7 Extremal Graph Theory

In this chapter we study how global parameters of a graph, such as its edge density or chromatic number, can influence its local substructures. How many edges, for instance, do we have to give a graph on n vertices to be sure that, no matter how these edges are arranged, the graph will contain a K^r subgraph for some given r? Or at least a K^r minor? Will some sufficiently high average degree or chromatic number ensure that one of these substructures occurs?

Questions of this type are among the most natural ones in graph theory, and there is a host of deep and interesting results. Collectively, these are known as *extremal graph theory*.

Extremal graph problems in this sense fall neatly into two categories, as follows. If we are looking for ways to ensure by global assumptions that a graph G contains some given graph H as a minor (or topological minor), it will suffice to raise $\|G\|$ above the value of some linear function of $|G|$, i.e., to make $\varepsilon(G)$ large enough. The precise value of ε needed to force a desired minor or topological minor will be our topic in Section 7.2. Graphs whose number of edges is about[1] linear in their number of vertices are called *sparse*, so Section 7.2 is devoted to 'sparse extremal graph theory'. *sparse*

A particularly interesting way to force an H minor is to assume that $\chi(G)$ is large. Recall that if $\chi(G) \geqslant k+1$, say, then G has a subgraph G' with $2\varepsilon(G') \geqslant \delta(G') \geqslant k$ (Lemma 5.2.3). The question here is whether the effect of large χ is limited to this indirect influence via ε, or whether an assumption of $\chi \geqslant k+1$ can force bigger minors than

[1] Formally, the notions of sparse and dense (below) make sense only for classes of graphs whose order tends to infinity, not for individual graphs.

the assumption of $2\varepsilon \geqslant k$ can. *Hadwiger's conjecture*, which we meet in Section 7.3, asserts that χ has this quality. The conjecture can be viewed as a generalization of the four colour theorem, and is regarded by many as the most challenging open problem in graph theory.

On the other hand, if we ask what global assumptions might imply the existence of some given graph H as a *subgraph*, it will not help to raise invariants such as ε or χ, let alone any of the other invariants discussed in Chapter 1. For as soon as H contains a cycle, there are graphs of arbitrarily large chromatic number not containing H as a subgraph (Theorem 5.2.5). In fact, unless H is bipartite, any function f such that $f(n)$ edges on n vertices force an H subgraph must grow quadratically with n (why?).

dense

edge density

Graphs with a number of edges about quadratic in their number of vertices are usually called *dense*; the number $\|G\|/\binom{|G|}{2}$, the proportion of its potential edges that G actually has, is the *edge density* of G. The question of exactly which edge density is needed to force a given subgraph is the archetypal extremal graph problem, and it is our first topic in this chapter (Section 7.1). Rather than attempting to survey the wide field of 'dense extremal graph theory', however, we shall concentrate on its two most important results: we first prove Turán's classical extremal graph theorem for $H = K^r$ – a result that has served as a model for countless similar theorems for other graphs H – and then state the fundamental Erdős-Stone theorem, which gives precise asymptotic information for all H at once.

Although the Erdős-Stone theorem can be proved by elementary means, we shall use the opportunity of its proof to portray a powerful modern proof technique that has transformed much of extremal graph theory in recent years: Szemerédi's *regularity lemma*. This lemma is presented and proved in Section 7.4. In Section 7.5, we outline a general method for applying it, and illustrate this in the proof of the Erdős-Stone theorem. Another application of the regularity lemma will be given in Chapter 9.2.

7.1 Subgraphs

Let H be a graph and $n \geqslant |H|$. How many edges will suffice to force an H subgraph in any graph on n vertices, no matter how these edges are arranged? Or, to rephrase the problem: which is the greatest possible number of edges that a graph on n vertices can have *without* containing a copy of H as a subgraph? What will such a graph look like? Will it be unique?

A graph $G \not\supseteq H$ on n vertices with the largest possible number of edges is called *extremal* for n and H; its number of edges is denoted by

extremal

$\text{ex}(n, H)$. Clearly, any graph G that is extremal for some n and H will also be edge-maximal with $H \not\subseteq G$. Conversely, though, edge-maximality does not imply extremality: G may well be edge-maximal with $H \not\subseteq G$ while having fewer than $\text{ex}(n, H)$ edges (Fig. 7.1.1). $\text{ex}(n, H)$

Fig. 7.1.1. Two graphs that are edge-maximal with $P^3 \not\subseteq G$; is the right one extremal?

As a case in point, we consider our problem for $H = K^r$ (with $r > 1$). A moment's thought suggests some obvious candidates for extremality here: all complete $(r-1)$-partite graphs are edge-maximal without containing K^r. But which among these have the greatest number of edges? Clearly those whose partition sets are as equal as possible, i.e. differ in size by at most 1: if V_1, V_2 are two partition sets with $|V_1| - |V_2| \geqslant 2$, we may increase the number of edges in our complete $(r-1)$-partite graph by moving a vertex from V_1 to V_2.

The unique complete $(r-1)$-partite graphs on $n \geqslant r-1$ vertices whose partition sets differ in size by at most 1 are called *Turán graphs*; we denote them by $T^{r-1}(n)$ and their number of edges by $t_{r-1}(n)$ (Fig. 7.1.2). For $n < r-1$ we shall formally continue to use these definitions, with the proviso that – contrary to our usual terminology – the partition sets may now be empty; then, clearly, $T^{r-1}(n) = K^n$ for all $n \leqslant r-1$. $T^{r-1}(n)$ $t_{r-1}(n)$

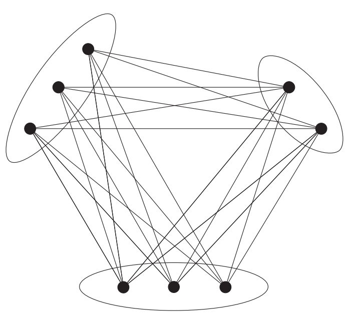

Fig. 7.1.2. The Turán graph $T^3(8)$

The following theorem tells us that $T^{r-1}(n)$ is indeed extremal for n and K^r, and as such unique; in particular, $\text{ex}(n, K^r) = t_{r-1}(n)$.

Theorem 7.1.1. (Turán 1941) [7.1.2] [9.2.2]
For all integers r, n with $r > 1$, every graph $G \not\supseteq K^r$ with n vertices and $\text{ex}(n, K^r)$ edges is a $T^{r-1}(n)$.

We give two proofs: one using induction, the other by a short and direct local argument. While that argument in the second proof is perhaps clever, the first proof is particularly natural and offers more insight.

First proof. We apply induction on n. For $n \leqslant r-1$ we have $G = K^n = T^{r-1}(n)$ as claimed. For the induction step, let now $n \geqslant r$.

K Since G is edge-maximal without a K^r subgraph, it has a subgraph $K = K^{r-1}$, with vertices $x_1, \ldots, x_{r-1}$ say. From $K^r \not\subseteq G$ alone we have

$$\|G - K\| \leqslant \mathrm{ex}(n-r+1, K^r) \tag{1}$$

and

Every vertex in $G-K$ sends at most $r-2$ edges to K. (2)

$V_1, \ldots, V_{r-1}$ For $i = 1, \ldots, r-1$ let $V_i := \{\, v \in V(G) \mid vx_i \notin E(G) \,\}$ (Fig. 7.1.3).

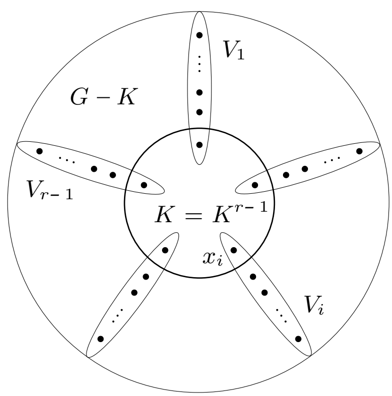

Fig. 7.1.3. Are the sets $V_1, \ldots, V_{r-1}$ covering $G-K$ disjoint? Are they each independent?

Let G' be obtained from K and an $(r-1)$-partite Turán graph on $V(G-K)$ by joining its ith partition class completely to $K - \{x_i\}$ for each i. This G' contains no K^r and satisfies both (1) and (2) with equality, in the case of (1) by the induction hypothesis. Hence so does G, because G cannot have fewer edges than G'.

Equality in (1) means, by the induction hypothesis, that $G-K$ is indeed an $(r-1)$-partite Turán graph. And its partition classes must be the sets $V_i \setminus \{x_i\}$, since equality in (2) implies with $G \not\supseteq K^r$ that every V_i is independent, and $(r-1)$-partite Turán graphs cannot be covered by $r-1$ sets of independent vertices other than their partition classes. In particular, these classes differ in size by at most 1. Hence G is a Turán graph, too: with classes $V_1, \ldots, V_{r-1}$. □

In our second proof of Turán's theorem we shall use an operation called *vertex duplication*. By *duplicating* a vertex $v \in G$ we mean adding to G a new vertex v' and joining it to exactly the neighbours of v (but not to v itself). *vertex duplication*

Second proof. We have already seen that among the complete k-partite graphs on n vertices the Turán graphs $T^k(n)$ have the most edges, and their degrees show that $T^{r-1}(n)$ has more edges than any $T^k(n)$ with $k < r-1$. So it suffices to show that G is complete multipartite.

If not, then non-adjacency is not an equivalence relation on $V(G)$, and so there are vertices y_1, x, y_2 such that $y_1x, xy_2 \notin E(G)$ but $y_1y_2 \in E(G)$. If $d(y_1) > d(x)$, then deleting x and duplicating y_1 yields another K^r-free graph with more edges than G, contradicting the choice of G. So $d(y_1) \leqslant d(x)$, and similarly $d(y_2) \leqslant d(x)$. But then deleting both y_1 and y_2 and duplicating x twice yields a K^r-free graph with more edges than G, again contradicting the choice of G. $\square$

The Turán graphs $T^{r-1}(n)$ are dense: in order of magnitude, they have about n^2 edges. More exactly, for every n and r we have

$$t_{r-1}(n) \leqslant \tfrac{1}{2}n^2\,\frac{r-2}{r-1}\,,$$

with equality whenever $r-1$ divides n (Exercise 7). It is therefore remarkable that just ϵn^2 more edges (for any fixed $\epsilon > 0$ and n large) give us not only a K^r subgraph (as does Turán's theorem) but a K^r_s for any given integer s – a graph itself teeming with K^r subgraphs:

Theorem 7.1.2. (Erdős & Stone 1946)
For all integers $r \geqslant 2$ and $s \geqslant 1$, and every $\epsilon > 0$, there exists an integer n_0 such that every graph with $n \geqslant n_0$ vertices and at least

$$t_{r-1}(n) + \epsilon n^2$$

edges contains K^r_s as a subgraph.

A proof of the Erdős-Stone theorem will be given in Section 7.5, as an illustration of how the regularity lemma may be applied. But the theorem can also be proved directly; see the notes for references.

The Erdős-Stone theorem is interesting not only in its own right: it also has a most interesting corollary. In fact, it was this entirely unexpected corollary that established the theorem as a kind of meta-theorem for the extremal theory of dense graphs, and thus made it famous.

Given a graph H and an integer n, consider the number $h_n := \mathrm{ex}(n,H)/\binom{n}{2}$: the maximum edge density that an n-vertex graph can have without containing a copy of H. Could it be that this critical density is essentially just a function of H, that h_n converges as $n \to \infty$? Theorem 7.1.2 implies this, and more: the limit of h_n is determined by a very simple function of a natural invariant of H – its chromatic number!

Corollary 7.1.3. *For every graph H with at least one edge,*

$$\lim_{n\to\infty} \mathrm{ex}(n,H)\binom{n}{2}^{-1} = \frac{\chi(H)-2}{\chi(H)-1}.$$

For the proof of Corollary 7.1.3 we need as a lemma that $t_{r-1}(n)$ never deviates much from the value it takes when $r-1$ divides n (see above), and that $t_{r-1}(n)/\binom{n}{2}$ converges accordingly. The proof of the lemma is left as an easy exercise with hint (Exercise 8).

[7.1.2] **Lemma 7.1.4.**

$$\lim_{n\to\infty} t_{r-1}(n)\binom{n}{2}^{-1} = \frac{r-2}{r-1}.$$

□

r **Proof of Corollary 7.1.3.** Let $r := \chi(H)$. Since H cannot be coloured with $r-1$ colours, we have $H \not\subseteq T^{r-1}(n)$ for all $n \in \mathbb{N}$, and hence

$$t_{r-1}(n) \leqslant \mathrm{ex}(n,H).$$

On the other hand, $H \subseteq K^r_s$ for all sufficiently large s, so

$$\mathrm{ex}(n,H) \leqslant \mathrm{ex}(n,K^r_s)$$

s for all those s. Let us fix such an s. For every $\epsilon > 0$, Theorem 7.1.2 implies that eventually (i.e. for large enough n)

$$\mathrm{ex}(n,K^r_s) < t_{r-1}(n) + \epsilon n^2.$$

Hence for n large,

$$\begin{aligned} t_{r-1}(n)/\tbinom{n}{2} &\leqslant \mathrm{ex}(n,H)/\tbinom{n}{2} \\ &\leqslant \mathrm{ex}(n,K^r_s)/\tbinom{n}{2} \\ &< t_{r-1}(n)/\tbinom{n}{2} + \epsilon n^2/\tbinom{n}{2} \\ &= t_{r-1}(n)/\tbinom{n}{2} + 2\epsilon/(1-\tfrac{1}{n}) \\ &\leqslant t_{r-1}(n)/\tbinom{n}{2} + 4\epsilon \qquad \text{(assume } n \geqslant 2\text{)}. \end{aligned}$$

Therefore, since $t_{r-1}(n)/\binom{n}{2}$ converges to $\frac{r-2}{r-1}$ (Lemma 7.1.4), so does $\mathrm{ex}(n,H)/\binom{n}{2}$. □

For bipartite graphs H, Corollary 7.1.3 says that substantially fewer than $\binom{n}{2}$ edges suffice to force an H subgraph. It turns out that

$$c_1 n^{2-\frac{2}{r+1}} \leqslant \mathrm{ex}(n, K_{r,r}) \leqslant c_2 n^{2-\frac{1}{r}}$$

for suitable constants c_1, c_2 depending on r; the lower bound is obtained by random graphs,[2] the upper bound is calculated in Exercise 12. If H is a forest, then $H \subseteq G$ as soon as $\varepsilon(G)$ is large enough, so $\mathrm{ex}(n, H)$ is at most linear in n (Exercise 14). Erdős and Sós conjectured in 1963 that $\mathrm{ex}(n, T) \leqslant \frac{1}{2}(k-1)n$ for all trees with $k \geqslant 2$ edges; as a general bound for all n, this is best possible for every T (Exercises 15–17). *Erdős-Sós conjecture*

A related but rather different question is whether large values of ε or χ can force a graph G to contain a given tree T as an *induced* subgraph. Of course, we need some additional assumption for this to make sense – for example, to prevent G from just being a large complete graph. The weakest sensible such assumption is that G has bounded clique number, i.e., that $G \not\supseteq K^r$ for some fixed integer r. Then large average degree still does not force an induced copy of T – consider complete bipartite graphs – but large chromatic number might: according to the Gyárfás-Sumner conjecture from Chapter 5.6, there exists for every $r \in \mathbb{N}$ and every tree T an integer $k = k(T, r)$ such that every graph G with $\omega(G) \leqslant r$ and $\chi(G) > k$ contains T as an induced subgraph.

7.2 Minors

In this section and the next we ask to what extent assumptions about invariants of a graph such as average degree, chromatic number, or girth can force it to contain another given graph as a minor or topological minor.

As a starting question, let us consider the analogue of Turán's theorem: how many edges on n vertices force a K^r minor or topological minor? The qualitative answer is that, unlike for K^r subgraphs where we might need as many as $\frac{1}{2}\frac{r-2}{r-1}n^2$ edges, a number of edges linear in n is enough: it suffices to assume that the graph has large enough average degree (depending on r).

Proposition 7.2.1. *Every graph of average degree at least 2^{r-2} has a K^r minor, for all $r \in \mathbb{N}$.*

Proof. We apply induction on r. For $r \leqslant 2$ the assertion is trivial. For the induction step let $r \geqslant 3$, and let G be any graph of average degree at least 2^{r-2}. Then $\varepsilon(G) \geqslant 2^{r-3}$; let H be a minimal minor of G with $\varepsilon(H) \geqslant 2^{r-3}$. Pick a vertex $x \in H$. By the minimality of H, x is not iso-

[2] see Chapter 11

lated. And each of its neighbours y has at least 2^{r-3} common neighbours with x: otherwise contracting the edge xy would lose us one vertex and at most 2^{r-3} edges, yielding a smaller minor H' with $\varepsilon(H') \geqslant 2^{r-3}$. The subgraph induced in H by the neighbours of x therefore has minimum degree at least 2^{r-3}, and hence has a K^{r-1} minor by the induction hypothesis. Together with x this yields the desired K^r minor of G. □

In Proposition 7.2.1 we needed an average degree of 2^{r-2} to force a K^r minor by induction on r. Forcing a topological K^r minor is a little harder: we shall fix its branch vertices in advance and then construct its subdivided edges inductively, which requires an average degree of $2^{\binom{r}{2}}$ to start with. Apart from this difference, the proof follows the same idea:

Proposition 7.2.2. *Every graph of average degree at least $2^{\binom{r}{2}}$ has a topological K^r minor, for every integer $r \geqslant 2$.*

(1.2.2)
(1.3.1)

Proof. The assertion is clear for $r = 2$, so let us assume that $r \geqslant 3$. We show by induction on $m = r, \dots, \binom{r}{2}$ that every graph G of average degree $d(G) \geqslant 2^m$ has a topological minor X with r vertices and m edges.

If $m = r$ then, by Propositions 1.2.2 and 1.3.1, G contains a cycle of length at least $\varepsilon(G) + 1 \geqslant 2^{r-1} + 1 \geqslant r + 1$, and the assertion follows with $X = C^r$.

Now let $r < m \leqslant \binom{r}{2}$, and assume the assertion holds for smaller m. Let G with $d(G) \geqslant 2^m$ be given; thus, $\varepsilon(G) \geqslant 2^{m-1}$. Since G has a component C with $\varepsilon(C) \geqslant \varepsilon(G)$, we may assume that G is connected.

U Consider a maximal set $U \subseteq V(G)$ such that U is connected in G and $\varepsilon(G/U) \geqslant 2^{m-1}$; such a set U exists, because G itself has the form G/U with $|U| = 1$. Since G is connected, we have $N(U) \neq \emptyset$.

H Let $H := G[N(U)]$. If H has a vertex v of degree $d_H(v) < 2^{m-1}$, we may add it to U and obtain a contradiction to the maximality of U: when we contract the edge vv_U in G/U, we lose one vertex and $d_H(v) + 1 \leqslant 2^{m-1}$ edges, so ε will still be at least 2^{m-1}. Therefore $d(H) \geqslant \delta(H) \geqslant 2^{m-1}$. By the induction hypothesis, H contains a TY with $|Y| = r$ and $\|Y\| = m - 1$. Let x, y be two branch vertices of this TY that are non-adjacent in Y. Since x and y lie in $N(U)$ and U is connected in G, G contains an x–y path whose inner vertices lie in U. Adding this path to the TY, we obtain the desired TX. □

In Chapter 3.5 we used the TK^r from Proposition 7.2.2 (stated there as Lemma 3.5.1) for a first proof that large enough connectivity $f(k)$ implies that a graph is k-linked. Later, in Theorem 3.5.3, we saw that connectivity as low as $2k$, coupled with an average degree of at least $16k$, is enough to imply this.

Conversely, we can use the more involved Theorem 3.5.3 to reduce the bound in Proposition 7.2.2 from exponential to quadratic, which is best possible up to a multiplicative constant (Exercise 24):

Theorem 7.2.3. *There is a constant $c \in \mathbb{R}$ such that, for every $r \in \mathbb{N}$, every graph G of average degree $d(G) \geqslant cr^2$ contains K^r as a topological minor.*

Proof. We prove the theorem with $c = 10$. Let G with $d(G) \geqslant 10r^2$ be given. By Theorem 1.4.3 for $k := r^2$, G has a subgraph H with $\kappa(H) \geqslant r^2$ and $\varepsilon(H) > \varepsilon(G) - r^2 \geqslant 4r^2$. For a TK^r in H, pick a set X of r vertices in H as branch vertices, and a set Y of $r(r-1)$ neighbours of X in H, $r-1$ for each vertex in X, as initial subdividing vertices. These are r^2 vertices altogether; they can be chosen distinct, since $\delta(H) \geqslant \kappa(H) \geqslant r^2$. (1.4.3) (3.5.3)

It remains to link up the vertices of Y in pairs, by disjoint paths in $H' := H - X$ corresponding to the edges of K^r. This can be done if Y is linked in H'. We show more generally that H' is $\frac{1}{2}r(r-1)$-linked, by checking that H' satisfies the premise of Theorem 3.5.3 for $k = \frac{1}{2}r(r-1)$. We have $\kappa(H') \geqslant \kappa(H) - r \geqslant r(r-1) = 2k$. And as H' was obtained from H by deleting at most $r|H|$ edges (as well as some vertices), we also have $\varepsilon(H') \geqslant \varepsilon(H) - r \geqslant 4r(r-1) = 8k$. □

For small r one can try to determine the exact number of edges needed to force a TK^r subgraph on n vertices. For $r = 4$, this number is $2n-2$; see Corollary 7.3.2. For $r = 5$, plane triangulations yield a lower bound of $3n-5$ (Corollary 4.2.10). The converse, that $3n-5$ edges do force a TK^5 – not just either a TK^5 or a $TK_{3,3}$, as they do by Corollary 4.2.10 and Kuratowski's theorem – is already a difficult theorem (Mader 1998).

The average degree needed to force an arbitrary K^r minor is less than that for a TK^r, and it is known very precisely; see the notes for the value of c in the following result.

Theorem 7.2.4. (Kostochka 1982)
There exists a constant $c \in \mathbb{R}$ such that, for every $r \in \mathbb{N}$, every graph G of average degree $d(G) \geqslant c\,r\sqrt{\log r}$ contains K^r as a minor. Up to the value of c, this bound is best possible as a function of r.

The easier implication of the theorem, the fact that in general an average degree of $c\,r\sqrt{\log r}$ is needed to force a K^r minor, follows from considering random graphs as introduced in Chapter 11. The converse implication, that this average degree suffices, is proved by methods not dissimilar to the proof of Theorem 3.5.3.

Rather than proving Theorem 7.2.4, therefore, we devote the remainder of this section to another striking aspect of forcing minors: that we can force a K^r minor in a graph simply by raising its girth (as long as we do not merely subdivide edges). At first glance, this may seem almost paradoxical. But it looks more plausible if, rather than trying to force a K^r minor directly, we instead try to force a minor just of large minimum or average degree – which suffices by Theorem 7.2.4. For if the girth g of a graph is large then the ball $\{ v \mid d(x,v) < \lfloor g/2 \rfloor \}$ around a vertex x induces a tree with many leaves, each of which sends all but one of its incident edges away from the tree. Contracting enough disjoint such trees we can thus hope to obtain a minor of large average degree, which in turn will have a large complete minor.

The following lemma realizes this idea.

Lemma 7.2.5. *Let $d,k \in \mathbb{N}$ with $d \geqslant 3$, and let G be a graph of minimum degree $\delta(G) \geqslant d$ and girth $g(G) \geqslant 8k+3$. Then G has a minor H of minimum degree $\delta(H) \geqslant d(d-1)^k$.*

Proof. Let $X \subseteq V(G)$ be maximal with $d(x,y) > 2k$ for all distinct X $x,y \in X$. For each $x \in X$ put $T_x^0 := \{x\}$. Given $i < 2k$, assume that we have defined disjoint trees $T_x^i \subseteq G$ (one for each $x \in X$) whose vertices together are precisely the vertices at distance at most i from X in G. Joining each vertex at distance $i+1$ from X to a neighbour at distance i, we obtain a similar set of disjoint trees T_x^{i+1}. As every vertex of G has distance at most $2k$ from X (by the maximality of X), the trees T_x $T_x := T_x^{2k}$ obtained in this way partition the entire vertex set of G. Let H be the minor of G obtained by contracting every T_x.

To prove that $\delta(H) \geqslant d(d-1)^k$, note first that the T_x are induced subgraphs of G, because $\operatorname{diam}(T_x) \leqslant 4k$ and $g(G) > 4k+1$. Similarly, there is at most one edge in G between any two trees T_x and T_y: two such edges, together with the paths joining their ends in T_x and T_y, would form a cycle of length at most $8k+2 < g(G)$. So all the edges leaving T_x are preserved in the contraction.

How many such edges are there? Note that, for every vertex $u \in T_x^{k-1}$, all its $d_G(u) \geqslant d$ neighbours v also lie in T_x: since $d(v,x) \leqslant k$ and $d(x,y) > 2k$ for every other $y \in X$, we have $d(v,y) > k \geqslant d(v,x)$, so v was added to T_x rather than to T_y when those trees were defined. Therefore T_x^k, and hence also T_x, has at least $d(d-1)^{k-1}$ leaves. But every leaf of T_x sends at least $d-1$ edges away from T_x, so T_x sends at least $d(d-1)^k$ edges to (distinct) other trees T_y. □

Lemma 7.2.5 provides Theorem 7.2.4 with the following corollary:

Theorem 7.2.6. (Thomassen 1983)
There exists a function $f\colon \mathbb{N} \to \mathbb{N}$ such that every graph of minimum degree at least 3 and girth at least $f(r)$ has a K^r minor, for all $r \in \mathbb{N}$.

Proof. We prove the theorem with $f(r) := 8\log r + 4\log\log r + c$, for some constant $c \in \mathbb{R}$. Let $k = k(r) \in \mathbb{N}$ be minimal with $3 \cdot 2^k \geqslant c' r\sqrt{\log r}$, where $c' \in \mathbb{R}$ is the constant from Theorem 7.2.4. Then for a suitable constant $c \in \mathbb{R}$ we have $8k+3 \leqslant 8\log r + 4\log\log r + c$, and the result follows by Lemma 7.2.5 and Theorem 7.2.4. □

Large girth can also be used to force a topological K^r minor. We now need some vertices of degree at least $r-1$ to serve as branch vertices, but if we assume a minimum degree of $r-1$ to secure these, we can even get by with a girth bound that is independent of r:

Theorem 7.2.7. (Kühn & Osthus 2002) [7.3.9]
There exists a constant g such that $G \supseteq TK^r$ for every graph G satisfying $\delta(G) \geqslant r-1$ and $g(G) \geqslant g$.

7.3 Hadwiger's conjecture

As we saw in Section 7.2, an average degree of $c\,r\sqrt{\log r}$ suffices to force an arbitrary graph to have a K^r minor, and an average degree of cr^2 forces it to have a topological K^r minor. If we replace 'average degree' above with 'chromatic number' then, with almost the same constants c, the two assertions remain true: this is because every graph with chromatic number k has a subgraph of average degree at least $k-1$ (Lemma 5.2.3).

Although both functions above, $c\,r\sqrt{\log r}$ and cr^2, are best possible (up to the constant c) for the said implications with 'average degree', the question arises whether they are still best possible with 'chromatic number' – or whether some slower-growing function would do in that case. What lies hidden behind this problem about growth rates is a fundamental question about the nature of the invariant χ: can this invariant have some direct *structural* effect on a graph in terms of forcing concrete substructures, or is its effect no greater than that of the 'unstructural' property of having lots of edges somewhere, which it implies trivially?

Neither for general nor for topological minors is the answer to this question known. For general minors, however, the following conjecture of Hadwiger suggests a positive answer:

Conjecture. (Hadwiger 1943)
The following implication holds for every integer $r > 0$ and every graph G:

$$\chi(G) \geqslant r \;\;\Rightarrow\;\; G \succcurlyeq K^r.$$

Hadwiger's conjecture is trivial for $r \leqslant 2$, easy for $r = 3$ and $r = 4$ (exercises), and equivalent to the four colour theorem for $r = 5$ and $r = 6$. For $r \geqslant 7$ the conjecture is open, but it is true for line graphs (Exercise 34) and for graphs of large girth (Exercise 32; see also Corollary 7.3.9). Rephrased as $G \succcurlyeq K^{\chi(G)}$, it is true for almost all graphs.[3] In general, the conjecture for $r+1$ implies it for r (exercise).

[3] See Chapter 11 for the notion of 'almost all'.

The Hadwiger conjecture for any fixed r is equivalent to the assertion that every graph without a K^r minor has an $(r-1)$-colouring. In this reformulation, the conjecture raises the question of what the graphs without a K^r minor look like: any sufficiently detailed structural description of those graphs should enable us to decide whether or not they can be $(r-1)$-coloured.

For $r = 3$, for example, the graphs without a K^r minor are precisely the forests (why?), and these are indeed 2-colourable. For $r = 4$, there is also a simple structural characterization of the graphs without a K^r minor:

[12.6.2] **Proposition 7.3.1.** *A graph with at least three vertices is edge-maximal without a K^4 minor if and only if it can be constructed recursively from triangles by pasting[4] along K^2s.*

(1.7.3) (4.4.4) *Proof.* Recall first that every IK^4 contains a TK^4, because $\Delta(K^4) = 3$ (Proposition 1.7.3); the graphs without a K^4 minor thus coincide with those without a topological K^4 minor. The forward implication is left as an easy exercise. To prove Hadwiger's conjecture for $r = 4$ we only need the backward implication, which we prove by induction on $|G|$.

Let G be edge-maximal without a TK^4. Then $\kappa(G) \leqslant 2$, as at the start of the proof of Lemma 3.2.2. Hence either G is a triangle, or the assertion follows from the induction hypothesis by Lemma 4.4.4.[5] □

One of the interesting consequences of Proposition 7.3.1 is that all the edge-maximal graphs without a K^4 minor have the same number of edges, and are thus all 'extremal':

Corollary 7.3.2. *Every edge-maximal graph G without a K^4 minor has $2\,|G| - 3$ edges.*

Proof. Induction on $|G|$. □

Corollary 7.3.3. *Hadwiger's conjecture holds for $r = 4$.*

Proof. If G arises from G_1 and G_2 by pasting along a complete graph, then $\chi(G) = \max\{\chi(G_1), \chi(G_2)\}$ (see the proof of Proposition 5.5.2). Hence, Proposition 7.3.1 implies by induction on $|G|$ that all edge-maximal (and hence all) graphs without a K^4 minor can be 3-coloured. □

It is also possible to prove Corollary 7.3.3 by a simple direct argument (Exercise 33).

[4] This was defined formally in Chapter 5.5.

[5] The proof of this lemma is elementary and can be read independently of the rest of Chapter 4.

By the four colour theorem, Hadwiger's conjecture for $r = 5$ follows from the following structure theorem for the graphs without a K^5 minor, just as it follows from Proposition 7.3.1 for $r = 4$. The proof of Theorem 7.3.4 is similar to that of Proposition 7.3.1, but considerably longer. We therefore state the theorem without proof:

Theorem 7.3.4. (Wagner 1937)
Let G be an edge-maximal graph without a K^5 minor. If $|G| \geqslant 4$ then G can be constructed recursively, by pasting along triangles and K^2s, from plane triangulations and copies of the graph W (Fig. 7.3.1).

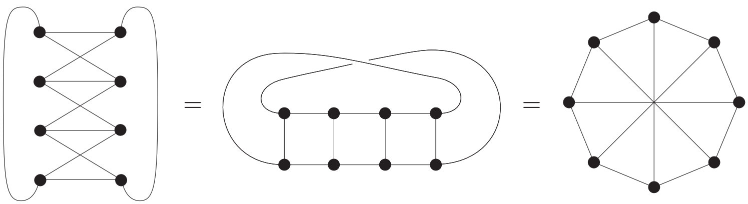

Fig. 7.3.1. Three representations of the *Wagner graph W*

Using Corollary 4.2.10, one can easily compute which of the graphs constructed as in Theorem 7.3.4 have the most edges. It turns out that these *extremal* graphs without a K^5 minor have no more edges than those that are extremal with respect to $\{IK^5, IK_{3,3}\}$, i.e. the maximal planar graphs: (4.2.10)

Corollary 7.3.5. *A graph with n vertices and no K^5 minor has at most $3n - 6$ edges.* □

Since $\chi(W) = 3$, Theorem 7.3.4 and the four colour theorem imply Hadwiger's conjecture for $r = 5$:

Corollary 7.3.6. *Hadwiger's conjecture holds for $r = 5$.* □

The Hadwiger conjecture for $r = 6$ is again substantially more difficult than the case $r = 5$, and again it relies on the four colour theorem. The proof shows (without using the four colour theorem) that any minimal-order counterexample arises from a planar graph by adding one vertex – so by the four colour theorem it is not a counterexample after all.

Theorem 7.3.7. (Robertson, Seymour & Thomas 1993)
Hadwiger's conjecture holds for $r = 6$.

As mentioned earlier, the challenge posed by Hadwiger's conjecture is to devise a proof technique that makes better use of the assumption of $\chi \geqslant r$ than just using its consequence of $\delta \geqslant r-1$ in a suitable subgraph, which we know cannot force a K^r minor (Theorem 7.2.4). So far, no such technique is known.

If we resign ourselves to using just $\delta \geqslant r-1$, we can still ask what additional assumptions might help in making this force a K^r minor. Theorem 7.2.7 says that an assumption of large girth has this effect; see also Exercise 32. In fact, a much weaker assumption suffices: for any fixed $s \in \mathbb{N}$ and all large enough d depending only on s, the graphs $G \not\supseteq K_{s,s}$ of average degree at least d can be shown to have K^r minors for r considerably larger than d. For Hadwiger's conjecture, this implies the following:

Theorem 7.3.8. (Kühn & Osthus 2005)
For every integer s there is an integer r_s such that Hadwiger's conjecture holds for all graphs $G \not\supseteq K_{s,s}$ and $r \geqslant r_s$.

The strengthening of Hadwiger's conjecture that graphs of chromatic number at least r contain K^r as a *topological* minor has become known as *Hajós's conjecture*. It is false in general, but Theorem 7.2.7 implies it for graphs of large girth:

Corollary 7.3.9. *There is a constant g such that all graphs G of girth at least g satisfy the implication $\chi(G) \geqslant r \Rightarrow G \supseteq TK^r$ for all r.*

(5.2.3)
(7.2.7)

Proof. Let g be the constant from Theorem 7.2.7. If $\chi(G) \geqslant r$ then, by Lemma 5.2.3, G has a subgraph H of minimum degree $\delta(H) \geqslant r-1$. As $g(H) \geqslant g(G) \geqslant g$, Theorem 7.2.7 implies that $G \supseteq H \supseteq TK^r$. □

7.4 Szemerédi's regularity lemma

Some 50 years ago, in the course of the proof of a theorem about arithmetic progressions of integers, Szemerédi developed a graph-theoretical tool that has since come to dominate methods in extremal graph theory like none other: his *regularity lemma*. Very roughly, the lemma says that all graphs can be approximated by random graphs in the following sense: every graph can be partitioned, into a bounded number of equal parts, so that most of its edges run between different parts and the edges between any two parts are distributed fairly uniformly – just as we would expect it if they had been generated at random.

In order to state the regularity lemma precisely, we need some definitions. Let $G = (V, E)$ be a graph, and let $X, Y \subseteq V$ be disjoint. We denote by $\|X, Y\|$ the number of X–Y edges of G, and call $\|X, Y\|$

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

$d(X,Y)$

the *density* of the pair (X, Y). (This is a real number between 0 and 1.) Given any $\epsilon > 0$, we call a pair (A, B) of disjoint sets $A, B \subseteq V$ *ϵ-regular* if all $X \subseteq A$ and $Y \subseteq B$ with *density* *ϵ-regular pair*

$$|X| \geqslant \epsilon\,|A| \quad \text{and} \quad |Y| \geqslant \epsilon\,|B|$$

satisfy

$$\big|d(X,Y) - d(A,B)\big| \leqslant \epsilon\,.$$

The edges in an ϵ-regular pair are thus distributed fairly uniformly, the more so the smaller the ϵ we started with.

Consider a partition $\{V_0, V_1, \ldots, V_k\}$ of V in which one set V_0 has been singled out as an *exceptional set*. (This exceptional set V_0 may be empty.[6]) We call such a partition an *ϵ-regular* partition of G if it satisfies the following three conditions: *exceptional set*

(i) $|V_0| \leqslant \epsilon\,|V|$; *$\epsilon$-regular partition*

(ii) $|V_1| = \ldots = |V_k|$;

(iii) all but at most ϵk^2 of the pairs (V_i, V_j) with $1 \leqslant i < j \leqslant k$ are ϵ-regular.

The role of the exceptional set V_0 is one of pure convenience: it makes it possible to require that all the other partition sets have exactly the same size. Since condition (iii) affects only the sets $V_1, \ldots, V_k$, we may think of V_0 as a kind of bin: its vertices are disregarded when the regularity of the partition is assessed, but there are only few such vertices.

Theorem 7.4.1. (Regularity Lemma) [7.1.2] [9.2.2]
For every $\epsilon > 0$ and every integer $m \geqslant 1$ there exists an integer M such that every graph of order at least m admits an ϵ-regular partition $\{V_0, V_1, \ldots, V_k\}$ with $m \leqslant k \leqslant M$.

---

[6] So V_0 may be an exception also to our terminological rule that partition sets are not normally empty.

The regularity lemma thus says that, given any $\epsilon > 0$, every graph has an ϵ-regular partition into a bounded number of sets. The upper bound M on the number of partition sets ensures that for large graphs the partition sets are large too; note that ϵ-regularity is trivial when the partition sets are singletons, and a powerful property when they are large. The lemma also allows us to specify a lower bound m for the number of partition sets. This can be used to increase the proportion of edges running between different partition sets (i.e., of edges governed by the regularity assertion) over edges inside partition sets (about which we know nothing). See Exercise 39 for an example.

Note that the regularity lemma in this form is designed for use with dense graphs:[7] for sparse graphs it becomes trivial, because all densities of pairs – and hence their differences – tend to zero (Exercise 40).

The remainder of this section is devoted to the proof of the regularity lemma. Although the proof is not difficult, any reader meeting the regularity lemma here for the first time is likely to draw more insight from seeing how the lemma is typically applied than from studying the technicalities of its proof. Any such reader is encouraged to skip to the start of Section 7.5 now and come back to the proof at his or her leisure.

We shall need the following inequality for reals $\mu_1, \dots, \mu_k > 0$ and $e_1, \dots, e_k \geqslant 0$:

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

This follows from the Cauchy-Schwarz inequality $\sum a_i^2 \sum b_i^2 \geqslant \left(\sum a_i b_i\right)^2$ by taking $a_i := \sqrt{\mu_i}$ and $b_i := e_i/\sqrt{\mu_i}$.

$G = (V, E)$ n Let $G = (V, E)$ be a graph and $n := |V|$. For disjoint sets $A, B \subseteq V$ we define

$q(A, B)$

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

For partitions $\mathcal{A}$ of A and $\mathcal{B}$ of B we set

$q(\mathcal{A}, \mathcal{B})$

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

and for a partition $\mathcal{P} = \{C_1, \dots, C_k\}$ of V we let

$q(\mathcal{P})$

$$q(\mathcal{P}) := \sum_{i<j} q(C_i, C_j)\,.$$

[7] Sparse versions were developed later; see the notes.

However, if $\mathcal{P} = \{C_0, C_1, \ldots, C_k\}$ is a partition of V *with exceptional set* C_0, we treat C_0 as a set of singletons and define

$$q(\mathcal{P}) := q(\tilde{\mathcal{P}}),$$

where $\tilde{\mathcal{P}} := \{C_1, \ldots, C_k\} \cup \big\{\{v\} : v \in C_0\big\}$. $\tilde{\mathcal{P}}$

The function $q(\mathcal{P})$ plays a pivotal role in our proof of the regularity lemma. On the one hand, it measures the regularity of the partition $\mathcal{P}$: if $\mathcal{P}$ has too many irregular pairs (A, B), we can take some of the pairs (X, Y) of subsets violating the regularity of the pairs (A, B) and make those sets X and Y into partition sets of their own; as we shall prove, this can refine $\mathcal{P}$ into a partition for which q is substantially greater than for $\mathcal{P}$. Here, 'substantial' means that the increase of $q(\mathcal{P})$ is bounded below by some constant depending only on ϵ. On the other hand,

$$\begin{aligned} q(\mathcal{P}) &= \sum_{i<j} q(C_i, C_j) \\ &= \sum_{i<j} \frac{|C_i|\,|C_j|}{n^2}\, d^2(C_i, C_j) \\ &\leqslant \frac{1}{n^2} \sum_{i<j} |C_i|\,|C_j| \ \leqslant\ 1\,. \end{aligned}$$

The number of times that $q(\mathcal{P})$ can be increased by a constant is thus also bounded by a constant – in other words, after some bounded number of refinements our partition will be ϵ-regular! To complete the proof of the regularity lemma, all we have to do then is to note how many sets that last partition can possibly have if we start with a partition into m sets, and to choose this number as our promised bound M.

Let us make all this precise. We begin by showing that, when we refine any partition of V, its q-value will not decrease:

Lemma 7.4.2.

(i) *Let $C, D \subseteq V$ be disjoint. If $\mathcal{C}$ is a partition of C and $\mathcal{D}$ is a partition of D, then $q(\mathcal{C}, \mathcal{D}) \geqslant q(C, D)$.*

(ii) *If $\mathcal{P}, \mathcal{P}'$ are partitions of V and $\mathcal{P}'$ refines $\mathcal{P}$, then $q(\mathcal{P}') \geqslant q(\mathcal{P})$.*

Proof. (i) Let $\mathcal{C} =: \{C_1, \ldots, C_k\}$ and $\mathcal{D} =: \{D_1, \ldots, D_\ell\}$. Then

$$\begin{aligned} q(\mathcal{C}, \mathcal{D}) &= \sum_{i,j} q(C_i, D_j) \\ &= \frac{1}{n^2} \sum_{i,j} \frac{\|C_i, D_j\|^2}{|C_i|\,|D_j|} \\ &\underset{(1)}{\geqslant} \frac{1}{n^2} \frac{\left(\sum_{i,j} \|C_i, D_j\|\right)^2}{\sum_{i,j} |C_i|\,|D_j|} \end{aligned}$$

$$= \frac{1}{n^2} \frac{\|C, D\|^2}{\left(\sum_i |C_i|\right)\left(\sum_j |D_j|\right)}$$
$$= q(C, D)\,.$$

(ii) Let $\mathcal{P} =: \{C_1, \dots, C_k\}$, and for $i = 1, \dots, k$ let $\mathcal{C}_i$ be the partition of C_i induced by $\mathcal{P}'$. Then

$$\begin{aligned} q(\mathcal{P}) &= \sum_{i<j} q(C_i, C_j) \\ &\underset{\text{(i)}}{\leqslant} \sum_{i<j} q(\mathcal{C}_i, \mathcal{C}_j) \\ &\leqslant q(\mathcal{P}')\,, \end{aligned}$$

since $q(\mathcal{P}') = \sum_i q(\mathcal{C}_i) + \sum_{i<j} q(\mathcal{C}_i, \mathcal{C}_j)$. □

Next, we show that by subpartitioning an irregular pair of partition sets we can increase the q-value of that partition a little. Since we are dealing with a single pair only, however, we cannot guarantee yet that the amount of this increase exceeds any constant that depends only on ϵ.

Lemma 7.4.3. *Let $\epsilon > 0$, and let $C, D \subseteq V$ be disjoint. If (C, D) is not ϵ-regular, then there are partitions $\mathcal{C} = \{C_1, C_2\}$ of C and $\mathcal{D} = \{D_1, D_2\}$ of D such that*

$$q(\mathcal{C}, \mathcal{D}) \geqslant q(C, D) + \epsilon^4 \frac{|C|\,|D|}{n^2}\,.$$

Proof. Suppose (C, D) is not ϵ-regular. Then there are sets $C_1 \subseteq C$ and $D_1 \subseteq D$ with $|C_1| \geqslant \epsilon\,|C|$ and $|D_1| \geqslant \epsilon\,|D|$ such that

$$|\eta| > \epsilon \tag{2}$$

η for $\eta := d(C_1, D_1) - d(C, D)$. Let $\mathcal{C} := \{C_1, C_2\}$ and $\mathcal{D} := \{D_1, D_2\}$, where $C_2 := C \setminus C_1$ and $D_2 := D \setminus D_1$.

Let us show that $\mathcal{C}$ and $\mathcal{D}$ satisfy the conclusion of the lemma. We
c_i, d_i, e_{ij} shall write $c_i := |C_i|$, $d_i := |D_i|$, $e_{ij} := \|C_i, D_j\|$, $c := |C|$, $d := |D|$
c, d, e and $e := \|C, D\|$. As in the proof of Lemma 7.4.2,

$$\begin{aligned} q(\mathcal{C}, \mathcal{D}) &= \frac{1}{n^2} \sum_{i,j} \frac{e_{ij}^2}{c_i d_j} \\ &= \frac{1}{n^2} \left(\frac{e_{11}^2}{c_1 d_1} + \sum_{i+j>2} \frac{e_{ij}^2}{c_i d_j} \right) \\ &\underset{(1)}{\geqslant} \frac{1}{n^2} \left(\frac{e_{11}^2}{c_1 d_1} + \frac{(e - e_{11})^2}{cd - c_1 d_1} \right). \end{aligned}$$

By definition of η, we have $e_{11} = c_1 d_1 e/cd + \eta c_1 d_1$, so

$$\begin{aligned}
n^2\, q(\mathcal{C},\mathcal{D}) &\geqslant \frac{1}{c_1 d_1}\left(\frac{c_1 d_1 e}{cd} + \eta c_1 d_1\right)^2 \\
&\qquad + \frac{1}{cd - c_1 d_1}\left(\frac{cd - c_1 d_1}{cd} e - \eta c_1 d_1\right)^2 \\
&= \frac{c_1 d_1 e^2}{c^2 d^2} + \frac{2e\eta c_1 d_1}{cd} + \eta^2 c_1 d_1 \\
&\qquad + \frac{cd - c_1 d_1}{c^2 d^2} e^2 - \frac{2e\eta c_1 d_1}{cd} + \frac{\eta^2 c_1^2 d_1^2}{cd - c_1 d_1} \\
&\geqslant \frac{e^2}{cd} + \eta^2 c_1 d_1 \\
&\underset{(2)}{\geqslant} \frac{e^2}{cd} + \epsilon^4 cd \;=\; n^2 q(C,D) + \epsilon^4 |C|\,|D|
\end{aligned}$$

as claimed, since $c_1 \geqslant \epsilon c$ and $d_1 \geqslant \epsilon d$ by the choice of C_1 and D_1. □

Finally, we show that if a partition has enough irregular pairs of partition sets to fall short of the definition of an ϵ-regular partition, then subpartitioning all those pairs at once can increase q by a constant:

Lemma 7.4.4. *Let $0 < \epsilon \leqslant 1/4$, and let $\mathcal{P} = \{C_0, C_1, \dots, C_k\}$ be a partition of V, with exceptional set C_0 of size $|C_0| \leqslant \epsilon n$ and $|C_1| = \dots = |C_k| =: c$. If $\mathcal{P}$ is not ϵ-regular, then there is a partition* c *$\mathcal{P}' = \{C_0', C_1', \dots, C_\ell'\}$ of V with exceptional set C_0', where $k \leqslant \ell \leqslant k4^{k+1}$, such that $|C_0'| \leqslant |C_0| + n/2^k$, all other sets C_i' have equal size, and either $\mathcal{P}'$ is ϵ-regular or*

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

Proof. For all $1 \leqslant i < j \leqslant k$, let us define a partition $\mathcal{C}_{ij}$ of C_i and a $\mathcal{C}_{ij}$ partition $\mathcal{C}_{ji}$ of C_j, as follows. If the pair (C_i, C_j) is ϵ-regular, we let $\mathcal{C}_{ij} := \{C_i\}$ and $\mathcal{C}_{ji} := \{C_j\}$. If not, then by Lemma 7.4.3 there are partitions $\mathcal{C}_{ij}$ of C_i and $\mathcal{C}_{ji}$ of C_j with $|\mathcal{C}_{ij}| = |\mathcal{C}_{ji}| = 2$ and

$$q(\mathcal{C}_{ij}, \mathcal{C}_{ji}) \geqslant q(C_i, C_j) + \epsilon^4 \frac{|C_i|\,|C_j|}{n^2} = q(C_i, C_j) + \frac{\epsilon^4 c^2}{n^2}\,. \qquad (3)$$

For each $i = 1, \dots, k$, let $\mathcal{C}_i$ be the unique minimal partition of C_i that $\mathcal{C}_i$ refines every partition $\mathcal{C}_{ij}$ with $j \neq i$. (In other words, if we consider two elements of C_i as equivalent whenever they lie in the same partition set of $\mathcal{C}_{ij}$ for every $j \neq i$, then $\mathcal{C}_i$ is the set of equivalence classes.) Thus, $|\mathcal{C}_i| \leqslant 2^{k-1}$. Now consider the partition

$\mathcal{C}$

$$\mathcal{C} := \{C_0\} \cup \bigcup_{i=1}^{k} \mathcal{C}_i$$

of V, with C_0 as exceptional set. Then $\mathcal{C}$ refines $\mathcal{P}$ and $|\mathcal{C} \smallsetminus \{C_0\}| \leqslant k2^{k-1}$, so

$$k \leqslant |\mathcal{C}| \leqslant k2^k. \tag{4}$$

$\mathcal{C}_0$ Let $\mathcal{C}_0 := \big\{\{v\} \mid v \in C_0\big\}$. Now if $\mathcal{P}$ is not ϵ-regular, then for more than ϵk^2 of the pairs (C_i, C_j) with $1 \leqslant i < j \leqslant k$ the partition $\mathcal{C}_{ij}$ is non-trivial. Hence, by our definition of q for partitions with exceptional set, and Lemma 7.4.2 (i),

$$\begin{aligned} q(\mathcal{C}) &= \sum_{1\leqslant i<j} q(\mathcal{C}_i, \mathcal{C}_j) + \sum_{1\leqslant i} q(\mathcal{C}_0, \mathcal{C}_i) + \sum_{0\leqslant i} q(\mathcal{C}_i) \\ &\geqslant \sum_{1\leqslant i<j} q(\mathcal{C}_{ij}, \mathcal{C}_{ji}) + \sum_{1\leqslant i} q\big(\mathcal{C}_0, \{C_i\}\big) + q(\mathcal{C}_0) \\ &\underset{(3)}{\geqslant} \sum_{1\leqslant i<j} q(C_i, C_j) + \epsilon k^2 \frac{\epsilon^4 c^2}{n^2} + \sum_{1\leqslant i} q\big(\mathcal{C}_0, \{C_i\}\big) + q(\mathcal{C}_0) \\ &= q(\mathcal{P}) + \epsilon^5 \left(\frac{kc}{n}\right)^2 \\ &\geqslant q(\mathcal{P}) + \epsilon^5/2\,. \end{aligned}$$

(For the last inequality, recall that $|C_0| \leqslant \epsilon n \leqslant \frac{1}{4}n$, so $kc \geqslant \frac{3}{4}n$.)

In order to turn $\mathcal{C}$ into our desired partition $\mathcal{P}'$, all that remains to do is to cut its sets up into pieces of some common size so that all remaining vertices can be collected into the exceptional set without making this too large.

If $c < 4^k$, the ϵ-regular partition $\mathcal{P}'$ into $C_0' := C_0$ and the singletons $\{v\}$ with $v \in V \smallsetminus C_0$ is as desired, since there are ℓ such singletons for $k \leqslant \ell = kc < k4^k$.

Assume now that $c \geqslant 4^k$. Let $C_1', \dots, C_\ell'$ be a maximal collection of

d disjoint sets of size $d := \lfloor c/4^k \rfloor \geqslant 1$ such that each C_i' is contained in some

$\mathcal{P}'$ $C \in \mathcal{C} \smallsetminus \{C_0\}$, and put $C_0' := V \smallsetminus \bigcup C_i'$. Then $\mathcal{P}' = \{C_0', C_1', \dots, C_\ell'\}$ is indeed a partition of V. Moreover, $\tilde{\mathcal{P}}'$ refines $\tilde{\mathcal{C}}$, so

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

by Lemma 7.4.2 (ii). Since each set $C_i' \neq C_0'$ is also contained in one of the sets $C_1, \dots, C_k$, but no more than $c/d \leqslant 4^{k+1}$ sets C_i' can lie inside the same C_j (by the choice of d), we also have $k \leqslant \ell \leqslant k4^{k+1}$ as required. Finally, the sets $C'_1, \dots, C'_\ell$ use all but at most d vertices from each set $C \neq C_0$ of $\mathcal{C}$. Hence,

$$\begin{aligned} |C'_0| &\leqslant |C_0| + d\,|\mathcal{C}| \\ &\underset{(4)}{\leqslant} |C_0| + \frac{c}{4^k} k 2^k \\ &= |C_0| + ck/2^k \\ &\leqslant |C_0| + n/2^k. \end{aligned}$$

□

We are no ready to prove the regularity lemma. Our plan is, roughly, to start for fixed ϵ with any partition $\mathcal{P}$ of V to which we can apply Lemma 7.4.4. We shall iterate this until the partition the lemma returns is ϵ-regular. This should happen after at most $2/\epsilon^5$ iterations, since the q-value of our partition grows by at least $\epsilon^5/2$ at each step but cannot exceed 1. The number of sets in our final partition will therefore also be bounded in terms of ϵ, and we can take this bound as the value for M required in the assertion of the regularity lemma.

Proof of Theorem 7.4.1. Let $\epsilon > 0$ and $m \geqslant 1$ be given, and assume without loss of generality that $\epsilon \leqslant 1/4$. Let $s := 2/\epsilon^5$. ϵ, m

There is one critical requirement which a partition $\{C_0, C_1, \dots, C_k\}$ with $|C_1| = \ldots = |C_k|$ has to satisfy before Lemma 7.4.4 can be (re-) applied: the size $|C_0|$ of its exceptional set must not exceed ϵn. As the size of the exceptional set can grow with each iteration of the lemma, see there, we thus have to start with C_0 small enough that after up to $s-1$ increases its size is still less than ϵn.

The increase to $|C_0|$ effected by one application of Lemma 7.4.4 is stated there as being at most $n/2^k$. This k is the number of non-exceptional sets in the partition to which we apply the lemma, so $n/2^k$ decreases with each step. Hence we only need to start with k large enough that s increments even of our starting value of $n/2^k$ add up to at most $\epsilon n/2$ and ensure, if we can, that the initial size of C_0 is also at most $\epsilon n/2$.

Pick $k \geqslant m$ large enough that $2^{k-1} \geqslant s/\epsilon$. Then $s/2^k \leqslant \epsilon/2$, and hence k

$$k + sn/2^k \leqslant \epsilon n \tag{5}$$

whenever $k \leqslant \epsilon n/2$, i.e. for all $n \geqslant 2k/\epsilon$.

Let us now choose M. This should be an upper bound on the number of non-exceptional sets in our partition after up to s iterations of Lemma 7.4.4, where in each iteration this number may grow from its current value r to at most $r4^{r+1}$. So let f be the function $x \mapsto x4^{x+1}$,

M and take $M := \max\{f^s(k), 2k/\epsilon\}$; the second term in the maximum ensures that any $n \geqslant M$ is large enough to satisfy (5).

We finally have to show that every graph $G = (V, E)$ of order at least m has an ϵ-regular partition $\{V_0, V_1, \ldots, V_{k'}\}$ with $m \leqslant k' \leqslant M$. So n let G be given, and let $n := |G|$. If $n \leqslant M$, we partition G into $k' := n$ singletons, choosing $V_0 := \emptyset$ and $|V_1| = \ldots = |V_{k'}| = 1$. This partition of G is clearly ϵ-regular. Suppose now that $n > M$. Let $C_0 \subseteq V$ be minimal such that k divides $|V \setminus C_0|$, and let $\{C_1, \ldots, C_k\}$ be any partition of $V \setminus C_0$ into sets of equal size. Then $|C_0| \leqslant k-1$, and hence $|C_0| \leqslant \epsilon n$ by (5). Starting with $\{C_0, C_1, \ldots, C_k\}$ we apply Lemma 7.4.4 again and again, until the partition of G obtained is ϵ-regular; this will happen after at most s iterations, since by (5) the size of the exceptional set in the partitions stays below ϵn, so the lemma could indeed be reapplied up to the theoretical maximum of s times. □

7.5 Applying the regularity lemma

The purpose of this section is to illustrate how the regularity lemma is typically applied in the context of (dense) extremal graph theory. Suppose we are trying to prove that a certain edge density of a graph G suffices to force the occurrence of some given subgraph H, and that we have an ϵ-regular partition of G. For most of the pairs (V_i, V_j) of partition sets, the edges between V_i and V_j are distributed fairly uniformly; their density, however, may depend on the pair. But since G has many edges, this density cannot be too small for too many pairs: some sizeable proportion of the pairs will have at least a certain positive density. Moreover if G is large, then so are the pairs: recall that the number of partition sets is bounded, and they have equal size. But any large enough bipartite graph with equal partition sets, fixed positive edge density (however small) and a uniform distribution of edges will contain any given bipartite subgraph;[8] this will be made precise below. Writing H as a union of bipartite subgraphs, say those induced by pairs of colour classes of some vertex colouring of H, we shall obtain $H \subseteq G$ as desired.

These ideas will be formalized by the so-called 'blow-up lemma', Lemma 7.5.2 below. We shall then use this and the regularity lemma to prove the Erdős-Stone theorem from Section 7.1; another application will be given later, in the proof of Theorem 9.2.2. We wind up the section with an informal review of the application of the regularity lemma that we have seen, summarizing what it can teach us for similar applications. In particular, we look at how the various parameters involved depend

[8] Readers already acquainted with random graphs may find it instructive to compare this statement with Proposition 11.3.1.

on each other, and in which order they have to be chosen to make the lemma work.

Let us begin by noting a simple consequence of the ϵ-regularity of a pair (A, B). For any subset $Y \subseteq B$ that is not too small, most vertices of A have about the expected number of neighbours in Y:

Lemma 7.5.1. *Let (A, B) be an ϵ-regular pair, of density d say, and let $Y \subseteq B$ have size $|Y| \geqslant \epsilon\,|B|$. Then all but fewer than $\epsilon\,|A|$ of the vertices in A have (each) at least $(d-\epsilon)|Y|$ neighbours in Y.*

Proof. Let $X \subseteq A$ be the set of vertices with fewer than $(d-\epsilon)|Y|$ neighbours in Y. Then $\|X, Y\| < |X|(d-\epsilon)|Y|$, so

$$d(X, Y) = \frac{\|X, Y\|}{|X|\,|Y|} < d - \epsilon = d(A, B) - \epsilon\,.$$

As (A, B) is ϵ-regular and $|Y| \geqslant \epsilon\,|B|$, this implies that $|X| < \epsilon\,|A|$. $\square$

Let G be a graph with an ϵ-regular partition $\{V_0, V_1, \ldots, V_k\}$, with exceptional set V_0 and $|V_1| = \ldots = |V_k| =: \ell$. Given $d \in [0, 1]$, let R be the graph on $\{1, \ldots, k\}$ in which two vertices i, j are adjacent if and only if (V_i, V_j) is ϵ-regular of density $\geqslant d$. We shall call R a *regularity graph* of G with parameters ϵ, ℓ and d. Given $s \in \mathbb{N}$, let us now replace every vertex i of R by a set V_i^s of s vertices, and every edge by a complete bipartite graph between the corresponding s-sets. The resulting graph will be denoted by R_s. (For $R = K^r$, for example, we have $R_s = K^r_s$.)

R

regularity graph

V_i^s

R_s

Lemma 7.5.2, a simple version of the blow-up lemma that suffices for our intended applications of the regularity lemma, says that subgraphs of R_s can also be found in G, provided that $d > 0$, that ϵ is small enough, and that the V_i are large enough. In fact, the values of ϵ and ℓ required depend only on (d and) the maximum degree of the subgraph:

Lemma 7.5.2. (Blow-up Lemma)
For all $d \in (0, 1]$ and $\Delta \geqslant 1$ there exists an $\epsilon_0 > 0$ with the following property: if G is any graph, H is a graph with $\Delta(H) \leqslant \Delta$, $s \in \mathbb{N}$, and R is any regularity graph of G with parameters $\epsilon \leqslant \epsilon_0$, $\ell \geqslant 2s/d^\Delta$ and d, then [9.2.2]

$$H \subseteq R_s \;\Rightarrow\; H \subseteq G\,.$$

(See the notes for a the original more general blow-up lemma.)

Proof. Given d and Δ, choose $\epsilon_0 > 0$ small enough that $\epsilon_0 < d$ and

d, Δ, ϵ_0

$$(d-\epsilon_0)^\Delta - \Delta\epsilon_0 \geqslant \tfrac{1}{2}d^\Delta\,; \qquad (1)$$

G, H, R, R_s V_i ϵ, k, ℓ u_i, h σ v_i

such a choice is possible, since $(d-\epsilon)^\Delta - \Delta\epsilon \to d^\Delta$ as $\epsilon \to 0$. Now let G, H, s and R be given as stated. Let $\{V_0, V_1, \dots, V_k\}$ be the ϵ-regular partition of G that gave rise to R; thus, $\epsilon \leqslant \epsilon_0$ and $V(R) = \{1, \dots, k\}$ and $|V_1| = \ldots = |V_k| = \ell \geqslant 2s/d^\Delta$. Let us assume that H is actually a subgraph of R_s (not just isomorphic to one), with vertices $u_1, \dots, u_h$ say. Each vertex u_i lies in one of the s-sets V_j^s of R_s, which defines a map $\sigma: i \mapsto j$. Our aim is to define an embedding $u_i \mapsto v_i \in V_{\sigma(i)}$ of H in G as a subgraph; thus, $v_1, \dots, v_h$ will be distinct, and $v_i v_j$ will be an edge of G whenever $u_i u_j$ is an edge of H.

Our plan is to choose the vertices $v_1, \dots, v_h$ inductively. Throughout the induction, we shall have a 'target set' $Y_i \subseteq V_{\sigma(i)}$ assigned to each u_i; this contains the vertices that are still candidates for the choice of v_i. Initially, Y_i is the entire set $V_{\sigma(i)}$. As the embedding proceeds, Y_i will get smaller and smaller (until it collapses to $\{v_i\}$ when v_i is chosen): whenever we choose a vertex v_j with $j < i$ and $u_j u_i \in E(H)$, we delete all those vertices from Y_i that are not adjacent to v_j. The set Y_i thus evolves as

$$V_{\sigma(i)} = Y_i^0 \supseteq \ldots \supseteq Y_i^i = \{v_i\},$$

where Y_i^j denotes the version of Y_i current after the definition of v_j and the resulting deletion of vertices from Y_i^{j-1}.

In order to make this approach work, we have to ensure that the target sets Y_i do not get too small. When we come to embed a vertex u_j, we consider all the indices $i > j$ with $u_j u_i \in E(H)$; there are at most Δ such i. For each of these i, we wish to select v_j so that

$$Y_i^j = N(v_j) \cap Y_i^{j-1} \tag{2}$$

is still relatively large: smaller than Y_i^{j-1} by no more than a constant factor such as $(d-\epsilon)$. Now this can be done by Lemma 7.5.1 (with $A = V_{\sigma(j)}$, $B = V_{\sigma(i)}$ and $Y = Y_i^{j-1}$): provided that Y_i^{j-1} still has size at least $\epsilon\ell$ (which induction will ensure), all but at most $\epsilon\ell$ choices of v_j will be such that the new set Y_i^j as in (2) satisfies

$$|Y_i^j| \geqslant (d-\epsilon)|Y_i^{j-1}|. \tag{3}$$

Excluding the bad choices for v_j for all the relevant values of i simultaneously, we find that all but at most $\Delta\epsilon\ell$ choices of v_j from $V_{\sigma(j)}$, and in particular from $Y_j^{j-1} \subseteq V_{\sigma(j)}$, satisfy (3) for all i.

It remains to show that the sets Y_i^{j-1} considered above as Y for Lemma 7.5.1 never fall below the size of $\epsilon\ell$, and that when we come to select $v_j \in Y_j^{j-1}$ we have a choice of at least s suitable candidates: since before u_j at most $s-1$ vertices u were given an image in $V_{\sigma(j)}$, we can then choose v_j distinct from these.

But all this follows from our choice of ϵ_0. Indeed, the initial target sets Y_i^0 have size ℓ, and each Y_i shrinks at most Δ times by a factor of $(d-\epsilon)$ when some v_j with $j < i$ and $u_j u_i \in E(H)$ is defined. Thus,

$$|Y_i^{j-1}| - \Delta\epsilon\ell \underset{(3)}{\geqslant} (d-\epsilon)^\Delta \ell - \Delta\epsilon\ell \geqslant (d-\epsilon_0)^\Delta \ell - \Delta\epsilon_0\ell \underset{(1)}{\geqslant} \tfrac{1}{2} d^\Delta \ell \geqslant s$$

for all $j \leqslant i$; in particular, we have $|Y_i^{j-1}| \geqslant \epsilon\ell$ and $|Y_j^{j-1}| - \Delta\epsilon\ell \geqslant s$ as desired. □

We are now ready to prove the Erdős-Stone theorem.

(7.1.1)
Proof of Theorem 7.1.2. Let $r \geqslant 2$ and $s \geqslant 1$ be given as in the (7.1.4)
statement of the theorem. For $s = 1$ the assertion follows from Turán's (7.4.1)
theorem, so we assume that $s \geqslant 2$. Let $\gamma > 0$ be given; this γ will play r, s
the role of the ϵ of the theorem. If any graph G with $|G| =: n$ has γ

$$\|G\| \geqslant t_{r-1}(n) + \gamma n^2$$ $\|G\|$

edges, then $\gamma < 1$. We want to show that $K_s^r \subseteq G$ if n is large enough.

Our plan is to use the regularity lemma to show that G has a regularity graph R dense enough to contain a K^r by Turán's theorem. Then R_s contains a K_s^r, so we may hope to use our blow-up lemma, Lemma 7.5.2, to deduce that $K_s^r \subseteq G$.

On input $d := \gamma$ and $\Delta := \Delta(K_s^r)$ Lemma 7.5.2 returns an $\epsilon_0 > 0$. d, Δ, ϵ_0
To apply the regularity lemma, let $m > 1/\gamma$ and choose $\epsilon > 0$ small m, ϵ
enough that $\epsilon \leqslant \epsilon_0$,

$$\epsilon < \gamma/2 < 1\,, \tag{1}$$

and

$$\delta := 2\gamma - \epsilon^2 - 4\epsilon - d - \frac{1}{m} > 0\,;$$ δ

this is possible, since $2\gamma - d - \frac{1}{m} > 0$. On input ϵ and m, the regularity lemma returns an integer M. Let us assume that M

$$n \geqslant \frac{2Ms}{d^\Delta(1-\epsilon)}\,.$$ n

Since this number is at least m, the regularity lemma provides us with an ϵ-regular partition $\{V_0, V_1, \ldots, V_k\}$ of G, where $m \leqslant k \leqslant M$; let k
$|V_1| = \ldots = |V_k| =: \ell$. Then ℓ

$$n \geqslant k\ell\,, \tag{2}$$

and

$$\ell = \frac{n - |V_0|}{k} \geqslant \frac{n - \epsilon n}{M} = n\,\frac{1-\epsilon}{M} \geqslant \frac{2s}{d^\Delta}$$

R by the choice of n. Let R be the regularity graph of G with parameters ϵ, ℓ, d corresponding to the above partition. Then Lemma 7.5.2 will imply $K^r_s \subseteq G$ as desired if $K^r \subseteq R$ (and hence $K^r_s \subseteq R_s$).

Our plan was to show $K^r \subseteq R$ by Turán's theorem. We thus have to check that R has enough edges, i.e. that enough ϵ-regular pairs (V_i, V_j) have density at least d. This should follow from our assumption that G has at least $t_{r-1}(n) + \gamma n^2$ edges, i.e. an edge density of about $\frac{r-2}{r-1} + 2\gamma$: this lies substantially above the approximate density of $\frac{r-2}{r-1}$ of the Turán graph $T^{r-1}(k)$, and hence substantially above any density that G could derive from $t_{r-1}(k)$ dense pairs alone, even if all these had density 1.

Let us then estimate $\|R\|$ more precisely. How many edges of G lie outside ϵ-regular pairs? At most $\binom{|V_0|}{2}$ edges lie inside V_0, and by condition (i) in the definition of ϵ-regularity these are at most $\frac{1}{2}(\epsilon n)^2$ edges. At most $|V_0|k\ell \leqslant \epsilon n^2$ edges join V_0 to other partition sets. The at most ϵk^2 other pairs (V_i, V_j) that are not ϵ-regular contain at most ℓ^2 edges each, together at most $\epsilon k^2 \ell^2$. The ϵ-regular pairs of insufficient density ($< d$) each contain no more than $d\ell^2$ edges, altogether at most $\frac{1}{2}k^2 d\ell^2$ edges. Finally, there are at most $\binom{\ell}{2}$ edges inside each of the partition sets $V_1, \dots, V_k$, together at most $\frac{1}{2}\ell^2 k$ edges. All *other* edges of G lie in ϵ-regular pairs of density at least d, and thus contribute to edges of R. Since each edge of R corresponds to at most ℓ^2 edges of G, we thus have in total

$$\|G\| \leqslant \tfrac{1}{2}\epsilon^2 n^2 + \epsilon n^2 + \epsilon k^2 \ell^2 + \tfrac{1}{2}k^2 d\ell^2 + \tfrac{1}{2}\ell^2 k + \|R\|\,\ell^2 .$$

Hence, for all sufficiently large n,

$$\begin{aligned}
\|R\| &\geqslant \tfrac{1}{2}k^2\, \frac{\|G\| - \frac{1}{2}\epsilon^2 n^2 - \epsilon n^2 - \epsilon k^2\ell^2 - \frac{1}{2}dk^2\ell^2 - \frac{1}{2}k\ell^2}{\frac{1}{2}k^2\ell^2} \\
&\underset{(1,2)}{\geqslant} \tfrac{1}{2}k^2 \left(\frac{t_{r-1}(n) + \gamma n^2 - \frac{1}{2}\epsilon^2 n^2 - \epsilon n^2}{n^2/2} - 2\epsilon - d - \frac{1}{k} \right) \\
&\geqslant \tfrac{1}{2}k^2 \left(\frac{t_{r-1}(n)}{n^2/2} + 2\gamma - \epsilon^2 - 4\epsilon - d - \frac{1}{m} \right) \\
&= \tfrac{1}{2}k^2 \left(t_{r-1}(n) \binom{n}{2}^{-1} \left(1 - \frac{1}{n}\right) + \delta \right) \\
&> \tfrac{1}{2}k^2\, \frac{r-2}{r-1} \\
&\geqslant t_{r-1}(k)\,.
\end{aligned}$$

(The strict inequality follows from Lemma 7.1.4.) Therefore $K^r \subseteq R$ by Theorem 7.1.1, as desired. $\square$

Having seen a typical application of the regularity lemma in full detail, let us now step back and try to separate the wheat from the chaff: what were the main ideas, how do the various parameters depend on each other, and in which order were they chosen?

The task was to show that γn^2 more edges than can be accommodated on n vertices without creating a K^r force a K^r_s subgraph, provided that G is large enough. The plan was to do this using our blow-up lemma, which asks for the input of two parameters: d and Δ. As we wish to find a copy of $H = K^r_s$ in G, it is clear that we must choose $\Delta := \Delta(K^r_s)$. We shall return to the question of how to choose d in a moment.

Given d and Δ, our blow-up lemma tells us how small we must choose ϵ to make the regularity lemma provide us with a suitable partition. The regularity lemma also requires the input of a lower bound m for the number of partition classes; we shall discuss this below, together with d.

All that remains now is to choose G large enough that the partition classes have size at least $2s/d^\Delta$, as required by the blow-up lemma. (The s here depends on the graph H we wish to embed, and $s := |H|$ would certainly be big enough. In our case, we can use the s from our $H = K^r_s$.) How large is 'large enough' for $|G|$ follows straight from the upper bound M on the number of partition classes returned by the regularity lemma: roughly, i.e. disregarding V_0, an assumption of $|G| \geqslant 2Ms/d^\Delta$ suffices.

So far, everything was entirely straightforward, and standard for any application of the regularity lemma of this kind. But now comes the interesting bit, the part specific to this proof: the observation that, if only d is small enough, our γn^2 'additional edges' force an 'additional dense ϵ-regular pair' of partition sets, giving us more than $t_{r-1}(k)$ dense ϵ-regular pairs in total (where 'dense' means 'of density at least d'), thus forcing R to contain a K^r and hence R_s to contain a K^r_s.

Let us examine why this is so. Suppose we have at most $t_{r-1}(k)$ dense ϵ-regular pairs . Inside these, G has at most

$$\tfrac{1}{2}k^2\frac{r-2}{r-1}\,\ell^2 \leqslant \tfrac{1}{2}n^2\frac{r-2}{r-1}$$

edges, even if we use those pairs to their full capacity of ℓ^2 edges each (where ℓ is again the common size of the partition sets other than V_0, so that $k\ell$ is nearly n). Thus, we have almost exactly our γn^2 additional edges left to accommodate elsewhere in the graph: either in ϵ-regular pairs of density less than d, or in some exceptional way, i.e. in irregular pairs, inside a partition set, or with an end in V_0. Now the number of edges in low-density ϵ-regular pairs is less than

$$\tfrac{1}{2}k^2 d\ell^2 \leqslant \tfrac{1}{2}dn^2,$$

and hence less than half of our extra edges if $d \leqslant \gamma$. The other half, the remaining $\frac{1}{2}\gamma n^2$ edges, are more than can be accommodated in exceptional ways, provided we choose m large enough and ϵ small enough (giving an additional upper bound for ϵ). It is now a routine matter to compute the values of m and ϵ that will work.

7.6 Two regularity tools

Our acquaintance with the regularity lemma so far already includes the essentials: the notion of regularity, a statement of the regularity lemma, its proof, and a typical application of the lemma in our proof of Theorem 7.1.2 in Section 7.5. Another application, also based on no more than we have covered so far, will be given in the proof of Theorem 9.2.2. Both these applications illustrate techniques typically involved in many uses of the regularity lemma.

Our aim in this section is to extract some of these techniques and bundle them into formal tools that can be used out of the box: the 'counting lemma', the 'removal lemma', and some immediate consequences of these. Their proofs will not be short, but the ideas they involve will look familiar by now. Conversely, our proofs of Theorems 7.1.2 and 9.2.2 could be rewritten to use some of these formal tools (Exercise 45).

To facilitate the proof of the counting lemma, it will be convenient to use a slightly different notion of ϵ-regularity. To avoid confusion, we shall call this 'ϵ-uniformity' here; but in the literature it is usually referred to as (another version of) 'ϵ-regularity' as well.

(ϵ, d)-uniform

Call a pair (A, B) of disjoint sets of vertices in a graph *(ϵ, d)-uniform* for $\epsilon > 0$ and $0 \leqslant d \leqslant 1$ if their non-empty subsets $X \subseteq A$ and $Y \subseteq B$ satisfy

$$\big|\|X, Y\| - d\,|X||Y|\big| \;\leqslant\; \epsilon\,|A||B|\,,$$

or equivalently,

$$\big|d(X, Y) - d\big| \;\leqslant\; \epsilon\,\frac{|A||B|}{|X||Y|}\,.$$

ϵ-uniform

We may think of the real number d as a 'target density' with which the densities $d(X, Y) = \|X, Y\|/|X||Y|$ of the pairs (X, Y) are compared. If this is chosen as $d = d(A, B)$, the (ϵ, d)-uniform pair (A, B) is simply called *ϵ-uniform*.

This differs from our definition of ϵ-regularity in Section 7.4 in two ways. One is that all subsets $X \subseteq A$ and $Y \subseteq B$ are required to satisfy the inequality, including those of size less than $\epsilon|A|$ or $\epsilon|B|$. The other difference is best seen in the second of the two inequalities. Its left-hand side measures how much the density of (X, Y) differs from that of (A, B). Our definition of ϵ-regularity requires that this should be at most ϵ. This is relaxed here by the factor $|A||B|/|X||Y| \geqslant 1$: the density of small pairs (X, Y) is allowed to deviate from d more than that of bigger ones.

The difference between ϵ-regularity and ϵ-uniformity, however, is qualitatively irrelevant: ϵ-regular pairs are also ϵ-uniform, which we use freely in this section, and ϵ-uniform pairs are $\sqrt[3]{\epsilon}$-regular (Exercise 42).

The purpose of the counting lemma is to count in how many ways a given graph H can be embedded as a subgraph, not necessarily induced, in a larger graph G that comes with an ϵ-regular partition. In our version we pre-assign to each vertex of H the partition class of G in which we want to embed it. We then count only those embeddings $H \to G$ that conform to this assignment, and assume that the pairs of partition classes due to accommodate the edges of H form ϵ-regular pairs in G. The lemma then says, roughly, that the number of such embeddings is close to what their expected number would be if G was a random graph that had edges only between distinct classes V_i, V_j and these edges were chosen independently with probabilities d_{ij} close to $d(V_i, V_j)$.[9]

It turns out that, for G large enough, there are at least $c\,|G|^{|H|}$ such embeddings $H \to G$, where c depends only on H, ϵ, and the densities $d(V_i, V_j)$ between the partition classes needed to accommodate the edges of H. If all these densities are positive we have $c > 0$, and can therefore find at least one copy of H in any large enough G that has such an ϵ-regular partition.

Before we state and prove the counting lemma, let us note what happens in the special case that $H = K^k$ and G is k-partite with classes $V_1, \dots, V_k$, for some $k \in \mathbb{N}$. Since the vertices of H are pairwise adjacent, any embedding in G sends them to distinct classes, and there are $\prod_{1 \leqslant i \leqslant k} |V_i|$ ways of doing that. Not all of them will induce a K^k in G, but we can estimate how many of them might: a pair (v_i, v_j) from (V_i, V_j) induces an edge in G with a 'probability' of $d(V_i, V_j)$. If these are close to given parameters d_{ij}, we can expect that about $\prod_{1 \leqslant i < j \leqslant k} d_{ij}$ of our $\prod_{1 \leqslant i \leqslant k} |V_i|$ maps $V(K^k) \to V(G)$ induce a K^k also in G.

k

d_{ij}

Proposition 7.6.1. *Let $\epsilon > 0$, and let G be a multipartite graph with vertex classes $V_1, \dots, V_k$ whose pairs (V_i, V_j) are (ϵ, d_{ij})-uniform in G. Then*

$$\Big|\, |\mathcal{K}^k(G)| \;-\; \prod_{1 \leqslant i < j \leqslant k} d_{ij} \cdot \prod_{1 \leqslant i \leqslant k} |V_i| \,\Big| \;\leqslant\; \epsilon \tbinom{k}{2} \prod_{1 \leqslant i \leqslant k} |V_i|\,,$$

where $\mathcal{K}^k(G)$ denotes the set of complete subgraphs of order k in G.

Proposition 7.6.1 will follow from Lemma 7.6.2 (Exercise 44).

For the more general counting lemma we need some notation. Let us denote by $\mathrm{Hom}(H, G)$ the set, and by $\mathrm{hom}(H, G)$ the number, of homomorphisms from H to G. Recall that these are the maps $V(H) \to V(G)$

$\mathrm{Hom}(H, G)$
$\mathrm{hom}(H, G)$

[9] Both uses of 'close' here mean, essentially, 'up to an ϵ-portion'.

that preserve adjacency, but not necessarily non-adjacency. Vertices of H can map to the same vertex of G, but only if they are non-adjacent.[10]

Injective homomorphisms $\varphi\colon H \to G$ embed H in G as a subgraph, though not necessarily as an induced subgraph. Note that there can be more such homomorphisms than subgraphs of G isomorphic to H, because there can be many ways of mapping H to a fixed copy of H in G. We may think of each φ as specifying a 'labelled copy' of H in G: one whose vertices are each 'labelled' by the vertices of H that φ maps there.

Consider a fixed homomorphism $\varphi\colon H \to R$ with $V(R) = \{1, \dots, k\}$, and let G be a k-partite graph with vertex classes $V_1, \dots, V_k$. Given any subgraph $H' \subseteq H$ (often, but not always, $H' = H$) we write $\mathrm{Hom}_\varphi(H', G)$ for the set, and $\mathrm{hom}_\varphi(H', G)$ for the number, of homomorphisms $\psi\colon H' \to G$ that satisfy $\psi(u) \in V_{\varphi(u)}$ for all $u \in V(H')$; we call these the *φ-partite* homomorphisms from H' to G.[11]

$\mathrm{Hom}_\varphi(H,G)$ $\mathrm{hom}_\varphi(H,G)$

Lemma 7.6.2. (Counting Lemma)
Let R be a graph on $\{1, \dots, k\} \subseteq \mathbb{N}$, and $\varphi\colon H \to R$ a homomorphism for some graph H. Let G be a multipartite graph with vertex classes $V_1, \dots, V_k$, whose pairs (V_i, V_j) with $ij \in E(R)$ are (ϵ, d_{ij})-uniform in G for some $\epsilon > 0$ and various $d_{ij} \geqslant 0$. Then

$$\Big|\mathrm{hom}_\varphi(H,G) \;-\; \prod_{e \in E(H)} d_e^\varphi \cdot \prod_{u \in V(H)} |V_{\varphi(u)}|\Big| \;\leqslant\; \epsilon\,\|H\| \prod_{u \in V(H)} |V_{\varphi(u)}|\,,$$

where $d_e^\varphi = d_{\varphi(u)\varphi(w)}$ for edges $e = uw$ of H.

A typical application of the counting lemma, which we shall prove below, says that there are many φ-partite homomorphisms $H \to G$, at least

$$\mathrm{hom}_\varphi(H,G) \;\geqslant\; \Big(\prod d_e^\varphi \,-\, \epsilon\,\|H\|\Big) \cdot \prod |V_{\varphi(u)}|$$

many, of which comparatively few are non-injective. The conclusion, then, is that there are many injective homomorphisms $H \to G$. In particular, there is at least one, so G contains H as a subgraph.

For example, if $\chi(H) = k$, say, let us show that we can find a copy of H in any graph G that has a vertex partition $\{V_1, \dots, V_k\}$ whose pairs (V_i, V_j) are ϵ-uniform with positive density, at least $d > 0$ say, as soon as the classes V_i have some common size ℓ that is large enough in terms of d and ϵ, and ϵ is small enough that $d^{\|H\|} - \epsilon\|H\| > 0$.[12]

d ℓ

[10] ... unless G has loops, which in this book happens only in Chapters 4.6 and 6.

[11] Fixing $\varphi\colon H \to R$ and then looking for injective φ-partite homomorphisms $H \to G$ is similar to finding embeddings $H \to G$ based on a specified embedding $H \to R_s$, as we did in our version of the blow-up lemma, Lemma 7.5.2.

[12] In an application, these V_i might be only some of the partition classes of a bigger graph to which we have applied the regularity lemma. And only some of their pairs may be ϵ-uniform with density at least d, just enough pairs for $\varphi\colon H \to R$ to exist.

As $\chi(H) = k$, there exists a homomorphism $\varphi\colon H \to R := K^k$. We apply the counting lemma with $\epsilon < d^{\|H\|}/\|H\|$ and $d_{ij} := d(V_i, V_j) \geqslant d$. It tells us that

$$\hom_\varphi(H, G) \geqslant \left(d^{\|H\|} - \epsilon\,\|H\|\right) \ell^{|H|}.$$

Let us think of this lower bound as $c\,\ell^{|H|}$ in which c is a 'constant': it depends only on H, d and our choice of ϵ (which ensures that $c > 0$), but not on ℓ. Our aim now is to show that if ℓ is large enough there are more injective homomorphism than non-injective ones.

How many non-injective φ-partite homomorphisms $\psi\colon H \to G$ are there? Any such ψ maps two vertices u, w of H to the same vertex of G, and for every choice of u and w there are at most $\ell^{|H|-1}$ such ψ. Hence for $c' = \binom{|H|}{2}$ there are at most $c'\ell^{|H|-1}$ non-injective φ-partite homomorphisms $H \to G$ altogether. As $c\,\ell^{|H|} - c'\ell^{|H|-1} \to \infty$ as $\ell \to \infty$, we find that for ℓ large enough there exists an injective φ-partite homomorphism $H \to G$. In particular, G has a subgraph isomorphic to H.

Proof of Lemma 7.6.2. We apply induction on $\|H\|$. The induction starts trivially with $\|H\| = 0$, where we follow the convention that the empty product $\prod_{e\in\emptyset} d_e^\varphi$ equals 1. For the induction step, pick an edge $ab \in H$ and let $H' := H - ab$. a, b, H'

The homomorphisms in $\mathrm{Hom}_\varphi(H, G)$, which we wish to count, are precisly the homomorphisms in $\mathrm{Hom}(H', G)$ that map a, b to adjacent vertices of G. Let us write $\mathbb{1}_G$ for the function that sends a pair of vertices of G to 1 if they are adjacent, and to 0 if not; in particular, $\mathbb{1}_G(v, v) = 0$ for all $v \in V(G)$. Then $\mathbb{1}_G$

$$\hom_\varphi(H, G) \;=\; \sum_{\psi\in\mathrm{Hom}_\varphi(H',G)} \mathbb{1}_G\big(\psi(a), \psi(b)\big)\,.$$

In order to apply the induction hypothesis, we rewrite this sum as

$$\begin{aligned}\hom_\varphi(H, G) \;&=\; \sum_{\psi\in\mathrm{Hom}_\varphi(H',G)} \Big(\mathbb{1}_G\big(\psi(a), \psi(b)\big) - d_{ab}^\varphi + d_{ab}^\varphi\Big)\\ &=\; \sum_{\psi\in\mathrm{Hom}_\varphi(H',G)} \Big(\mathbb{1}_G\big(\psi(a), \psi(b)\big) - d_{ab}^\varphi\Big) + d_{ab}^\varphi\cdot \hom_\varphi(H', G),\end{aligned}$$

where d_{ab}^φ is defined as in the statement of the counting lemma. d_{ab}^φ

Qualitatively, the induction hypothesis tells that the last term above, $\hom_\varphi(H', G)$, is close to $\prod_{e\in E(H')} d_e^\varphi \cdot \prod_{u\in V(H')} |V_{\varphi(u)}|$. As $V(H') = V(H)$, this differs from the corresponding target value for $\hom_\varphi(H, G)$ only by a factor of d_{ab}^φ: exactly the factor that precedes $\hom_\varphi(H', G)$ in the equation above. It thus remains to show that the big sum in the last equation is small.

Quantitatively, the last summand $d^{\varphi}_{ab} \cdot \hom_{\varphi}(H', G)$ in our rewriting of $\hom_{\varphi}(H, G)$ differs from the target value for $\hom_{\varphi}(H, G)$ as stated in the counting lemma, the number $\prod_{e \in E(H)} d^{\varphi}_{e} \cdot \prod_{u \in V(H)} |V_{\varphi(u)}|$, by

$$\begin{aligned} \Big| d^{\varphi}_{ab} \cdot \hom_{\varphi}(H', G) \;-\; & \prod_{e \in E(H)} d^{\varphi}_{e} \cdot \prod_{u \in V(H)} |V_{\varphi(u)}| \Big| \\ &= \Big| d^{\varphi}_{ab} \cdot \hom_{\varphi}(H', G) \;-\; d^{\varphi}_{ab} \cdot \prod_{e' \in E(H')} d^{\varphi}_{e'} \cdot \prod_{u \in V(H)} |V_{\varphi(u)}| \Big| \\ &\leqslant d^{\varphi}_{ab} \cdot \epsilon \cdot \|H'\| \cdot \prod_{u \in V(H)} |V_{\varphi(u)}| \\ &\leqslant \epsilon \cdot \big(\|H\| - 1\big) \prod_{u \in V(H)} |V_{\varphi(u)}| \,. \end{aligned}$$

This difference is smaller than the difference of $\epsilon \|H\| \prod_{u \in V(H)} |V_{\varphi(u)}|$ that the counting lemma allows between $\hom_{\varphi}(H, G)$ and its target value of $\prod_{e \in E(H)} d^{\varphi}_{e} \cdot \prod_{u \in V(H)} |V_{\varphi(u)}|$: smaller by the quantity of $\epsilon \cdot \prod |V_u|$ that corresponds to the -1 in the last line above.

To complete the induction step, it thus remains to show that this margin of $\epsilon \cdot \prod |V_u|$ can accommodate the first summand in our earlier estimate for $\hom_{\varphi}(H, G)$, i.e. that

$$\Big| \sum_{\psi \in \mathrm{Hom}_{\varphi}(H', G)} \Big(\mathbb{1}_G\big(\psi(a), \psi(b)\big) - d^{\varphi}_{ab} \Big) \Big| \;\leqslant\; \epsilon \cdot \prod_{u \in V(H)} |V_{\varphi(u)}| \,. \qquad (*)$$

H'' For a proof of $(*)$, let $H'' := H - a - b$ be the subgraph of H obtained by deleting the vertices a and b. Then

$$\begin{aligned} \Big| \sum_{\psi \in \mathrm{Hom}_{\varphi}(H', G)} & \Big(\mathbb{1}_G\big(\psi(a), \psi(b)\big) - d^{\varphi}_{ab} \Big) \Big| \\ &= \Big| \sum_{\psi' \in \mathrm{Hom}_{\varphi}(H'', G)} \; \sum_{\substack{\psi \in \mathrm{Hom}_{\varphi}(H', G) \\ \psi|_{H''} = \psi'}} \Big(\mathbb{1}_G\big(\psi(a), \psi(b)\big) - d^{\varphi}_{ab} \Big) \Big| \\ &\leqslant \sum_{\psi' \in \mathrm{Hom}_{\varphi}(H'', G)} \Big| \sum_{\substack{\psi \in \mathrm{Hom}_{\varphi}(H', G) \\ \psi|_{H''} = \psi'}} \Big(\mathbb{1}_G\big(\psi(a), \psi(b)\big) - d^{\varphi}_{ab} \Big) \Big| . \end{aligned}$$

The first summation here, over all $\psi \in \mathrm{Hom}_{\varphi}(H', G)$, is thus performed in steps: we first sum over all the possible ways ψ' in which ψ can act on the vertices other than a and b, and then sum for each of these ψ' over all the ways in which we can extend it to a and b to obtain a homomorphism ψ from all of H' to G. In particular, the ψ in the last sum have to map a to vertices of G that are adjacent to all the previously chosen images $\psi'(u)$

of neighbours u that a has in $V(H'')$, and likewise for b:

$$\psi(a) \in V_{\varphi(a)} \cap \bigcap_{u \in N_{H'}(a)} N_G\big(\psi'(u)\big) \;=:\; W_a$$

W_a

and

$$\psi(b) \in V_{\varphi(b)} \cap \bigcap_{u \in N_{H'}(b)} N_G\big(\psi'(u)\big) \;=:\; W_b\,.$$

W_b

Conversely, every pair in $W_a \times W_b$, chosen as $\psi(a)$ and $\psi(b)$, extends a given $\psi' \in \mathrm{Hom}_\varphi(H'', G)$ to a homomorphism $\psi \in \mathrm{Hom}_\varphi(H', G)$. Thus,

$$\begin{aligned}
&\Big|\sum_{\psi \in \mathrm{Hom}_\varphi(H',G)} \Big(\mathbb{1}_G\big(\psi(a),\psi(b)\big) - d^\varphi_{ab}\Big)\Big| \\
&\qquad\leqslant \sum_{\psi' \in \mathrm{Hom}_\varphi(H'',G)} \Big|\sum_{\substack{\psi \in \mathrm{Hom}_\varphi(H',G)\\ \psi|_{H''} = \psi'}} \Big(\mathbb{1}_G\big(\psi(a),\psi(b)\big) - d^\varphi_{ab}\Big)\Big| \\
&\qquad= \sum_{\psi' \in \mathrm{Hom}_\varphi(H'',G)} \Big|\sum_{\substack{w_a \in W_a\\ w_b \in W_b}} \Big(\mathbb{1}_G(w_a, w_b) - d^\varphi_{ab}\Big)\Big| \\
&\qquad= \sum_{\psi' \in \mathrm{Hom}_\varphi(H'',G)} \Big|\|W_a, W_b\|_G - d^\varphi_{ab}\,|W_a||W_b|\Big| \\
&\qquad\leqslant \mathrm{hom}_\varphi(H'', G) \cdot \epsilon\,|V_{\varphi(a)}||V_{\varphi(b)}|
\end{aligned}$$

by the $(\epsilon, d^\varphi_{ab})$-uniformity of $(V_{\varphi(a)}, V_{\varphi(b)})$ for their subsets W_a and W_b. Since

$$\mathrm{hom}_\varphi(H'', G) \leqslant \prod_{u \in V(H'')} |V_{\varphi(u)}|\,,$$

assertion $(*)$ follows. $\square$

We now turn to our second regularity tool, the removal lemma. Qualitatively,[13] the removal lemma says that if a large graph G is so sparse that most of the embeddings of the vertices of some fixed graph H in G fail to induce a copy (or supergraph) of H, we can delete *all* the copies of H in G by removing few of its edges:

Lemma 7.6.3. (Removal Lemma)
For every graph H and every $\varrho > 0$ there exist $\eta > 0$ and n_0 such that the following holds for all graphs G of order $n \geqslant n_0$: if there are at most $\eta n^{|H|}$ injective homomorphisms $H \to G$, then G contains a set E of at most ϱn^2 edges such that $H \not\subseteq G - E$.

---

[13] Quantitatively: if only $o(|G|^{|H|})$ of the maps $V(H) \to V(G)$ induce a supergraph of H in G, we can remove all the copies of H in G by deleting $o(|G|^2)$ edges.

Informally, the removal lemma says that if at most an η-fraction of all the maps $V(H) \to V(G)$ are injective $H \to G$ homomorphisms, then all these can be destroyed by deleting only a ϱ-fraction of the possible edges on $V(G)$. We can wish for ϱ to be as small as we like, but have to pay for it by having a smaller η and a bigger n_0 returned by the lemma, so the assertion will apply to fewer G and work only if there are fewer injective $H \to G$ homomorphisms.

The proof of the removal lemma will be along the following lines. The first thing we shall note is that we can allow non-injective $H \to G$ homomorphisms too in our count: there are too few of those to matter. All we shall need is that there are no more non-injective homomorphisms than injective ones, which will be easy to prove; then the number of all of them together is still bounded by $2\eta n^{|H|}$.

The desired few edges of G that hit all the copies of H in it will be all the edges outside ϵ-regular pairs of partition classes of density at least ϱ, for any ϵ-regular partition of G with ϵ chosen small enough. An easy count will show that there are at most ϱn^2 such edges, which we delete. The interesting bit then is why that destroys *all* the copies of H in G.

This is where ϵ-regularity and the counting lemma come in. If there is even a single copy H_0 of H left in G, its edges must lie in ϵ-regular pairs of partition classes of density at least ϱ. By the counting lemma, however, this means that there is not just one copy of H in G but many: more than our assumptions allow. Indeed, the existence of H_0 as above means that there exists a homomorphism $\varphi\colon H \to R$, where R is the regularity graph for precisely those ϵ-regular pairs that accommodate H_0. The counting lemma, then, concludes that there should be more than $2\eta n^{|H|}$ homomorphisms $H \to G$ if η is small enough in terms of ϱ. But, as we have seen, this is ruled out by our assumption about η.

Proof of Lemma 7.6.3. If H has no edges, we can make the premise of the lemma void by suitable choices of η and n_0, regardless of the value of ϱ. If H has an edge but $\varrho \geqslant 1/2$, the conclusion of the lemma holds
ϱ with $E = E(G)$. We shall therefore assume that $\|H\| \geqslant 1$ and $\varrho < 1/2$.

On input

ϵ, m
$$\epsilon := \min\left\{\tfrac{1}{8}\varrho\,,\,\tfrac{1}{2}\varrho^{\|H\|}/\|H\|\right\} \qquad \text{and} \qquad m := \lceil 2/\varrho \rceil \tag{1}$$

M the Regularity Lemma 7.4.1 returns an integer M. We use this to set

η, n_0
$$\eta := \tfrac{1}{5}\varrho^{\|H\|}/(2M)^{|H|} \qquad \text{and} \qquad n_0 := \max\left\{M, \left\lceil \tbinom{|H|}{2}/\eta \right\rceil\right\}. \tag{2}$$

Now let G be any graph of order $n \geqslant n_0$ as considered in the lemma: such that there are at most $\eta n^{|H|}$ injective homomorphisms $H \to G$.

It is easy to see – compare our discussion just before the proof of the counting lemma – that there are also at most

$$n^{|H|-1}\tbinom{|H|}{2} \leqslant n^{|H|}\tbinom{|H|}{2}/n_0 \leqslant \eta n^{|H|}$$

non-injective homomorphisms $H \to G$. The number of all homomorphisms $H \to G$, injective or not, is therefore at most

$$\hom(H,G) \leqslant 2\eta n^{|H|}. \tag{3}$$

The regularity lemma yields an ϵ-regular partition $\{V_0, V_1, \dots, V_k\}$ k
of G with $|V_1| = \dots = |V_k| := \ell$ and ℓ

$$2/\varrho \underset{(1)}{\leqslant} m \leqslant k \leqslant M. \tag{4}$$

Let us use this partition to choose our set E to satisfy the lemma. Let E be the set of all edges xy of G that satisfy one of the following statements: E

(a) x or y lies in V_0;

(b) x and y lie in the same V_i with $i \geqslant 1$;

(c) $x \in V_i$ and $y \in V_j$ with (V_i, V_j) not ϵ-regular and $i, j \geqslant 1$;

(d) $x \in V_i$ and $y \in V_j$ with $d(V_i, V_j) < \varrho$ and $i, j \geqslant 1$.

To prove our lemma, we have to show that

$$|E| \leqslant \varrho n^2 \qquad \text{and} \qquad H \nsubseteq G - E\,. \tag{5}$$

For a proof of the first statement in (5) note that

- at most $|V_0| \cdot n \leqslant \epsilon n^2$ edges of G satisfy (a);
- at most $k \cdot \binom{n/k}{2} < n^2/2k$ edges of G satisfy (b);
- at most $\epsilon k^2 \cdot (n/k)^2 = \epsilon n^2$ edges of G satisfy (c);
- at most $\binom{k}{2} \cdot \varrho \cdot (n/k)^2 < \varrho n^2/2$ edges of G satisfy (d).

Hence

$$|E| \leqslant \left(\epsilon + \tfrac{1}{2k} + \epsilon + \tfrac{1}{2}\varrho\right) n^2 \underset{(4)}{\leqslant} \left(\epsilon + \tfrac{1}{4}\varrho + \epsilon + \tfrac{1}{2}\varrho\right) n^2 \underset{(1)}{\leqslant} \varrho n^2$$

as desired.

For a proof of the second statement in (5) suppose that $H \subseteq G - E$. By our choice of E, every edge of this copy of H lies in some ϵ-regular pair of partition classes of density at least ϱ. This defines a homomorphism $\varphi \colon H \to R$, where R is the regularity graph of G for our partition with φ
parameters ϵ, ℓ and ϱ.[14] R

[14] The regularity graph for an ϵ-regular partition is defined before Lemma 7.5.2.

Since ϵ-regularity implies ϵ-uniformity (Exercise 42), φ and $G - E$ satisfy the premise for the counting lemma with $d_{ij} := d(V_i, V_j) \geqslant \varrho$ for the edges ij of R. The counting lemma yields

$$\begin{aligned}
\hom(H,G) &\geqslant \hom_\varphi(H, G-E)\\
&\geqslant \Big(\prod_{e \in E(H)} d_e^\varphi - \epsilon \, \|H\| \Big) \cdot \prod_{u \in V(H)} |V_{\varphi(u)}|\\
&\geqslant \big(\varrho^{\|H\|} - \epsilon \, \|H\|\big) \big(n/2M\big)^{|H|}\\
&\underset{(1)}{\geqslant} \tfrac{1}{2} \varrho^{\|H\|} \big(n/2M\big)^{|H|}\\
&\underset{(2)}{>} 2\eta n^{|H|}.
\end{aligned}$$

This contradicts (3), and thereby completes the proof of (5). □

The number of copies of a fixed graph H in some dense graphs G tells us something about the structure of those G when they get large. Assume, for simplicity, that we have an infinite class of graphs G each
n, α with αn^2 edges, where $n = |G|$ and $\alpha > 0$ is some fixed parameter telling
k, m us how dense all those G are. Let $k := |H| \geqslant 3$ and $m := \|H\| \geqslant 1$. Let
γ, θ $\gamma > 0$ be any constant, and assume that the number $\theta = \theta(G)$ of copies of H in our graphs G satisfies

$$\gamma n^2 \leqslant \theta m \leqslant \eta n^k \tag{$*$}$$

for any given $\eta > 0$ when n is large enough.[15] On average, every edge of G lies in $\theta m / \alpha n^2$ copies of H. The removal lemma tells us that we can hit all those copies of H with only ϱn^2 edges, for any $\varrho > 0$ we may choose, when n is large enough. These special edges each lie in at least $\theta / \varrho n^2$ copies of H on average: more than our overall average by a factor
c of at least $c := \alpha / \varrho m$. And since we may choose ϱ, while α and m are constant, we can make c as large as we like, just by making ϱ small and n large enough that the removal lemma applies.

So the removal lemma tells us that the copies of H in G cannot be spread evenly over G, in the sense that the maximum number of copies in which a given edge of G lies differs from the average such number by no more than a constant factor as n grows. The copies of H in G must instead be lumped together at a few edges that lie in many more such copies than an edge of G does on average: in at least c times as many, where $c \to \infty$ as $n \to \infty$. In absolute terms, these special edges lie in at least

$$c \cdot \theta m / \alpha n^2 \geqslant c\, \gamma / \alpha$$

[15] In complexity-theoretic notation: $\theta(G) = \Omega(n^2)$ but $\theta(G) = o(n^k)$. The η here could be any returned by the removal lemma on input some ϱ we may later choose.

copies of H on average, for n large enough. Since γ/α is a positive constant, this number will get arbitrarily large if we choose ϱ small enough. This bites most if the average number $\theta m/\alpha n^2$ of copies of H in which an edge of G lies is itself bounded as n grows. We shall return to this case in a moment.

By contraposition, our considerations above also tell us something about the density of the graphs G in a given class, and about the number of copies of H they contain, if those copies are *not* lumped together at a few edges but are evenly spread over G. If these G are dense, we must also have $\theta \geqslant \gamma n^2$ for some $\gamma > 0$ (unless $\theta = 0$), since otherwise the copies of H would technically be 'lumped together' (as defined earlier) at every edge of G that lies in at least one of them. So our graphs satisfy the left inequality in $(*)$, which means that the right one must fail: they must contain at least ηn^k copies of H for some fixed $\eta > 0$. In particular, for some mysterious reason exposed by the removal lemma, *we cannot have quadratically*[16] *many copies of H spread evenly over dense graphs G.*

Let us say that the copies of H are *spread thinly* over the graphs G in some infinite class $\mathcal{G}$ if, for every $G \in \mathcal{G}$, every edge of G lies in at least one copy of H in G but only in a bounded number: in at most d copies, say, with the same $d \in \mathbb{N}$ for all $G \in \mathcal{G}$. Then θ and $\hom(H, G)$ grow at most quadratically in n, more slowly than ηn^k for any η returned by the removal lemma. Hence those G will satisfy the premise of the removal lemma as soon as they are big enough. Our considerations above imply that those G cannot be dense. Let us show this more formally: *spread thinly*

Theorem 7.6.4. *Let H be any graph with at least 3 vertices. Let $\mathcal{G}$ be a class of graphs over which their subgraphs isomorphic to H are spread thinly. Then for every $\delta > 0$ there exists an integer n_0 such that every $G \in \mathcal{G}$ of order $n \geqslant n_0$ has at most δn^2 edges and contains at most δn^2 copies of H.*

Proof. Let $k := |H| \geqslant 3$ and $m := \|H\|$; we may assume that $\|H\| \geqslant 1$. *k, m*
Let d be the constant from the definition of 'thinly spread': an integer such that every edge of any $G \in \mathcal{G}$ lies in at most d copies of H in G. *d*
For each G, write $h = h(G)$ for the number of injective homomorphisms $H \to G$. *h*

We apply the removal lemma with $\varrho = \delta/(m\,k!\,d)$. It returns some $\eta > 0$ and an integer 'n_0', which we now call n_1. We use these to define *ϱ, η*

$$n_0 := \max\{n_1, \lceil k! \cdot d/\eta \rceil\}.$$

Let $G \in \mathcal{G}$ of order $n \geqslant n_0$ be given. Each copy of H in G is the

[16] In complexity-theoretic notation: $\Theta(n^2)$ many.

image of at most $k!$ injective homomorphisms $H \to G$, so

$$h(G) \leqslant k! \cdot d\,\|G\| \leqslant k! \cdot d \cdot n^2 \leqslant \eta n_0 \cdot n^2 \leqslant \eta n^3.$$

The removal lemma therefore yields a set E of at most ϱn^2 edges in G such that $H \not\subseteq G - E$. Like every edge of G, each of the edges in E lies in at most d copies of H in G. So we also have

$$h(G) \leqslant k! \cdot d\,|E|.$$

As the copies of H are thinly spread over G, every edge of G lies in the image of some injective homomorphism $H \to G$. Hence $h(G) \geqslant \|G\|/m$, so

$$\|G\| \leqslant m\,h(G) \leqslant m\,k!\,d\,|E| \leqslant m\,k!\,d\,\varrho n^2 = \delta n^2.$$

As the number h of copies of H in G is no bigger than $h(G)$, the above inequality also shows $h \leqslant h(G) \leqslant \delta n^2$, as claimed. □

[7.7.3] **Corollary 7.6.5.** *For every $\delta > 0$ there exists an integer n_0 such that the following holds for every graph G of order $n \geqslant n_0$: if every edge of G lies in exactly one triangle, then G has at most δn^2 edges.* □

7.7 Szemerédi's theorem

When Szemerédi proved his first version of the regularity lemma, he did not think of it as a precursor to the fundamental structure theorem for all of extremal graph theory that it has since become. He thought of it, well, as a lemma: a lemma he needed in a proof of a conjecture of Erdős and Turán about 'arithmetic progressions' of integers. We shall describe this theorem in this section, and then use the tools from Section 7.6 to derive its simplest case in just over a page.

arithmetic progression An *arithmetic progression* is an increasing sequence of integers of the form $a, a + d, a + 2d, \ldots$, one in which adjacent terms always differ by the same amount. The number of terms in the sequence is its *length*. *length* One of the earliest results in combinatorics is *van der Waerden's theorem* that, no matter how we partition the natural numbers into finitely many classes, one of these will contain arbitrarily long arithmetic progressions.[17] Szemerédi's theorem explores the deeper reasons why this is so.

upper density In any partition $\{A_1, \ldots, A_p\}$ of $\mathbb{N}$ into finitely many classes, one of these A_i will have positive *upper density*: for some $\delta > 0$, in our case certainly for $\delta = 1/p$, arbitrarily large $n \in \mathbb{N}$ will satisfy $|A_i \cap n| \geqslant \delta n$

[17] Compare Exercise 5 in Chapter 9.

(where $n = \{0, \dots, n-1\}$). Szemerédi's theorem says that the latter is enough to assume to get arbitrarily long arithmetic progressions in A_i. We do not need a choice between A_i and other partition sets, as in van der Waerden's theorem, we just need a single set that is dense enough:

Theorem 7.7.1. (Szemerédi 1975)
Every set $A \subseteq \mathbb{N}$ of positive upper density contains arbitrarily long arithmetic progressions.

Szemerédi's theorem is often expressed in more compact notation. For all $k, n \in \mathbb{N}$, and abbreviating 'arithmetic progression' as 'AP', let AP

$$r_k(n) := \max\left\{\, |D| : D \subseteq n \text{ and } D \text{ contains no AP of length } k \,\right\}.$$

Thus, if $k \in \mathbb{N}$ is such that $\lim_{n\to\infty} r_k(n)/n = 0$ then, for any sets $D_n \subseteq n$ that contain no arithmetic progression of length k, the values of $|D_n|/n$ will tend to zero as n grows. But if $A \subseteq \mathbb{N}$ has positive upper density, the values of $|A \cap n|/n$ exceed some $\delta > 0$ again and again. For those n, the sets $A \cap n$ will eventually contain arithmetic progressions of length k. This happens for every k, just for later n when k is bigger. Hence the assertions that $\lim_{n\to\infty} r_k(n)/n = 0$, one for each k, together imply Theorem 7.7.1.

For this reason, Szemerédi's theorem is often stated as follows:

Theorem 7.7.2. *For all integers $k \geqslant 3$ we have* $\lim_{n\to\infty} r_k(n)/n = 0$.

Even given the regularity lemma, Szemerédi's theorem is still a deep result far beyond the scope of this book. Indeed, it is surprising that an arithmetic statement such as this is susceptible to a purely combinatorial proof at all. The case of $k = 3$, however, which is known as Roth's theorem (1953), follows easily from Corollary 7.6.5:

Proposition 7.7.3. $\lim_{n\to\infty} r_3(n)/n = 0$.

Proof. We show that, for every $\epsilon > 0$, we have $r_3(n)/n \leqslant \epsilon$ for every n greater than the n_0 returned by Corollary 7.6.5 on input $\delta := \epsilon/12$. To prove this for given $n \geqslant n_0$, consider any set $D \subseteq n$ that contains no arithmetic progression of length 3. We have to show that $|D| \leqslant \epsilon n$. [7.6.5] ϵ, δ, n_0 D

We shall apply Corollary 7.6.5 to the tripartite graph G with vertex classes $X := \{1, \dots, n\}$ and $Y := \{1, \dots, 2n\}$ and $Z := \{1, \dots, 3n\}$, viewed as disjoint sets. As the edges of G we take those of its *defining triangles*, the complete graphs $K(x,d) = K^3$ with vertices X, Y, Z *defining triangle* $K(x,d)$

$$x \in X \quad\text{and}\quad x + d \in Y \quad\text{and}\quad 2x + d \in Z$$ $E(G)$

for all $x \in X$ and $d \in D$.

Since any two vertices of a defining triangle determine its third vertex, the defining triangles are edge-disjoint. Thus, every edge of G lies in a unique defining triangle, and

$$|G| = 6n \quad \text{and} \quad \|G\| = 3n|D|\,.$$

Let us show that G has no other triangles than the defining ones, and therefore satisfies the premise of Corollary 7.6.5.

T — T_1, T_2, T_3

Suppose $T \subseteq G$ is a non-defining triangle. Let T_1, T_2, T_3 be the defining triangles that contributed its edges. They are distinct, because any triangle that contains two of the edges of T contains all its vertices, and hence equals T. Since there are no edges inside the partition classes X, Y, Z, our T has one vertex in each class (Fig. 7.7.1, left).

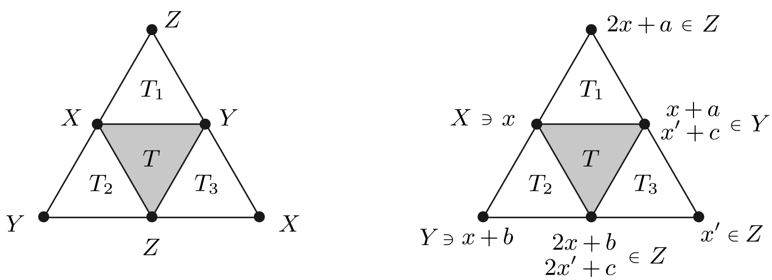

Fig. 7.7.1. Three defining triangles creating another triangle

Assume that the X–Y edge of T was contributed by T_1, and its X–Z edge by T_2. The vertex x of T in X then lies in $T_1 \cap T_2$, so

$$T_1 = K(x, a) \quad \text{and} \quad T_2 = K(x, b) \quad \text{and} \quad T_3 = K(x', c)$$

a, b, c — for some $a, b, c \in D$ and $x' \in X$. Since T_3 contributes the Y–Z edge of T but not its other edges, we have $x' \neq x$. The vertices of T, then, are

$$x \in X \quad \text{and} \quad x + a = x' + c \in Y \quad \text{and} \quad 2x + b = 2x' + c \in Z$$

(Fig. 7.7.1, right), so

$$a = c + (x' - x) \quad \text{and} \quad b = c + 2(x' - x)\,.$$

Hence either (c, a, b) or (b, a, c) is an arithmetic progression of length 3 in D, with difference $|x' - x| \neq 0$ since $x' \neq x$. This contradicts the choice of D, so we conclude that G contains no triangles other than the defining ones. In particular, every edge of G lies in exactly one triangle.

Corollary 7.6.5 now yields

$$\|G\| \leqslant \delta \cdot |G|^2 = \delta \cdot (6n)^2.$$

As $\|G\| = 3n|D|$, we conclude that

$$|D| = \|G\|/3n \leqslant \delta \cdot 12n = \epsilon n$$

as desired. $\square$

Exercises

1.$^-$ Show that $K_{1,3}$ is extremal without a P^3.

2.$^-$ Given $k > 0$, determine the extremal graphs of chromatic number at most k.

3.$^-$ Is there a graph that is edge-maximal without a K^3 minor but not extremal?

4. Determine the value of $\mathrm{ex}(n, K_{1,r})$ for all $r, n \in \mathbb{N}$.

5.$^+$ Given $k > 0$, determine the extremal graphs without a matching of size k.

   (Hint. Theorem 2.2.3 and Ex. 20, Ch. 2.)

6. Without using Turán's theorem, show that the maximum number of edges in a triangle-free graph of order $n > 1$ is $\lfloor n^2/4 \rfloor$.

7. Show that

$$t_{r-1}(n) \leqslant \tfrac{1}{2} n^2 \, \frac{r-2}{r-1} \, ,$$

   with equality whenever $r - 1$ divides n.

8. Show that $t_{r-1}(n)/\binom{n}{2}$ converges to $(r-2)/(r-1)$ as $n \to \infty$.
   (Hint. $t_{r-1}\big((r-1)\lfloor \frac{n}{r-1} \rfloor\big) \leqslant t_{r-1}(n) \leqslant t_{r-1}\big((r-1)\lceil \frac{n}{r-1} \rceil\big)$.)

9. Does every large enough graph G with at most $c\,|G|$ edges, where c is any constant, contain a set of 100 independent vertices?

10. Show that deleting at most $(m-s)(n-t)/s$ edges from a $K_{m,n}$ will never destroy all its $K_{s,t}$ subgraphs.

11. For $0 < s \leqslant t \leqslant n$ let $z(n, s, t)$ denote the maximum number of edges in a bipartite graph whose partition sets both have size n, and which does not contain a $K_{s,t}$. Show that $2\,\mathrm{ex}(n, K_{s,t}) \leqslant z(n, s, t) \leqslant \mathrm{ex}(2n, K_{s,t})$.

12.$^+$ Let $1 \leqslant r \leqslant n$ be integers. Let G be a bipartite graph with bipartition $\{A, B\}$, where $|A| = |B| = n$, and assume that $K_{r,r} \not\subseteq G$. Show that

$$\sum_{x \in A} \binom{d(x)}{r} \leqslant (r-1)\binom{n}{r}.$$

   Using the previous exercise, deduce that $\mathrm{ex}(n, K_{r,r}) \leqslant c n^{2-1/r}$ for some constant c depending only on r.

13. The *upper density* of an infinite graph G is the lim sup of the maximum edge densities of its (finite) n-vertex subgraphs as $n \to \infty$.

    (i) Show that, for every $r \in \mathbb{N}$, every infinite graph of upper density $> \frac{r-2}{r-1}$ has a K^r_s subgraph for every $s \in \mathbb{N}$.

    (ii) Deduce that the upper density of infinite graphs can only take the countably many values of $0, 1, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots$.

14. Given a tree T, find an upper bound for $\mathrm{ex}(n, T)$ that is linear in n and independent of the structure of T, i.e. depends only on $|T|$.

15. Show that the Erdős-Sós conjecture is best possible in the sense that, for every k and infinitely many n, there is a graph on n vertices and with $\frac{1}{2}(k-1)n$ edges that contains no tree with k edges.

16.$^-$ Prove the Erdős-Sós conjecture for the case when the tree considered is a star.

17. Prove the Erdős-Sós conjecture for the case when the tree considered is a path.

    (Hint. Use Exercise 9 of Chapter 1.)

18. Can large average degree force the chromatic number up if we exclude some tree as an induced subgraph? More precisely: For which trees T is there a function $f : \mathbb{N} \to \mathbb{N}$ such that, for every $k \in \mathbb{N}$, every graph of average degree at least $f(k)$ either has chromatic number at least k or contains an induced copy of T?

19. Given two numerical graph invariants i_1 and i_2, write $i_1 \leqslant i_2$ if we can force i_2 to be arbitrarily high on some subgraph of G by assuming that $i_1(G)$ is large enough. (Formally: write $i_1 \leqslant i_2$ if there exists a function $f : \mathbb{N} \to \mathbb{N}$ such that, given any $k \in \mathbb{N}$, every graph G with $i_1(G) \geqslant f(k)$ has a subgraph H with $i_2(H) \geqslant k$.) If $i_1 \leqslant i_2$ as well as $i_1 \geqslant i_2$, write $i_1 \sim i_2$. Show that this is an equivalence relation for graph invariants, and sort the following invariants into equivalence classes ordered by $<$: minimum degree; average degree; connectivity; arboricity; chromatic number; colouring number; choice number; $\max \{ r \mid K^r \subseteq G \}$; $\max \{ r \mid TK^r \subseteq G \}$; $\max \{ r \mid K^r \preccurlyeq G \}$; $\min \max d^+(v)$, where the maximum is taken over all vertices v of the graph, and the minimum over all its orientations.

20.$^+$ Prove, from first principles and without using average or minimum degree arguments, the existence of a function $f : \mathbb{N} \to \mathbb{N}$ such that every graph of chromatic number at least $f(r)$ has a K^r minor.

    (Hint. Use induction on r. For the induction step $(r-1) \to r$ try to find a connected set U of vertices whose neighbours induce a subgraph that needs enough colours to contract to K^{r-1}. If no such set U exists, show that the given graph can be coloured with fewer colours than assumed.)

21. Given a graph G with $\varepsilon(G) \geqslant k \in \mathbb{N}$, find a minor $H \preccurlyeq G$ such that $\delta(H) \geqslant k \geqslant |H|/2$.

22.$^+$ Find a constant c such that every graph with n vertices and at least $n + 2k(\log k + \log\log k + c)$ edges contains k edge-disjoint cycles (for all $k \in \mathbb{N}$). Deduce an edge-analogue of the Erdős-Pósa theorem (2.3.2).

(Hint. Assuming $\delta \geqslant 3$, delete the edges of a short cycle and apply induction. The calculations are similar to the proof of Lemma 2.3.1.)

23. Simplify the proof of Theorem 7.2.3 by using Exercise 36 of Chapter 3.

24.$^+$ Show that any function h as in Lemma 3.5.1 satisfies the inequality $h(r) > \frac{1}{8}r^2$ for all even r, and hence that Theorem 7.2.3 is best possible up to the value of the constant c.

25. Characterize the graphs with n vertices and more than $3n - 6$ edges that contain no $TK_{3,3}$. In particular, determine $\mathrm{ex}(n, TK_{3,3})$.

(Hint. You may use the theorem of Wagner that every edge-maximal graph without a $K_{3,3}$ minor can be constructed recursively from maximal planar graphs and copies of K^5 by pasting along K^2s.)

26.$^-$ Derive the four colour theorem from Hadwiger's conjecture for $r = 5$.

27.$^-$ Show that Hadwiger's conjecture for $r + 1$ implies the conjecture for r.

28.$^-$ Deduce the following weakening of Hadwiger's conjecture from known results: given any $\epsilon > 0$, every graph of chromatic number at least $r^{1+\epsilon}$ has a K^r minor, provided that r is large enough.

29.$^-$ Show that any graph constructed as in Proposition 7.3.1 is edge-maximal without a K^4 minor.

30. Prove the implication $\delta(G) \geqslant 3 \Rightarrow G \supseteq TK^4$.

(Hint. You may use any result from Section 7.3.)

31. A multigraph is called *series-parallel* if it can be constructed recursively from a K^2 by the operations of subdividing and of doubling edges. Show that a 2-connected multigraph is series-parallel if and only if it has no (topological) K^4 minor.

32. Without using Theorem 7.3.8, prove Hadwiger's conjecture for all graphs of girth at least 11 and r large enough. Without using Corollary 7.3.9, show that there is a constant $g \in \mathbb{N}$ such that all graphs of girth at least g satisfy Hadwiger's conjecture, irrespective of r.

33.$^+$ Prove Hadwiger's conjecture for $r = 4$ from first principles.

34.$^+$ Prove Hadwiger's conjecture for line graphs.

35. Prove Corollary 7.3.5.

36.$^-$ In the definition of an ϵ-regular pair, what is the purpose of the requirement that $|X| \geqslant \epsilon\,|A|$ and $|Y| \geqslant \epsilon\,|B|$?

37.$^-$ Show that any ϵ-regular pair in G is also ϵ-regular in $\overline{G}$.

38. Does the regularity lemma for $m = 1$ imply the general version?

39. Consider a partition of a finite set V into k equally sized subsets. Show that the complete graph on V has about $k-1$ as many edges between different partition sets as edges inside partition sets. Explain how this leads to the choice of $m := 1/\gamma$ in the proof of the Erdős-Stone theorem.

40. (i) Deduce the regularity lemma from the assumption that it holds, given $\epsilon > 0$ and $m \geqslant 1$, for all graphs of order at least some $n = n(\epsilon, m)$.

(ii) Prove the regularity lemma for sparse graphs – more precisely, for every sequence $(G_n)_{n \in \mathbb{N}}$ of graphs G_n of order n such that $\|G_n\|/n^2 \to 0$ as $n \to \infty$.

41.$^-$ Consider a pair (A, B) of disjoint vertex sets in a graph, and a pair of subsets $X \subseteq A$ and $Y \subseteq B$ that are small in that $|X||Y| \leqslant \epsilon\,|A||B|$. For which d does (X, Y) satisfy the requirement of (ϵ, d)-uniformity for (A, B)?

42. Compare the definitions of ϵ-regularity and ϵ-uniformity for pairs of disjoint vertex sets in a graph:

(i)$^-$ Show that every ϵ-regular pair is also ϵ-uniform.

(ii) Show that every ϵ-uniform pair is $\sqrt[3]{\epsilon}$-regular.

43. Discuss the factor of $\binom{k}{2}$ in Proposition 7.6.1. Is it natural? Is it important? Why is it there?

44. Deduce Proposition 7.6.1 from Lemma 7.6.2.

45. Replace the use of the blow-up lemma with the counting lemma...

(i) ...in the proof of Theorem 7.1.2;

(ii) ...in the proof of Theorem 9.2.2.

46. A consequence of the Erdős-Stone theorem is that γn^2 more edges than are needed to force a K^r subgraph will force many such subgraphs, as many as a K^r_s contains. Can you prove this directly using the removal lemma? Can you find even more K^r subgraphs in this way?

47. Let H be any graph, of order k say. Does the removal lemma say, essentially, that we can hit all the copies of H in arbitrary graphs G by $o(n^2)$ edges, where $n := |G|$, if H is such that no more than $o(n^k)$ of the n^k maps $V(H) \to V(G)$ induce homomorphisms $H \to G$? If not, what is missing?

48.$^-$ Find a class of 3-connected planar graphs over which triangles are spread thinly.

49. (i) In the definition of 'thinly spread', and again in Corollary 7.6.5, there is a condition requiring that every edge of G lie in at least one copy of H. If we drop this assumption, will Theorem 7.6.4 and Corollary 7.6.5 fail?

(ii) In our earlier informal discussion of a typical scenario for the removal lemma, which begins with the inequalities $(*)$ and concludes that many copies of H are lumped together at just a few edges of G, there is no such condition as in (i). Why was it not needed there?

50. In the definition of 'thinly spread', weaken the condition that every edge of G lies in at least one copy of H to the condition that the edges of G lie in at least β copies of H on average, for some $\beta > 0$ that is the same for all $G \in \mathcal{G}$. Does Theorem 7.6.4 still hold?

51. We saw that Theorem 7.7.2 implies Theorem 7.7.1. Does the converse implication hold, too?

Notes

The standard reference work for results and open problems in extremal graph theory (in a very broad sense) is still B. Bollobás, *Extremal Graph Theory*, Academic Press 1978. A kind of update on the book is given by its author in his chapter of the *Handbook of Combinatorics* (R.L. Graham, M. Grötschel & L. Lovász, eds.), North-Holland 1995. An instructive survey of extremal graph theory in the narrower sense of Section 7.1 is given by M. Simonovits in (L.W. Beineke & R.J. Wilson, eds.) *Selected Topics in Graph Theory 2*, Academic Press 1983. This paper focuses among other things on the particular role played by the Turán graphs. A more recent survey by the same author can be found in (R.L. Graham & J. Nešetřil, eds.) *The Mathematics of Paul Erdős*, Vol. 2, Springer 1996.

Turán's theorem is not merely one extremal result among others: it is the result that sparked off the entire line of research. Our first proof of Turán's theorem is essentially the original one; the second is a version of a proof of Zykov due to Brandt.

Turán's theorem has been generalized as follows. Suppose that, for some fixed $r \geqslant 3$, we wish to construct a graph on n vertices with at least γn^2 edges, where now $\frac{1}{2}\frac{r-2}{r-1} < \gamma < \frac{1}{2}$, in such a way as to create as few K^r subgraphs as possible. The *clique density theorem* says that, for fixed γ, the asymptotically best way to do this is to form a complete multipartite graph in which all classes have the same size except for one, which may be smaller. How many such classes there are depends on γ, but not on n: as in Turán's theorem, s classes will always give about γn^2 edges for $\gamma = \frac{1}{2}\frac{s-1}{s}$. The clique density theorem had been conjectured by Lovász and Simonovits in 1983, and was finally proved for all r by C. Reiher, The clique density theorem, *Ann. Math.* **184** (2016), 683–707, arXiv:1212.2454.

Our version of the Erdős-Stone theorem is a slight simplification of the original. A direct proof, not using the regularity lemma, is given in L. Lovász, *Combinatorial Problems and Exercises* (2nd edn.), North-Holland 1993. Its most fundamental application, Corollary 7.1.3, was only found 20 years after the theorem, by Erdős and Simonovits (1966).

Of our two bounds on $\mathrm{ex}(n, K_{r,r})$ the upper one is thought to give the correct order of magnitude. For vastly off-diagonal complete bipartite graphs this was verified by J. Kollár, L. Rónyai & T. Szabó, Norm-graphs and bipartite Turán numbers, *Combinatorica* **16** (1996), 399–406, who proved that $\mathrm{ex}(n, K_{r,s}) \geqslant c_r n^{2-\frac{1}{r}}$ when $s > r!$.

Details about the Erdős-Sós conjecture, including an approximate solution for large k, can be found in the survey by Komlós and Simonovits cited below. The case where the tree T is a path (Exercise 17) was proved by

Erdős & Gallai in 1959. It was this result, together with the easy case of stars (Exercise 16) at the other extreme, that inspired the conjecture as a possible unifying result. A proof of the precise conjecture for large graphs was announced in 2009 by Ajtai, Komlós, Simonovits and Szemerédi, but has not been made publicly available.

The Erdős-Sós conjecture says that graphs of average degree greater than $k-1$ contain every tree with k edges. Loebl, Komlós and Sós have conjectured a 'median' version, which appears to be easier: that if at least half the vertices of a graph have degree greater than $k-1$ it contains every tree with k edges. An approximate version of this conjecture has been proved by Hladký, Komlós, Piguet, Simonovits, Stein and Szemerédi in arXiv:1408.3870.

Theorem 7.2.3 was first proved by B. Bollobás & A.G. Thomason, Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs, *Eur. J. Comb.* **19** (1998), 883–887, and independently by J. Komlós & E. Szemerédi, Topological cliques in graphs II, *Comb. Probab. Comput.* **5** (1996), 79–90. For large G, the latter authors show that the constant c in the theorem can be brought down to about $\frac{1}{2}$, which is not far from the lower bound of $\frac{1}{8}$ given in Exercise 24.

Theorem 7.2.4 was first proved in 1982 by Kostochka, and in 1984 with a better constant by Thomason. For references and more insight, also in these early proofs, see A.G. Thomason, The extremal function for complete minors, *J. Comb. Theory, Ser. B* **81** (2001), 318–338. There, Thomason determines the smallest possible value of the constant c in Theorem 7.2.4 asymptotically for large r. It can be written as $c = \alpha + o(1)$, where $\alpha = 0.53131\ldots$ is an explicit constant and $o(1)$ stands for a function of r tending to zero as $r \to \infty$.

Surprisingly, the average degree needed to force an *in*complete minor H of order r remains at $cr\sqrt{\log r}$, with $c = \alpha\gamma(H) + o(1)$, where γ is a graph invariant $H \mapsto [0,1]$ that is bounded away from 0 for dense H, and $o(1)$ is a function of $|H|$ tending to 0 as $|H| \to \infty$. See J.S. Myers & A.G. Thomason, The extremal function for noncomplete minors, *Combinatorica* **25** (2005), 725–753.

As Theorem 7.2.4 is best possible, there is no constant c such that all graphs of average degree at least cr have a K^r minor. Strengthening this assumption to $\kappa \geqslant cr$, however, can force a K^r minor in all large enough graphs; this was proved by T. Böhme, K. Kawarabayashi, J. Maharry and B. Mohar, Linear connectivity forces large complete bipartite minors, *J. Comb. Theory, Ser. B* **99** (2009), 557–582. Their proof rests on a structure theorem for graphs of large tree-width not containing a given minor, which was proved only later by R. Diestel, K. Kawarabayashi, Th. Müller & P. Wollan, On the excluded minor structure theorem for graphs of large tree-width, *J. Comb. Theory, Ser. B* **102** (2012), 1189–1210, arXiv:0910.0946. A simple direct argument that bypasses the use of this structure theorem was found by J.-O. Fröhlich and Th. Müller, Linear connectivity forces large complete bipartite minors: an alternative approach, *J. Comb. Theory, Ser. B* **101** (2011), 502–508, arXiv:0906.2568.

The fact that large enough girth can force minors of arbitrarily high minimum degree, and hence large complete minors, was discovered by Thomassen in 1983. The reference can be found in W. Mader, Topological subgraphs in graphs of large girth, *Combinatorica* **18** (1998), 405–412, from which our Lemma 7.2.5 is extracted. Our girth assumption of $8k+3$ has been reduced to about $4k$ by D. Kühn and D. Osthus, Minors in graphs of large girth, *Random Struct. Alg.* **22** (2003), 213–225, which is conjectured to be best possible.

The original reference for Theorem 7.2.7 can be found in D. Kühn and D. Osthus, Improved bounds for topological cliques in graphs of large girth, *SIAM J. Discrete Math.* **20** (2006), 62–78, where they re-prove their theorem with $g \leqslant 27$. See also D. Kühn & D. Osthus, Subdivisions of K_{r+2} in graphs of average degree at least $r+\varepsilon$ and large but constant girth, *Comb. Probab. Comput.* **13** (2004), 361–371.

The proof of Hadwiger's conjecture for $r = 4$ hinted at in Exercise 33 was given by Hadwiger himself, in the 1943 paper containing his conjecture. Like Hadwiger's conjecture, Hajós's conjecture has (later) been proved for graphs of large girth (Corollary 7.3.9) and for line graphs; see C. Thomassen, Hajós' conjecture for line graphs, *J. Comb. Theory, Ser. B* **97** (2007), 156–157. A counterexample to the general Hajós conjecture was found as early as 1979 by Catlin. A little later, Erdős and Fajtlowicz proved that Hajós's conjecture is false for 'almost all' graphs, while Bollobás, Catlin and Erdős showed that Hadwiger's conjecture is true for 'almost all graphs' (see Chapter 11). Proofs of Wagner's Theorem 7.3.4 (with Hadwiger's conjecture for $r = 5$ as a corollary) can be found in Bollobás's *Extremal Graph Theory* (see above) and in Halin's *Graphentheorie* (2nd ed.), Wissenschaftliche Buchgesellschaft 1989. Hadwiger's conjecture for $r = 6$ was proved by N. Robertson, P.D. Seymour and R. Thomas, Hadwiger's conjecture for K_6-free graphs, *Combinatorica* **13** (1993), 279–361.

V. Dujmović, L. Esperet, P. Morin and D.R. Wood proved the following weakening of Hadwiger's conjecture: for every integer $r > 0$, the vertices of any graph without a K^r minor can be coloured with at most $r-1$ colours so that the colour classes induce subgraphs with bounded-sized components, the bound being a function only of r. See arXiv:2306.06224.

For infinite graphs, the following weakening of the assertion of Hadwiger's conjecture is true: every graph of chromatic number $\alpha \geqslant \aleph_0$ contains every K^β with $\beta < \alpha$ as a minor, even as a topological minor. This was proved by R. Halin, Unterteilungen vollständiger Graphen in Graphen mit unendlicher chromatischer Zahl, *Abh. Math. Sem. Univ. Hamburg* **31** (1967), 156–165. The case of $\alpha = \aleph_0$ is Exercise 16 in Chapter 8; the proof for $\alpha > \aleph_0$ is included in R. Diestel, *Graph Decompositions*, Oxford University Press 1990.

The investigation of graphs not containing a given graph as a minor, or topological minor, has a long history. It probably started with Wagner's 1935 PhD thesis, in which he sought to 'detopologize' the four colour problem by classifying the graphs without a K^5 minor. His hope was to be able to show abstractly that all those graphs were 4-colourable; since the graphs without a K^5 minor include the planar graphs, this would amount to a proof of the four colour conjecture involving no topology whatsoever. The result of Wagner's efforts, Theorem 7.3.4, falls tantalizingly short of this goal: although it succeeds in classifying the graphs without a K^5 minor in structural terms, planarity re-emerges as one of the criteria used in the classification. From this point of view, it is instructive to compare Wagner's K^5 theorem with similar classification theorems, such as his analogue for K^4 (Proposition 7.3.1), where the graphs are decomposed into parts from a *finite* set of irreducible graphs. See R. Diestel, *Graph Decompositions*, Oxford University Press 1990, for more such classification theorems.

Despite its failure to resolve the four colour problem, Wagner's K^5 structure theorem had consequences for the development of graph theory like few others. To mention just two: it prompted Hadwiger to make his famous conjecture; and it inspired much of the work of Robertson and Seymour on minors (Chapter 12), in particular the notion of a tree-decomposition and the structure theorem for graphs without a K^n minor (Theorem 12.6.6). Wagner himself responded to Hadwiger's conjecture with a proof in 1964 that, to force a K^r minor, it does suffice to raise the chromatic number of a graph to *some* value depending only on r (Exercise 20). This theorem, along with its analogue for topological minors proved independently by Dirac and by Jung, prompted the question which average degree suffices to force the desired minor. This was first addressed by Mader, whose seminal proofs of Propositions 7.2.1 and 7.2.2 were part of his PhD thesis in 1967.

Theorem 7.3.8 is a consequence of the more fundamental result of D. Kühn and D. Osthus, Complete minors in $K_{s,s}$-free graphs, *Combinatorica* **25** (2005) 49–64, that every graph without a $K_{s,s}$ subgraph that has average degree $r \geqslant r_s$ has a K^p minor for $p = \lfloor r^{1+\frac{1}{2(s-1)}} / (\log r)^3 \rfloor$. This was improved further by M. Krivelevich and B. Sudakov, Minors in expanding graphs, *Geom. Funct. Anal.* **19** (2009), 294–331, arXiv:0707.0133.

As in the Gyárfás-Sumner conjecture one may ask under what additional assumptions large average degree forces an *induced* subdivision of a given graph H. This was answered for arbitrary H by D. Kühn and D. Osthus, Induced subdivisions in $K_{s,s}$-free graphs of large average degree, *Combinatorica* **24** (2004) 287–304, who proved that for all $r, s \in \mathbb{N}$ there exists $d \in \mathbb{N}$ such that every graph $G \not\supseteq K_{s,s}$ with $d(G) \geqslant d$ contains a TK^r as an induced subgraph.

The regularity lemma is proved in E. Szemerédi, Regular partitions of graphs, *Colloques Internationaux CNRS* **260** – *Problèmes Combinatoires et Théorie des Graphes, Orsay* (1976), 399–401. Our proof follows an account by Alex Scott (personal communication). A broad survey on the regularity lemma and its applications is given by J. Komlós & M. Simonovits in (D. Miklós, V.T. Sós & T. Szőnyi, eds.) *Paul Erdős is 80*, Vol. 2, Proc. Colloq. Math. Soc. János Bolyai (1996). The concept of a regularity graph and Lemma 7.5.2 are taken from this paper. The latter is a simplified version, for embedding small graphs H, of a more general *blow-up lemma* for embedding larger graphs H (even as spanning subgraphs), from J. Komlós, G.N. Sárkősy and E. Szemerédi, Blow-up lemma, *Combinatorica* **17** (1997), 109–123. The regularity lemma was adapted to sparse graphs by A.D. Scott, Szemerédi's regularity lemma for matrices and sparse graphs, *Comb. Probab. Comput.* **20** (2011), 455–466.

Sections 7.6 and 7.7 follow unpublished lecture notes by Mathias Schacht. The natural analogue of the removal lemma for induced subgraphs was proved by N. Alon, E. Fischer, M. Krivelevich and M. Szegedy, Efficient testing of large graphs, *Combinatorica* **20** (2000), 451–476. The $k = 3$ case of Szemerédi's theorem, Proposition 7.7.3, had been conjectured by Erdős and Turán in 1936 and was proved by K.F. Roth, On certain sets of integers, *J. Lond. Math. Soc.* **28** (1953), 104–109. The general case, Theorem 7.7.2, was proved by E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.