartite graphs

Recall that in Chapter 3 we proved that

$$\mathrm{ex}\big(n, K(s,t)\big) \; = \; O\big(n^{2-1/s}\big).$$

The following result is a significant generalization of that theorem.

Theorem 7.1.1. *Let H be a bipartite graph on $A \cup B$, and suppose that each vertex of B has degree at most s. Then*

$$\mathrm{ex}\big(n, H\big) \;=\; O\big(n^{2-1/s}\big),$$

where the implicit constant depends only on H.

In order to prove Theorem 7.1.1, we shall need the following straightforward embedding lemma.

Lemma 7.1.2. *Let H be a bipartite graph on $A \cup B$, and suppose that each vertex of B has degree at most s. If a graph G contains a set $U \subset V(G)$ with the following properties:*

(*a*) $|U| = |A|$,

(*b*) *All subsets of U of size s have at least $|H|$ common neighbours,*

then $H \subset G$.

Proof. Choose an arbitrary bijection between U and A, and label the vertices of $B = \{v_1, \ldots, v_b\}$; we shall embed them one at a time. Indeed, suppose $v_1, \ldots, v_{j-1}$ have already been embedded in G, and consider the neighbours of v_j in A. Since v_j has degree at most s, the corresponding vertices of U have at least $|H|$ common neighbours. Thus, at least one of these has not been used yet (i.e., it is not in U or $\{v_1, \ldots, v_{j-1}\}$), and so we may embed v_j in G as required. This proves the lemma. □

It is now straightforward to deduce Theorem 7.1.1.

Proof of Theorem 7.1.1. Suppose that G is a graph on n vertices, with $m \geqslant cn^{2-1/s}$ edges. We wish to find a subset $U \subset V(G)$ with $a := |A|$ vertices, such that each s-subset of U has at least $|H|$ common neighbours. By the Dependent Random Choice Lemma (with $t = s$ and $k = |H|$), such a set U exists if

$$\frac{(2m)^s}{n^{2s-1}} - \binom{n}{s}\left(\frac{|H|}{n}\right)^s \;\geqslant\; (2c)^s - \frac{|H|^s}{s!} \;\geqslant\; c^s \;\geqslant\; a,$$

where the last two inequalities hold if $c = c(H)$ is chosen to be sufficiently large.

By Lemma 7.1.2, if such a set $U \subset V(G)$ exists, then $H \subset G$. Thus $\mathrm{ex}(n, H) \leqslant m$, as required. □

We remark that Theorem 7.1.1 was first proved by Füredi in 1991; the proof above is due to Alon, Krivelevich and Sudakov.

7.2 Ramsey-Turán revisited

Recall that in Section 6.3.4 we proved that $RT(n, K_4, o(n)) = \frac{n^2}{8} + o(n)$. It is natural to ask, how small does $f(n)$ need to be to make $RT(n, K_4, f(n)) = o(n^2)$? Indeed, it was suggested by several authors that this might only be true when $f(n) \leqslant n^\alpha$ for some $\alpha < 1$.

The following theorem shows that this is not the case, and follows almost immediately from the Dependent Random Choice Lemma.

Theorem 7.2.1 (Sudakov, 2003). *Let $f(n) = 2^{-\omega\sqrt{\log n}}n$, where $\omega = \omega(n)$ is any function which tends to infinity as $n \to \infty$. Then*

$$RT(n, K_4, f(n)) \;=\; o(n^2).$$

Proof. Let G be a graph with n vertices and $m \geqslant 2^{-\omega^2/2}n^2 = o(n^2)$ edges, and suppose that $K_4 \not\subset G$ and $\alpha(G) \leqslant f(n)$. Let $k = f(n)$, and apply Dependent Random Choice. We easily see that

$$\frac{(2m)^t}{n^{2t-1}} - \binom{n}{2}\left(\frac{k}{n}\right)^t \;\geqslant\; 2^{t-\omega^2 t/2}n \;-\; 2^{-\omega t\sqrt{\log n}}n^2,$$

since $m \geqslant 2^{-\omega^2/2}n^2$, and $k = f(n) = 2^{-\omega\sqrt{\log n}}n$. Thus, choosing $t = \dfrac{2}{\omega}\sqrt{\log n}$, we obtain

$$\frac{(2m)^t}{n^{2t-1}} - \binom{n}{2}\left(\frac{k}{n}\right)^t \;\geqslant\; 2^t 2^{-\omega\sqrt{\log n}}n \;-\; 2^{-2\log n}n^2 \;\geqslant\; f(n).$$

So, by the DRC Lemma, there exists a set $U \subset V(G)$, with $|U| \geqslant f(n)$, such that every pair of vertices in U has at least $f(n)$ common neighbours.

Since $\alpha(G) \leqslant f(n)$, there is an edge $\{u, v\}$ in U, and u and v have at least $f(n)$ common neighbours. Let $A = N(u) \cap N(v)$; since $\alpha(G) \leqslant f(n)$, there is an edge $\{x, y\}$ in A. But now the set $\{u, v, x, y\}$ induces a K_4, so we have a contradiction. $\square$

Problem 7.2.2. *For which function $f(n)$ can you prove that*

$$RT(n, K_5, f(n)) = o(n^2)?$$

7.3 The Balog-Szemerédi-Gowers Theorem

Our final application of Dependent Random Choice is possibly its most important consequence, and was also the reason it was first introduced by Gowers in 1998. One of the central objects of study in Additive Combinatorics is the sumset of two sets,

$$A + B := \{a + b : a \in A, b \in B\}.$$

For example, a fundamental result in the area, Freiman's Theorem, says that if $|A + A| \leqslant C|A|$, then A is contained in a (bounded dimension) generalized arithmetic progression of size at most $C'|A|$.

If G is a bipartite graph, with parts A and B, then we define the *partial sumset* by

$$A +_G B \;=\; \{a + b : a \in A, b \in B, ab \in E(G)\}.$$

In many applications, instead of knowing $|A + B|$, we only know the size of the partial sumset for some dense graph G. Moreover, it is easy to construct a set A and a dense graph G such that $|A +_G A| = O(n)$ but $|A + A| = \Theta(n^2)$: simply take an arithmetic progression of length $n/2$, plus $n/2$ random elements, and let G be the complete graph on the arithmetic progression.

Despite this, the following theorem, first proved by Balog and Szemerédi in 1994 using the Regularity Lemma, allows us to draw a useful conclusion.

The Balog-Szemerédi-Gowers Theorem. *Let $|A| = |B| = n$, let G be a bipartite graph on $A \cup B$ with at least cn^2 edges, and suppose that $|A +_G B| \leqslant Cn$. Then there exist sets $A' \subset A$ and $B' \subset B$, with $|A'| \geqslant c'|A|$ and $|B'| \geqslant c'|B|$, such that*

$$|A' + B'| \;\leqslant\; C'n,$$

where c' and C' depend on c and C but not on n.

The Regularity Lemma proof gives tower-type bounds on c' and C', whereas Gowers' approach, using Dependent Random Choice, gives polynomial-type bounds. This is a common theme: whenever it can be used, Dependent Random Choice tends to give better bounds and simpler proofs than other methods.

The first key idea in the proof is to relate the partial sumset to paths of length three in G. Let (a, b', a', b) be such a path, where $a, a' \in A$ and $b, b' \in B$, and observe that

$$y = a + b = (a + b') - (a' + b') + (a' + b) = x - x' + x'',$$

where x, x' and x'' are members of $A +_G B$. Thus, we have a injective map from pairs (y, P) to triples $(x, x', x'') \in (A +_G B)^3$, where $y \in A + B$ and P is a path of length three in G whose endpoints sum to y. This observation motivates the following graph-theoretic lemma.

Lemma 7.3.1. *For every $c > 0$ there exists a $c' > 0$ such that, if $|A| = |B| = n$, and G is a bipartite graph on $A \cup B$ with at least cn^2 edges, then the following holds.*

There exist subsets $A' \subset A$ and $B' \subset B$, with $|A'|, |B'| \geqslant c'n$, such that for every $a \in A'$ and $b \in B'$, there are at least $c'n^2$ paths of length three in G from a to b.

We shall prove Lemma 7.3.1 using the following variant of the Dependent Random Choice Lemma. It allows us to find a much larger subset U (in fact, linear size), at the cost of not *every* pair of vertices (but still *almost* every pair) having a large common neighbourhood.

Lemma 7.3.2. *Let G be a bipartite graph on $A \cup B$, let $c > 0$ and $\varepsilon > 0$, and suppose that $e(G) = 2c|A||B|$. There exists a subset $U \subset A$ such that*

(*a*) $|U| \geqslant c|A|$,

(*b*) *At most $\varepsilon|U|^2$ of the pairs in U have at most $\varepsilon c^2|B|$ common neighbours in B.*

Proof. Pick a vertex $v \in B$ uniformly at random. We shall show that, for some choice of v, the set $N(v)$ of neighbours of v has the desired properties.

Indeed, let $X = |N(v)|$, and let Y denote the number of pairs in $N(v)$ with at most $\varepsilon c^2|B|$ common neighbours in B. Then, by Cauchy-Schwarz, we have

$$\mathbb{E}[X^2] \geqslant \mathbb{E}[X]^2 = \left(\frac{e(G)}{|B|}\right)^2 = (2c)^2|A|^2,$$

and since $\{x, y\} \subset N(v)$ if and only if $v \in N(x) \cap N(y)$, we have

$$\mathbb{E}[Y] \leqslant \varepsilon c^2 |A|^2 \leqslant \frac{\varepsilon}{4} \cdot \mathbb{E}[X^2].$$

Thus

$$\mathbb{E}\big[X^2 - Y/\varepsilon\big] = \mathbb{E}\big[X^2\big] - \frac{1}{\varepsilon}\mathbb{E}\big[Y\big] \geqslant 3c^2|A|^2.$$

Hence there must exist $v \in B$ such that $X^2 - Y/\varepsilon \geqslant c^2|A|^2$.

Set $U = N(v)$; we claim that U has the desired properties. Indeed, (a) follows from $X^2 \geqslant c^2|A|^2$, and (b) holds since $Y \leqslant \varepsilon X^2 = \varepsilon |U|^2$. □

We can now prove Lemma 7.3.1. For simplicity, let's assume that c is sufficiently small.

Proof of Lemma 7.3.1. We choose subsets $A' \subset A$ and $B' \subset B$ as follows. First let

$$A_1 := \big\{ v \in A : |N(v) \cap B| \geqslant c^2 n \big\},$$

and apply Lemma 7.3.2 to the pair (A_1, B), to obtain a set $U \subset A_1$. Now set

$$A' := \big\{ v \in U : v \text{ is in at most } c^2|U| \text{ bad pairs in } U \big\},$$

where a pair is bad $\{u, v\}$ if u and v have at most $c^6 n$ common neighbours in B, and set

$$B' := \big\{ v \in B : |N(v) \cap U| \geqslant c^3|U| \big\}.$$

We claim that A' and B' are the required sets.

We show first that $|A'| \geqslant c^4 n$. Easy edge-counting shows that $|A_1| \geqslant c^2 n$, and so, by Lemma 7.3.2, it follows that $|U| \geqslant c^3 n$, and at most $c^3|U|^2$ pairs in U have fewer than $c^6 n$ common neighbours in B. Thus $|A'| \geqslant |U|/2 \geqslant c^4 n$, as claimed.

Next we show that $|B'| \geqslant c^2 n$. Indeed, there are at least $c^2 n|U|$ edges from U to B, since $U \subset A_1$, and at most $c^3 n|U|$ of these miss B'. The bound follows, since each vertex of B' sends at most $|U|$ edges into U.

Finally, given $a \in A'$ and $b \in B'$, we shall show there are at least $c^{11}n^2$ paths of length three between a and b. Indeed, b has at least $c|U|/2$ neighbours in U, and a is a bad pair with at most $c^2|U|$ of these. Thus there are at least $c^8 n|U| \geqslant c^{11}n^2$ paths of length two from a to a neighbour of b in U, as required. □

Finally, let's deduce the Balog-Szemerédi-Gowers Theorem. We have already sketched the proof above.

Proof of the Balog-Szemerédi-Gowers Theorem. Let A' and B' be the sets given by Lemma 7.3.1, and recall that $|A +_G B| \leqslant Cn$. For each $a \in A'$, $b \in B'$ and path P of length three from a to b in G, we obtain a triple $(x, x', x'') \in (A +_G B)^3$, since

$$a + b = (a + b') - (a' + b') + (a' + b),$$

as noted earlier. There are at most C^3n^3 such triples, and each $y \in A' + B'$ corresponds to at least $c^{11}n^2$ of them, by the mapping above. By the pigeonhole principle, this implies that

$$|A' + B'| \leqslant \frac{C^3 n^3}{c^{11} n^2} = C'n,$$

as required. □

Problem 7.3.3. *What properties of the group in which A and B live did we assume in the proof above?*

7.4 Recommended further reading

The survey by Fox and Sudakov [3], on which this chapter is based, is highly recommended. For more on Additive Combinatorics, see the book by Tao and Vu [5], and also the course notes (and many beautiful exercises) in:

A. Geroldinger and I.Z. Ruzsa, Combinatorial Number Theory and Additive Group Theory, Advanced Courses in Mathematics, CRM Barcelona, Birkhäuser (2009).

7.5 Exercises

The Ramsey number $R(G)$ of a graph G is the smallest n such that any 2-colouring of the edges of K_n contains a monochromatic copy of G. The k-cube Q_k is the graph with vertex set $\{0,1\}^k$, and an edge between vertices that differ in exactly one direction.

1. Show that $R(Q_k) \leqslant 2^{3k}$.

A topological copy of a graph H is a graph H' obtained by replacing edges of H by paths. If each path has length $t+1$, then H' is called a t-subdivision of H.

The following problem is due to Erdős.

2. Show that, if G has n vertices and cn^2 edges, then it contains a 1-subdivision of a clique on $c'\sqrt{n}$ vertices.

3. Let G be a graph with n vertices and m edges, in which every pair of adjacent vertices has at most a common neighbours. Using Dependent Random Choice, show that G has an induced subgraph on $X \geqslant \dfrac{m^t}{n^{2t-1}}$ vertices, with $Y \leqslant a^t \left(\dfrac{n}{m}\right)^{t-1} X$ edges.

4. (*a*) What is the largest sum-free subset of $[n]$?
   (*b*) What is the largest sum-free subset of $\mathbb{Z}_p$, for p prime?
   (*c*) What about for an arbitrary abelian group of even order?
   (*d*) Show that every $A \subset \mathbb{N}$ has a sum-free subset of size $|A|/3$.

Define the lower density of a set $A \subset \mathbb{N}$ to be

$$\underline{d}(A) \;=\; \liminf_{n\to\infty} \frac{|A \cap [n]|}{n}.$$

5. (*a*) If $\underline{d}(A) + \underline{d}(B) > 1$, show that then $A+B$ contains all but finitely many elements of $\mathbb{N}$.
   (*b*) If $\underline{d}(A) + \underline{d}(B) = 1$, show that $A+B$ has an asymptotic density, and determine its possible values.

Definition (Ruzsa distance)**.** *Given sets A and B in a group G, define*

$$d_R(A,B) \;:=\; \log\left(\frac{|A-B|}{\sqrt{|A||B|}}\right).$$

The following result justifies the name 'Ruzsa distance'.

The Ruzsa Triangle Inequality. *The Ruzsa distance is non-negative, symmetric, and satisfies*

$$d(A,C) \;\leqslant\; d(A,B) + d(B,C)$$

for every A, B and C.

6. Prove the Ruzsa triangle inequality.

7. Show (without using Szemerédi's Theorem), that if $\underline{d}(A) > 0$, then $A - A$ contains arbitrarily long APs.

Bibliography

[1] N. Alon and J. Spencer, The Probabilistic Method, 3rd edition, Wiley, 2008.

[2] B. Bollobás, Modern Graph Theory, 2nd edition, Springer, 2002.

[3] J. Fox and B. Sudakov, Dependent Random Choice, *Random Structures Algorithms*, **38** (2011), 68–99.

[4] J. Matoušek, Using the Borsuk-Ulam Theorem: Lectures on Topological Methods in Combinatorics and Geometry, Springer, 2003.

[5] T. Tao and V. Vu, Additive Combinatorics, Cambridge University Press, 2006.