The Probabilistic Method and the Regularity Lemma

Habilitationsschrift
zur Erlangung der Lehrbefähigung
für das Fach Informatik

vorgelegt dem Rat der Mathematisch-Naturwissenschaftlichen Fakultät II
der Humboldt-Universität zu Berlin

von

Dr. Deryk Simeon Osthus
geb. am 10.10.1974 in Cardiff

Prof. Dr. Jürgen Mlynek, Präsident
der Humboldt-Universität zu Berlin

Prof. Dr. Elmar Kulke, Dekan

Zulassung zum Habilitationsverfahren: 3.11.2003
Datum des öffentlichen Fachgesprächs: 18.2.2004
Berlin, den 26.4.2004 (Abschluß des Verfahrens)

Gutachter:
1. Prof. Dr. Martin Grohe
2. Prof. Dr. Hans Jürgen Prömel
3. Prof. Dr. Andrew Thomason

Contents

Chapter 1

Introduction

The purpose of this chapter is firstly to give a brief introduction to the probabilistic method as well as to the Regularity lemma and secondly to give a summary of the results obtained in this thesis.

1.1 Some background

1.1.1 The probabilistic method

Roughly speaking, the probabilistic method can be described as follows: the existence of some desired object is demonstrated by considering some suitably defined probability space. Such an object might be a satisfying assignment for a Boolean function, a graph with some special properties or a certain substructure within a given graph. A large number of such examples is described in the textbook by Alon and Spencer [7] which is devoted to this method.

In many cases, the probabilistic existence argument can also be made constructive, which means that the method is strongly linked to the design and analysis of randomized algorithms (see the textbook by Motwani and Raghavan [83]). In fact, the resulting algorithms can sometimes even be derandomized, i.e. they can be turned into purely deterministic algorithms (see [83] again or [7]). Thus, the probabilistic method is also a useful tool to develop such algorithms.

Erdős was the first to apply and develop the method in a systematic way. He inititally applied it in order to obtain a lower bound on the Ramsey number $R(t,t)$ [31]. Here $R(t,t)$ is the smallest number n so that for every 2-colouring of the edges of the complete graph on n vertices, one can find a monochromatic complete graph on t vertices. Erdős considered a *random* 2-colouring of a complete graph on n vertices and showed that if n is not too large, then the expected number of monochromatic complete graphs of order t is smaller than one. This of course immediately implies the *existence* of a 2-colouring without a monochromatic complete graph on t vertices. So far no explicit construction is known which gives a comparable lower bound on $R(t,t)$.

The above elementary argument can also be rephrased as a (double) counting argument without making any explicit reference to probability. This is in fact also the case for several other applications of the probabilistic method.

However, the probabilistic viewpoint has two advantages: firstly (and arguably), it often provides extra intuition and thus provides the key to solving the problem in question. Secondly (and more importantly), most applications of the probabilistic method involve more sophisticated tools (like Chernoff's bound or the Lovász local lemma for instance) for which nonprobabilistic formulations are much harder or even impossible to obtain.

Counting arguments and the probabilistic method are also closely linked to the study of the typical properties of random structures. One such example is the celebrated result of Erdős [32] that there exist graphs which have both high girth and high chromatic number: the probabilistic proof was obtained by considering random graphs. We will refine this argument slightly in Chapter 8.

1.1.2 The Regularity lemma

The Regularity lemma [97] was developed in the 1970's by Szemerédi as a tool in Ramsey theory and combinatorial number theory. He used it to prove the famous conjecture of Erdős and Turán from 1936 that every dense subset of the integers must contain arbitrarily long arithmetic progressions. Roughly speaking, the assertion of the Regularity lemma is that the vertices of any dense graph can be partitioned into a bounded number of clusters so that most of the bipartite subgraphs between the clusters look like random graphs. Thus even more roughly, it states that every dense graph can in some sense be approximated by random graphs. The exact statement is rather technical, and we defer it to Chapter 2. As Komlós [48] puts it in his survey: "This is not a very transparent theorem, but it grows on you with time". Indeed, gradually the Regularity lemma was recognized to be an extremely important tool also beyond Ramsey theory. Its impact on computer science has increased considerably since Alon et al. [4] discovered an algorithmic version (i.e. an algorithm which constructs the above partition in polynomial time).

More recently, Komlós, Sárközy and Szemerédi developed another powerful tool (called the Blow-up lemma [51]), which is very useful in conjunction with the Regularity lemma. The Blow-up lemma makes it possible to find special spanning subgraphs in the subgraphs between the clusters of the regularity partition (and thus with some further work also in the original graph). For instance the above authors applied it to prove the long-standing conjecture of Pósa (for large graphs) that every graph on n vertices whose minimum degree is at least $2n/3$ contains the square of a Hamilton cycle [50]. We will also apply this tool in the final three chapters of this thesis.

The Regularity lemma can often be combined with the probabilistic method. In fact, with the exception of Chapter 12, this will always be the case when we apply the Regularity lemma. One such example occurs in Chapter 10, where we need (and prove) an auxiliary result which states that (roughly speaking) with very high probability pseudo-randomness of bipartite graphs is inherited by subgraphs induced by random subsets. We apply this to find large topological cliques in graphs but hope that it will also be applicable elsewhere. In fact, a similar statement was proved independently of us by Gerke et al. [34] in the context of extremal subgraphs of random graphs.

1.2 Summary of results

A more detailed discussion of the results stated below is deferred to the relevant chapters. The results in Chapters 8 and 9 are joint work with Hans Jürgen Prömel and Anusch Taraz [88, 89]. The results in Chapter 12 are joint work with Daniela Kühn and Anusch Taraz [73]. The results in the remaining chapters are joint work with Daniela Kühn [63]–[66], [68]–[71].

I. The probabilistic method

Almost all of the proofs of the main results in this part make use of the probabilistic method.

Minors in graphs of large girth

We say that a graph H is a *minor* of some other graph G if H can be obtained from a subgraph of G by contracting edges. A fundamental result of Mader [75] on subdivisions implies that minors can be forced by large average degree: for every natural number r there exists a smallest number $f(r)$ such that every graph G of average degree larger than $f(r)$ contains the complete graph K_r of order r as minor. Kostochka [59] and Thomason [99] independently showed that there exists a constant c such that $f(r) \leq cr\sqrt{\log r}$. Recently, Thomason [100] was able to determine $f(r)$ asymptotically: he showed that $f(r) = (1 + o(1))cr\sqrt{\log r}$ for an explicit constant c. Dense random graphs are extremal in the sense that they provide the lower bound for this result.

The question of which chromatic number guarantees a K_r minor is still wide open. In 1943 Hadwiger conjectured that a chromatic number of at least r suffices. Since K_r has chromatic number r but does not contain a K_{r+1} minor, the conjecture would be best possible. As every graph of chromatic number r has a subgraph of minimum degree at least $r-1$, the above results about average degree imply that if c is a sufficiently large positive constant, then every graph of chromatic number at least $cr\sqrt{\log r}$ contains a K_r minor. But it is not known whether graphs of large chromatic number contain larger complete minors than those forced by the average degree of their subgraphs. For example, it is still open whether one can always find a complete minor whose order is linear in the chromatic number.

Thomassen [102] observed that large complete minors are also forced by large girth. More precisely, he showed that if the girth of a graph G is large and its minimum degree is at least 3, then G contains minors whose minimum degree is much larger than that of G itself and thus G also contains large complete minors. Here the *girth* of a graph is the length of its shortest cycle. So if a graph G has large girth then it looks locally like a tree. Thus Thomassen's observation is the surprising fact that if the minimum degree of such a 'locally sparse' graph is at least 3, i.e. if its large girth is not merely obtained by subdividing edges, then it must contain a large 'dense' substructure, namely a large complete minor. Diestel and Rempel [29] gave a better bound on the girth required to force a K_r minor in a graph of minimum degree at least 3: they showed that there exists a constant c such that a girth of at least $6\log_2 r + 3\log_2\log_2 r + c$ will do.

In Chapter 3 we will use probabilistic arguments to give more precise asymptotic bounds on the girth required to force a minor of given minimum degree:

Theorem A *For every odd integer $g \geq 3$ there exists a positive constant $c = c(g)$ such that for all $r \geq 3$ every graph G of minimum degree at least r and girth at least g contains a minor of minimum degree at least $c(r-1)^{(g+1)/4}$.*

Together with the result of Kostochka and Thomason mentioned earlier this implies the following.

Corollary B *For every odd integer $g \geq 3$ there exists a positive constant $c = c(g)$ such that every graph of minimum degree at least r and girth at least g contains a complete graph as minor whose order is at least $cr^{\frac{g+1}{4}}/\sqrt{\log r}$.*

As every graph of chromatic number r contains a subgraph of minimum degree at least $r-1$, this implies Hadwiger's conjecture for graphs of girth at least five and sufficiently large chromatic number. (We will improve on this result in the following section).

For fixed girth, Theorem A and Corollary B would give the right order of magnitude (as a function of r) provided that a well known conjecture about the existence of small graphs of given minimum degree and given girth is true. Moreover, if $g = 4k+3$ for some $k \in \mathbb{N}$, then the constant c in Theorem A is in fact independent of g and thus Theorem A improves the cited bound by Diestel and Rempel: it implies that there exists a constant c such that every graph of minimum degree at least 3 and girth at least $4\log_2 r + 2\log_2\log_2 r + c$ has a K_r minor. If the above-mentioned conjecture holds, the leading constant 4 would be correct.

Minors in $K_{s,s}$-free graphs

In Chapter 4 we shall see that not only graphs of large girth contain large minors, but also graphs which are locally sparse in the much weaker sense that they do not contain a fixed complete bipartite graph $K_{s,s}$ as a subgraph:

Theorem C *For each integer $s \geq 2$ every $K_{s,s}$-free graph of average degree at least r contains a complete graph as minor whose order is at least*

$$r^{1+\frac{1}{2(s-1)}-o(1)}.$$

Again, as every graph of chromatic number r contains a subgraph of minimum degree at least $r-1$, this implies Hadwiger's conjecture for $K_{s,s}$-free graphs of sufficiently large chromatic number. (Note that the condition here is weaker than that in the previous section.)

Similarly as in the case of large girth, Theorem C would be best possible up to the error term $o(1)$, provided that a well known conjecture about the existence of $K_{s,s}$-free graphs with many edges is true. Moreover, it is easy to see that it does not make sense to forbid a non-bipartite graph instead of a $K_{s,s}$ since there are bipartite graphs which have no larger complete minors than those guaranteed by the result of Kostochka and Thomason cited above.

Large topological cliques in graphs without a 4-cycle

A *subdivision* TH of a graph H is a graph obtained from H by replacing the edges of H with internally disjoint paths. We say that a graph G contains H as a *topological minor* if G contains a subdivision of H. So H is a minor of G if H is a topological minor of G, but the converse is not necessarily true. Similarly as minors, topological minors can be forced by large average degree: Mader [75] showed that for every $r \in \mathbb{N}$ there exists a smallest number $d(r)$ such that every graph G of average degree larger than $d(r)$ contains a subdivision of K_r. Bollobás and Thomason [20] as well as Komlós and Szemerédi [58] independently proved that there is a constant c such that $d(r) \leq cr^2$. As was first observed by Jung [43], complete bipartite graphs with vertex classes of equal size show that this gives the correct order of magnitude.

Jung's observation implies that in general a minimum degree of order r^2 is needed to force a subdivision of K_{r+1}. However, Mader [79] showed that if we only consider graphs of large girth as host graphs then a minimum degree of r already suffices.

Mader's bound on the girth required is linear in r. In Kühn and Osthus [62] it is shown that in fact the necessary girth does not depend on r. More precisely, every graph of minimum degree r and girth at least 186 contains a subdivided K_{r+1} and a girth of at least 15 will do if $r \geq 435$. This implies the conjecture of Hajós that every graph of chromatic number at least r contains a subdivision of K_r (which is false in general) for graphs of girth at least 186.

Since complete bipartite graphs are triangle-free but they do not contain subdivisions of large cliques, the constant 15 in the result of [62] cannot be replaced by anything less than 5. Mader [80] asked whether a girth of at least 5 is already sufficient to force a subdivision of K_{r+1}. In Chapter 5 we will see that in graphs G of girth at least 5 one can find a subdivision of a clique whose order is at least almost linear in the average degree of G:

Theorem D *Every graph G of girth at least 5 and average degree at least r contains a clique of order at least $r^{1-o(1)}$ as subdivision.*

Moreover, we will show that complete bipartite graphs are in a sense the only counterexamples: if we only consider $K_{s,s}$-free graphs G as host graphs (for fixed s), then these graphs contain significantly larger subdivided cliques than those guaranteed by the average degree of G.

Induced subdivisions in $K_{s,s}$-free graphs

In Chapter 6 we show that if we consider $K_{s,s}$-free graphs as host graphs, then we can require our subdivisions to be induced:

Theorem E *For all natural numbers s and r there exists a number $d = d(s, r)$ such that every $K_{s,s}$-free graph of average degree at least d contains an induced* *subdivision of K_r.*

Of course, an induced subdivision need not exist in general. Indeed, if G is a complete bipartite graph, then it does not even contain an induced subdivision of K_4. Furthermore, forbidding a non-bipartite graph instead of a $K_{s,s}$ makes no sense as then G could be a complete bipartite graph and so the result is best possible in this sense.

Theorem E was motivated by an analogous conjecture of Scott [95] about induced subdivisions in K_s-free graphs of large chromatic number.

Forcing unbalanced complete bipartite minors

Recently, Myers and Thomason [87] determined the asymptotics of the average degree which is necessary to force a given (large) dense graph H as a minor. They also raised the question of what happens for sparse graphs H. In particular, Myers [86] posed the conjecture that for every integer s there exists a positive constant C such that for all integers t every graph of average degree at least Ct contains the complete bipartite graph $K_{s,t}$ as a minor. In Chapter 7 we prove the following strengthened version of this conjecture. (In fact, we prove an even stronger but slightly more technical result.)

Theorem F *For every $\varepsilon > 0$ and every integer s there exists a number $t_0 = t_0(\varepsilon, s)$ such that for all integers $t \geq t_0$ every graph of average degree at least $(1+\varepsilon)t$ contains $K_{s,t}$ as a minor.*

Asymptotically, the bound on the average degree is obviously best possible.

Graphs of high girth and high chromatic number

As mentioned in Section 1.1.1, the existence of graphs which have both high girth and high chromatic number was first demonstrated by Erdős using the probabilistic method. In Chapter 8, we observe that a simple refinement of his argument actually shows that such graphs are quite common – almost all graphs of high girth and suitable density have high chromatic number:

Theorem G *For all $\ell \geq 3$ and $k \in \mathbb{N}$, there are constants C_1 and C_2 such that almost all graphs with n vertices and m edges whose girth is greater than ℓ have chromatic number at least k, provided that $C_1 n \leq m \leq C_2 n^{\ell/(\ell-1)}$ (i.e. the proportion of such graphs tends to 1 as n tends to ∞).*

For $\ell = 3$ the bound on the number of edges is close to best possible: In Osthus, Prömel and Taraz [90] it is proved that almost all triangle-free graphs with n vertices and m edges are in fact bipartite if m is a little larger or smaller. (The precise statement of the latter result is given in the introduction to Chapter 8.) Thus Theorem G leads to a more complete picture of the likely chromatic number of random triangle-free graphs of given density.

Random planar graphs

In Chapter 9 we investigate the properties of random planar graphs and apply counting arguments to give bounds on the number of planar graphs. As indicated in Section 1.1.1, counting arguments, the probabilistic method and the study of random graphs are all closely related. Thus while Chapter 9 does not involve the probabilistic method directly, the type of arguments used here is related to those employed elsewhere. Moreover, triangulations play a prominent role both here and in the final two chapters of this thesis.

The typical properties of random planar graphs (and the related problem of estimating the number of planar graphs) were first investigated by Denise, Vasconcellos and Welsh [27]. In particular, they asked about the likely number of edges of a random planar graph on n vertices. By Euler's theorem, this is of course at most $3n-6$. Gerke and McDiarmid [35] proved that almost all labelled planar graphs on n vertices have at least $13n/7$ edges. We prove the following upper bound:

Theorem H *Almost all labelled planar graphs on n vertices have at most $2.56n$ edges (i.e. the proportion of such graphs tends to 1 as n tends to ∞).*

We also improve bounds of [27] on the number of planar graphs. Very recently, some improvements to our above results were obtained by Bonichon, Gavoille and Hanusse [22].

Our proofs are based on the following result, which states that the number of triangulations of a planar graph is exponential in the number of edges which are needed to triangulate it. The bound is best possible up to the value of the constant ε.

Theorem I *Every labelled planar graph G with n vertices and m edges is contained in at least $\varepsilon 3^{(3n-m)/2}$ labelled triangulations on n vertices, where ε is an absolute constant.*

II. The Regularity lemma

The proofs of the main results in this part are all based on the Regularity lemma.

Extremal connectivity for topological cliques in bipartite graphs

Recall that subdivisions of complete graphs are forced by large average degree. However, the correct asymptotics for the average degree $d(r)$ which is necessary to force a subdivided K_r is not yet known. So far, the best known bounds are

$$(1+o(1))\frac{9r^2}{64} \leq d(r) \leq (1+o(1))\frac{r^2}{2}. \tag{1.1}$$

The upper bound is due to Komlós and Szemerédi [58]. As observed by Łuczak, the lower bound is obtained by considering a random subgraph of a complete bipartite graph with edge probability $3/4$. With high probability the connectivity of these random graphs is about the same as their average degree. Thus, a connectivity of $(1+o(1))9r^2/64$ is necessary to guarantee a subdivided K_r, even if we only consider bipartite graphs as host graphs. The main result of Chapter 10 states that for bipartite graphs this gives the correct asymptotics:

Theorem J *Given $r \in \mathbb{N}$, let $c_{bip}(r)$ be the smallest integer such that every $c_{bip}(r)$-connected bipartite graph contains a subdivided K_r. Then*

$$c_{bip}(r) = (1 + o(1))\frac{9r^2}{64}.$$

Moreover, the proof of this result shows that in the non-bipartite case a connectivity of $(1 + o(1))r^2/4$ suffices to force a subdivided K_r. Thus in general the connectivity which is necessary to guarantee a subdivided K_r lies between $(1 + o(1))9r^2/64$ and $(1 + o(1))r^2/4$. We also improve the constant $1/2$ in the upper bound in (1.1) slightly to $10/23$.

Packings in dense regular graphs

In Chapter 11 we not only seek a single subdivision of a graph H in some graph G, but we want to cover (almost) all of the vertices of G by disjoint subdivided copies of a given graph H. Let us call a collection of disjoint subdivisions of H in G a *TH-packing in G*. (These subdivisions need not necessarily be isomorphic.) Clearly, we cannot always find a TH-packing which covers almost all of the vertices of G, not even if G is dense. Indeed, if G is a large complete bipartite graph whose vertex classes have very different sizes, then for example any TK_4-packing misses a large number of vertices in the larger vertex class of G. However, if G is regular and dense, then we can even cover all but a constant number of vertices of G:

Theorem K *For every graph H without isolated vertices which is not a union of cycles and every positive c there exists a constant $C = C(H, c)$ such that every cn-regular graph G of order n has a TH-packing which covers all but at most C of its vertices.*

In fact, this result remains true if G is 'almost-regular'. Moreover, we prove that for the cases $H = K_4$ and $H = K_5$ one can even cover all vertices of G if the order of G is sufficiently large. For the case when H is a cycle, Alon [3] proved that every r-regular graph G contains a TH-packing which covers all but an ε-fraction of the vertices of G, provided that r is sufficiently large compared with ε and $|H|$. All these results were motivated (and give support to) a conjecture of Verstraëte [106] (see also Alon [3]). We also obtain related results about packings of subgraphs.

Large planar subgraphs in dense graphs

In Chapters 12 and 13 we prove sufficient and essentially necessary conditions in terms of the minimum degree for a graph to contain planar subgraphs with many edges. More precisely, we study the following extremal question: Given a function $m = m(n)$, how large does the minimum degree of a graph G of order n have to be in order to guarantee a planar subgraph with at least $m(n)$ edges?

Since the facial cycles of any planar subgraph of a bipartite graph G have length at least 4, Euler's formula implies that no planar subgraph of G has more than $2n-4$ edges. Thus, as long as the minimum degree is at most $n/2$, we cannot guarantee a planar subgraph with more than $2n-4$ edges. Our first result in Chapter 12 shows that a significantly smaller minimum degree already forces a planar subgraph with 'roughly' $2n$ edges.

Theorem L *For every $0 < \varepsilon < 1$ there exists a constant $c = c(\varepsilon)$ such that every graph G of order n and minimum degree at least $c\sqrt{n}$ contains a planar subgraph with at least $(2-\varepsilon)n$ edges.*

It turns out that the condition on the minimum degree is best possible up to the value of c. Moreover, as long as the minimum degree is $o(n)$, one cannot hope for a planar subgraph with $2n-C$ edges, where C does not depend on n. However, if the minimum degree is linear in n, then a planar subgraph with $2n-C$ edges can be guaranteed:

Theorem M *For every $\gamma > 0$ there is a constant $C = C(\gamma)$ such that every graph G of order n and minimum degree at least γn contains a planar subgraph with at least $2n-C$ edges and such that every graph G of order n and minimum degree at least $(1/2+\gamma)n$ contains a planar subgraph with at least $3n-C$ edges.*

This is best possible in the sense that in both cases the constant C has to depend on γ. Moreover, also in the second part of the statement the additional term γn in the bound on the minimum degree cannot be replaced by a sublinear one.

Spanning triangulations in graphs

In Chapter 13 we answer the question of which minimum degree is needed to force a triangulation of the plane as a spanning subgraph, i.e. a planar subgraph with $3n-6$ edges:

Theorem N *There exists an integer n_0 such that every graph G of order $n \geq n_0$ and minimum degree at least $2n/3$ contains some triangulation of the plane as a spanning subgraph.*

This is best possible: for all integers n there are graphs of order n and minimum degree $\lceil 2n/3 \rceil - 1$ without a spanning triangulation.

All proofs in the final three chapters of this thesis are algorithmic, i.e. the structures guaranteed by the respective results can be found in polynomial time. This implies for instance that the Maximum Planar Subgraph Problem is solvable in polynomial time for graphs of minimum degree at least $2n/3$ (while it is NP-hard in general).

Chapter 2

Basic definitions and tools

In this chapter, we collect some basic and well-known definitions and tools which are needed in several places of this thesis. More specialized notions are introduced in the relevant chapters.

We often omit floors and ceilings if this does not affect the argument. We usually write 'disjoint' instead of 'vertex-disjoint'. We write $e(G)$ for the number of edges of a graph G, $|G|$ for its order, $\delta(G)$ for its minimum degree, $\Delta(G)$ for its maximum degree, $d(G) := 2e(G)/|G|$ for its average degree and $\chi(G)$ for its chromatic number. We denote the degree of a vertex $x \in G$ by $d_G(x)$ or, if this is unambiguous, by $d(x)$ and the set of its neighbours by $N_G(x)$ or $N(x)$. Given a set A of vertices of G, we write $N_G(A)$ for the set of all those neighbours of vertices in A which lie outside A. Given disjoint $A, B \subseteq V(G)$, an *A–B edge* is an edge of G with one endvertex in A and the other in B, the number of these edges is denoted by $e_G(A, B)$ or $e(A, B)$ if this is unambiguous. We write $(A, B)_G$ for the bipartite subgraph of G whose vertex classes are A and B and whose edges are all A–B edges in G. More generally, we often write (A, B) for a bipartite graph with vertex classes A and B. Given graphs G and H we say that G is *H-free* if G does not contain H as a subgraph.

A *subdivision* TH of a graph H is a graph obtained from H by replacing the edges of H with internally disjoint paths. The *branch vertices* of TH are all those vertices that correspond to vertices of H. We say that H is a *topological minor* of a graph G if G contains a subdivision of H as a subgraph.

A graph H is a *minor* of G if H can be obtained from a subgraph of G by contracting edges. Thus, H is a minor of G if for every vertex $h \in H$ there is a connected subgraph G_h of G such that all the G_h are disjoint and G contains a G_h–$G_{h'}$ edge whenever hh' is an edge in H. We also say that H is the *minor of G obtained by contracting the G_h*. (The vertex set of) G_h is called the *branch set* corresponding to h.

We shall frequently use the following easy facts.

Proposition 2.1 *Every graph G with at least one edge contains a subgraph of average degree at least $d(G)$ and minimum degree greater than $d(G)/2$.*

Proposition 2.2 *The vertex set of every graph G can be partitioned into disjoint sets A, B such that the minimum degree of $(A, B)_G$ is at least $\delta(G)/2$.*

In the remainder of this chapter we collect some of the information we need about Szemerédi's Regularity lemma [97] (see [56] for a survey). Let us start with some more notation. The *density* of a bipartite graph $G = (A, B)$ is defined to be

$$d(A, B) := \frac{e(A, B)}{|A||B|}.$$

Given $\varepsilon > 0$, we say that G is *ε-regular* if for all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ we have $|d(A, B) - d(X, Y)| < \varepsilon$. We will often use the following simple fact.

Proposition 2.3 *Given an ε-regular bipartite graph (A, B) of density at least d and a set $X \subseteq A$ with $|X| \geq \varepsilon|A|$, there are less than $\varepsilon|B|$ vertices in B which have at most $(d - \varepsilon)|X|$ neighbours in X.*

Given $d \in [0, 1]$, we say that $G = (A, B)$ is *(ε, d)-super-regular* if all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ satisfy $d(X, Y) > d$ and, furthermore, if $d_G(a) > d|B|$ for all $a \in A$ and $d_G(b) > d|A|$ for all $b \in B$.

Proposition 2.4 *Every ε-regular graph $G = (A, B)$ of density at least d can be made into an $(\varepsilon/(1 - \varepsilon), d - 2\varepsilon)$-super-regular graph by deleting $\varepsilon|A|$ vertices of A and $\varepsilon|B|$ vertices of B.*

Proof. By Proposition 2.3, there are at most $\varepsilon|A|$ vertices in A whose degree is at most $(d - \varepsilon)|B|$; and similarly there are at most $\varepsilon|B|$ vertices in B whose degree is at most $(d - \varepsilon)|A|$. It can be easily checked that the graph obtained from G by deleting these vertices of small degree (as well as possibly some other vertices to make up the required number) is $(\varepsilon/(1 - \varepsilon), d - 2\varepsilon)$-super-regular. □

We will use the following degree form of Szemerédi's Regularity lemma which can be easily derived from the classical version. Proofs of the latter are for example included in [14] and [28].

Lemma 2.5 (Regularity lemma) *For all $\varepsilon > 0$ and all integers k_0 there is an $N = N(\varepsilon, k_0)$ such that for every number $d \in [0, 1]$ and for every graph G there exist a partition of $V(G)$ into $V_0, V_1, \dots, V_k$ and a spanning subgraph G' of G such that the following holds:*

- $k_0 \leq k \leq N$,
- $|V_0| \leq \varepsilon|G|$,
- $|V_1| = \cdots = |V_k| =: L$,
- $d_{G'}(x) > d_G(x) - (d + \varepsilon)|G|$ *for all vertices* $x \in G$,
- *for all* $i \geq 1$ *the graph* $G'[V_i]$ *is empty,*
- *for all* $1 \leq i < j \leq k$ *the graph* $(V_i, V_j)_{G'}$ *is ε-regular and has density either* 0 *or* $> d$.

The sets V_i ($i \geq 1$) are called *clusters*, V_0 is called the *exceptional set*. Given clusters and G' as in Lemma 2.5, the *reduced graph* R is the graph whose vertices are $V_1, \dots, V_k$ and in which V_i is joined to V_j whenever $(V_i, V_j)_{G'}$ is ε-regular and has density $> d$. Thus V_iV_j is an edge of R if and only if G' has an edge between V_i and V_j.

The proof of the next proposition is similar to that of Proposition 2.4.

Proposition 2.6 *Let H be a subgraph of the reduced graph R with $\Delta(H) \leq \Delta$. Then each vertex V_i of H contains a subset V_i' of size $(1 - \varepsilon\Delta)L$ such that for every edge V_iV_j of H the graph $(V_i', V_j')_{G'}$ is $(\varepsilon/(1 - \varepsilon\Delta), d - (1 + \Delta)\varepsilon)$-super-regular.*

Part I

The probabilistic method and counting

Chapter 3

Minors in graphs of large girth

3.1 Introduction

For every $r > 0$ define $p = p(r)$ to be the largest integer such that all graphs G of average degree at least r contain the complete graph K_p on p vertices as a minor. Kostochka [59] and Thomason [99] independently proved that there exists a positive constant c such that

$$p(r) \geq c\frac{r}{\sqrt{\log r}}, \tag{3.1}$$

which improved a bound of Mader. Random graphs show that (3.1) gives the correct order of magnitude. Recently, Thomason [100] showed that $p(r) = (1 + o(1))\gamma r/\sqrt{\log r}$ for an explicit constant γ.

On the other hand, Thomassen [102] observed that if the girth of a graph G is large, then G contains (complete) minors whose minimum degree is much larger than that of G itself. In this chapter we use probabilistic arguments to give more precise asymptotic bounds: for example, we show that every graph G of girth at least five and minimum degree r contains a minor of minimum degree $c_1 r^{3/2}$ and that if the girth of G is at least seven then G contains a minor of minimum degree at least $c_2 r^2$. In both cases the bound on the minimum degree is best possible up to the value of the constant. More generally, the main results of this chapter are as follows:

Theorem 3.1 *Let $k \geq 1$ and $r \geq 3$ be integers and put $g := 4k + 3$. Then every graph G of minimum degree r and girth at least g contains a minor of minimum degree at least $(r-1)^{k+1}/48 = (r-1)^{(g+1)/4}/48$.*

Theorem 3.2 *Let $k \geq 1$ and $r \geq \max\{5k, 2 \cdot 10^6\}$ be integers and put $g := 4k + 1$. Then every graph G of minimum degree at least $4r$ and girth at least g contains a minor of minimum degree at least $r^{k+1/2}/288 = r^{(g+1)/4}/288$.*

In addition to the two cases mentioned above, this is also best possible up to the value of the constant for graphs G of girth 11. In fact, we will see in Section 3.4 that the above results would give the correct order of magnitude (as a function of r) for arbitrary girth g if there exist graphs of minimum degree r

and odd girth g whose order is at most $c(r-1)^{(g-1)/2}$. The minimum order of such graphs is known to lie between $(r-1)^{(g-1)/2}$ and $4(r-1)^{g-2}$, and it has been conjectured (see e.g. Bollobás [12, p. 164]) that the lower bound gives the proper order of magnitude.

An application of (3.1) to the minors obtained in Theorems 3.1 and 3.2 immediately yields the following.

Corollary 3.3 *For all odd integers $g \geq 3$ there exists $c = c(g) > 0$ such that every graph of minimum degree r and girth at least g contains a K_t minor for some*

$$t \geq \frac{cr^{\frac{g+1}{4}}}{\sqrt{\log r}}.$$

□

If the conjecture mentioned above holds, then Corollary 3.3 would be best possible up to the value of the constant c (see Proposition 4.14).

As every graph of chromatic number at least r contains a subgraph of minimum degree at least $r-1$ and every such graph contains a bipartite subgraph of minimum degree at least $(r-1)/2$, Corollary 3.3 in turn implies that Hadwiger's conjecture (that every graph of chromatic number r contains K_r as minor) is true for C_4-free graphs of sufficiently large chromatic number:

Corollary 3.4 *There exists an integer r_0 such that every C_4-free graph of chromatic number $r \geq r_0$ contains a K_r minor.* □

In fact, in Chapter 4 we show that similar results (with weaker bounds) even hold for $K_{s,s}$-free graphs whose minimum degree (respectively chromatic number) is sufficiently large compared with s. In Section 3.2 we also give a simple argument which implies that Hadwiger's conjecture holds for all graphs of girth at least 19 (Corollary 3.9).

At the other extreme, given an integer t, Theorem 3.1 with $r = 3$ shows that every graph G of minimum degree at least three contains a minor of minimum degree at least t if its girth is sufficiently large. This fact was first observed by Thomassen [102], who obtained a bound on the girth linear in t. Diestel and Rempel [29] reduced it to $6\log_2 t + 4$. Theorem 3.1 applied with $r = 3$ and $k = \lceil \log_2 t + 5 \rceil$ shows that the constant 6 can be reduced to 4:

Corollary 3.5 *Let $t \geq 3$ be an integer. Then every graph of minimum degree at least 3 and girth at least $4\log_2 t + 27$ contains a minor of minimum degree at least t. Hence there exists a constant c such that every graph of minimum degree at least 3 and girth at least $4\log_2 t + 2\log_2\log_2 t + c$ contains a K_t minor.* □

(The second part of Corollary 3.5 immediately follows from the first by an application of (3.1).) As already observed in [29], the existence of 3-regular graphs of girth at least g and order at most $c2^{g/2}$ (which is a special case of the conjecture mentioned earlier) would show that Corollary 3.5 is asymptotically best possible in the sense that the constant 4 in the leading terms cannot be reduced any further (see Section 3.4). The minimal order of such 3-regular graphs is known to lie between $c_1 2^{g/2}$ and $c_2 2^{3g/4}$.

Mader [81] proved that for every $\varepsilon > 0$ and every graph H with $\Delta(H) \geq 3$ there exists an integer g such that every graph G of average degree at least $\Delta(H) - 1 + \varepsilon$ and girth at least g contains H as a topological minor. (His bound on g is at least linear in $|H|$ and also depends on ε.) This implies that for every $\varepsilon > 0$ and every integer t there exists an integer g such that every graph of average degree at least $2 + \varepsilon$ and girth at least g contains a minor of minimum degree t. Indeed, first apply the special (and much easier) case $\Delta(H) = 3$ of Mader's result to obtain a 3-regular graph of large girth as a minor and then the observation of Thomassen mentioned before Corollary 3.5 to this minor. In [67] we strengthen Mader's result for the case when H is a large clique: for all $\varepsilon > 0$ every graph of average degree at least $t - 2 + \varepsilon$ and girth at least 1000 contains a topological K_t minor if t is sufficiently large compared with ε. Also, based on techniques of Mader [79], in [62] we proved that for large t every graph of minimum degree at least $t - 1$ and girth at least 15 contains a topological K_t minor. This implies the conjecture of Hajós for all graphs of girth at least 15 and sufficiently large chromatic number. See also Chapter 5 for related results.

This chapter is organized as follows. In Section 3.2 we introduce necessary definitions and collect some tools which we will need later on. We will also apply an idea of Mader to prove Corollary 3.9. In Section 3.3 we then prove Theorems 3.1 and 3.2. In the final section we show that Theorem 3.1 and 3.2 are best possible up to the value of the constant provided that the conjecture mentioned above is true (which is known to be the case for girth $5, 7$ and 11).

3.2 Notation, tools and preliminary observations

The *length* of a cycle C or a path P is the number of its edges. The *girth* of a graph G is the length of its shortest cycle and denoted by $g(G)$. The *distance* between two vertices x, y of a graph G is the length of the shortest path joining x to y and denoted by $d_G(x, y)$. Given $\ell \in \mathbb{N}$, the *ℓ-ball $B^\ell_G(x)$ in G around a vertex x* is the subgraph of G induced by all its vertices of distance at most ℓ from x. If $P = x_1 \dots x_\ell$ is a path and $1 \leq i \leq j \leq \ell$, we write $x_i P x_j$ for its subpath $x_i \dots x_j$.

In the proof of Theorem 3.2 we will need the following Chernoff type bound (see [7, Thm. A.13]).

Lemma 3.6 *Let $X_1, \dots, X_n$ be independent 0-1 random variables with $\mathbb{P}(X_i = 1) = p$ for all $i \leq n$, and let $X := \sum_{i=1}^n X_i$. Then for all $0 < \varepsilon < 1$ we have*

$$\mathbb{P}(X \leq (1 - \varepsilon)\mathbb{E}X) \leq \mathrm{e}^{-\varepsilon^2 \mathbb{E}X/2}.$$

Let us now present a simple proposition which shows that if G is a graph of large girth, then G contains minors whose minimum degree is much larger than that of G itself. Its proof is the same as the beginning of Mader's proof of his main result of [79]. We include it here as it implies a counterpart (Corollary 3.9) to Corollary 3.4 for graphs of small chromatic number. Moreover, it should help to illustrate the basic ideas underlying the proofs of Theorems 3.1 and 3.2, which use a probabilistic version of Mader's argument.

Proposition 3.7 *Let $k \geq 1$ and $r \geq 3$ be integers. Then every graph of girth at least $8k+3$ and minimum degree r contains a minor of minimum degree at least $r(r-1)^k$.*

Proof. Let X be a maximal set of vertices of G that have pairwise distance at least $2k+1$ from each other. Thus for distinct $x, y \in X$ the balls $B^k_G(x)$ and $B^k_G(y)$ are disjoint. Extend the $B^k_G(x)$ $(x \in X)$ to disjoint connected subgraphs of G by first adding each vertex of distance $k+1$ from X to one of the $B^k_G(x)$ to which it is adjacent. Then add each vertex of distance $k+2$ from X to one of the subgraphs constructed in the previous step to which it is adjacent. Continue in this fashion until each vertex of G lies in one of the constructed subgraphs and denote the subgraph obtained from $B^k_G(x)$ in this way by T_x. The choice of X implies that each vertex of G has distance at most $2k$ from X. So each vertex of T_x has distance at most $2k$ from x in T_x. Therefore, as $g(G) \geq 4k+2$, each T_x is an induced subtree of G. In particular $B^k_G(x)$ is a tree in which every vertex that is not a leaf has degree at least r and in which every leaf has distance k from x. So $B^k_G(x)$ (and thus also T_x) has at least $r(r-1)^{k-1}$ leaves. Hence T_x sends at least $r(r-1)^k$ edges to vertices outside T_x. As $g(G) \geq 8k+3$, at most one edge of G joins T_x to a given other tree T_y $(y \in X \setminus \{x\})$. Thus the graph obtained from G by contracting the trees T_x $(x \in X)$ has minimum degree at least $r(r-1)^k$, as required. □

An application of the bound (3.1) of Kostochka and Thomason to the minor obtained in Proposition 3.7 for $k=1$ shows that for sufficiently large r every graph G of minimum degree r and girth at least 11 contains a K_{r+1} minor. For small r, we will apply the following result of Mader (see [76] or [12, Ch. VII.1]).

Theorem 3.8 *For all integers $t \geq 4$ every graph of average degree $> 16(t-2)\log_2(t-2)$ contains a K_t minor. Moreover, every graph of average degree > 10 contains a K_7 minor.*

As above, combining this with Proposition 3.7 leads to the observation that Hadwiger's conjecture is true for all graphs of girth at least 19:

Corollary 3.9 *Every graph of girth at least 19 and minimum degree r contains a K_{r+1} minor. In particular, every graph of girth at least 19 and chromatic number r contains a K_r minor.* □

3.3 Proof of Theorems 3.1 and 3.2

In the proof of Proposition 3.7 we covered the entire vertex set of our graph G with suitable disjoint rooted trees T_x and considered the minor M obtained by contracting these trees. Amongst other properties, these trees had radius between k and $2k$. If we could choose them all of radius at most k while still maintaining sufficiently many edges between the trees, this would reduce the bound on the girth from $8k+3$ to $4k+3$. We will achieve this in the proof of Theorem 3.1 by choosing the roots of the trees at random, albeit at the expense that there will be a small number of vertices which do not lie in any of the trees. The case when $k=1$ (i.e. when the trees are stars) of the first part of the proof is similar to an argument of Alon which shows the existence of small dominating sets (i.e. sets of vertices to which every vertex has distance at most one) in graphs of large minimum degree (see [7, Thm. 2.2]).

Proof of Theorem 3.1. We may assume that $(r-1)^k \geq 48$. Consider a random subset X of $V(G)$ which is obtained by including each vertex in X with probability $p := 4/(r-1)^k$ independently of all other vertices. The branch sets of our minor will be trees of radius at most k whose roots are the elements of X. As $g(G) \geq 2k+2$ and $\delta(G)=r$, for each vertex $x \in G$ the graph $B^k_G(x)$ is a tree with at least $(r-1)^k$ leaves. Call an edge $e=xy$ of G *bad* if $d(x,X)>k$ or $d(y,X)>k$. Then

$$\begin{aligned}\mathbb{P}(xy \text{ is bad}) &\leq \mathbb{P}(B^k_G(x)\cap X=\emptyset)+\mathbb{P}(B^k_G(y)\cap X=\emptyset)\\ &= (1-p)^{|B^k_G(x)|}+(1-p)^{|B^k_G(y)|}\leq 2(1-p)^{(r-1)^k}\\ &\leq 2\mathrm{e}^{-p(r-1)^k}=2/\mathrm{e}^4,\end{aligned}$$

and so

$$\mathbb{E}(\text{number of bad edges}) \leq 2e(G)/\mathrm{e}^4.$$

Markov's inequality now implies

$$\mathbb{P}(> e(G)/9 \text{ edges are bad}) \leq 18/\mathrm{e}^4 \leq 1/3.$$

Moreover, the expected size of X is $p|G|$, and so again, by Markov's inequality,

$$\mathbb{P}(|X|>2p|G|)\leq 1/2.$$

Thus with probability at least $1-1/2-1/3>0$ there is an outcome X with $|X|\leq 2p|G|$ and so that at most $e(G)/9$ edges of G are bad.

Extend the vertices in X to disjoint connected subgraphs G_x $(x\in X)$ of G with $x\in G_x$ by first adding each vertex of distance one from X to a vertex in X to which it is adjacent, then adding each vertex of distance two from X to one of the subgraphs constructed in the previous step to which it is now adjacent etc. Continue in this fashion until each vertex of G of distance at most k from X is contained in one of the graphs G_x thus obtained. Then each vertex of G_x has distance at most k from x. As $g(G)\geq 2k+2$, every G_x is an induced subtree of G. So each edge of G that is not bad and does not lie in $\bigcup_{x\in X}E(G_x)$ joins distinct G_x. Moreover, since $g(G)\geq 4k+3$, there is at most one edge of G joining a given pair of graphs G_x. Thus for the minor M of G whose branch sets are the G_x we have

$$\begin{aligned}d(M) &\geq \frac{2(e(G)-|\bigcup_{x\in X}E(G_x)|-e(G)/9)}{|X|}\geq\frac{16e(G)-18|G|}{9\cdot 2p|G|}\\ &\geq \frac{4r-9}{9p}\geq\frac{4(r-1)}{9p}\cdot\frac{3}{8}\geq\frac{(r-1)^{k+1}}{24}.\end{aligned}$$

(The fourth inequality holds since $r \geq 3$.) By Proposition 2.1 the graph M contains a subgraph of minimum degree at least $(r-1)^{k+1}/48$, as desired.

□

A result of Györi [37] states that every C_6-free bipartite graph can be made into a graph of girth at least 7 by deleting at most half of its edges. This implies that the assertion of the $g = 7$ case of Theorem 3.1 remains true for C_6-free graphs (with a modified constant).

To prove Theorem 3.2, we will again cover a large part of our graph G with disjoint trees of radius k whose roots are chosen at random (similarly as in the proof of Theorem 3.1). However, this time the girth is not large enough to ensure that between every pair of these trees there is at most one edge. To deal with such multiple edges we choose the trees more carefully and prove that firstly there are still many (good) edges joining leaves of distinct trees and secondly that only a small fraction of these edges is redundant in the sense that there are many additional (good) edges joining the same pair of trees.

Proof of Theorem 3.2. First apply Proposition 2.2 to obtain a bipartite subgraph $G_1 = (A, B)_G$ of G of minimum degree at least $2r$. We may assume that $|A| \geq |B|$. Delete edges if necessary to obtain a bipartite subgraph G_2 of G_1 in which the degree of every vertex in A is precisely $2r$. Thus $d(G_2) \geq 2r$. Now apply Proposition 2.1 to obtain a subgraph $H = (C, D)_{G_2}$ of G_2 of minimum degree at least $r + 1$ and average degree at least $2r$ and where every vertex in C has degree at most $2r$.

We now assign orientations to the edges of H as follows. For every vertex $x \in H$ choose any $r + 1$ of its neighbours in H and orient the edges between x and these neighbours from x towards these. We thus obtain a graph $\vec{H}$ in which every edge has either none, one or two orientations and in which the outdegree of every vertex is precisely $r + 1$. We say that a path $x_0 \dots x_\ell$ in $\vec{H}$ is *directed* from x_0 to x_ℓ if each edge $x_i x_{i+1}$ is oriented from x_i towards x_{i+1}. So $x_i x_{i+1}$ may additionally be oriented from x_{i+1} to x_i. Given two vertices x and y of $\vec{H}$ we write $\vec{d}(x, y)$ for the length of the shortest directed path from x to y (and set $\vec{d}(x, y) := \infty$ if such a path does not exist). Given $\ell \in \mathbb{N}$, we write $\vec{S}^\ell(x)$ for the set of all those vertices $y \in \vec{H}$ with $\vec{d}(x, y) = \ell$. We define $\vec{B}^\ell(x)$ to be the subgraph of $\vec{H}$ which consists of all directed paths of length at most ℓ starting at x. Note that if $\ell < 2k$, then, as $g(H) \geq 4k$, the graph $\vec{B}^\ell(x)$ is an induced subtree of $\vec{H}$ with root x in which every edge is oriented away from the root (and possibly also towards it). As the outdegree of every vertex of $\vec{H}$ is $r + 1$, every vertex of $\vec{B}^\ell(x)$ which is not a leaf has either $r + 1$ or $r + 2$ neighbours in $\vec{B}^\ell(x)$ and every leaf has distance precisely ℓ from x. In particular,

$$r^\ell \leq |\vec{S}^\ell(x)| \leq (r + 1)^\ell.$$

Consider a random subset X of $V(H)$ which is obtained by including each vertex of H in X with probability

$$p := \frac{1}{4(r + 1)^{k-1/2}}$$

independently of all other vertices. For some suitable outcome X, the branch sets of the desired minor in G will be subtrees of H of radius k and with roots in X. Call a vertex $v \in H$ *good* if it satisfies the following three conditions.

(i) $|\vec{S}^k(v) \cap X| \geq \sqrt{r}/6$.

(ii) $\vec{B}^{k-1}(v) \cap X = \emptyset$.

(iii) Each component of $\vec{B}^k(v) - v$ contains at most one vertex of $\vec{S}^k(v) \cap X$.

We will now show that the probability that a given vertex v is good is quite large. First note that, as $r \geq 5k$,

$$\mathbb{E}(|\vec{S}^k(v) \cap X|) = p|\vec{S}^k(v)| \geq pr^k = \frac{\sqrt{r}}{4} \cdot \left(\frac{r}{r+1}\right)^{k-1/2}$$
$$\geq \frac{\sqrt{r}}{4}\left(1 - \frac{1}{r}\right)^k \geq \frac{\sqrt{r}}{4}\left(1 - \frac{k}{r}\right) \geq \frac{\sqrt{r}}{5}.$$

Since $r \geq 2 \cdot 10^6$, Lemma 3.6 with $\varepsilon := 1/6$ implies that

$$\mathbb{P}(|\vec{S}^k(v) \cap X| \leq \sqrt{r}/6) \leq 1/25. \tag{3.2}$$

Secondly,

$$\mathbb{E}(|\vec{B}^{k-1}(v) \cap X|) = p|\vec{B}^{k-1}(v)| \leq 2p|\vec{S}^{k-1}(v)| \leq \frac{1}{2\sqrt{r+1}},$$

and hence, as $r \geq 625$, Markov's inequality implies

$$\mathbb{P}(|\vec{B}^{k-1}(v) \cap X| \geq 1) \leq 1/50. \tag{3.3}$$

Finally, given a component L of $\vec{B}^k(v) - v$, let $S(v, L) := L \cap \vec{S}^k(v)$. Writing $\sum_{x,y}$ for the sum over all unordered pairs $x \neq y$ of vertices in $S(v, L)$, we have

$$\mathbb{P}(|S(v, L) \cap X| \geq 2) \leq \sum_{x,y} \mathbb{P}(x, y \in X) \leq \binom{|S(v, L)|}{2} p^2$$
$$\leq \frac{((r+1)^{k-1}p)^2}{2} = \frac{1}{32(r+1)}.$$

As the outdegree of v is $r+1$ and so $\vec{B}^k(v) - v$ has precisely $r+1$ components, it follows that

$$\mathbb{E}(\text{number of components } L \text{ of } \vec{B}^k(v) - v \text{ for which } |S(v, L) \cap X| \geq 2) \leq 1/32.$$

Hence Markov's inequality implies that

$$\mathbb{P}(\text{there is a component } L \text{ of } \vec{B}^k(v) - v \text{ for which } |S(v, L) \cap X| \geq 2) \leq 1/32. \tag{3.4}$$

From (3.2), (3.3) and (3.4) together it now follows that

$$\mathbb{P}(v \text{ is not good}) \leq 1/10.$$

Call an edge e of H *good* if both of its endvertices are good. Thus

$$\mathbb{P}(e \text{ is not good}) \le 1/5,$$

and therefore

$$\mathbb{E}(\text{number of edges of } H \text{ which are not good}) \le e(H)/5.$$

Hence Markov's inequality shows that

$$\mathbb{P}(\ge e(H)/2 \text{ edges of } H \text{ are not good}) \le 2/5. \tag{3.5}$$

Moreover, $\mathbb{E}(|X|) = p|H|$, and so Markov's inequality implies that

$$\mathbb{P}(|X| \ge 2p|H|) \le 1/2. \tag{3.6}$$

Now (3.5) and (3.6) show that with probability at least $1-2/5-1/2>0$ there is an outcome X with $|X| \le 2p|H|$ and such that at least half of the edges of H are good. Let U be the set of all good vertices of H. We say that a vertex $x \in X$ *belongs to* a vertex $u \in U$ if $\vec{d}(u,x) = k$. So condition (i) in the definition of a good vertex implies that at least $\sqrt{r}/6$ vertices in X belong to u. As $g(H) > 2k$, there exists precisely one directed path P_{ux} of length k from u to a vertex x belonging to u. Given $x \in X$, let U_x denote the set of all the good vertices to which x belongs, and let H_x be the union of all paths P_{ux} over all $u \in U_x$. If $U_x = \emptyset$, we put $H_x := x$. As $g(H) \ge 2k+2$, each H_x is an induced subtree of H and U_x is the set of its leaves.

Let us now prove the following claim.

If $x, y \in X$ are distinct, x belongs to $u \in U$, y belongs to $u' \in U$ and $u \ne u'$ then P_{ux} and $P_{u'y}$ are disjoint. $(*)$

Suppose not and let z be the first vertex on P_{ux} that is contained in $P_{u'y}$. Note that $|uP_{ux}z| \ne |u'P_{u'y}z|$ would imply the existence of either a directed u'–x path or a directed u–y path of length $< k$. Hence $|uP_{ux}z| = |u'P_{u'y}z|$ (and thus in particular $z \ne u$), as both u and u' are good vertices (cf. condition (ii)). So if L is the component of $\vec{B}^k(u) - u$ containing z, then both x and y lie in $L \cap \vec{S}^k(u) \cap X$, contradicting condition (iii) for u.

For every $u \in U$ choose a vertex $x_u \in X$ which belongs to u uniformly at random independently of the other elements of U. For every $x \in X$ we then define T_x to be the subtree of H_x consisting of the paths P_{ux} for all those $u \in U_x$ with $x_u = x$. If there are no such paths we set $T_x := x$. So every choice of the x_u ($u \in U$) yields a family T_x ($x \in X$) of trees. Note that $(*)$ implies that T_x and T_y are disjoint whenever $x \ne y$. We will show that with non-zero probability the x_u will have the property that the minor M of $H \subseteq G$ whose branch sets are the T_x ($x \in X$) thus defined has large average degree. To do this, we will show that with non-zero probability there are only a few pairs T_x, T_y such that H contains many good T_x–T_y edges. Then a large fraction of the good edges of H will join different pairs T_x, T_y and thus will correspond injectively to edges

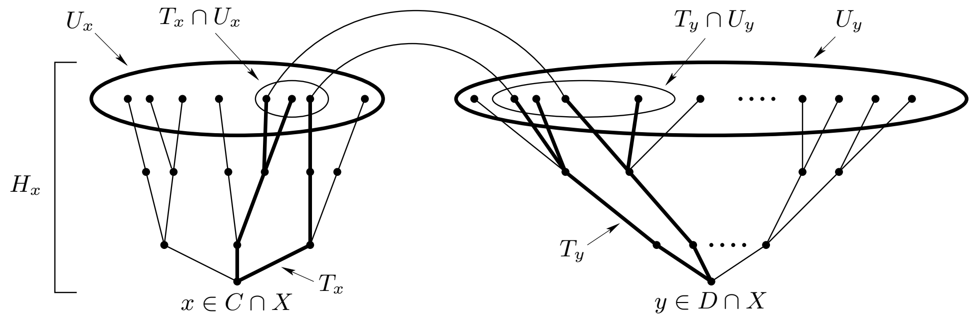


Figure 3.1: Illustrating a cycle of length 14 in the final part of the proof of Theorem 3.2 for $k = 3$ and $g = 13$.

of M. As X is relatively small, this will imply that M has large average degree.

Suppose that $x, y \in X$ are given, and let us first estimate the expected number of good edges of H joining T_x to T_y. Recall that by definition, every good T_x–T_y edge joins $T_x \cap U_x$ to $T_y \cap U_y$. As $g(G) \geq 4k+1$, for every component L of $H_x - x$ there is at most one edge in H joining L to H_y. Similarly, for every component L of $H_y - y$ there is at most one edge in H joining L to H_x. So in particular the U_x–U_y edges in H are independent and their number is at most $\min\{d_{H_x}(x), d_{H_y}(y)\} \leq \min\{d_H(x), d_H(y)\}$. But as $g(H) \geq 2k$, every vertex in U_x has distance precisely k from x in $H = (C, D)_{G_2}$. Thus either $U_x \subseteq C$ or $U_x \subseteq D$, and the same is true for U_y. So if H contains a U_x–U_y edge, then one of U_x, U_y must be contained in C while the other one is contained in D. Hence one of x, y lies in C. As every vertex in C has degree at most $2r$ in H, it follows that H contains at most $2r$ edges joining U_x to U_y.

Consider a U_x–U_y edge u_1u_2 with $u_1 \in U_x$ and $u_2 \in U_y$. Then $u_1 \notin U_y$ and $u_2 \notin U_x$, since $g(H) \geq 2k+2$. So the probability that u_1u_2 is a T_x–T_y edge equals the probability that $x_{u_1} = x$ and $x_{u_2} = y$ which in turn is the inverse of the product of the number of vertices in X belonging to u_1 with the number of vertices in X belonging to u_2; so by (i) this probability is at most $(6/\sqrt{r})^2$. Hence

$$\begin{aligned}\mathbb{E}(\text{number of good } T_x\text{–}T_y \text{ edges in } H) &\leq e_H(U_x, U_y) \cdot \left(\frac{6}{\sqrt{r}}\right)^2 \\ &\leq 2r \cdot \frac{36}{r} = 72.\end{aligned}$$

So Markov's inequality implies that

$$\mathbb{P}(H \text{ contains at least 144 good } T_x\text{–}T_y \text{ edges}) \leq 1/2. \tag{3.7}$$

Given a good edge u_1u_2, call it *overloaded* if there are at least 144 good edges of H which are distinct from u_1u_2 and join $T_{x_{u_1}}$ to $T_{x_{u_2}}$. For $i = 1, 2$ let X_i be the set of all vertices in X belonging to u_i. (So $X_1 \cap X_2 = \emptyset$.) For all $x \in X_1$, $y \in X_2$ let $\mathcal{A}_{xy}$ be the event that there are at least 144 good T_x–T_y edges which are distinct from u_1u_2. As the U_x–U_y edges of H are independent, and thus the event that $x_{u_1} = x$ and $x_{u_2} = y$ is independent from $\mathcal{A}_{xy}$, we have

$$\begin{aligned}\mathbb{P}(u_1u_2 \text{ is overloaded}) &\le \sum_{x\in X_1,\ y\in X_2} \mathbb{P}(x_{u_1} = x,\ x_{u_2} = y \text{ and } \mathcal{A}_{xy} \text{ is true})\\ &\overset{(3.7)}{\le} \sum_{x\in X_1,\ y\in X_2} \frac{1}{|X_1|}\cdot\frac{1}{|X_2|}\cdot\frac{1}{2} = \frac{1}{2}.\end{aligned}$$

Thus

$$\begin{aligned}\mathbb{E}(\text{number of overloaded edges}) &= \sum_{u_1u_2\in E(H[U])} \mathbb{P}(u_1u_2 \text{ is overloaded})\\ &\le e_H(U,U)/2.\end{aligned}$$

But this means that for all $u \in U$ the vertices x_u can be chosen in such a way that for the trees T_x ($x \in X$) thus defined at most half of the good edges of H are overloaded. Let F be the subgraph of H which consists of all those good edges that are not overloaded. Thus

$$e(F) \ge e_H(U,U)/2 \ge e(H)/4 \ge r|H|/4.$$

Consider the minor M of H whose branch sets are the T_x and let $e = u_1u_2$ be an edge of F. Recall that as $g(H) \ge 2k+2$, the endpoints of e must lie in distinct T_x, i.e. $x_{u_1} \ne x_{u_2}$. As e is not overloaded, there are less than 144 other edges of F joining $T_{x_{u_1}}$ to $T_{x_{u_2}}$. Thus to each edge of M there correspond at most 144 edges of F, i.e. $e(M) \ge e(F)/144$. Hence

$$d(M) \ge \frac{2e(F)}{144|X|} \ge \frac{r|H|}{144\cdot 2\cdot 2p|H|} \ge \frac{r^{k+1/2}}{144}.$$

Proposition 2.1 implies that M contains a subgraph of minimum degree at least $r^{k+1/2}/288$, as desired. $\square$

We remark that the constants in Theorems 3.1 and 3.2 could be improved a little by more careful calculations. Furthermore, the proof of the case $k = 1$ (i.e. $g = 5$) of Theorem 3.2 can easily be modified to give the following.

Theorem 3.10 *There exists a constant $c > 0$ such that for all integers $t \ge 2$ every $K_{2,t}$-free graph G of minimum degree d contains a minor of minimum degree at least $cd^{3/2}/t$.*

Proof. By choosing c sufficiently small, we may assume that $r := \lfloor d/4\rfloor \ge 2\cdot 10^6$. It then suffices to make the following minor changes in the proof of the case $k = 1$ of Theorem 3.2. Define H, p, X, U, H_x and T_x as before. For every vertex $x \in X$ there are now less than t edges (instead of at most one) joining a given leaf of the star H_x to leaves of a given other star H_y ($y \in X$). So H contains at most $2rt$ edges joining U_x to U_y. Similarly as before, this shows that with probability at most $1/2$ the graph H contains at least $144t$ good T_x–T_y edges. This time we call a good edge u_1u_2 overloaded if there are at least $144t$

good edges of H which are *disjoint* from u_1u_2 and join $T_{x_{u_1}}$ to $T_{x_{u_2}}$. Again, it follows that for all $u \in U$ the vertices x_u can be chosen so that at most half of the good edges are overloaded. But for each good edge u_1u_2 which is not overloaded there are at most $144t + 2t$ other good edges joining $T_{x_{u_1}}$ to $T_{x_{u_2}}$ (as there are at most t edges joining u_1 to leaves of $T_{x_{u_2}}$ and vice versa). Thus the minor M of H whose branch sets are the T_x has average degree at least $r^{3/2}/146t$. By Proposition 2.1, M contains a subgraph of minimum degree at least $r^{3/2}/292t$, as desired. □

More generally, in Chapter 4 we prove that for all $t \geq s \geq 2$ every $K_{s,t}$-free graph of minimum degree at least r contains a graph of minimum degree $r^{1+\frac{1}{2(s-1)}-o(1)}$ as minor. This implies that for sufficiently large r every $2r$-connected $K_{s,t}$-free graph is r-linked (see Chapter 4).

3.4 Upper bounds

The following simple proposition (which generalizes [29, Prop. 2.2]) shows that the existence of small graphs of large girth can be used to prove upper bounds on the minimum degree of minors in graphs of large girth.

Proposition 3.11 *Let $c, \ell > 0$ and let d, r be integers such that $r \geq 2$. Suppose that G is a graph of maximum degree at most cr and order at most $c(r-1)^\ell$ which contains a minor of minimum degree d. Then $d < 2c(r-1)^{(\ell+1)/2}$.*

Proof. Suppose that H is a minor of G of minimum degree d. Let $W \subseteq V(G)$ be a branch set corresponding to a vertex of H. As each vertex of W sends at most cr edges to other branch sets, $|W| \geq d/cr$. Hence

$$c(r-1)^\ell \geq |G| \geq \frac{d|H|}{cr} > \frac{d^2}{cr} \geq \frac{d^2}{2c(r-1)}.$$

This shows that $d < 2c(r-1)^{(\ell+1)/2}$, as required. □

We will now use Proposition 3.11 to observe that the truth of the following well-known conjecture (see e.g. Bollobás [12, p. 164]) would show that for fixed girth Theorems 3.1 and 3.2 are best possible up to the value of the constant and also that the constant 4 in Corollary 3.5 cannot be replaced by a smaller one.

Conjecture 3.12 *There exists a constant c such that for all integers $r, g \geq 3$ there is a graph of minimum degree at least r and girth at least g whose order is at most $c(r-1)^{\lfloor \frac{g-1}{2} \rfloor}$.*

An observation of Tutte (see [12, Ch. III, Thm. 1.2]) shows that this would be close to best possible: Consider any vertex x in a graph G of minimum degree at least r and girth at least g. Then the graph obtained from the $\lfloor \frac{g-1}{2} \rfloor$-ball around x by deleting any edges between vertices of distance $\lfloor \frac{g-1}{2} \rfloor$ from x is a tree. Since $\delta(G) \geq r$, this tree (and so also G) has at least $(r-1)^{\lfloor \frac{g-1}{2} \rfloor}$

vertices. This argument also shows that any graph G demonstrating the truth of Conjecture 3.12 must have maximum degree at most cr. (Indeed, take for x a vertex of maximum degree in G.) Thus by Proposition 3.11 with $\ell := \lfloor \frac{g-1}{2} \rfloor$, such a graph G has no minor of minimum degree at least $2c(r-1)^{\frac{1}{2}\lfloor \frac{g+1}{2} \rfloor}$. In other words, the truth of Conjecture 3.12 would imply that Theorem 3.1 is best possible up to the value of the constant, and so is Theorem 3.2 if the girth g is fixed. It would also imply that Theorems 3.1 and 3.2 give the correct order of magnitude even for graphs of fixed even girth. Furthermore, as we will see in Chapter 4 (Proposition 4.14), if Conjecture 3.12 holds then Corollary 3.3 would also be best possible up to the value of the constant c (for fixed g).

There are several constructions which show that for infinitely many values of r there are graphs of girth at least 5 and minimum degree r whose order is at most $3(r-1)^2$ (see e.g. Brown [23, Thm. 3.4(b)] or the proof of [15, Thm. 1.3.3]). For $g = 7, 11$ Benson [10] showed that for infinitely many integers r there are graphs of minimum degree r and girth at least g whose order is at most $3(r-1)^{\frac{g-1}{2}}$. Together with the above this implies the following

Proposition 3.13 *For $g = 5, 7$ and 11 there are infinitely many integers r for which there exists a graph of minimum degree r and girth at least g that does not contain a minor of minimum degree at least $6(r-1)^{\frac{g+1}{4}}$.* □

The best known general upper bound for the minimal order of graph of minimum degree at least r and girth at least g was proved by Sauer. It implies that for $r \geq 3$ and odd $g \geq 3$ the minimal order of such graphs is at most $4(r-1)^{g-2}$ (see [12, Ch. III, Thm. 1.4]).

Turning to the case $r = 3$, Weiss [107] proved that a construction of Biggs and Hoare [11] yields infinitely many integers g for which there are 3-regular graphs of girth g and order at most $c2^{3g/4}$. Together with Proposition 3.11 this implies that for infinitely many integers t there are 3-regular graphs of girth at least $\frac{8}{3}\log t - c'$ that have no minor of minimum degree t (as was already observed by Diestel and Rempel [29]). In particular, the constant 4 in Corollary 3.5 cannot be replaced by a number smaller than 8/3. Again, the constant 4 in Corollary 3.5 would be best possible if Conjecture 3.12 holds for $r = 3$.

Chapter 4

Minors in $K_{s,s}$-free graphs

4.1 Introduction

In Chapter 3 we have seen that if the girth of a graph G is sufficiently large then G has a K_p minor where p is asymptotically much larger than the average degree of G. Here we prove that the same is true if G is locally sparse in the much weaker sense that it does not contain a fixed complete bipartite graph $K_{s,s}$ as a subgraph:

Theorem 4.1 *For every integer $s \geq 2$ there exists an r_s such that every $K_{s,s}$-free graph of average degree at least $r \geq r_s$ contains a K_p minor for all*

$$p \leq \frac{r^{1+\frac{1}{2(s-1)}}}{(\log r)^3}.$$

As every graph of chromatic number k contains a subgraph of minimum degree at least $k-1$, this implies Hadwiger's conjecture for $K_{s,s}$-free graphs of sufficiently large chromatic number:

Corollary 4.2 *For every integer $s \geq 2$ there exists an integer k_s such that every $K_{s,s}$-free graph of chromatic number $k \geq k_s$ contains a K_k minor.* □

In Section 4.3 we will see that there exists an absolute constant α so that we can take $k_s := s^{\alpha s}$ in Corollary 4.2.

A probabilistic argument (Proposition 4.13) shows that the bound on p in Theorem 4.1 is best possible up to the logarithmic term, provided that there exist $K_{s,s}$-free graphs G with at least $c_s|G|^{2-1/s}$ edges. These are known to exist for $s = 2, 3$ and have been conjectured to exist also in general (see e.g. [12, p. 362] or [25, p.36]). More precisely, the above conjecture would imply that the exponent of the logarithmic term has to be at least 1/2. We believe that this is the correct order of magnitude.

We now turn to an application of Theorem 4.1 to highly connected graphs. Improving an earlier bound of Bollobás and Thomason [19], Thomas and Wollan [98] proved that every $16k$-connected graph is k-linked. As is well known and easy to see, the graph obtained from K_{3k-1} by deleting k independent edges shows that the function $16k$ cannot be replaced by anything smaller than $3k-2$. On the other hand, a result in [19] states that if a graph G is $2k$-connected and contains a minor H with $2\delta(H) \geq |H| + 4k - 2$ then G is k-linked. Together with Theorem 4.1 this immediately implies the following.

Corollary 4.3 *For every integer $s \geq 2$ there exists an integer k_s such that for all $k \geq k_s$ every $2k$-connected $K_{s,s}$-free graph is k-linked.* □

Mader [79, Cor. 1] showed that for $k \geq 2$ one cannot replace $2k$ by $2k-1$.

Note that Theorem 4.1 is far from being true if we forbid a non-bipartite graph H instead of a $K_{s,s}$. Indeed, recall that there are graphs of average degree r containing no complete graph of order at least $c'r/\sqrt{\log r}$ as minor. These graphs can be made bipartite (and thus H-free) by deleting at most half of their edges. In particular, the resulting graphs G contain no complete graph as minor whose order exceeds the average degree of G. However, replacing average degree with chromatic number might help:

Problem 4.4 *Given an integer $s \geq 3$, does there exist a function $\omega_s(k)$ tending to infinity such that every K_s-free graph of chromatic number k contains a K_p minor for all $p \leq k \cdot \omega_s(k)$?*

In other words, the question is whether for K_s-free graphs of sufficiently large chromatic number Hadwiger's conjecture is true with room to spare. For a survey on Hadwiger's conjecture and related questions see e.g. [40].

This chapter is organized as follows. In Section 4.2 we introduce some notation and state several results which we will need later on. Theorem 4.1 is then proved in Section 4.3. The methods are related to those in Chapter 3. The final section is concerned with upper bounds for the size of the complete minor in Theorem 4.1.

4.2 Notation and tools

All logarithms in this chapter are base e, where e denotes the Euler number. We will use the following Chernoff bound (see e.g. [39, Cor. 2.3]).

Lemma 4.5 *Let $X_1, \dots, X_n$ be independent 0-1 random variables with $\mathbb{P}(X_i = 1) = p$ for all $i \leq n$, and let $X := \sum_{i=1}^n X_i$. Then*

$$\mathbb{P}(X \leq \mathbb{E}X/2 \text{ or } X \geq 2\mathbb{E}X) \leq 2\mathrm{e}^{-\mathbb{E}X/12}.$$

A proof of the next lemma can be found in [14, Ch. IV, Lemma 9].

Lemma 4.6 *Let (A, B) be a bipartite graph that does not contain a $K_{s,t}$ with t vertices in A and s vertices in B. Suppose that on average each vertex in A has d neighbours in B. Then*

$$|A|\binom{d}{s} \leq t\binom{|B|}{s}.$$

Lemma 4.6 can be used to prove the following upper bound on the number of edges of a $K_{s,t}$-free graph (see e.g. [12, Ch. VI, Thm. 2.3]).

Theorem 4.7 *Let $t \geq s \geq 2$ be integers. Then every $K_{s,t}$-free graph G has at most $t|G|^{2-1/s}$ edges and thus satisfies*

$$|G| \geq \left(\frac{\delta(G)}{2t}\right)^{1+\frac{1}{s-1}}. \tag{4.1}$$

Finally, we will need the following consequence of Corollary 6.16.

Lemma 4.8 *Let ℓ, t be integers with $\ell \geq 8t$. Let $G = (A, B)$ be a $K_{t,t}$-free bipartite graph such that $|A| \geq \ell^{12t}|B|$ and $d_G(a) = \ell$ for every vertex $a \in A$. Then G contains a subdivision of some graph of average degree at least $\ell^9/2^{14}$.*

4.3 Dense Minors in $K_{s,t}$-free graphs

Instead of proving Theorem 4.1, we will prove the following slightly more general result on the existence of dense minors in $K_{s,t}$-free graphs.

Theorem 4.9 *For all integers $t \geq s \geq 2$ and all $r \geq (100t)^{16s}$ every $K_{s,t}$-free graph G of average degree r contains a minor of average degree at least*

$$d := \frac{r^{1+\frac{1}{2(s-1)}}}{10^9 t^4 (\log r)^{2+\frac{1}{s+1}}}. \tag{4.2}$$

Note that asymptotically the restriction on the range of r is not too severe: if $r \leq t^s$, then (4.2) is already smaller than the trivial lower bound of r on the average degree of the densest minor of G.

Proof of Theorem 4.1. Theorem 4.1 immediately follows by an application of (3.1) to the minor obtained from the $s = t$ case of Theorem 4.9. $\square$

Furthermore, Theorem 4.9 shows that there exists an absolute constant α so that we can take $k_s := s^{\alpha s}$ in Corollary 4.2. (Indeed, given a $K_{s,s}$-free graph G of chromatic number $r + 1$, apply Theorem 4.9 to a subgraph H of G of minimum degree at least r. If $r \geq s^{\alpha s}$ where α is sufficiently large compared with the constant c appearing in (3.1), then this shows that H contains a K_{r+1} minor, since then the value d in (4.2) satisfies $cd/\sqrt{\log d} \geq r + 1$.)

Our aim in the proof of Theorem 4.9 is to find disjoint stars in G such that a large fraction of the edges of G joins two distinct stars. If the number of these stars is not too large and if only a few edges join the same pair of stars, then the minor of G obtained by contracting the stars (and deleting all other vertices) has large average degree, as desired. We will find such stars by first choosing the set X of their centres at random and then assigning vertices $v \in G$ with distance one to X to one of the centres adjacent to v in a suitable way. For this to work we need that G is 'almost regular'. The following lemma allows us to assume this at the expense of only a small loss of the average degree.

Lemma 4.10 *For all integers $t \geq 2$ and all $r \geq 10^9 t^4$ every $K_{t,t}$-free graph G of average degree at least r either contains a subdivision of some graph of average degree at least r^3 or a bipartite subgraph H such that $\delta(H) \geq \frac{r}{400t \log r}$ and $\Delta(H) \leq r$.*

Proof. Apply Propositions 2.1 and 2.2 to obtain a bipartite subgraph G' of G of minimum degree at least $d := \lceil r/4 \rceil$. Let A be the larger vertex class of G' and delete edges if necessary to obtain a (bipartite) subgraph G'' with $d_{G''}(a) = d$ for all $a \in A$. Let B be the set of all vertices in $G'' - A$ that are not isolated and put $G^* := (A, B)_{G''}$. So $d_{G^*}(a) = d$ for all vertices $a \in A$ and thus $d(G^*) \geq d$ (since $|A| \geq |B|$). Put $N := \lceil 1 + (6t + 1) \log d \rceil$ and note that

$$\frac{d}{N} \geq \frac{d}{8t \log d} \geq 10^5 t^2. \tag{4.3}$$

Partition B into N disjoint sets $B_1, \dots, B_N$ such that

$$\begin{aligned} \mathrm{e}^{i-1} &\leq d_{G^*}(x) < \mathrm{e}^i \quad && \forall x \in B_i,\ i = 1, \dots, N-1 \\ \mathrm{e}^{N-1} &\leq d_{G^*}(x) && \forall x \in B_N. \end{aligned}$$

Then there exists an index i such that $e_{G^*}(A, B_i) \geq e(G^*)/N$. First assume that $i \leq \log d$. Then Proposition 2.1 implies that $(A, B_i)_{G^*}$ contains a subgraph H with $\delta(H) \geq d((A, B_i)_{G^*})/2 \geq d/2N$. As $\Delta(H) \leq \Delta((A, B_i)_{G^*}) \leq d$, H is as required in the lemma.

Next assume that $i = N$. Let A^* be the set of all those vertices in A which send at least $\lfloor \sqrt{d}/(2N)^{1/9} \rfloor =: \ell$ edges in G^* to B_N. Then

$$d|A^*| + \ell |A| \geq e_{G^*}(A, B_N) \geq \frac{e(G^*)}{N} = \frac{d|A|}{N},$$

and therefore

$$|A^*| \geq \left(\frac{d}{N} - \ell \right) \frac{|A|}{d} \geq \frac{|A|}{2N}. \tag{4.4}$$

Moreover, $d|A| = e(G^*) \geq \mathrm{e}^{N-1} |B_N| \geq d^{6t+1} |B_N|$. Together with (4.4) this implies that

$$|A^*| \geq \frac{d^{6t} |B_N|}{2N} \geq \ell^{12t} |B_N|.$$

Let H^* be the graph obtained from $(A^*, B_N)_{G^*}$ by deleting edges if necessary such that $d_{H^*}(a) = \ell$ for all $a \in A^*$. Since $\ell \geq 8t \geq 2$ by (4.3), Lemma 4.8 implies that H^* (and hence G) contains a subdivision of some graph of average degree at least

$$\frac{\ell^9}{2^{14}} \geq r^3 \cdot \frac{d^{3/2}}{4^3 \cdot 2^{15+9} N} \overset{(4.3)}{\geq} r^3.$$

So we may assume that $\log d < i < N$. Set $k := \lfloor d/2N \rfloor$ and let A_p be a random subset of A which is obtained by including every vertex into A_p with probability $p := 2k/\mathrm{e}^{i-1}$ independently of all other vertices. Then for every vertex $b \in B_i$ we have

$$2k \leq d_{G^*}(b) p = \mathbb{E}(|N_{G^*}(b) \cap A_p|) \leq 2\mathrm{e}k \leq d/2. \tag{4.5}$$

Let us call a vertex $b \in B_i$ *bad* if $|N_{G^*}(b) \cap A_p| \le k$ or $|N_{G^*}(b) \cap A_p| \ge d$. So (4.3), (4.5) and Lemma 4.5 together imply that the probability that a given vertex $b \in B_i$ is bad is at most $2\mathrm{e}^{-k/6} \le 1/24$. So the expected number of bad vertices in B_i is at most $|B_i|/24$. Hence Markov's inequality implies that

$$\mathbb{P}(\ge |B_i|/6 \text{ vertices of } |B_i| \text{ are bad}) \le 1/4. \tag{4.6}$$

Moreover

$$2k|A| \le \frac{d|A|}{N} \le e_{G^*}(A, B_i) \le \mathrm{e}^i |B_i|,$$

and so $|A| \le \mathrm{e}^i|B_i|/2k$. Hence

$$\mathbb{E}(|A_p|) = p|A| \le \frac{p\mathrm{e}^i|B_i|}{2k} = \mathrm{e}|B_i|.$$

Thus Markov's inequality shows that

$$\mathbb{P}(|A_p| \ge 4|B_i|) \le \mathrm{e}/4.$$

Together with (4.6) this implies that with probability at least $1 - 1/4 - \mathrm{e}/4 > 0$ there exists an outcome A_p such that $|A_p| \le 4|B_i|$ and at most $|B_i|/6$ vertices of B_i are bad. Let H' be the subgraph of G^* induced by A_p and those vertices in B_i that are not bad. Then $\Delta(H') \le d$ and $e(H') \ge 5k|B_i|/6$. Moreover, $|H'| \le |A_p| + |B_i| \le 5|B_i|$, and so the average degree of H' is at least $k/3$. By Proposition 2.1, H' has a subgraph H with

$$\delta(H) \ge \frac{k}{6} \overset{(4.3)}{\ge} \frac{d}{100t\log d} \ge \frac{r}{400t\log r}.$$

So H is as required in the lemma. □

Proof of Theorem 4.9. Apply Lemma 4.10 to G to obtain (without loss of generality) a bipartite subgraph H with $\Delta(H) \le r$ and

$$\delta := \delta(H) \ge \frac{r}{400t\log r}. \tag{4.7}$$

Define ε by

$$r^\varepsilon = \frac{r^{\frac{1}{2(s-1)}}}{32t(r/\delta)^{\frac{1}{s+1}}}. \tag{4.8}$$

Put $\ell := r^{1-\varepsilon}$ and let X be a random subset of $V(H)$ which is obtained by including each vertex into X with probability $p := 2\ell/\delta$ independently of all other vertices. The branch sets of our minor of large average degree will consist of stars whose centres are precisely the vertices in X. Since $r \ge (100t)^{16(s-1)}$, for every vertex $v \in H$ we have

$$\begin{aligned}\mathbb{P}(v \in X) = p = \frac{2r}{r^\varepsilon\delta} &\le \frac{2\cdot 400t\log r}{r^\varepsilon} \\ &\le \frac{2\cdot 32t\cdot(400t)^2}{r^{\frac{1}{4(s-1)}}}\cdot\frac{(\log r)^2}{r^{\frac{1}{4(s-1)}}} \le 1/20. \end{aligned}\tag{4.9}$$

Call a vertex $v \in H$ *good* if it satisfies the following two conditions.

(i) $v \notin X$.

(ii) $|N_H(v) \cap X| \geq \ell$.

We will now show that with large probability a given vertex $v \in H$ is good. First note that

$$\mathbb{E}(|N_H(v) \cap X|) = d_H(v) \cdot p \geq 2\ell.$$

As $\ell \geq \sqrt{r}$, Lemma 4.5 implies that

$$\mathbb{P}(|N_H(v) \cap X| < \ell) \leq 2\mathrm{e}^{-\sqrt{r}/6} \leq 1/20.$$

Together with (4.9) this implies that the probability that a given vertex $v \in H$ is not good is at most 1/10. Call an edge $uv \in H$ *good* if both u and v are good. So the probability that a given edge $uv \in H$ is not good is at most 1/5 and therefore

$$\mathbb{E}(\text{number of edges which are not good}) \leq e(H)/5.$$

So Markov's inequality implies that

$$\mathbb{P}(\geq e(H)/2 \text{ edges are not good}) \leq 2/5.$$

Using Markov's inequality once more, we see that

$$\mathbb{P}(|X| \geq 2p|H|) \leq 1/2.$$

Thus with probability at least $1 - 2/5 - 1/2 > 0$ there is an outcome X with $|X| \leq 2p|H|$ and for which at least half of the edges of H are good. Let U be the set of good vertices of H. So $e_H(U, U)$ is precisely the number of good edges of H. For every $x \in X$ put $U_x := U \cap N_H(x)$. Note that, since H is bipartite, $H[U_x]$ consists of isolated vertices. Given a vertex $u \in U$, let $X_u := X \cap N_H(u)$. So condition (ii) implies that $|X_u| \geq \ell$.

For every vertex $u \in U$ choose a vertex $x_u \in X_u$ uniformly at random, independently of all other vertices in U. For all $x \in X$, let S_x be the set of all those $u \in U_x$ with $x_u = x$. Note that the S_x are disjoint and their union is U. Moreover, every good edge of H joins vertices in distinct S_x. We will now show that with positive probability the minor M of H whose branch sets are the $S_x \cup \{x\}$ ($x \in X$) has large average degree. For this, we will show that with positive probability a large fraction of good edges joins different pairs S_x, S_y and thus corresponds to different edges of M. As $|X|$ (i.e. the number of vertices of M) is relatively small, this will imply that M has large average degree. Thus, given a good edge $uv \in H$, we say that

- *uv is of type I* if there exists a good edge $ab \neq uv$ joining S_{x_u} to S_{x_v} such that ab and uv are disjoint,

- *uv is of type II* if there exists a good edge $ab \neq uv$ joining S_{x_u} to S_{x_v} such that a is an endvertex of uv and $|N_H(a) \cap U_{x_w}| \leq \ell/30$, where w is the endvertex of uv distinct from a (Fig. 4.1),

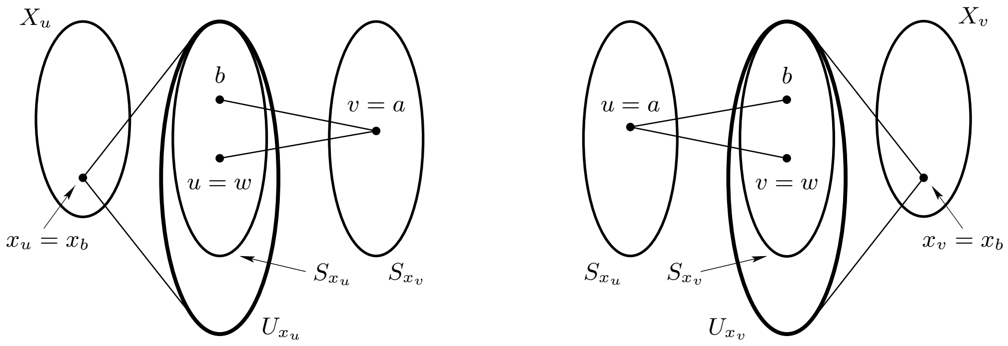


Figure 4.1: The two possibilities for an edge uv to be of type II

- *uv is of type III* if there exists a good edge $ab \neq uv$ joining S_{x_u} to S_{x_v} such that a is an endvertex of uv and $|N_H(a) \cap U_{x_w}| > \ell/30$, where w is the endvertex of uv distinct from a.

Note that for all distinct $x, y \in X$ the graph $H[U_x \cup U_y]$ does not contain a $K_{s-1,t}$ (since this would form a $K_{s,t}$ together with either x or y). So Theorem 4.7 implies that

$$e_H(U_x, U_y) \leq 4tr^{2-\frac{1}{s-1}}. \tag{4.10}$$

Recall that the S_x–S_y edges are precisely those U_x–U_y edges uv (with $u \in U_x$ and $v \in U_y$) for which u has chosen x and v has chosen y, i.e. for which $x = x_u$ and $y = x_v$. Since the probability that $x = x_u$ and $y = x_v$ is $|X_u|^{-1}|X_v|^{-1} \leq \ell^{-2}$, it follows that

$$\begin{aligned}\mathbb{P}(\text{there is a good } S_x\text{–}S_y \text{ edge}) &\leq e_H(U_x, U_y) \cdot \left(\frac{1}{\ell}\right)^2 \\ &\overset{(4.10)}{\leq} 4tr^{2\varepsilon - \frac{1}{s-1}} \overset{(4.8)}{\leq} \frac{1}{60}.\end{aligned} \tag{4.11}$$

So given a good edge uv we have

$$\begin{aligned}\mathbb{P}(uv \text{ is of type I}) &= \sum_{x \in X_u,\, y \in X_v} \mathbb{P}(uv \text{ is of type I and } x = x_u \text{ and } y = x_v) \\ &= \sum_{x \in X_u,\, y \in X_v} \mathbb{P}(\text{there is a good } S_x\text{–}S_y \text{ edge disjoint from } uv) \cdot \frac{1}{|X_u|} \cdot \frac{1}{|X_v|} \\ &\overset{(4.11)}{\leq} \frac{1}{60}.\end{aligned}$$

Moreover, given x_u and x_v, in the definition of a type II edge uv there are at most two possibilities for a and at most $\ell/30$ candidates for b and $\mathbb{P}(x_b = x_w) \leq 1/\ell$. Thus

$$\begin{aligned}\mathbb{P}(uv \text{ is of type II}) &= \sum_{x \in X_u,\, y \in X_v} \mathbb{P}(uv \text{ is of type II and } x = x_u \text{ and } y = x_v) \\ &\leq \sum_{x \in X_u,\, y \in X_v} 2 \cdot \frac{\ell}{30} \cdot \frac{1}{\ell} \cdot \frac{1}{|X_u|} \cdot \frac{1}{|X_v|} = \frac{4}{60}.\end{aligned}$$

Hence

$$\mathbb{E}(\text{number of good edges which are type I or II}) \leq e_H(U,U)/12,$$

and so Markov's inequality implies that

$$\mathbb{P}(\geq e_H(U,U)/4 \text{ good edges are of type I or II}) \leq 1/3. \tag{4.12}$$

It remains to show that also with only small probability a large fraction of the good edges is of type III. This trivially holds for $s = 2$. Indeed, as $\ell/30 \geq t$, the vertices a and x_w in the definition of a type III edge form a $K_{2,t}$ together with any t vertices in $N_H(a) \cap U_{x_w}$. Thus there are no good edges of type III in this case. So suppose that $s \geq 3$. Given a vertex $y \in X$, let V_y be the set of all those vertices in U which have at least $\ell/30$ neighbours in U_y. So $V_y \subseteq U \setminus U_y$. As H is $K_{s,t}$-free, Lemma 4.6 implies that

$$|V_y|\binom{\ell/30}{s} \leq t\binom{|U_y|}{s}.$$

Thus

$$|V_y| \leq \left(\frac{32}{\ell}\right)^s t \cdot |U_y|^s \leq (32r^{\varepsilon})^s t. \tag{4.13}$$

Given distinct good edges uv and ub and vertices $x, y \in X$, we say that the ordered quadruple uv, ub, x, y *forms a configuration of type III* if $u \in U_x$, $v, b \in U_y$ and if u has at least $\ell/30$ neighbours in U_y. So each configuration of type III can be obtained by first selecting a vertex $v \in U$, then selecting a vertex $y \in X_v$, then selecting a neighbour u of v which lies in V_y (i.e. which lies in U and sends at least $\ell/30$ edges to U_y), then we select a vertex $x \in X_u$ and finally we select a neighbour b of u in $U_y \setminus v$. We say that a configuration of type III *survives* if u has chosen x and both v and b have chosen y, i.e. if $x = x_u$ and $y = x_v = x_b$. Thus the probability that it survives is precisely $|X_u|^{-1}|X_v|^{-1}|X_b|^{-1} \leq |X_u|^{-1}|X_v|^{-1}/\ell$. Hence

$$\begin{aligned}
& \mathbb{E}(\text{number of good edges which are of type III}) \\
\leq\; & \mathbb{E}(\text{number of surviving configurations of type III}) \\
\leq\; & \sum_{v \in U} \sum_{y \in X_v} \sum_{u \in N_H(v) \cap V_y} \sum_{x \in X_u} \sum_{b \in N_H(u) \cap U_y \setminus v} \frac{1}{|X_u||X_v|\ell} \\
\overset{(4.13)}{\leq}\; & \frac{|H|(32r^{\varepsilon})^s tr}{r^{1-\varepsilon}} \\
\overset{(4.8)}{=}\; & |H| 32^s t \left(\frac{r^{\frac{1}{2(s-1)}}}{32t}\right)^{s+1} \cdot \frac{\delta}{r} \\
\overset{(s\geq 3)}{\leq}\; & \frac{\delta|H|}{32} \leq \frac{e(H)}{16} \leq \frac{e_H(U,U)}{8}.
\end{aligned}$$

Hence Markov's inequality implies that also for $s \geq 3$

$$\mathbb{P}(\geq e_H(U,U)/4 \text{ good edges are of type III}) \leq 1/2.$$

Together with (4.12) this shows that for every $u \in U$ there exists a choice of x_u such that at most $e_H(U,U)/2$ good edges are of type I, II or III. Let F be the set of all good edges which are not of type I, II or III.

Consider the minor M of H whose branch sets are the sets $S_x \cup \{x\}$ (for all $x \in X$). As H is bipartite, every edge in F joins distinct branch sets and, by definition of F, no two edges in F join the same pair of branch sets. Thus $e(M) \geq |F|$ and so

$$\begin{aligned} d(M) &\geq \frac{2|F|}{|X|} \geq \frac{e_H(U,U)}{|X|} \geq \frac{e(H)}{2 \cdot 2p|H|} \geq \frac{\delta}{8p} = \frac{\delta^2}{16r^{1-\varepsilon}} \\ &\overset{(4.8)}{=} \frac{r^{1+\frac{1}{2(s-1)}}}{16 \cdot 32t} \cdot \left(\frac{\delta}{r}\right)^{2+\frac{1}{s+1}} \\ &\overset{(4.7)}{\geq} \frac{r^{1+\frac{1}{2(s-1)}}}{16 \cdot 32t \cdot (400t \log r)^{2+\frac{1}{s+1}}} \geq \frac{r^{1+\frac{1}{2(s-1)}}}{10^9 t^4 (\log r)^{2+\frac{1}{s+1}}}, \end{aligned} \tag{4.14}$$

as required. □

Note that for regular graphs G the logarithmic term in (4.2) is not necessary. Indeed, we only have to replace the graph H in the proof of Theorem 4.9 with a bipartite subgraph obtained from G by an application of Proposition 2.2, and then (4.14) shows that this subgraph contains a minor of the required average degree. Moreover, for non-regular graphs the exponent $2+\frac{1}{s+1}$ of the logarithmic term can be reduced to $1+\frac{1}{2(s-1)}$. However, we do not give the details as we conjecture that (as in the case $s=2$, see Theorem 3.2) the logarithmic term in (4.2) can be removed altogether. This would then match the upper bound implied by Proposition 4.12.

4.4 Upper bounds

In this section we observe that the truth of the following well-known conjecture about the existence of dense $K_{s,t}$-free graphs would imply that for fixed s and t Theorems 4.1 and 4.9 are best possible up to the logarithmic term (and that this term cannot be omitted completely in Theorem 4.1).

Conjecture 4.11 *For all integers $t \geq s \geq 2$ there exists a positive constant $c = c(s,t)$ such that for all integers n there is a $K_{s,t}$-free graph G of order n with at least $cn^{2-1/s}$ edges.*

(See e.g. [12, p. 362] or [25, p. 36] for the case $s = t$ which of course would already imply the general case.) In other words, the conjecture states that the upper bound on the number of edges of a $K_{s,t}$-free graph in Theorem 4.7 gives the correct order of magnitude. Conjecture 4.11 is known to be true for all $t \geq s$ with $s = 2, 3$ (see [12, Ch. VI]). Furthermore, Alon, Rónyai and Szabó [6] proved the conjecture for all $t \geq s \geq 2$ with $t > (s-1)!$ by modifying a construction of [47]. The following proposition immediately implies that Theorems 4.1 and 4.9 are best possible up to the logarithmic term, provided that Conjecture 4.11 holds.

Proposition 4.12 *For every $c > 0$ and every $s \geq 2$ there exists a constant $C = C(c, s)$ such that whenever G is a graph with $e(G) \geq c|G|^{2-1/s}$ then every minor H of G satisfies*

$$d(H) \leq C \cdot d(G)^{1+\frac{1}{2(s-1)}}.$$

Proof. Put $n := |G|$, $r := d(G)$ and $d := d(H)$. For every vertex $h \in H$ let $V_h \subseteq V(G)$ be the branch set corresponding to h. Then

$$nr = 2e(G) \geq \sum_{h \in H} \sum_{v \in V_h} d_G(v) \geq \sum_{h \in H} d_H(h) = 2e(H) \geq d^2,$$

and so $d \leq \sqrt{nr}$. But $r \geq 2cn^{1-1/s}$, i.e. $n \leq (r/2c)^{\frac{s}{s-1}}$. Therefore

$$d \leq \frac{r^{\frac{1}{2}(1+\frac{s}{s-1})}}{(2c)^{\frac{s}{2(s-1)}}},$$

as required. $\square$

In general, for $s \geq 4$ the best known lower bound on the maximum number of edges of a $K_{s,s}$-free graph G is $c|G|^{2-2/(s+1)}$ (see e.g. [12, Ch. VI, Thm. 2.10]). Together with Proposition 4.12, this still yields an upper bound of $c'r^{1+\frac{1}{s-1}}$ for the order of the complete minor in Theorem 4.1 and the average degree of the minor in Theorem 4.9.

Finally, the next proposition shows that if we ask for a complete graph instead of just a graph of large average degree as minor, then we lose an extra $\sqrt{\log r}$ factor. In particular, if Conjecture 4.11 holds, then the logarithmic term in Theorem 4.1 cannot be omitted completely. The proof of Proposition 4.13 is an extension of the well-known probabilistic argument which shows that (3.1) gives the correct order of magnitude for the function $p(r)$ (see Bollobás, Catlin and Erdős [17]).

Proposition 4.13 *Suppose that Conjecture 4.11 holds. Then for all integers $t \geq s \geq 2$ there exists a constant $C' = C'(s, t)$ such that for every integer r_0 there is a $K_{s,t}$-free graph G' of average degree $r \geq r_0$ which does not contain a complete graph of order at least $C'r^{1+\frac{1}{2(s-1)}}/\sqrt{\log r}$ as minor.*

Proof. Let $c = c(s, t)$ be the constant in Conjecture 4.11 and let G be a $K_{s,t}$-free graph whose order n is sufficiently large compared with c and r_0 and such that $cn^{2-1/s} \leq e(G) \leq 2cn^{2-1/s}$. Throughout the proof we will also assume that n is sufficiently large for our estimates to hold. Set $\ell := d(G)$,

$$k := \sqrt{\frac{32c(\ell/2c)^{\frac{2s-1}{s-1}}}{\log \ell}}$$

and put $h := \lceil k \rceil + 1$. Consider a random spanning subgraph G_p which is obtained from G by including every edge of G into G_p with probability $p := 1 - 1/\mathrm{e}$ independently of all other edges of G. Then $\mathbb{E}(e(G_p)) = p \cdot e(G)$ and thus Lemma 4.5 implies that with probability at least 1/2 we have $e(G_p) \geq p \cdot e(G)/2 > e(G)/4$, i.e. $d(G_p) > \ell/4 \geq r_0$. Hence it suffices to show that with probability at most 1/4 the graph G_p contains a K_h minor. (Indeed, if C' is sufficiently large, then this would show that with probability at least $1/2 - 1/4 > 0$ the graph G_p is as required in Proposition 4.13.) Consider any family $\mathcal{V} = \{V_1, \dots, V_h\}$ of disjoint subsets of $V(G)$ and let $\mathcal{W}$ be the set of all such families. Call $\mathcal{V}$ *admissible* if, for all $i < j$, the graph G_p contains at least one edge between V_i and V_j. Thus if $\mathcal{V}$ consists of the branch sets of a K_h minor, then $\mathcal{V}$ is admissible. Therefore, it suffices to show that with probability at least 3/4 no $\mathcal{V} \in \mathcal{W}$ is admissible. For this, we first estimate the probability that a given family $\mathcal{V}$ is admissible. For all $1 \leq i < j \leq h$ call the pair V_i, V_j *thin* if the bipartite subgraph $(V_i, V_j)_G$ of G between V_i and V_j contains at most $(\log \ell)/4$ edges. We claim that at least half of the pairs V_i, V_j are thin. Indeed, suppose not. Then

$$
\begin{aligned}
e(G) &\geq \frac{1}{2}\binom{h}{2} \cdot \frac{\log \ell}{4} \\
&> \frac{k^2}{4} \cdot \frac{\log \ell}{4} = 2c\left(\frac{\ell}{2c}\right)^{\frac{2s-1}{s-1}} \geq 2c\left(n^{\frac{s-1}{s}}\right)^{\frac{2s-1}{s-1}} = 2cn^{2-\frac{1}{s}},
\end{aligned}
\tag{4.15}
$$

a contradiction. Thus

$$
\begin{aligned}
\mathbb{P}(\mathcal{V} \text{ is admissible}) &\leq \mathbb{P}(e_{G_p}(V_i, V_j) \geq 1 \text{ for all thin pairs } V_i, V_j) \\
&= \prod_{V_i, V_j \text{ thin}} \left(1 - (1-p)^{e_G(V_i, V_j)}\right) \leq \prod_{V_i, V_j \text{ thin}} (1 - \ell^{-1/4}) \\
&\leq (1 - \ell^{-1/4})^{\frac{1}{2}\binom{h}{2}} \leq \exp(-\ell^{-1/4} \cdot k^2/4) \\
&\overset{(4.15)}{\leq} \exp\left(-\ell^{-1/4} \cdot cn^{2-1/s} \cdot \frac{8}{\log \ell}\right) \leq \exp\left(\frac{-8c}{\log n} \cdot n^{\frac{7}{4}-\frac{1}{s}}\right) \\
&\leq \mathrm{e}^{-2n \log n}.
\end{aligned}
$$

As $|\mathcal{W}| \leq n^n$, this implies that

$$
\mathbb{P}(\text{some } \mathcal{V} \in \mathcal{W} \text{ is admissible}) \leq n^n \cdot \mathrm{e}^{-2n \log n} = \mathrm{e}^{-n \log n} \leq \frac{1}{4},
$$

as required. □

If we replace the graph G in the proof of Proposition 4.13 by a graph as in Conjecture 3.12, then we obtain the following.

Proposition 4.14 *Suppose that Conjecture 3.12 holds. Then for all odd integers g there exists a constant $C' = C'(g)$ such that for every integer r_0 there is a graph G' of minimum degree $r \geq r_0$ and girth at least g which does not contain a complete graph of order at least $C' r^{(g+1)/4}/\sqrt{\log r}$ as minor.* □

Thus, the truth of Conjecture 3.12 would imply that, for fixed odd g, Corollary 3.3 is best possible up to the value of the constant c.

Chapter 5

Large topological cliques in graphs without a 4-cycle

5.1 Introduction

Bollobás and Thomason [20] as well as Komlós and Szemerédi [58] independently proved the following result, which improved an earlier bound of Mader.

Theorem 5.1 [20, 58] *There exists a positive constant c such that every graph of average degree r contains a subdivision of a complete graph of order at least $c\sqrt{r}$.*

It is easy to see (and was first observed by Jung [43]) that the complete bipartite graph $K_{r,r}$ contains no subdivision of a complete graph K_ℓ with $\ell \geq \sqrt{8r}$. So in general Theorem 5.1 is best possible up to the value of the constant c. However, it turns out that dense bipartite graphs are the only counterexamples in the sense that we can improve Theorem 5.1 if we forbid a fixed complete bipartite subgraph $K_{s,t}$:

Theorem 5.2 *For all integers $t \geq s \geq 2$ there exists an $r_0 = r_0(s,t)$ such that every $K_{s,t}$-free graph G of average degree $r \geq r_0$ contains a subdivision of a complete graph of order at least*

$$\frac{r^{\frac{1}{2}+\frac{1}{2(s-1)}}}{(\log r)^{12}}. \tag{5.1}$$

By Jung's observation, clearly we cannot hope for a similar result if we forbid a non-bipartite graph H instead of a $K_{s,t}$ since then complete bipartite graphs would be H-free.

In the C_4-free case $s = t = 2$ the bound (5.1) is 'almost linear' and thus best possible up to the logarithmic term. For arbitrary $t \geq s \geq 2$ a classical conjecture on the existence of dense $K_{s,t}$-free graphs (see e.g. [12, p. 362] or [25, p. 36]) would also imply that the bound (5.1) is best possible up to the logarithmic term. We will give the details in Section 5.4.

Up to the logarithmic term, the special case $s = t = 2$ of Theorem 5.2 gives an affirmative answer to a question of Mader [80], who asked whether every graph G of girth at least 5 (and hence also every C_4-free graph) contains a subdivision of a complete graph whose order is at least linear in the average degree of G. As remarked in Section 3.1, this is true if the girth is at least 15.

We remark that the C_4-free case of Theorem 5.2 implies that also for all $\ell \geq 2$ every $C_{2\ell}$-free graph G contains a subdivision of a complete graph whose order is 'almost linear' in the average degree of G. Indeed, this follows immediately from a result of [72] that every $C_{2\ell}$-free graph can be made C_4-free by deleting a constant fraction of its edges (the case $\ell = 3$ of the latter result is due to Györi [37]).

The proof of Theorem 5.2 uses results of Komlós and Szemerédi [57, 58]. In fact, Theorems 2.1 and 2.2 of [58] together with Theorem 4.7 below already imply the weaker bound $r^{\frac{1}{2}+\frac{1}{6(s-1)}-o(1)}$ instead of (5.1) in Theorem 5.2.

This chapter is organized as follows. In Section 5.2 we state several results which we will need later on. We prove Theorem 5.2 in Section 5.3. In the final section we derive the upper bounds mentioned above.

5.2 Notation and tools

All logarithms in this chapter are base e, where e denotes the Euler number. We will now collect some results which we need in our proof of Theorem 5.2. Lemma 4.10 allows us to assume that in the proof of Theorem 5.2 our given graph G is 'almost regular' in the sense that its maximum degree is not much larger than its minimum degree.

Throughout this chapter, we fix a constant κ such that

$$1 < \kappa < 6/5 \quad \text{and} \quad \kappa^2 + 3\kappa + 3 < 8. \tag{5.2}$$

Given positive constants d and ε_0, let

$$\varepsilon(x) := \begin{cases} 0 & \text{if } x < d/4 \\ \varepsilon_0/(\log(8x/d))^\kappa & \text{if } x \geq d/4. \end{cases} \tag{5.3}$$

Note that $\varepsilon(x)x$ is monotone increasing for all $x \geq d/2$. We call a graph H a *$(\kappa, d, \varepsilon_0)$-expander for sets of size at least* x_0 if every $X \subseteq V(H)$ with $x_0 \leq |X| \leq |H|/2$ satisfies $|N_H(X)| \geq \varepsilon(|X|)|X|$, where ε is the function defined in (5.3). H is a *$(\kappa, d, \varepsilon_0)$-expander* if we can take $x_0 = 0$.

The following result of Komlós and Szemerédi [57, Thm. 2.2] shows that every graph G contains an expander whose average degree is not much smaller than that of G.

Theorem 5.3 *Let $d, \varepsilon_0 > 0$ and suppose that the function ε defined in* (5.3) *satisfies $\sum_{x=1}^{\infty} \varepsilon(x)/x \leq 1/6$ (which holds if ε_0 is sufficiently small compared with κ). Then every graph G has a subgraph H with $d(H) \geq d(G)/2$ and $\delta(H) \geq d(H)/2$ which is a $(\kappa, d, \varepsilon_0)$-expander for sets of size at least* $3d/4$.

Corollary 5.4 *There is a positive $\varepsilon_0 = \varepsilon_0(\kappa) < 1$ such that every graph G has a subgraph H with $d(H) \geq d(G)/2$ and $\delta(H) \geq d(H)/2$ which is a $(\kappa, d(H), \varepsilon_0)$-expander.*

Proof. Let G' be a subgraph of G which maximizes $d(G')$. Put $d' := d(G')/6$. If ε_0 is sufficiently small, we may apply Theorem 5.3 to G' to obtain a graph H which is a $(\kappa, d', 8\varepsilon_0)$-expander for sets of size at least $3d'/4$. Using that $d(H) \le 6d'$, it is easy to check that for $x \ge d(H)/4$ we have

$$\frac{8\varepsilon_0}{(\log(8x/d'))^\kappa} \ge \frac{\varepsilon_0}{(\log(8x/d(H)))^\kappa}.$$

Since $d(H)/4 \ge 3d'/4$ this shows that H is a $(\kappa, d(H), \varepsilon_0)$-expander. □

The following simple consequence of expansion is implicit in [57]. It shows that expanders have 'robustly small diameter'. A proof is included in [58, Lemma 2.1].

Lemma 5.5 *Let $d > 0$, $1 > \varepsilon_0 > 0$ and let G be a $(\kappa, d, \varepsilon_0)$-expander. Let ε be as defined in* (5.3) *and suppose that $X, Y, Z \subseteq V(G)$ such that $|X|, |Y| \ge x \ge d$, $|Z| \le \varepsilon(x)x/4$ and $(X \cup Y) \cap Z = \emptyset$. Then the distance between X and Y in $G - Z$ is at most*

$$\frac{2\log(|G|/x)}{\log(1 + \varepsilon(|G|)/2)} \le \frac{8(\log(8|G|/d))^{1+\kappa}}{\varepsilon_0}.$$

In the proof of Theorem 5.2 we will first replace our given graph G with an 'almost regular' subgraph obtained by Lemma 4.10. Then we apply Corollary 5.4 to this subgraph to obtain an expander H which is still 'almost regular'. The following result of Komlós and Szemerédi [57, Thm. 3.1] implies that we are already done if the order of H is sufficiently large compared with the average degree of H.

Theorem 5.6 *Let $\varepsilon_0 > 0$ and let $\alpha > \kappa^2 + 3\kappa + 3 > 7$. Then there exists a positive constant c such that every graph G which is a $(\kappa, d(G), \varepsilon_0)$-expander satisfying $d(G)/2 \le \delta(G) \le \Delta(G) \le 72(d(G))^2$ and $\log|G| \ge (\log d(G))^\alpha$ contains a subdivision of a complete graph of order at least $cd(G)$.*

In the remainder of this section we collect some other results which we will use in our proof of Theorem 5.2. Wwill use the following variant of Lemma 4.5 (see e.g. [39, Cor. 2.3 and 2.4]).

Lemma 5.7 *Let $X_1, \dots, X_n$ be independent* 0-1 *random variables with $\mathbb{P}(X_i = 1) = p$ and let $X := \sum_{i=1}^n X_i$. Then*

$$\mathbb{P}(X \le \mathbb{E}X/2) \le 2\mathrm{e}^{-\mathbb{E}X/12}, \tag{5.4}$$

$$\mathbb{P}(X \ge x) \le \mathrm{e}^{-x} \quad \textit{for all } x \ge 7\mathbb{E}(X). \tag{5.5}$$

The next result is an easy consequence of Hall's matching theorem (see e.g. [14, Ch. III, Thm. 7] or [28, Thm. 2.1.2]).

Corollary 5.8 *Let $G = (A, B)$ be a bipartite graph such that $d_G(a) \ge d_A$ for all $a \in A$ and $d_G(B) \le d_B$ for all $b \in B$. Then G contains $|A|$ disjoint stars with centres in A and such that each of them has $\lfloor d_A/d_B \rfloor$ leaves.*

Proof. Form a new bipartite graph $G' = (A', B)$ by replacing every vertex $a \in A$ with $\tau := \lfloor d_A/d_B \rfloor$ new vertices and joining each such vertex to all the neighbours of a. For every $A^* \subseteq A'$ we have

$$|A^*|d_A \leq e_{G'}(A^*, N_{G'}(A^*)) \leq |N_{G'}(A^*)|\tau d_B$$

and thus $|N_{G'}(A^*)| \geq |A^*|$. So by Hall's theorem there exists a matching of A' in G'. But this corresponds to the required disjoint stars in G. $\square$

5.3 Proof of Theorem 5.2

As indicated in Section 5.2, in the proof of Theorem 5.2 we may assume the we are given a graph H which is an 'almost regular' expander such that $\log |H| \leq (\log d(H))^\alpha$. But then by Lemma 5.5, the distance in H between any two sufficiently large sets is small in terms of $d(H)$ and this remains true if we delete a few vertices of H. Roughly, we shall use this property as follows. Let ℓ be the value of (5.1) in Theorem 5.2. So we are seeking a subdivision TK_ℓ of K_ℓ in H. Lemma 5.9 below implies that we can find ℓ disjoint stars in H such that the neighbourhood in H of each star is large even if we delete a small but arbitrary subset of the leaves. The centres of these stars will form the branch vertices of our TK_ℓ. To find the subdivided edges, we will apply Lemma 5.5 to obtain for every pair of stars a short path joining the neighbourhoods of the stars. All these paths will be disjoint, will avoid the stars themselves and they can be extended to subdivided edges of the TK_ℓ.

Given a star S, we denote by $L(S)$ the set of its leaves.

Lemma 5.9 *For all integers $t \geq s \geq 2$ there exists an $r_0 = r_0(s,t)$ such that for each $r \geq r_0$ every $K_{s,t}$-free graph G with $\delta(G) \geq r/1600t \log r$ and $\Delta(G) \leq r$ contains at least*

$$k := \left\lfloor \frac{r^{\frac{1}{2}+\frac{1}{2(s-1)}}}{t(1600 \log r)^2} \right\rfloor \tag{5.6}$$

disjoint stars where each such star S satisfies the following two conditions.

(i) $|L(S)| = k$.

(ii) *For every $v \in L(S)$ there is a set N_v of k neighbours of v outside $V(S)$ such that $N_v \cap N_w = \emptyset$ for distinct $v, w \in L(S)$.*

As described in Section 5.4, it is believed that for $t \geq s \geq 2$ there are $K_{s,t}$-free graphs of average degree r and order at most $c_s r^{1+1/(s-1)}$. Note that for such graphs G the union of the stars in Lemma 5.9 (and thus the subdivision of the complete graph which we will construct in our proof of Theorem 5.2) would cover a significant portion of $V(G)$.

Proof of Lemma 5.9. Throughout the proof of the lemma we will assume that r is sufficiently large compared with s and t. Put $n := |G|$, $\delta := \delta(G)$ and $f := 2(\log r)r^{\frac{1}{2}-\frac{1}{2(s-1)}}$. Consider a random subset X_p of $V(G)$ which is obtained by including each vertex of G with probability $p := f/\mathrm{e}^2 r$ in X_p, independently of all other vertices of G. Call a vertex $v \in G$ *good* if it has at most f neighbours in X_p. Then Stirling's inequality (see e.g. [14, p. 216]) implies that

$$\mathbb{P}(v \text{ is not good}) \leq \binom{d_G(v)}{f} p^f \leq \left(\frac{\mathrm{e}r}{f} \cdot p\right)^f = \mathrm{e}^{-f} \leq r^{-2}. \tag{5.7}$$

Let n_p denote the number of vertices in G which are not good or have a neighbour that is not good. Then (5.7) implies that $\mathbb{E}(n_p) \leq (r+1)n/r^2$. So writing $m_p := |X_p| - n_p$, we have

$$\mathbb{E}(m_p) \geq pn - \frac{(r+1)n}{r^2} \geq \frac{pn}{2} \overset{(4.1)}{\geq} \frac{f}{2\mathrm{e}^2 r} \cdot \left(\frac{\delta}{2t}\right)^{1+\frac{1}{s-1}} \geq \frac{r^{\frac{1}{2}+\frac{1}{2(s-1)}}}{(2\mathrm{e}1600t^2)^2 \log r} \geq 2k.$$

Hence there is an outcome X_p which contains least $2k$ vertices that are good and have only good neighbours in G. Let X denote the set of all these vertices.

We remark that for the case $t = s = 2$ the lemma now follows easily. Indeed, since every vertex $x \in X$ is good, it has at least $\delta - f \geq \delta/2$ neighbours outside X and, since each such neighbour is good, it sends at most f edges to X. As

$$\frac{\delta}{2f} \geq 100k \tag{5.8}$$

we can apply Corollary 5.8 to $(X, N_G(X))_G$ to obtain $|X|$ disjoint stars whose centres are the vertices in X and where each such star has k leaves. Then these stars S are as required in the lemma. (Given $v \in L(S)$, we can take for N_v any set of k neighbours of v outside S. As $t = s = 2$ these sets are disjoint for distinct $v \in L(S)$.) The argument easily extends to the case $t \geq s = 2$ but not to the general case. However, we will show that a random assignment of leaves (these will be the vertices in $N_G(X)$) to star centres (which will be the vertices in X) works for all $t \geq s \geq 2$.

Given a vertex $v \in N_G(X)$, with probability $|N_G(v) \cap X|/f$ choose one of the vertices $x \in N_G(v) \cap X$. Here each of these vertices is equally likely to be chosen and so the corresponding probability is $1/f$. Choose no vertex at all with the remaining probability $1 - |N_G(v) \cap X|/f$. (Recall that $|N_G(v) \cap X| \leq f$ since $N_G(X) \ni v$ consists of good vertices. So the probability is well defined.) Do this independently for all vertices $v \in N_G(X)$. Let S_x denote the star in G whose centre is x and whose leaves are the vertices in $N_G(X)$ that have chosen x. Thus the S_x are disjoint for distinct x. We will now show that with positive probability at least half of the stars S_x $(x \in X)$ contain a substar which satisfies (i) and (ii). So call S_x *useful* if there is a set $L_x \subseteq L(S_x)$ satisfying the following two conditions.

(a) $|L_x| = k$.

(b) For every $v \in L_x$ there is a set N_v of k neighbours of v outside $L_x \cup \{x\}$ such that $N_v \cap N_w = \emptyset$ for distinct $v, w \in L_x$.

Call S_x *useless* if it is not useful. Fix a set A_x of $\lfloor \delta/2 \rfloor$ neighbours of x in G that lie outside X. For each $v \in A_x$ fix a set V_v of $\lceil \delta/2 \rceil$ neighbours of v in

G outside $A_x \cup \{x\}$. Let G_x denote the bipartite subgraph of G whose vertex classes are A_x and $B_x := \bigcup_{v \in A_x} V_v$ and in which each vertex $v \in A_x$ is joined to precisely the vertices in V_v. So $e(G_x) = \lceil \delta/2 \rceil |A_x|$. Denote by B_x^1 the set of all vertices in B_x whose degree in G_x is at most f^2 and let $B_x^2 := B_x \setminus B_x^1$.

We now claim that $e_{G_x}(A_x, B_x^2) \le e(G_x)/2$. Suppose not. Then on average each vertex in A_x has at least $\delta/4$ neighbours in B_x^2. Since $(A_x, B_x^2)_{G_x}$ does not contain a $K_{s-1,t}$ with t vertices in A_x and $s-1$ vertices in B_x^2 (such a $K_{s-1,t}$ would yield a $K_{s,t}$ in G together with x), Lemma 4.6 implies that

$$|A_x| \binom{\delta/4}{s-1} \le t \binom{|B_x^2|}{s-1}.$$

As $|A_x| \ge \lfloor \delta/2 \rfloor \ge \delta/4$ it follows that

$$|B_x^2| \ge \frac{\delta |A_x|^{\frac{1}{s-1}}}{8t} \ge \frac{\delta^{1+\frac{1}{s-1}}}{32t}. \tag{5.9}$$

On the other hand, we have $f^2|B_x^2| \le e_{G_x}(A_x, B_x^2) \le \delta^2$, and thus

$$|B_x^2| \le \frac{\delta^2}{f^2} \le \frac{\delta^{1+\frac{1}{s-1}}}{(2\log r)^2},$$

contradicting (5.9). So $e_{G_x}(A_x, B_x^1) \ge e(G_x)/2$. Let A_x' be the set of all those vertices in A_x which have at least $\delta/8$ neighbours in G_x that lie inside B_x^1. Then

$$|A_x'|\delta + |A_x|\delta/8 \ge e_{G_x}(A_x, B_x^1) \ge e(G_x)/2 \ge |A_x|\delta/4$$

and thus $|A_x'| \ge |A_x|/8$. We claim that S_x is useful if $|A_x' \cap L(S_x)| \ge k$ and if in G_x each vertex in B_x^1 has at most $7f$ neighbours lying inside $L(S_x)$. To see this, apply Corollary 5.8 with $A := A_x' \cap L(S_x)$, $B := B_x^1$, $d_A := \delta/8$ and $d_B := 7f$ to the graph $(A,B)_{G_x}$ to obtain $|A_x' \cap L(S_x)| \ge k$ disjoint stars with centres in $A_x' \cap L(S_x)$ and such that each star has at least $\lfloor \delta/56f \rfloor$ leaves. Since $\lfloor \delta/56f \rfloor \ge k$ by (5.8), we can take for the set L_x in the definition of a useful star S_x any set of k centres of these stars. Hence S_x is useful.

So it remains to estimate the probability that $|A_x' \cap L(S_x)| \le k$ or that B_x^1 contains a vertex with more than $7f$ neighbours in $L(S_x)$. As each vertex in $A_x \supseteq A_x'$ chooses x with probability $1/f$, we have

$$\mathbb{E}(|A_x' \cap L(S_x)|) = |A_x'|/f \ge |A_x|/8f \ge \delta/32f \overset{(5.8)}{\ge} 2k.$$

Together with inequality (5.4) of Lemma 5.7 this implies

$$\mathbb{P}(|A_x' \cap L(S_x)| \le k) \le 2\mathrm{e}^{-2k/12} \le 1/4. \tag{5.10}$$

Furthermore, the definition of B_x^1 implies that for every vertex $b \in B_x^1$

$$\mathbb{E}(|N_{G_x}(b) \cap L(S_x)|) \le f,$$

and thus from (5.5) it follows that

$$\mathbb{P}(|N_{G_x}(b) \cap L(S_x)| \ge 7f) \le \mathrm{e}^{-7f}.$$

Hence

$$\mathbb{P}(\exists\, b \in B_x^1 \text{ with } |N_{G_x}(b) \cap L(S_x)| \geq 7f) \leq \mathrm{e}^{-7f}|B_x^1| \leq \mathrm{e}^{-7f}\delta^2 \leq 1/4.$$

Together with (5.10) this implies that with probability at most $1/2$ the star S_x is useless. Hence the expected number of useless stars S_x is at most $|X|/2$, and therefore for some outcome at least $|X|/2 \geq k$ of the stars S_x $(x \in X)$ are useful. For each such S_x let $S'_x \subseteq S_x$ be the star whose centre is x and whose leaves are the vertices in a set L_x satisfying (a) and (b). Then the S'_x are stars as required in the lemma. □

Proof of Theorem 5.2. Throughout the proof we assume that r is sufficiently large compared with s and t. Let k be as defined in (5.6) and put

$$\ell := \left\lfloor \frac{k}{(\log r)^{4+5\kappa}} \right\rfloor \overset{(5.2)}{\geq} \frac{r^{\frac{1}{2}+\frac{1}{2(s-1)}}}{(\log r)^{12}}.$$

We will show that G contains a subdivision of K_ℓ. First we apply Lemma 4.10 to G. Since by Theorem 5.1 every graph of average degree r^3 contains a subdivision of a complete graph of order $r \geq \ell$, we may assume that the lemma returns a subgraph G' with $\delta(G') \geq \frac{r}{400t\log r}$ and $\Delta(G') \leq r$. Apply Corollary 5.4 to G' to obtain a positive constant $\varepsilon_0 = \varepsilon_0(\kappa) < 1$ and a subgraph H which is a $(\kappa, d(H), \varepsilon_0)$-expander and satisfies $d := d(H) \geq d(G')/2$ and $\delta(H) \geq d(H)/2$. Since $72d^2 \geq r \geq \Delta(H)$, Theorem 5.6 with $\alpha := 8$ shows that H contains a subdivision of a complete graph of order at least $cd \geq \ell$, provided that $\log|H| \geq (\log d)^8$. Thus, setting $n := |H|$, we may assume that

$$\log n < (\log d)^8 \leq (\log r)^8. \tag{5.11}$$

Apply Lemma 5.9 to H to obtain k disjoint stars as described there. Pick ℓ of these stars, $S_1, \dots, S_\ell$ say. For all leaves v of S_i fix a set N_v^i satisfying condition (ii) of Lemma 5.9 and let $A_i := \bigcup_{v \in L(S_i)} N_v^i$. So $|A_i| = k^2$. The branch vertices of our subdivision of K_ℓ in G will be the centres of the S_i and each edge ij of K_ℓ will correspond to a path joining a leaf of S_i to a leaf of S_j. We will find disjoint such paths as follows. For each edge $ij \in K_\ell$ in turn we use Lemma 5.5 to find a short A_i–A_j path in the graph obtained from H by deleting $S_1, \dots, S_\ell$ as well as all previously constructed paths. We have to take care that for every vertex $i \in K_\ell$ the paths that correspond to the edges of K_ℓ which are incident with i start in distinct sets $N_v^i \subseteq A_i$ and thus can be joined by independent edges to distinct leaves of S_i. Thus when defining the A_i–A_j path corresponding to the edge $ij \in K_\ell$, we will also delete all those sets N_v^i from H which contain the starting point of a previously constructed A_i–$A_{j'}$ path; and similarly for j.

More formally, we proceed as follows. Fix an enumeration $i_1j_1, \dots, i_{\binom{\ell}{2}}j_{\binom{\ell}{2}}$ of the edges of K_ℓ. We will show that for all $b \leq \binom{\ell}{2}$ there is a path P_b whose length is at most $2 + 8(\log n)^{1+\kappa}/\varepsilon_0 =: \mathrm{diam}$, which joins a leaf of S_{i_b} to a leaf of S_{j_b}, has no inner vertices in $\bigcup_{i=1}^{\ell} S_i$ and such that the P_b are disjoint for distinct $b \leq \binom{\ell}{2}$. Suppose inductively that for some $a \geq 1$ we have already defined P_b for all $b < a$.

To find P_a, let N_{i_a} be the union of all those $N_v^{i_a}$ for which the leaf v of S_{i_a} is an endpoint of a path P_b constructed previously. (In other words, these paths P_b are precisely the previously constructed paths that correspond to an edge of K_ℓ incident with i_a, i.e. for which $i_a \in \{i_b, j_b\}$.) Define N_{j_a} similarly. Let Z be the set consisting of the vertices in $N_{i_a} \cup N_{j_a}$ together with all vertices lying in some S_i ($i \le \ell$) and all vertices on the paths P_b already constructed. So

$$\begin{aligned} |Z| &\le 2\ell k + \ell(k+1) + (\mathrm{diam}+1) \cdot \binom{\ell}{2} \le 4\ell k + \frac{5\ell^2 (\log n)^{1+\kappa}}{\varepsilon_0} \\ &\overset{(5.11)}{\le} 4\ell k + \frac{5\ell^2 (\log r)^{8+8\kappa}}{\varepsilon_0} \le \frac{6k^2}{\varepsilon_0 (\log r)^{2\kappa}}. \end{aligned} \tag{5.12}$$

Let $A'_{i_a} := A_{i_a} \setminus Z$ and define A'_{j_a} similarly. Then

$$|A'_{i_a}| \ge |A_{i_a}| - |Z| \overset{(5.12)}{\ge} k^2 - \frac{6k^2}{\varepsilon_0 (\log r)^{2\kappa}} \ge \frac{k^2}{2}. \tag{5.13}$$

Let ε be as defined in (5.3). Using that $\varepsilon(x)x$ is monotone increasing for all $x \ge d/2$ and thus for $x \ge k^2/2$, it is easy to check that (5.12) and (5.13) imply $|Z| \le |A'_{i_a}| \varepsilon(|A'_{i_a}|)/4$. Similarly it follows that $|Z| \le |A'_{j_a}| \varepsilon(|A'_{j_a}|)/4$. Thus we may apply Lemma 5.5 to obtain an A'_{i_a}–A'_{j_a} path P in $H - Z$ of length at most

$$\frac{8(\log(8n/d))^{1+\kappa}}{\varepsilon_0} \le \mathrm{diam} - 2.$$

The definition of A'_{i_a} implies that the endpoint of P in A'_{i_a} can be joined by an edge to some leaf of S_{i_a} which is not already an endpoint of a path P_b constructed previously. The same is true for j_a. Altogether this shows that the S_{i_a}–S_{j_a} path P_a obtained from P in this way has the required properties. □

5.4 Upper bounds

The following proposition shows that the existence of sufficiently dense $K_{s,t}$-free graphs would imply that the bound (5.1) in Theorem 5.2 is best possible up to the logarithmic factor.

Proposition 5.10 *For every $c > 0$ and all $t \ge s \ge 2$ there is a constant $C = C(c, s, t)$ such that no $K_{s,t}$-free graph G with $e(G) \ge c|G|^{2-1/s}$ contains a subdivision of a complete graph of order at least $Cd(G)^{\frac{1}{2}+\frac{1}{2(s-1)}}$.*

Proof. We will show that $C := (16t)^s / c^{\frac{1}{2}+\frac{1}{2(s-1)}}$ works. Let $n := |G|$ and $r := d(G)$. Clearly, we may assume that G contains a subdivision TK_ℓ of K_ℓ for some $\ell \ge (16t)^s$. Recall that by Theorem 4.7, every subgraph H of G has at most $t|H|^{2-1/s}$ edges. In particular, the subgraph of G induced by the branch vertices of TK_ℓ contains at most $t\ell^{2-1/s} \le \ell^2/16 \le e(K_\ell)/4$ edges. So at least $3/4$ of the edges of K_ℓ correspond to paths in TK_ℓ of length at least two. Thus

$n \geq 3e(K_\ell)/4 \geq \ell^2/4$. On the other hand, our assumption on G implies that $n \leq (r/c)^{1+1/(s-1)}$. Hence

$$\ell \leq 2\left(\frac{r}{c}\right)^{\frac{1}{2}+\frac{1}{2(s-1)}},$$

as required. □

It is widely believed that $K_{s,t}$-free graphs as in the statement of Proposition 5.10 do exist (Conjecture 4.11).

For $s \geq 4$ the best known lower bound on the maximum number of edges of a $K_{s,s}$-free graph G is $c|G|^{2-2/(s+1)}$ (see e.g. [12, Ch. VI, Thm. 2.10]). Using this bound, the proof of Proposition 5.10 still yields an upper bound of $C'r^{\frac{1}{2}+\frac{1}{s-1}}$ for the order of the complete topological minor in Theorem 5.2.

Chapter 6

Induced subdivisions in $K_{s,s}$-free graphs of large average degree

6.1 Introduction

In Chapter 5 we showed that $K_{s,s}$-free graphs contain significantly larger cliques as subdivisions than those guaranteed by their average degree. Here we show that if we consider $K_{s,s}$-free graphs as host graphs, then we can require our subdivisions to be induced:

Theorem 6.1 *For every graph H and every $s \in \mathbb{N}$ there exists $d = d(H, s)$ such that every graph G of average degree at least d contains either a $K_{s,s}$ as a subgraph or an induced subdivision of H.*

Of course, Theorem 6.1 becomes false if we replace the $K_{s,s}$ by some non-bipartite graph G^* since then complete bipartite graphs would be G^*-free but they do not contain an induced subdivision of a path of length three. Moreover, one cannot replace 'subdivision' by 'subgraph', as for example there exist graphs which have both arbitrarily large average degree and arbitrarily large girth. On the other hand, Kierstead and Penrice [45] proved that if H is a tree then one can indeed find it as an induced subgraph in any $K_{s,s}$-free graph of sufficiently large average degree. They used this result to prove a special case of the conjecture of Gyárfás [36] and Sumner [96] that given a tree T and $s \in \mathbb{N}$, every K_s-free graph of sufficiently large chromatic number contains an induced copy of T. Scott [95] proved that this conjecture becomes true if we only require an induced subdivision of T. In [95] he also proposed a conjecture which is analogous to Theorem 6.1—replacing 'average degree' by 'chromatic number' and $K_{s,s}$ by K_s. In fact, Theorem 6.1 was motivated by this conjecture.

We now briefly outline the organization of this chapter and the strategy of our proof of Theorem 6.1. Consider a $K_{s,s}$-free graph G of large average degree. In Section 6.2 we prepare the ground for the proof by collecting some tools which we will need later on. In particular, it turns out that in order to find an induced subdivision of H in G, it suffices to prove the following

Theorem 6.2 *For all $k, s \in \mathbb{N}$ there exists $d = d(s,k)$ such that every $K_{s,s}$-free graph G of average degree at least d contains an induced subdivision of some graph H^* where the average degree of H^* is at least k and every edge of H^* is subdivided exactly once.*

We will call such a subdivision an induced *1-subdivision* of H^*. Note that both the set B of branch vertices and the set S of subdividing vertices have to be independent in G. The first step towards finding such a 1-subdivision of H^* is to find a large independent set I in G (Section 6.3). Ideally, we would like to find another independent set B^* such that the bipartite subgraph between I and B^* has large average degree. In this case, one can find B in the smaller of B^* and I and S in the larger of the two. Unfortunately, we cannot guarantee that such a set B^* always exists. However, in Section 6.4 we will show that one can come fairly close: we will find sets $I^* \subseteq I$ and B^* such that the bipartite subgraph between I^* and B^* has large average degree and $G[B^*]$ has small chromatic number. In Section 6.5, which constitutes the core of our proof, we then show how to find our induced 1-subdivision of H^* within $G[I^* \cup B^*]$. In Section 6.6 we then put everything together to complete the proof of Theorem 6.2 (and thus of Theorem 1). In the final section we mention some open problems.

Theorem 6.2 also implies induced analogues of a result of Thomassen on subdivisions and of a result of Häggkvist and Scott on cycles in graphs: Thomassen [103] proved that for every $k, \ell \in \mathbb{N}$ there exists $f = f(k,\ell)$ such that every graph of minimum degree at least f contains a subdivision of some graph H with minimum degree at least k in which every edge is subdivided exactly ℓ times. Combined with Theorem 6.2 this gives the following analogue for odd integers ℓ:

Corollary 6.3 *For all $k, s \in \mathbb{N}$ and every odd integer ℓ there exists $g = g(k,\ell,s)$ such that every $K_{s,s}$-free graph of minimum degree at least g contains an induced subdivision of some graph H with minimum degree at least k in which every edge is subdivided exactly ℓ times.*

Häggkvist and Scott [38] proved that every graph of minimum degree at least $300k^2$ contains k cycles of consecutive even lengths. (Verstraëte [105] improved the bound on the minimum degree to a linear one.) Applying this result to the graph H^* provided by Theorem 6.2 we obtain k induced cycles in G which are twice as long. In particular, we have

Corollary 6.4 *For all $k, s \in \mathbb{N}$ there exists $g = g(k,s)$ such that every $K_{s,s}$-free graph of minimum degree at least g contains k induced cycles whose lengths form an arithmetic progression.*

6.2 Notation and tools

In this chapter, all logarithms are base two. A *1-subdivision* of a graph H is the graph obtained from H by replacing the edges of H with internally disjoint paths of length two. If we say that a bipartite graph (A', B') is a subgraph of (A, B) then we tacitly assume that $A' \subseteq A$ and $B' \subseteq B$. We shall frequently consider the following class of graphs.

Definition. Given non-negative numbers d, i and $k \leq d/4$, we say that a bipartite graph (A, B) is a (d, i, k)*-graph* if $|A| \geq d^{12i}|B|$ and $d/4 - k \leq d(a) \leq 4d$ for all vertices $a \in A$. (Note that the order of A and B matters here.)

We now list some results which we need later on in the proof of Theorem 6.1. Since every graph of sufficiently large average degree contains a subdivision of K_r (see Chapter 5), Theorem 6.1 is a consequence of Theorem 6.2. Indeed, if the average degree of the graph H^* provided by Theorem 6.2 is sufficiently large, then H^* contains a subdivision of H; and it is easily checked that the corresponding subdivision of H in G is induced.

We shall frequently use the following simple observation. A proof is for example included in [28, Cor. 5.2.3].

Proposition 6.5 *Every graph G contains an induced subgraph of minimum degree at least $\chi(G) - 1$.*

Clearly, it suffices to prove Theorem 6.2 for graphs G which do not have subgraphs of average degree $> d(G)$. So the Propositions 2.1 and 6.5 enable us to assume that $\delta(G) \geq d(G)/2$ and $\chi(G) \leq d(G) + 1$.

The next lemma is a special case of Chernoff's inequality (see e.g. [7, Thm. A.1.12 and A.1.13]).

Lemma 6.6 *Let $X_1, \dots, X_n$ be independent 0-1 random variables with $\mathbb{P}(X_i = 1) = p$ for all $i \leq n$, and let $X := \sum_{i=1}^{n} X_i$. Then $\mathbb{P}(X \geq 2\mathbb{E}X) \leq (4/\mathrm{e})^{-\mathbb{E}X}$ and $\mathbb{P}(X \leq \mathbb{E}X/2) \leq \mathrm{e}^{-\mathbb{E}X/8}$.*

One case which arises in our proof of Theorem 6.2 is that we first find an induced bipartite subgraph (A, B) of large average degree in G and then find an induced subdivision of H in (A, B). To carry out this second step, it will turn out to be useful if the vertices in A have almost the same degree and $|B|$ is much smaller than $|A|$. The following lemma shows that by replacing (A, B) with an induced subgraph we can always satisfy these two additional conditions. The lemma is a slight extension of [93, Lemma 2.4]. Although the proof is almost the same, we include it here for completeness.

Lemma 6.7 *Let $r \geq 2^6$, $s \geq 1$ and $d \geq 8r^{12s+1}$. Then every bipartite graph of average degree d contains an induced copy of an $(r, s, 0)$-graph.*

Proof. Clearly, we may assume that our given bipartite graph has no subgraph of average degree $> d$. So by Proposition 2.1 this graph contains an induced subgraph $G = (A, B)$ such that $\delta(G) \geq d/2$, $d(G) = d$ and $|A| \geq |B|$. Thus at least half of the vertices of A have degree at most $2d$ in G. So, writing A' for the set of all vertices in A of degree at most $2d$, we have $|A'| \geq |A|/2 \geq |B|/2$.

Let us now consider a random subset B_p of B which is obtained by including each vertex of B independently with probability $p := r/d$. For every $a \in A'$ let $X_a := |N_G(a) \cap B_p|$. Then $r/2 \leq \mathbb{E}X_a \leq 2r$. Given B_p, let us call $a \in A'$ *useful* if $r/4 \leq X_a \leq 4r$. Lemma 6.6 implies that

$$\mathbb{P}(a \text{ is not useful}) \leq \mathbb{P}(X_a \geq 2\mathbb{E}X_a) + \mathbb{P}(X_a \leq \mathbb{E}X_a/2) \leq (4/\mathrm{e})^{-r/2} + \mathrm{e}^{-r/16} \leq \frac{1}{4}.$$

Hence the expected number of vertices in A' which are not useful is at most $|A'|/4$. So Markov's inequality (which states that $\mathbb{P}(X \geq c\,\mathbb{E}X) \leq 1/c$ for every $c \geq 1$) implies that

$$\mathbb{P}(\text{at least half of the vertices in } A' \text{ are not useful}) \leq \frac{1}{2}.$$

Moreover, using Lemma 6.6 again,

$$\mathbb{P}(|B_p| \geq 2p|B|) = \mathbb{P}(|B_p| \geq 2\mathbb{E}|B_p|) \leq (4/\mathrm{e})^{-p|B|} \leq \frac{1}{4}.$$

So the probability that both $|B_p| \leq 2p|B|$ and that at least half of the vertices in A' are useful is at least $1/2 - 1/4 > 0$. Hence there exists a choice B^* for B_p which has these two properties. Let A^* be the set of useful vertices in A'. Then $r/4 \leq d_{(A^*,B^*)_G}(a) \leq 4r$ for every vertex $a \in A^*$ and

$$|A^*| \geq \frac{|A'|}{2} \geq \frac{|B|}{4} \geq \frac{|B^*|}{8p} = \frac{d|B^*|}{8r} \geq r^{12s}|B^*|.$$

Thus $(A^*, B^*)_G$ is an induced $(r, s, 0)$-subgraph of G. □

6.3 Independent sets

Clearly, every graph G of maximum degree Δ has an independent set of size at least $|G|/\chi(G) \geq |G|/(\Delta + 1)$. Lemma 6.8 shows that we obtain a small but significant improvement if G is $K_{s,s}$-free. The proof is based on Alon's elegant proof of the result that any triangle-free graph H of maximum degree Δ contains an independent set of size $c|H| \log \Delta/\Delta$ (see e.g. [7], the result itself is due to Ajtai, Komlós and Szemerédi [2]).

Alternatively, we could have applied another result from [2]: for all ε there exists a constant c_0 so that every graph with maximum degree at most Δ which contains at most $|G|\Delta^{2-\varepsilon}$ triangles has an independent set of size at least $c_0|G| \log \Delta/\Delta$. But Theorem 4.7 implies that in a $K_{s,s}$-free graph G the neighbourhood of any vertex x can span at most $sd(x)^{2-1/s} \leq s\Delta^{2-1/s}$ edges and thus G contains at most $s|G|\Delta^{2-1/s}$ triangles. Although the proof of Lemma 6.8 given below yields a weaker bound, it is simpler and has the advantage of being self-contained.

Lemma 6.8 *For every $s \in \mathbb{N}$ there exists $c' = c'(s)$ such that for each $\Delta \geq 9$ every $K_{s,s}$-free graph G of maximum degree at most Δ has an independent set of size at least*

$$f := c'|G| \frac{(\log \Delta)^{1/s}}{\Delta \log \log \Delta}.$$

Proof. Let $n := |G|$. Let I be an independent set chosen uniformly at random from all independent sets of G. For every vertex $x \in G$ define

$$Z_x := \begin{cases} \Delta & \text{if } x \in I; \\ |N(x) \cap I| & \text{otherwise.} \end{cases}$$

Then

$$\sum_{x\in G} Z_x = \sum_{x\in I} Z_x + \sum_{x\notin I} Z_x \le \Delta|I| + e(I, V(G)\setminus I) \le 2\Delta|I|.$$

So it suffices to show that $\mathbb{E}(\sum_{x\in G} Z_x) \ge 2\Delta f$. Given any vertex $x \in G$, let $I_x := I \setminus (N(x) \cup \{x\})$. Rather than directly showing that $\mathbb{E}(\sum_{x\in G} Z_x)$ is large, we will show that $\mathbb{E}(Z_x|I_x)$ is large for every vertex x and every I_x.

Let N_x be the set of all neighbours of x which are not adjacent to a vertex in I_x. We will now show that if N_x is large then the average size of an independent subset of N_x is large as well. So suppose first that $|N_x| \ge 2$. Since $G[N_x]$ is $K_{s,s}$-free, it follows from Theorem 4.7 that every subgraph H of $G[N_x]$ has average degree at most $2s|H|^{1-1/s} \le 2s|N_x|^{1-1/s}$. Thus by Proposition 6.5 we have that $\chi(G[N_x]) \le 2s|N_x|^{1-1/s} + 1 \le 4s|N_x|^{1-1/s}$. So $G[N_x]$ has an independent set of size at least $|N_x|^{1/s}/(4s) =: \alpha$. Hence $G[N_x]$ contains at least $2^\alpha/2$ independent sets of size at least $\alpha/2$. Put $\beta := \alpha/(4\log|N_x|)$. Then the number of independent subsets of N_x of size at most β is at most

$$\binom{|N_x|}{0} + \cdots + \binom{|N_x|}{\lfloor\beta\rfloor} \le |N_x|^{2\beta} = 2^{2\beta\log|N_x|} = 2^{\alpha/2}.$$

If $|N_x| \ge (8s)^s$ then $2^\alpha/2 \ge 2^{\alpha/2}$ and $\alpha/2 \ge 2\beta$; and so in this case the average size ℓ_x of an independent subset of N_x is at least β.

Now note that, writing k_x for the number of independent sets in N_x, for every $|N_x| \ge 0$ we have

$$\mathbb{E}(Z_x|I_x) \ge \frac{\Delta + k_x\ell_x}{1+k_x} \ge \frac{\Delta}{2k_x} + \frac{\ell_x}{2}.$$

Thus, if $|N_x| \ge (\log\Delta)/2$ and if c' is sufficiently small compared with s, then

$$\mathbb{E}(Z_x|I_x) \ge \frac{\ell_x}{2} \ge \frac{\beta}{2} \ge \frac{|N_x|^{1/s}}{32s\log|N_x|} \ge \frac{2c'(\log\Delta)^{1/s}}{\log\log\Delta},$$

while if $0 \le |N_x| \le (\log\Delta)/2$ then

$$\mathbb{E}(Z_x|I_x) \ge \frac{\Delta}{2\cdot 2^{|N_x|}} \ge \frac{\Delta}{2\cdot 2^{(\log\Delta)/2}} = \frac{\sqrt{\Delta}}{2} \ge \frac{2c'(\log\Delta)^{1/s}}{\log\log\Delta}.$$

Hence we have $\mathbb{E}(Z_x) \ge 2\Delta f/n$ and so $\mathbb{E}(\sum_{x\in G} Z_x) = \sum_{x\in G}\mathbb{E}(Z_x) \ge 2\Delta f$, which completes the proof. □

Corollary 6.9 *For every $s \in \mathbb{N}$ there exists $d_0 = d_0(s)$ such that every $K_{s,s}$-free graph G of average degree $d \ge d_0$ contains an independent set of size at least $|G|(\log d)^{1/(s+1)}/d$.*

Proof. Let G' be the subgraph of G induced by the vertices of degree at most $2d$. Clearly, $|G'| \ge |G|/2$. If d is sufficiently large, then by Lemma 6.8, G' (and thus G) has an independent set of size at least $|G|(\log d)^{1/(s+1)}/d$. □

6.4 Finding a 'nearly' induced bipartite subgraph of large average degree

As remarked in the introduction, we would like to find an induced bipartite subgraph of large average degree in our original graph G. The aim of this section is to prove that if G does not contain such a subgraph, we can still come close to it: by Corollary 6.9 we may assume that G contains a large independent set I. We will use this to find a subgraph (A, B) of large average degree so that $A \subseteq I$ (so A is independent) and B has small chromatic number and is much smaller than A. The following lemma shows how to construct one colour class of B.

Lemma 6.10 *Let I be an independent set in a graph G such that $d(x) \geq d/2$ for every $x \in I$ and $|I| = 2c|G|/d$ for some $c \geq 2$. Suppose that $\chi(G) \leq 3d$. Then G has one of the following properties.*

(i) G contains an induced bipartite subgraph whose average degree is at least $(\log c)/24$.

(ii) There are a set $I' \subseteq I$ and an independent set J in $G - I$ such that in G every vertex of I' has exactly one neighbour in J, $|J| \leq |I| \log c/c$ and $|I'| \geq |I|/4(\log c)^2$.

Proof. Put $n := |G|$, $\overline{I} := V(G) \setminus I$ and let Y be the set of all vertices in $\overline{I}$ which have at least $c/2$ neighbours in I. Then $e(I, \overline{I} \setminus Y) \leq c|\overline{I} \setminus Y|/2 \leq cn/2$. On the other hand the degree of every vertex in I is at least $d/2$, and so we have that $e(I, \overline{I}) \geq cn$. Thus $e(I, Y) \geq cn/2$. As $\chi(G) \leq 3d$, there exists an independent set $A \subseteq Y$ such that

$$e(I, A) \geq \frac{e(I, Y)}{3d} \geq \frac{cn}{6d} = \frac{|I|}{12}. \tag{6.1}$$

Note also that

$$\frac{c}{2} \cdot |A| \leq e(I, A). \tag{6.2}$$

We may assume that the average degree of $(I, A)_G$ is at most $(\log c)/2$ (otherwise $(I, A)_G$ would be as desired in (i)). Since every vertex in A has at least $c/2 \geq (\log c)/2$ neighbours in I, this implies that $|I| \geq |A|$. Therefore

$$\frac{c}{2} \cdot |A| \leq e(I, A) = \frac{1}{2} \cdot d((I, A)_G)(|I| + |A|) \leq \frac{\log c}{2} \cdot |I|,$$

and hence

$$|A| \leq \frac{|I| \log c}{c}. \tag{6.3}$$

Using a probabilistic argument, we will show that there exist sets $J \subseteq A$ and $I' \subseteq I$ as desired in (ii). To make this work, we first need to replace I with the set $I_1 \subseteq I$ of all vertices which have at least one and at most $\log c$ neighbours in A. So let us first estimate the size of I_1. Denote by I_2 the set of all vertices in I which have no neighbours in A and put $I_3 := I \setminus (I_1 \cup I_2)$. We will show that we may assume that both $e(I_1, A) \geq e(I, A)/2$ and $|I_1| \geq |I|/\log c$. Suppose to the contrary that $e(I_1, A) \leq e(I, A)/2$. Then $e(I_3, A) \geq e(I, A)/2$ and so (6.2) implies that $e(I_3, A) \geq c|A|/4$. Thus on average, a vertex in A has at least $c/4$ neighbours in I_3. As every vertex in I_3 has at least $\log c$ neighbours in A, it follows that $(I_3, A)_G$ is as desired in (i). Hence we may assume that $e(I_1, A) \geq e(I, A)/2$. Next suppose that $|I_1| \leq |I|/\log c$. Then

$$e(I_1, A) \geq e(I, A)/2 \overset{(6.1)}{\geq} |I|/24 \geq |I_1|(\log c)/24$$

and

$$e(I_1, A) \geq e(I, A)/2 \overset{(6.2)}{\geq} c|A|/4.$$

Thus $(I_1, A)_G$ is as desired in (i). Therefore we may also assume that $|I_1| \geq |I|/\log c$.

Let us now consider a random subset A_p of A which is obtained by including each $a \in A$ independently with probability $p := 1/(2\log c)$. Call a vertex $x \in I_1$ *useful* if it has exactly one neighbour in A_p. Using the definition of I_1 it follows that for every $x \in I_1$

$$\begin{aligned}\mathbb{P}(x \text{ is useful}) &= |N(x) \cap A| \cdot p \cdot (1-p)^{|N(x)\cap A|-1} \geq 1 \cdot p \cdot (1-p)^{\lfloor \log c \rfloor} \\ &\geq p(1 - p\lfloor \log c \rfloor) \geq p/2.\end{aligned}$$

(The second inequality can be easily proved by induction.) Hence the expected number of useful vertices in I_1 is at least $p|I_1|/2$. So there exists a choice J for A_p such that at least $p|I_1|/2$ vertices in I_1 are useful. Let I' be the set of these useful vertices. Then

$$|I'| \geq \frac{p|I_1|}{2} = \frac{|I_1|}{4\log c} \geq \frac{|I|}{4(\log c)^2}$$

and

$$|J| \leq |A| \overset{(6.3)}{\leq} \frac{|I|\log c}{c}.$$

So I' and J are as desired in (ii). □

By repeated applications of Lemma 6.10 we obtain the following result.

Lemma 6.11 *Let $c \geq 2^{512}$, $d > 2c$ and let G be a graph of minimum degree at least $d/2$. Suppose that $\chi(G) \leq d+1$ and that G has an independent set I of size $2c|G|/d$. Put $r := \lfloor \log\log c \rfloor$. Then G has one of the following properties.*

- *(i) G contains an induced bipartite subgraph whose average degree is at least $(\log c)/48$.*
- *(ii) There are a set $I^* \subseteq I$ and disjoint independent subsets $J_1, \dots, J_r$ of $G - I^*$ such that every vertex of I^* has exactly one neighbour in each J_k, $|I^*| \geq |I|/4^r(\log c)^{2r}$ and $|J_k| \leq 4|I|\log c/c$ for every $k \leq r$.*

Proof. The proof follows from r applications of Lemma 6.10. Indeed, let $I_0 := I$ and suppose inductively that for some $0 \le \ell < r$ we already have obtained a set $I_\ell \subseteq I$ and disjoint independent sets $J_1, \dots, J_\ell$ in $G - I_\ell$ such that every vertex of I_ℓ has exactly one neighbour in each J_k, $|I_\ell| = \lceil |I|/4^\ell (\log c)^{2\ell} \rceil$ and $|J_k| \le 4|I| \log c/c$ for every $1 \le k \le \ell$. Put $n := |G|$, $G' := G - (J_1 \cup \dots \cup J_\ell)$, $n' := |G'|$ and $d' := d/2$. Thus $d_{G'}(x) \ge d/2 - \ell \ge d/4 = d'/2$ for every $x \in I_\ell$. Moreover, since $|J_k| \le 4n \log c/c$, we have that $n' \ge n/2$. Let c' be defined by $|I_\ell| = 2c'n'/d'$. Using $|I_\ell| \le |I|$ it follows that $c' \le c$. On the other hand

$$\frac{|I|}{4^\ell (\log c)^{2\ell}} \le |I_\ell| = \frac{2c'n'}{d'} \le \frac{4c'n}{d},$$

and so

$$c' \ge \frac{c}{2 \cdot 4^\ell (\log c)^{2\ell}} = \frac{c}{2(2\log c)^{2\ell}}.$$

In particular, $c' \ge 2$. Since also $\chi(G_\ell) \le d+1 \le 3d'$, we may apply Lemma 6.10 to the graph G' and the independent set I_ℓ. As

$$\frac{\log c'}{24} \ge \frac{\log c - 1 - \log((2\log c)^{2\ell})}{24} \ge \frac{\log c - 1 - 2r\log(2\log c)}{24} \ge \frac{\log c}{48}$$

we may assume that we have $I_{\ell+1} \subseteq I_\ell$ and $J_{\ell+1}$ satisfying condition (ii) of Lemma 6.10. Hence

$$|I_{\ell+1}| \ge \frac{|I_\ell|}{4(\log c')^2} \ge \frac{|I_\ell|}{4(\log c)^2} \ge \frac{|I|}{4^{\ell+1}(\log c)^{2(\ell+1)}},$$

and

$$|J_{\ell+1}| \le \frac{|I_\ell| \log c'}{c'} \le \frac{2 \cdot 4^\ell \cdot |I_\ell| (\log c)^{2\ell+1}}{c} \le \frac{4|I| \log c}{c}.$$

Note that we may assume that $|I_{\ell+1}| = \lceil |I|/4^{\ell+1}(\log c)^{2(\ell+1)} \rceil$ by making $I_{\ell+1}$ smaller if necessary. This completes the induction step. $\square$

Corollary 6.12 *For every $s \in \mathbb{N}$ there exists $c(s)$ such that the following holds. Let $c \ge c(s)$, $d > 2c$ and let G be a graph of minimum degree at least $d/2$. Suppose that G has an independent set I of size $2c|G|/d$ and that $\chi(G) \le d+1$. Put $r := \lfloor \log\log c \rfloor$. Then G has one of the following properties.*

(i) G contains an induced bipartite subgraph whose average degree is at least $(\log c)/48$.

(ii) There are disjoint vertex sets $A, B \subseteq V(G)$ such that A is independent, $\chi(G[B]) \le r$ and $(A, B)_G$ is an $(r, s, 0)$-graph.

Proof. Applying Lemma 6.11 we may assume that G contains independent sets I^* and $J_1, \dots J_r$ satisfying condition (ii) of Lemma 6.11. Let $A := I^*$ and $B := J_1 \cup \dots \cup J_r$. Clearly, every vertex of A has degree r in the bipartite graph $(A, B)_G$ and $\chi(G[B]) \le r$. Thus it remains to show that $|A| \ge r^{12s}|B|$. But

$$\frac{|A|}{|B|} \ge \frac{c}{4^{r+1} r (\log c)^{2r+1}} \ge r^{12s},$$

if c is sufficiently large. $\square$

6.5 Finding an induced 1-subdivision of a graph of large average degree

In the previous section we showed that we may assume that our original graph G contains a bipartite subgraph (A, B) of large average degree such that A is independent in G and $G[B]$ has small chromatic number (or is possibly independent as well). In this section we will show that this (A, B) contains a 1-subdivision of some graph H^* where H^* has large average degree and this 1-subdivision is induced in G.

To accomplish this, we first find a 1-subdivision of some graph H' of large average degree in (A, B) (Corollary 6.14). The branch vertices of this 1-subdivision are vertices in B, its subdivided edges are paths of length two in (A, B) and so the midpoints of the subdivided edges are vertices in A. In Lemma 6.15 we then show how to find a subgraph H'' of H' for which every midpoint of a subdivided edge is joined in G only to the two endpoints of this edge and to no other branch vertex. As A is independent, it follows that every edge of G which prevents the 1-subdivision of H'' from being induced must join two branch vertices, i.e. two vertices in B. So if B is also independent then this 1-subdivision is induced in G, as desired. The case when B is not independent is more difficult and dealt with in Lemma 6.17.

Let us now introduce some notation. A path P of length two in a bipartite graph (A, B) is called a *hat* of G if it begins and ends in B. A set $\mathcal{H}$ of hats of (A, B) is *uncrowded* if any two hats in $\mathcal{H}$ join distinct pairs of vertices and have distinct midpoints. (So the sets of subdivided edges of the 1-subdivisions of the graphs H' and H'' described above are both uncrowded; and conversely, an uncrowded set of hats can serve as the set of subdivided edges of a 1-subdivision whose set of branch vertices is B.)

Lemma 6.13 *Let $r, i \geq 1$ and $0 \leq k \leq r/8$. Let $G = (A, B)$ be an (r, i, k)-graph. Then either G has an uncrowded set of at least $r^{11}|B|/2^8$ hats or there are a vertex $b' \in B$ and an induced copy (A', B') of an $(r, i-1, k+1)$-graph in $G - b'$ such that $\emptyset \neq A' \subseteq N_G(b')$.*

Proof. Let us first suppose that every vertex $b \in B$ satisfies

$$|N^2(b)| \geq d(b)/r^{12(i-1)},$$

where $N^2(b)$ is the set of all vertices with distance two from b. In other words, for each $b \in B$ there is a set $\mathcal{H}_b$ of at least $d(b)/r^{12(i-1)}$ hats in G which have b as one endvertex, but whose other endvertices are distinct. Note that every pair of vertices in B belongs to at most two hats in $\bigcup_{b \in B} \mathcal{H}_b$. Hence there are at least $e(G)/2r^{12(i-1)}$ hats with distinct pairs of endpoints. Since the degree of every vertex $a \in A$ is at most $4r$, at most $(4r)^2$ of these hats have a as their midpoint. Thus G has a uncrowded set of at least

$$\frac{e(G)}{2 \cdot 16 r^{12(i-1)+2}} \geq \frac{(r/4 - k)|A|}{2^5 r^{12(i-1)+2}} \geq \frac{(r/4 - k) r^{12i} |B|}{2^5 r^{12(i-1)+2}} \geq \frac{r^{11}|B|}{2^8}$$

hats, as required.

So we may assume that there is a vertex $b' \in B$ with

$$|N^2(b')| < d(b')/r^{12(i-1)}.$$

Let $A' := N(b')$ and $B' := N^2(b')$. Then $(A', B')_G$ has the required properties. $\square$

The proof of the preceding lemma shows that in the case where we failed to find a large set of uncrowded hats (i.e. a 1-subdivision of some graph of large average degree), there must be a vertex b' so that the set of vertices with distance two from b' is much smaller that the neighbourhood of b'. However, if this happens we can reapply the lemma to the bipartite graph induced by these sets. In case of renewed failure, we can iterate the process—but if we encounter i successive failures, then this means that G contains contains a $K_{i,i}$:

Corollary 6.14 *Let $s \in \mathbb{N}$ and let $r \geq 8s$. Let $G = (A, B)$ be a $K_{s,s}$-free $(r, s, 0)$-graph. Then there exists $0 \leq i \leq s$ such that G contains an induced copy (A', B') of an $(r, s-i, i)$ graph which has an uncrowded set of at least $r^{11}|B'|/2^8$ hats.*

Proof. Applying Lemma 6.13 repeatedly, assume that there are sequences $(A, B) = (A_0, B_0) \supseteq (A_1, B_1) \supseteq \cdots \supseteq (A_s, B_s)$ of induced subgraphs of G and $b_1, b_2, \ldots, b_s$ of distinct vertices in B such that, for each $0 < i \leq s$, (A_i, B_i) is an $(r, s-i, i)$-graph and $\emptyset \neq A_i \subseteq N_G(b_i)$. Note that every vertex in A_s has degree at least $r/4 - s \geq r/8$ and so

$$s \leq \frac{r}{8} \leq |B_s| = r^{12(s-s)}|B_s| \leq |A_s|.$$

Thus together with any s vertices from A_s the vertices $b_1, \ldots, b_s$ induce a $K_{s,s}$ in G, a contradiction. $\square$

We say that an uncrowded set $\mathcal{H}$ of hats of a bipartite graph (A, B) is *induced* if $\bigcup \mathcal{H}$ is an induced subgraph of (A, B), i.e. if every midpoint of a hat in $\mathcal{H}$ has degree two in (A, B).

Lemma 6.15 *Let $r \geq 1$ and let $G = (A, B)$ be a bipartite graph with $d(a) \leq 4r$ for every vertex $a \in A$. Suppose that G has an uncrowded set $\mathcal{H}$ of at least $r^{11}|B|/2^8$ hats. Then there is an induced subgraph $G' = (A', B')$ of G which has an induced uncrowded set $\mathcal{H}'$ of at least $r^9|B'|/2^{15}$ hats.*

Proof. We may assume that A consists only of midpoints of hats in $\mathcal{H}$. Since $\mathcal{H}$ is uncrowded, every vertex $a \in A$ is the midpoint of exactly one hat in $\mathcal{H}$, and we say that a *owns* the endvertices of these hat. So every vertex in A owns exactly two vertices in B and

$$|A| = |\mathcal{H}| \geq \frac{r^{11}|B|}{2^8}.$$

Let us consider a random subset B_p of B which is obtained by including each vertex of B independently with probability $p := 1/(8r)$. Given B_p, let us call a vertex $a \in A$ *useful* if $N(a) \cap B_p$ consists precisely of the two vertices owned by a. Thus

$$\mathbb{P}(a \text{ is useful}) = p^2(1-p)^{d(a)-2} \ge p^2(1-p)^{\lfloor 4r \rfloor} \ge p^2(1 - \lfloor 4r \rfloor p) \ge p^2/2,$$

and so the expected number of useful vertices is at least $p^2|A|/2$. Hence there exists a choice B' for B_p such that at least $p^2|A|/2$ vertices in A are useful. Let A' denote the set of these vertices, and let $\mathcal{H}'$ be the set consisting of all hats in $\mathcal{H}$ whose midpoints lie in A'. Then

$$|\mathcal{H}| = |A'| \ge \frac{|A|}{2^7 r^2} \ge \frac{r^9|B|}{2^{15}} \ge \frac{r^9|B'|}{2^{15}},$$

and so $(A', B')_G$ and $\mathcal{H}'$ have the required properties. □

Corollary 6.16 *Let $s \in \mathbb{N}$ and $r \ge 8s$. Let $G = (A, B)$ be an $(r, s, 0)$ graph. Then either G contains a $K_{s,s}$ or an induced 1-subdivision of some graph H with $d(H) \ge r^9/2^{14}$.*

Proof. We may apply Corollary 6.14 and Lemma 6.15 to obtain an induced bipartite graph $G' = (A', B') \subseteq G$ and a set $\mathcal{H}'$ of hats as in Lemma 6.15. Let H be the graph whose vertex set is B' and in which $b, b' \in B'$ are joined by an edge if there is a hat in $\mathcal{H}'$ whose endvertices are b and b'. So every edge of H corresponds to a hat in $\mathcal{H}'$. As $\mathcal{H}'$ is induced, the 1-subdivision of H is induced in G' (and thus in G). Moreover $e(H) = |\mathcal{H}'| \ge r^9|B'|/2^{15}$, as desired. □

Lemma 6.17 *Let $r \ge 2^{25}$. Let A, B be a vertex partition of a graph G such that A is independent, $\chi(G[B]) \le r$ and $d(G') \le r^3$ for every $G' \subseteq G[B]$. Suppose that $(A, B)_G$ has an induced uncrowded set $\mathcal{H}$ of at least $r^9|B|/2^{15}$ hats. Then G contains an induced 1-subdivision of some graph H with $d(H) \ge r$.*

Proof. Let H_0 be the graph whose vertex set is B and in which $b, b' \in B$ are joined by an edge if they are the endpoints of a hat in $\mathcal{H}$. Hence G contains a 1-subdivision of H_0. Note that $e(H_0) = |\mathcal{H}|$ and so $d(H_0) \ge r^9/2^{14}$. Let H_1 be a subgraph of H_0 with

$$\delta(H_1) \ge \frac{r^9}{2^{15}}, \tag{6.4}$$

and put $B_1 := V(H_1)$ (where B_1 is thought of as a subset of B). Let G^* be the 1-subdivision of H_1 contained in G. Note that every edge which prevents G^* from being induced must join two branch vertices of G^*, i.e. vertices in B_1. Using a probabilistic argument, we will show that H_1 contains a subgraph H_2 of average degree at least r whose 1-subdivision in G is induced. In other words, we are given two graphs H_1 and $F := G[B_1]$ on the same vertex set such that H_1 has large average degree while every subgraph of F has small average degree. The desired subgraph H_2 of H_1 must avoid all edges of F.

Let B_1' denote the set of all vertices $b \in B_1$ with $d_F(b) \le 2r^3$. Then

$$2r^3|B_1 \setminus B_1'| \le 2e(F) = d(F)|F| \le r^3|B_1|$$

and thus

$$|B_1'| \geq \frac{|B_1|}{2}. \tag{6.5}$$

Consider a random subset B_p of B_1 which is obtained by including each vertex of B_1 independently with probability $p = 1/(4r^3)$. Given B_p, call a vertex $b \in B_1'$ *useful* if

(a) $b \in B_p$,

(b) $N_F(b) \cap B_p = \emptyset$,

(c) $|(N_{H_1}(b) \setminus N_F(b)) \cap B_p| \geq pr^9/2^{17}$.

Thus every useful vertex is isolated in $G[B_p]$ and in the graph H_1 it has many neighbours which are contained in B_p. The aim now is to show that with non-zero probability the set I_0 of useful vertices is large. As the chromatic number of $G[B_p]$ is small compared to $|N_{H_1}(b) \cap B_p|$ for any useful vertex b, there will be an independent set in $B_p \setminus I_0$ which together with I_0 induces a subgraph H_2 of H_1 with large average degree. Observe that the 1-subdivision of H_2 in G will be induced.

To prove that with non-zero probability B_1' contains many useful vertices, first note that for every $b \in B_1'$ the random variable $X := |(N_{H_1}(b) \setminus N_F(b)) \cap B_p|$ is binomially distributed with

$$\mathbb{E}X = p|N_{H_1}(b) \setminus N_F(b)| \geq p|\delta(H_1) - d_F(b)| \overset{(6.4)}{\geq} \frac{pr^9}{2^{16}} \geq 8.$$

So Lemma 6.6 implies that

$$\mathbb{P}(X \leq \frac{pr^9}{2^{17}}) \leq \mathbb{P}(X \leq \frac{\mathbb{E}X}{2}) \leq \mathrm{e}^{-\mathbb{E}X/8} \leq \frac{1}{2}.$$

Moreover, note that the events (a), (b) and (c) are mutually independent. Thus for every vertex $b \in B_1'$ we have that

$$\mathbb{P}(b \text{ is useful}) \geq p \cdot (1-p)^{d_F(b)} \cdot \frac{1}{2} \geq p \cdot (1-p)^{\lfloor 2r^3 \rfloor} \cdot \frac{1}{2} \geq \frac{p(1 - \lfloor 2r^3 \rfloor p)}{2} \geq \frac{p}{4}.$$

Hence by (6.5) the expected number of useful vertices is at least $p|B_1'|/4 \geq p|B_1|/8$. So there exists a choice B_2 for B_p such that at least $p|B_1|/8$ vertices in B_1' are useful. Let I_0 denote the set of these vertices. Every useful vertex is contained in B_2 and has at least $pr^9/2^{17}$ neighbours in H_1 which are contained in B_2. Thus there are at least

$$\frac{1}{2} \cdot \frac{pr^9}{2^{17}} \cdot \frac{p|B_1|}{8} = \frac{r^3|B_1|}{2^{25}}$$

edges of H_1 which emanate from vertices contained in I_0. Since $\chi(G[B]) \leq r$, we may partition $G[B_2 \setminus I_0]$ into r independent sets, $I_1, \ldots, I_r$ say. Then there exists $0 \leq i \leq r$ such that at least a $1/(r+1)$th of the edges of H_1 emanating

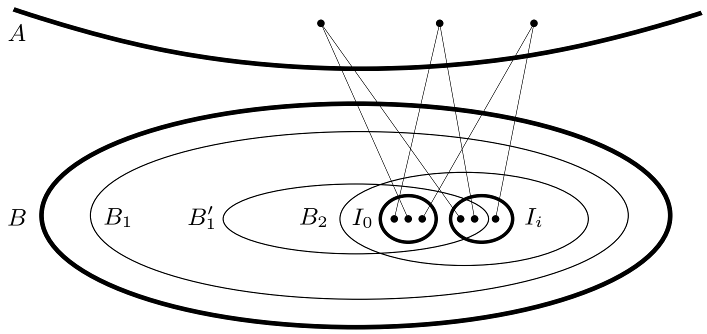


Figure 6.1: Finding an independent set of vertices in F which induces many hats

from I_0 ends in I_i (see Fig. 6.1). But then the subgraph H_2 of H_1 induced by $I_0 \cup I_i$ has at least

$$\frac{1}{r+1} \cdot \frac{r^3|B_1|}{2^{25}} \geq \frac{r|B_1|}{2}$$

edges and so it has average degree at least r. Moreover, since in F both I_0 and I_i are independent and no vertex in $B_2 \supseteq I_i$ is joined to a vertex in I_0, it follows that $I_0 \cup I_i$ is independent in G. As mentioned above, this implies that the 1-subdivision of H_2 is induced in G. □

By successively applying Corollary 6.14 and Lemmas 6.15 and 6.17 we obtain the following result.

Corollary 6.18 *Let $s \in \mathbb{N}$ and $r \geq \max\{8s, 2^{25}\}$. Let G be a $K_{s,s}$-free graph and let $A, B \subseteq V(G)$ be disjoint sets of vertices such that A is independent, $\chi(G[B]) \leq r$, $d(G') \leq r^3$ for every $G' \subseteq G[B]$ and so that $(A, B)_G$ is an $(r, s, 0)$ graph. Then G contains an induced 1-subdivision of some graph H with $d(H) \geq r$.* □

6.6 Proof of Theorem 6.2

We can now put everything together.

Proof of Theorem 6.2. Suppose that G is a $K_{s,s}$-free graph with $d(G) = d \geq d_0$ where d_0 is sufficiently large compared to k and s. Put $n := |G|$. Clearly, we may assume that G has no subgraph of average degree $> d$. So Propositions 2.1 and 6.5 enable us to assume that $\delta(G) \geq d/2$ and $\chi(G) \leq d+1$. Also Lemma 6.7 and Corollary 6.16 imply that Theorem 6.2 holds if G contains an induced bipartite subgraph of large average degree—we will make use of this fact twice in what follows.

Turning to the proof itself, we first apply Corollary 6.9 to G, which gives us an independent set I of size $2cn/d$ where $c \geq (\log d)^{1/(s+1)}/2$. We then apply Corollary 6.12 to obtain (without loss of generality) disjoint sets $A, B \subseteq V(G)$ as in condition (ii) of the corollary. In other words, A is independent, $\chi(G[B]) \leq r$

and $(A, B)_G$ is an $(r, s, 0)$-graph, where $r = \lfloor \log \log c \rfloor$. Now if $G[B]$ has an (induced) subgraph G' whose average degree is at least r^3 then, as $\chi(G') \leq r$, there must be two disjoint independent sets B_1, B_2 of G' such that

$$e((B_1, B_2)_{G'}) \geq \frac{e(G')}{\binom{r}{2}} \geq \frac{d(G')|G'|}{r^2} \geq r|G'| \geq r(|B_1| + |B_2|).$$

Hence $(B_1, B_2)_G$ is an induced bipartite subgraph of average degree at least $2r$. So we may assume that $d(G') \leq r^3$ for every $G' \subseteq G[B]$. But then Corollary 6.18 implies that G contains an induced 1-subdivision of some graph H^* which has average degree at least k, as desired. $\square$

6.7 Open problems

An obvious question is that of the growth of $d(s, k)$ in Theorem 6.2. The bounds which follow from our proof are quite large: k is about the 3-fold logarithm of d even for the case $s = 2$. Also, we are not aware of any nontrivial lower bound on d.

Our proof of Theorem 6.2 becomes easier if G contains an induced bipartite subgraph of large average degree. This raises the question whether there exists $d(s, k)$ such that every $K_{s,s}$-free graph of average degree at least $d(s, k)$ contains an induced bipartite subgraph with average degree at least k. The following result implies that much more is true for regular graphs: using a theorem of Johansson [41], Alon, Krivelevich and Sudakov [5, Corollary 2.4] proved that every $K_{s,s}$-free graph G with maximum degree Δ has chromatic number at most $c\Delta/\log \Delta$ for some constant c depending on s (and thus if G is regular, the largest colour class together with another one induce a bipartite graph of average degree at least $(\log \Delta)/c$). Of course the result of Alon, Krivelevich and Sudakov does not hold if we replace maximum degree by average degree: just consider a $K_{s,s}$-free graph G whose chromatic number is large and add sufficiently many isolated vertices to G.

Chapter 7

Forcing unbalanced complete bipartite minors

7.1 Introduction

Let $f(s)$ be the smallest number such that every graph of average degree greater than $f(s)$ contains the complete graph K_s as minor. Here we briefly review the facts which we know about $f(s)$. The existence of $f(s)$ was first proved by Mader [75]. Kostochka [59] and Thomason [99] independently showed that the order of magnitude of $f(s)$ is $s\sqrt{\log s}$. Later, Thomason [100] was able to prove that $f(s) = (\alpha + o(1))s\sqrt{\log s}$, where $\alpha = 0.638\ldots$ is an explicit constant. Here the lower bound on $f(s)$ is provided by random graphs. In fact, Myers [85] proved that all extremal graphs are essentially disjoint unions of pseudo-random graphs.

Recently, Myers and Thomason [87] extended the results of [100] from complete minors to H minors for arbitrary dense (and large) graphs H. The extremal function has the same form as $f(s)$, except that $\alpha \leq 0.638\ldots$ is now an explicit parameter depending on H and s is replaced by the order of H. They raised the question of what happens for sparse graphs H. One partial result in this direction was obtained by Myers [86]: he showed that every graph of average degree at least $t+1$ contains a $K_{2,t}$ minor. This is best possible as he observed that for all positive ε there are infinitely many graphs of average degree at least $t+1-\varepsilon$ which do not contain a $K_{2,t}$ minor. (These examples also show that random graphs are not extremal in this case.) More generally, Myers [86] conjectured that for fixed s the extremal function for a $K_{s,t}$ minor is linear in t:

Conjecture 7.1 (Myers) *Given $s \in \mathbb{N}$, there exists a positive constant C such that for all $t \in \mathbb{N}$ every graph of average degree at least Ct contains a $K_{s,t}$ minor.*

In this chapter we prove the following strengthened version of this conjecture. (It implies that asymptotically the influence of the number of edges on the extremal function is negligible.)

Theorem 7.2 *For every $0 < \varepsilon < 10^{-16}$ there exists a number $t_0 = t_0(\varepsilon)$ such that for all integers $t \geq t_0$ and $s \leq \varepsilon^6 t/\log t$ every graph of average degree at least $(1+\varepsilon)t$ contains a $K_{s,t}$ minor.*

Theorem 7.2 is essentially best possible in two ways. Firstly, the complete graph K_{s+t-1} shows that up to the error term εt the bound on the average degree cannot be reduced. Secondly, as we will see in Proposition 7.9 (applied with $\alpha := 1/3$), the result breaks down if we try to set $s \geq 18t/\log t$. Moreover, Proposition 7.9 also implies that if $t/\log t = o(s)$ then even a linear average degree (as in Conjecture 7.1) no longer suffices to force a $K_{s,t}$ minor.

The case where $s = ct$ for some constant $0 < c \leq 1$ is covered by the results of Myers and Thomason [87]. The extremal function in this case is $(\alpha\frac{2\sqrt{c}}{1+c} + o(1))r\sqrt{\log r}$ where $\alpha = 0.638\ldots$ again and $r = s+t$.

For fixed s, we obtain the following strengthening of Theorem 7.2:

Theorem 7.3 *For every $\varepsilon > 0$ and every integer s there exists a number $t_0 = t_0(\varepsilon, s)$ such that for all integers $t \geq t_0$ every graph of average degree at least $(1+\varepsilon)t$ contains $K_s + \overline{K}_t$ as a minor.*

This chapter is organized as follows. We first prove Theorem 7.2 for graphs whose connectivity is linear in their order (Lemma 7.8). We then use ideas of Thomason [100] to extend the result to arbitrary graphs. The proof of Theorem 7.3 is almost the same as that of Theorem 7.2 and so we only sketch it.

7.2 Notation and tools

If $P = x_1 \ldots x_\ell$ is a path and $1 \leq i \leq j \leq \ell$, we write $x_i P x_j$ for its subpath $x_i \ldots x_j$.

We will use the following result of Mader [77].

Theorem 7.4 *Every graph G contains a $\lceil d(G)/4 \rceil$-connected subgraph.*

Given $k \in \mathbb{N}$, we say that a graph G is *k-linked* if $|G| \geq 2k$ and for every $2k$ distinct vertices $x_1, \ldots, x_k$ and $y_1, \ldots, y_k$ of G there exist disjoint paths $P_1, \ldots, P_k$ such that P_i joins x_i to y_i. Jung as well as Larman and Mani independently proved that every sufficiently highly connected graph is k-linked. Later, Bollobás and Thomason [19] showed that a connectivity linear in k suffices. Simplifying the argument in [19], Thomas and Wollan [98] recently obtained an even better bound:

Theorem 7.5 *Every $16k$-connected graph is k-linked.*

Similarly as in [100], given positive numbers d and k, we shall consider the class $\mathcal{G}_{d,k}$ of graphs defined by

$$\mathcal{G}_{d,k} := \{G : |G| \geq d, \; e(G) > d|G| - kd\}.$$

We say that a graph G is *minor-minimal in* $\mathcal{G}_{d,k}$ if G belongs to $\mathcal{G}_{d,k}$ but no proper minor of G does. The following lemma states some properties of the minor-minimal elements of $\mathcal{G}_{d,k}$. The proof is simple, its counterpart for digraphs can be found in [100, Section 2]. (The first property follows by counting the number of edges of the complete graph on $\lfloor (2-\varepsilon)d \rfloor$ vertices.)

Lemma 7.6 *Given $0 < \varepsilon < 1/2$, $d \geq 2/\varepsilon$ and $1/d \leq k \leq \varepsilon d/2$, every minor-minimal graph in $\mathcal{G}_{d,k}$ satisfies the following properties:*

(i) $|G| \geq (2-\varepsilon)d$,

(ii) $e(G) \leq d|G| - kd + 1$,

(iii) *every edge of G lies in more than $d-1$ triangles,*

(iv) *G is $\lceil k \rceil$-connected.*

We will also use the following easy fact, see [100, Lemma 4.2] for a proof.

Lemma 7.7 *Suppose that x and y are distinct vertices of a k-connected graph G. Then G contains at least $k^2/4|G|$ internally disjoint x-y paths of length at most $2|G|/k$.*

7.3 Proof of theorems

The strategy of the proof of Theorem 7.2 is as follows. It is easily seen that to prove Theorem 7.2 for all graphs of average degree at least $(1+\varepsilon)t =: d$, it suffices to consider only those graphs G which are minor-minimal in the class $\mathcal{G}_{d/2,k}$ for some suitable k. In particular, together with Lemma 7.6 this implies that we only have to deal with k-connected graphs. If d (and so also k) is linear in the order of G, then a simple probabilistic argument gives us the desired $K_{s,t}$ minor (Lemma 7.8). In the other case we use that by Lemma 7.6 each vertex of G together with its neighbourhood induces a dense subgraph of G. We apply this to find 10 disjoint $K_{10s,\lceil d/9 \rceil}$ minors which we combine to a $K_{s,t}$ minor.

Lemma 7.8 *For all $0 < \varepsilon, c < 1$ there exists a number $k_0 = k_0(\varepsilon, c)$ such that for each integer $k \geq k_0$ every k-connected graph G whose order n satisfies $k \geq cn$ contains a $K_{s,t}$ minor where $t := \lceil (1-\varepsilon)n \rceil$ and $s := \lceil c^4 \varepsilon n/(32 \log n) \rceil$. Moreover, the branch sets corresponding to the vertices in the vertex class of the $K_{s,t}$ of size t can be chosen to be singletons whereas all the other branch sets can be chosen to have size at most $8 \log n/c^2$.*

Proof. Throughout the proof we assume that k (and thus also n) is sufficiently large compared with both ε and c for our estimates to hold. Put $a := \lfloor 4 \log s/c \rfloor$. Successively choose as vertices of G uniformly at random without repetitions. Let C_1 be the set of the first a of these vertices, let C_2 be the set of the next a vertices and so on up to C_s. Let C be the union of all the C_i. Given $i \leq s$, we call a vertex $x \in G - C$ *good for* i if x has at least one neighbour in C_i. Moreover, we say that x is *good* if it is good for every $i \le s$. Thus

$$\mathbb{P}(x \text{ is not good for } i) \le \left(1 - \frac{d_G(x) - as}{n}\right)^a \le \mathrm{e}^{-a(k-as)/n} \le \mathrm{e}^{-ac/2}$$

and so x is not good with probability at most $s\mathrm{e}^{-ac/2} < \varepsilon/2$. Therefore the expected number of good vertices outside C is at least $(1-\varepsilon/2)|G-C|$. Hence there exists an outcome $C_1, \dots, C_s$ for which at least $(1-\varepsilon/2)|G-C|$ vertices in $G - C$ are good.

We now extend all these C_i to disjoint connected subgraphs of G as follows. Let us start with C_1. Fix a vertex $x_1 \in C_1$. For each $x \in C_1 \setminus \{x_1\}$ in turn we apply Lemma 7.7 to find an x-x_1 path of length at most $2n/k \le 2/c$ which is internally disjoint from all the paths chosen previously and which avoids $C_2 \cup \dots \cup C_s$. Since Lemma 7.7 guarantees at least $k^2/4n \ge as \cdot 2/c$ short paths between a given pair of vertices, we are able to extend each C_i in turn to a connected subgraph in this fashion. Denote the graphs thus obtained from $C_1, \dots, C_s$ by $G_1, \dots, G_s$. Thus all the G_i are disjoint.

Note that at most $2as/c$ good vertices lie in some G_i. Thus at least $(1 - \varepsilon/2)|G-C| - 2as/c \ge (1-\varepsilon)n$ good vertices avoid all the G_i. Hence G contains a $K_{s,t}$ minor as required. (The good vertices avoiding all the G_i correspond to the vertices of the $K_{s,t}$ in the vertex class of size t. The branch sets corresponding to the vertices of the $K_{s,t}$ in the vertex class of size s are the vertex sets of $G_1, \dots, G_s$.) □

Proof of Theorem 7.2. Let $d := (1+\varepsilon)t$ and $s := \lfloor \varepsilon^6 d/\log d \rfloor$. Throughout the proof we assume that t (and thus also d) is sufficiently large compared with ε for our estimates to hold. We have to show that every graph of average degree at least d contains a $K_{s,t}$ minor. Put $k := \lceil \varepsilon d/4 \rceil$. Since $\mathcal{G}_{d/2,k}$ contains all graphs of average degree at least d, it suffices to show that every graph G which is minor-minimal in $\mathcal{G}_{d/2,k}$ contains a $K_{s,t}$ minor. Let $n := |G|$. As is easily seen, (i) and (iv) of Lemma 7.6 together with Lemma 7.8 imply that we may assume that $d \le n/600$. (Lemma 7.8 is applied with $c := \varepsilon/2400$ and with ε replaced by $\varepsilon/3$.) Let X be the set of all those vertices of G whose degree is at most $2d$. Since by Lemma 7.6 (ii) the average degree of G is at most d, it follows that $|X| \ge n/2$. Let us first prove the following claim.

> *Either G contains a $K_{s,t}$ minor or G contains* 10 *disjoint* $\lceil 3d/25 \rceil$-*connected subgraphs $G_1, \dots, G_{10}$ such that* $3d/25 \le |G_i| \le 3d$ *for each* $i \le 10$.

Choose a vertex $x_1 \in X$ and let G_1' denote the subgraph of G induced by x_1 and its neighbourhood. Then $|G_1'| = d_G(x_1) + 1 \le 2d + 1$. Since by Lemma 7.6 (iii) each edge between x_1 and $N_G(x_1)$ lies in at least $d/2 - 1$ triangles, it follows that the minimum degree of G_1' is at least $d/2 - 1$. Thus Theorem 7.4 implies that G_1' contains a $\lceil 3d/25 \rceil$-connected subgraph. Take G_1 to be this subgraph. Put $X_1 := X \setminus V(G_1)$ and let X_1' be the set of all those vertices in X_1 which have at least $d/500$ neighbours in G_1.

Suppose first that $|X_1'| \geq |X|/10$. In this case we will find a $K_{s,t}$ minor in G. Since the argument is similar to the proof of Lemma 7.8, we only sketch it. Set $a := \lfloor 10^4 \log s \rfloor$. This time, we choose the a-element sets $C_1, \dots, C_s$ randomly inside $V(G_1)$. Since every vertex in X_1' has at least $d/500$ neighbours in G_1, the probability that the neighbourhood of a given vertex $x \in X_1'$ avoids some C_i is at most $se^{-a/(3 \cdot 10^3)} < \varepsilon$. So the expected number of such bad vertices in X_1' is at most $\varepsilon|X_1'|$. Thus for some choice of $C_1, \dots, C_s$ there are at least $(1-\varepsilon)|X_1'| \geq (1-\varepsilon)n/20 \geq t$ vertices in X_1' which have a neighbour in each C_i. Since the connectivity of G_1 is linear in its order, we may again apply Lemma 7.7 to make the C_i into disjoint connected subgraphs of G_1 by adding suitable short paths from G_1. This shows that G contains a $K_{s,t}$ minor.

Thus we may assume that at least $|X_1| - |X|/10 \geq 9|X|/10 - 3d > 0$ vertices in X_1 have at most $d/500$ neighbours in G_1. Choose such a vertex x_2. Let G_2' be the subgraph of G induced by x_2 and all its neighbours outside G_1. Since by Lemma 7.6 (iii) every edge of G lies in at least $d/2-1$ triangles, it follows that the minimum degree of G_2' is at least $d/2 - 1 - d/500 > 12d/25$. Again, we take G_2 to be a $\lceil 3d/25 \rceil$-connected subgraph of G_2' obtained by Theorem 7.4.

We now put $X_2 := X_1 \setminus (X_1' \cup V(G_2))$ and define X_2' to be the set of all those vertices in X_2 which have at least $d/500$ neighbours in G_2. If $|X_2'| \geq |X|/10$, then as before, we can find a $K_{s,t}$ minor in G. If $|X_2'| \leq |X|/10$ we define G_3 in a similar way as G_2. Continuing in this fashion proves the claim. (Note that when choosing x_{10} we still have $|X_9| - |X|/10 \geq |X|/10 - 9 \cdot 3d > 0$ vertices at our disposal since $n \geq 600d$.)

Apply Lemma 7.8 with $c := 1/25$ to each G_i to find a $K_{10s, \lceil d/9 \rceil}$ minor. Let $C_1^i, \dots, C_s^i, D_1^i, \dots, D_{9s}^i$ denote the branch sets corresponding to the vertices of the $K_{10s,\lceil d/9 \rceil}$ in the vertex class of size $10s$. By Lemma 7.8 we may assume that all the C_j^i and all the D_j^i have size at most $8 \cdot 25^2 \log |G_i| \leq 10^5 \log d$ and that all the branch sets corresponding to the remaining vertices of the $K_{10s,\lceil d/9 \rceil}$ are singletons. Let $T^i \subseteq V(G_i)$ denote the union of all these singletons. Let C be the union of all the C_j^i, let D be the union of all the D_j^i and let T be the union of all the T^i.

We will now use these 10 $K_{10s,\lceil d/9 \rceil}$ minors to form a $K_{s,t}$ minor in G. Recall that by Lemma 7.6 (iv) the graph G is $\lceil \varepsilon d/4 \rceil$-connected and so by Theorem 7.5 it is $\lfloor \varepsilon d/64 \rfloor$-linked. Thus there exists a set $\mathcal{P}$ of $9s$ disjoint paths in G such that for all $i \leq 9$ and all $j \leq s$ the set C_j^i is joined to C_j^{i+1} by one of these paths and such that no path from $\mathcal{P}$ contains an inner vertex in $C \cup D$. (To see this, use that $\varepsilon d/64 \geq 100s \cdot 10^5 \log d \geq |C \cup D|$.)

The paths in $\mathcal{P}$ can meet T in many vertices. But we can reroute them such that every new path contains at most two vertices from each T^i. For every path $P \in \mathcal{P}$ in turn we will do this as follows. If P meets T^1 in more than 2 vertices, let t and t' denote the first and the last vertex from T^1 on P. Choose some set D_j^1 and replace the subpath tPt' by some path between t and t' whose interior lies entirely in $G[D_j^1]$. (This is possible since $G[D_j^1]$ is connected and since both t and t' have a neighbour in D_j^1.) Proceed similarly if the path thus obtained still meets some other T^i. Then continue with the next path from $\mathcal{P}$. (The sets D_j^i used for the rerouting are chosen to be distinct for different paths.) Note that the paths thus obtained are still disjoint since D was avoided by all the paths in $\mathcal{P}$.

We now have found our $K_{s,t}$ minor. Each vertex lying in the vertex class of size s of the $K_{s,t}$ corresponds to a set consisting of $C_j^1 \cup \cdots \cup C_j^{10}$ together with the (rerouted) paths joining these sets. For the remaining vertices of the $K_{s,t}$ we can take all the vertices in T which are avoided by the (rerouted) paths. There are at least t such vertices since these paths contain at most $20 \cdot 9s$ vertices from T and $|T| - 180s \geq 10d/9 - 180s \geq t$. $\square$

Proof of Theorem 7.3 (Sketch). Without loss of generality we may assume that $\varepsilon < 10^{-16}$. The proof of Theorem 7.3 is almost the same as that of Theorem 7.2. The only difference is that now we also apply Lemma 7.7 to find $\binom{s}{2}$ short paths connecting all the pairs of the C_i. This can be done at the point where we extend the C_i's to connected subgraphs. $\square$

The following proposition shows that the bound on s in Theorem 7.2 is essentially best possible. Its proof is an adaption of a well-known argument of Bollobás, Catlin and Erdős [17].

Proposition 7.9 *There exists an integer n_0 such that for each integer $n \geq n_0$ and each number $\alpha > 0$ there is a graph G of order n and with average degree at least $n/2$ which does not have a $K_{s,t}$ minor with $s := \lceil 2n/\alpha \log n \rceil$ and $t := \lceil \alpha n \rceil$.*

Proof. Let $p := 1 - 1/\mathrm{e}$. Throughout the proof we assume that n is sufficiently large for our estimates to hold. Consider a random graph G_p of order n which is obtained by including each edge with probability p independently from all other edges. We will show that with positive probability G_p is as required in the proposition. Clearly, with probability $> 3/4$ the average degree of G_p is at least $n/2$. Hence it suffices to show that with probability at most $1/2$ the graph G_p will have the property that its vertex set $V(G_p)$ can be partitioned into disjoint sets $S_1, \ldots, S_s$ and $T_1, \ldots, T_t$ such that G_p contains an edge between every pair S_i, T_j $(1 \leq i \leq s, 1 \leq j \leq t)$. Call such a partition of $V(G_p)$ *admissible.* Thus we have to show that the probability that G_p has an admissible partition is $\leq 1/2$. Let us first estimate the probability that a given partition $\mathcal{P}$ is admissible:

$$\begin{aligned}
\mathbb{P}(\mathcal{P} \text{ is admissible}) &= \prod_{i,j} \left(1 - (1-p)^{|S_i||T_j|}\right) \leq \exp\left(-\sum_{i,j}(1-p)^{|S_i||T_j|}\right) \\
&\leq \exp\left(-st\prod_{i,j}(1-p)^{|S_i||T_j|(st)^{-1}}\right) \leq \exp\left(-st(1-p)^{n^2(st)^{-1}}\right) \\
&\leq \exp\left(-\frac{2n^2}{\log n} \cdot n^{-\frac{1}{2}}\right) \leq \exp(-n^{\frac{4}{3}}).
\end{aligned}$$

(The first expression in the second line follows since the arithmetic mean is at least as large as the geometric mean.) Since the number of possible partitions is at most n^n, it follows that the probability that G_p has an admissible partition is at most $n^n \cdot \mathrm{e}^{-n^{4/3}} < 1/2$, as required. $\square$

Chapter 8

Almost all graphs with high girth and suitable density have high chromatic number

8.1 Introduction and Results

In 1959, Erdős [32] proved that there are graphs of arbitrarily large girth and arbitrarily large chromatic number. (Here the *girth* of a graph G is the length of its shortest cycle and is denoted by $\mathrm{girth}(G)$.) His proof is one of the first and most well-known examples of the probabilistic method: he showed that with high probability one can alter a random graph (with suitable edge probability) so that it has no short cycles and no large independent sets. Here we give a proof (also using the probabilistic method) which gives more information about the typical asymptotic structure of graphs of high girth and given density.

Let $\mathcal{F}_{n,m}(C_{\leq \ell})$ denote the set of all graphs with n vertices and m edges which contain no cycle whose length is at most ℓ, (writing $\mathcal{F}_{n,m}(K_3)$ instead of $\mathcal{F}_{n,m}(C_{\leq 3})$). We say that *almost all* graphs in $\mathcal{F}_{n,m}(C_{\leq \ell})$ have some property if the proportion of graphs in $\mathcal{F}_{n,m}(C_{\leq \ell})$ with this property tends to one as n tends to infinity.

Theorem 8.1 *For all $\ell \geq 3$ and $k \in \mathbb{N}$, there are constants C_1 and C_2 so that almost all graphs in $\mathcal{F}_{n,m}(C_{\leq \ell})$ have chromatic number at least k, provided that $C_1 n \leq m \leq C_2 n^{\ell/(\ell-1)}$.*

Let $G_{n,m}$ denote a graph chosen uniformly at random from the set of graphs with n vertices and m edges. We say that $G_{n,m}$ has some property $\mathcal{Q}$ *almost surely* if the probability that it has $\mathcal{Q}$ tends to one as n tends to infinity. The restriction that $m \geq C_1 n$ in Theorem 8.1 is clearly necessary, since for $m = o(n)$, $G_{n,m}$ almost surely contains no cycles at all. For the case $\ell = 3$, it turns out that the restriction that $m \leq C_2 n^{3/2}$ is also close to best possible. Indeed, building on earlier results, in [90] we showed the following. Set

$$t_3 = t_3(n) = \frac{\sqrt{3}}{4} n^{3/2} \sqrt{\log n}$$

and fix any $\varepsilon > 0$. Then if $m \geq (1+\varepsilon)t_3$, almost all graphs in $\mathcal{F}_{n,m}(K_3)$ are bipartite. This threshold is sharp in the sense that if $n/2 \leq m \leq (1-\varepsilon)t_3$, then almost no graph in $\mathcal{F}_{n,m}(K_3)$ is bipartite.

Instead of Theorem 8.1, we actually prove the following stronger result, which gives a lower bound on the chromatic number of almost all graphs in $\mathcal{F}_{n,m}(C_{\leq \ell})$ in terms of n and m.

Theorem 8.2 *For all $\ell > 3$ there exist constants d_1, d_2 and d_3 with the following properties. Let*

$$m_0 = d_1 n^{(\ell+2)/(\ell+1)} (\log n)^{2/(\ell+1)}.$$

If $2n \leq m \leq m_0$, then almost all graphs in $\mathcal{F}_{n,m}(C_{\leq \ell})$ have chromatic number at least

$$\frac{m}{2n \log(2m/n)}. \tag{8.1}$$

If $m_0 \leq m \leq d_2 n^{\ell/(\ell-1)}$, then almost all graphs in $\mathcal{F}_{n,m}(C_{\leq \ell})$ have chromatic number at least

$$d_3 \sqrt{n^\ell / m^{\ell-1}}. \tag{8.2}$$

We have made no attempt to find the best constants that can be obtained from our proof of Theorem 8.2. Note that for $m \leq m_0$, the bound is of the same order of magnitude as that which is known for $G_{n,p}$, where $p = m/\binom{n}{2}$ and $G_{n,p}$ is a random graph with n vertices with edge probability p. In fact Łuczak (see e.g. [39]) proved that if $pn \to \infty$ and $p \to 0$, then the chromatic number of $G_{n,p}$ is almost surely

$$(1+o(1)) \frac{pn}{2\log(pn)}.$$

It seems likely that the chromatic number of almost all graphs in $\mathcal{F}_{n,m}(C_{\leq \ell})$ is $\Theta(\frac{m}{n \log(m/n)})$ whenever $n \ll m \ll n^{\ell/(\ell-1)}$. However, this seems to be significantly more difficult to prove than Theorem 8.2 even for the triangle-free case.

Related to this is the question of how high the chromatic number of a graph can be if it has n vertices and girth greater than ℓ. Let $f(n,\ell)$ be the maximum chromatic number of such a graph. The proof of Erdős [32] shows that for fixed ℓ, $f(n,\ell) \geq n^{1/\ell + o(1)}$. For the triangle-free case $\ell = 3$ this was improved by Kim [46], who solved a longstanding open question by showing that $f(n,3) \geq \frac{1}{9} n^{1/2}/\sqrt{\log n}$, which (by a result of Ajtai, Komlós, and Szemerédi [1]) is best possible up to the value of the constant factor. It is well known (see e.g. Krivelevich [60, Lemma 6.1] or [91]) that $f(n,\ell) \geq n^{1/(\ell-1)+o(1)}$, which is the best known lower bound for $\ell > 3$. As pointed out to us by one of the journal referees, an upper bound on $f(n,\ell)$, where $\ell > 3$ is even, may be obtained as follows. For even ℓ, Bondy and Simonovits [21] showed that a C_ℓ-free graph has $\mathcal{O}(n^{1+2/\ell})$ edges. Thus it has an independent set of size $\Omega(n^{1-2/\ell})$. Removing this set and applying induction, it is easily seen that such a graph has chromatic number $\mathcal{O}(n^{2/\ell})$ and thus $f(n,\ell) = \mathcal{O}(n^{2/\ell})$ for even ℓ. This can be improved by a logarithmic factor using the results on independent sets in [1] (see also [13, Lemma XII.15]). The bounds obtained from Theorem 8.2

are much smaller than the lower bounds mentioned above: they achieve their maximum when $m = m_0 = n^{(\ell+2)/(\ell+1)+o(1)}$, where they imply that almost all graphs in $\mathcal{F}_{n,m}(C_{\leq \ell})$ have chromatic number at least $n^{1/(\ell+1)+o(1)}$.

In the remainder of this chapter, we prove Theorem 8.2. Although the proof is not quite as simple as that of the original existence result of Erdős, it turns out to be fairly straightforward. Indeed, for a graph G let $\alpha(G)$ denote the size of a largest independent set of vertices. Since for a graph G on n vertices, we have $\chi(G) \geq n/\alpha(G)$, it suffices to show that almost all graphs in $\mathcal{F}_{n,m}(C_{\leq \ell})$ have no large independent set (where m satisfies the conditions of the theorem). This is done by demonstrating that for suitable choices of parameters, the probability that there is a "large" independent set in $G_{n,m}$ is much smaller than the probability that $G_{n,m}$ has girth greater than ℓ.

8.2 Proof of Theorem 8.2

Throughout this section, we set $p = m/\binom{n}{2}$. Using the fact that $\chi(G) \geq n/\alpha(G)$ for any graph G on n vertices, Theorem 8.2 follows immediately from the following lemma. Throughout, we assume that n is large enough for our estimates to hold and we denote by $G_{n,m}$ a graph chosen uniformly at random from the set of graphs with n vertices and m edges.

Lemma 8.3 *For all $\ell > 3$ there exist constants c_1, c_2 and c_3 with the following properties. Let*

$$p_0 = c_1 n^{-\ell/(\ell+1)}(\log n)^{2/(\ell+1)}.$$

If $4/n \leq p \leq p_0$, then

$$\mathbb{P}[\,\alpha(G_{n,m}) \geq \frac{4}{p}\log(np) \mid \mathrm{girth}(G_{n,m}) > \ell\,] = o(1). \tag{8.3}$$

If $p_0 \leq p \leq c_2 n^{-(\ell-2)/(\ell-1)}$, then

$$\mathbb{P}[\,\alpha(G_{n,m}) \geq c_3\sqrt{p^{\ell-1}n^{\ell}} \mid \mathrm{girth}(G_{n,m}) > \ell\,] = o(1). \tag{8.4}$$

To prove Lemma 8.3, we shall need Lemma 8.4 (see also Prömel and Steger [92] and Theorem 3.11 in [39] for similar results), whose proof relies on the FKG-inequality (see e.g. [39]). For $i \geq 3$, let X_i denote the number of i-cycles in $G_{n,m}$. Note that

$$\mathbb{E}[X_i] = (1+o(1))\frac{(n)_i p^i}{2i} = \Theta(m^i/n^i).$$

Lemma 8.4 *For any $\ell \geq 3$, there are constants c, $c' > 0$ so that if $2n \leq m \leq c' n^{\ell/(\ell-1)}$,*

$$\mathbb{P}[\,\mathrm{girth}(G_{n,m}) > \ell\,] \geq \mathrm{e}^{-c\mathbb{E}[X_\ell]}.$$

Proof. We will make use of the inequality

$$1 - x \geq \mathrm{e}^{-x-x^2} \geq \mathrm{e}^{-2x}, \tag{8.5}$$

valid for $x \le 1/2$ (see e.g. page 5 of [13]). Since for $i \ge 3$, the number of i-cycles in the complete graph on n vertices is $\frac{(n)_i}{2i}$, the FKG-inequality implies that

$$\begin{aligned}\mathbb{P}[\,\mathrm{girth}(G_{n,2p}) > \ell\,] &\ge \prod_{i=3}^{\ell}(1-(2p)^i)^{\frac{(n)_i}{2i}} \overset{(8.5)}{\ge} \prod_{i=3}^{\ell} \mathrm{e}^{-2(2p)^i\frac{(n)_i}{2i}} \\ &\ge \prod_{i=3}^{\ell} \mathrm{e}^{-3\cdot 2^i\,\mathbb{E}[X_i]} \ge \mathrm{e}^{-3\ell\, 2^\ell\,\mathbb{E}[X_\ell]}, \end{aligned} \tag{8.6}$$

where the last line follows since $m \ge 2n$ implies that $\mathbb{E}[X_i] \le \mathbb{E}[X_\ell]$ for $3 \le i \le \ell$.

But since the property of containing no cycle of length at most ℓ is monotone decreasing, we have (denoting by $e(G)$ the number of edges of a graph G and letting $N = \binom{n}{2}$)

$$\begin{aligned}\mathbb{P}[\,\mathrm{girth}(G_{n,2p}) > \ell\,] &\le \mathbb{P}[\,\mathrm{girth}(G_{n,m}) > \ell\,] + \mathbb{P}[|e(G_{n,2p}) - 2pN| \ge pN] \\ &\le \mathbb{P}[\,\mathrm{girth}(G_{n,m}) > \ell\,] + \mathrm{e}^{-pN/12}, \end{aligned} \tag{8.7}$$

where the last line follows from standard tail estimates for the binomial distribution (see e.g. Theorem 7(i) in [13]). Thus (8.6) and (8.7) imply that

$$\mathbb{P}[\,\mathrm{girth}(G_{n,m}) > \ell\,] \ge \mathrm{e}^{-3\ell\, 2^\ell\,\mathbb{E}[X_\ell]} - \mathrm{e}^{-m/12}.$$

The result now follows immediately by observing that for c' sufficiently small, $m \le c' n^{\ell/(\ell-1)}$ implies that m is significantly larger than $\mathbb{E}[X_\ell]$. $\square$

We shall also need Pittel's inequality (see page 35 in [13]), which states that if $\mathcal{Q}$ is any property and $0 < p = m/\binom{n}{2} < 1$, then

$$\mathbb{P}[\,G_{n,m} \text{ has } \mathcal{Q}\,] \le 3\sqrt{m}\,\mathbb{P}[\,G_{n,p} \text{ has } \mathcal{Q}\,]. \tag{8.8}$$

Proof of Lemma 8.3. First note that for any $r = r(n)$ with $r \to \infty$,

$$\mathbb{P}[\alpha(G_{n,p}) \ge r] \le \binom{n}{r}(1-p)^{\binom{r}{2}} \le (\mathrm{e}n/r)^r \mathrm{e}^{-pr(r-1)/2} = \mathrm{e}^{-(1+o(1))\phi},$$

where for convenience we write

$$\phi = r(pr/2 - \log(n/r)).$$

Then by Lemma 8.4 and (8.8), there is a constant $c > \ell$ so that

$$\begin{aligned}\mathbb{P}[\,\alpha(G_{n,m}) \ge r \mid \mathrm{girth}(G_{n,m}) > \ell\,] &\le \frac{\mathbb{P}[\,\alpha(G_{n,m}) \ge r\,]}{\mathbb{P}[\,\mathrm{girth}(G_{n,m}) > \ell\,]} \\ &\le 3\sqrt{m}\,\mathbb{P}[\,\alpha(G_{n,p}) \ge r\,]\,\mathrm{e}^{c\mathbb{E}[X_\ell]} \\ &\le 3\sqrt{m}\,\mathrm{e}^{-(1+o(1))(\phi - c\mathbb{E}[X_\ell])}. \end{aligned}$$

Thus to prove (8.3), it suffices to prove that if $r = \frac{4}{p}\log(np)$ and $4/n \le p \le p_0$ (where c_1 in the definition of p_0 will be determined below), then $\phi \ge 4\log n$ and

$\phi/\mathbb{E}[X_\ell] \geq 2c$. Note that our choice of r implies that $\log(1/p) = (1+o(1))\log r$. This in turn implies that $pr/4 = (1+o(1))\log(n/r)$ and thus that

$$\phi = (1+o(1))pr^2/4 = (1+o(1))\frac{4}{p}\left(\log(np)\right)^2 \geq 4\log n,$$

with room to spare. Also

$$\frac{\phi}{\mathbb{E}[X_\ell]} = (1+o(1))\frac{p}{4}\left(\frac{4}{p}\log(np)\right)^2 \frac{2\ell}{n^\ell p^\ell} = (1+o(1))\frac{8\ell(\log(np))^2}{n^\ell p^{\ell+1}} \geq 2c,$$

as required. The final inequality holds if we choose c_1 (in the definition of p_0) sufficiently small compared to c.

Inequality (8.4), where $p_0 \leq p \leq c_2 n^{-(\ell-2)/(\ell-1)}$, is dealt with in a similar way. Indeed, setting $r = c_3\sqrt{p^{\ell-1}n^\ell}$, where c_3 is chosen to be sufficiently large compared to c_1, gives

$$pr/4 \geq \frac{c_3}{4}\sqrt{p_0^{\ell+1}n^\ell} \geq \log(n/r).$$

This in turn implies

$$\phi \geq pr^2/4 = (pn)^{\ell+o(1)} \geq 4\log n.$$

Also, we have

$$\frac{\phi}{\mathbb{E}[X_\ell]} = (1+o(1))\frac{p}{4}\,c_3^2 p^{\ell-1}n^\ell\,\frac{2\ell}{n^\ell p^\ell} = (1+o(1))\frac{c_3^2\ell}{2} \geq 2c,$$

as required, provided we choose c_3 sufficiently large compared to c. $\square$

Chapter 9

On random planar graphs, the number of planar graphs and their triangulations

9.1 Introduction

Compared to the wealth of knowledge one has about random graphs in general, rather little is known about the likely properties of a random planar graph on n vertices – not even the expected number of edges. Here we consider the uniform model: let $\mathcal{P}_n$ be the set of labelled planar graphs with vertex set $\{1, \ldots, n\}$. A random planar graph P_n is chosen uniformly from $\mathcal{P}_n$. This should not be confused with a random planar map, since a planar map is defined as a connected graph which is embedded in the plane, whereas a planar graph may have several embeddings. Moreover, we consider only simple graphs whereas maps may usually have multiple edges.

Random planar graphs were first investigated by Denise, Vasconcellos and Welsh [27]. They showed that $n!\,6^{n+o(n)} \leq |\mathcal{P}_n| \leq n!\,(75.8)^{n+o(n)}$, that the limiting probability that P_n is connected is greater than zero and that the expected number of edges of P_n is at least $3n/2$. They also introduced a Markov chain whose stationary distribution is the uniform measure on $\mathcal{P}_n$. This Markov chain was investigated in much more detail by Gerke and McDiarmid [35], who showed that almost surely P_n has at least $13n/7$ edges. Complementing the result of [27] on the connectivity of P_n, McDiarmid, Steger, and Welsh [82] proved (amongst other results) that the limiting probability that P_n is connected is less than one. Using generating function techniques, Bender, Gao and Wormald [9] proved that the number of 2-connected graphs in $\mathcal{P}_n$ is in fact asymptotically $Cn!\alpha^n n^{-7/2}$, where C is some positive constant and $\alpha \sim 26.1876$, which gives the best known lower bound on $|\mathcal{P}_n|$.

Concerning upper bounds on $|\mathcal{P}_n|$, in the final section we will prove upper bounds for the number of graphs in $\mathcal{P}_n$ with a given number of edges (Theorem 9.12), which will immediately imply the following result.

Theorem 9.1

$$|\mathcal{P}_n| \leq n!\,(37.3)^{n+o(n)}.$$

Theorem 9.12 will turn out to be an immediate consequence of a result of Tutte [104] on the number of planar triangulations and the following result, which states that the number of triangulations of *every* planar graph is exponential in the number of edges which are needed in order to triangulate the graph. We will prove this result in Section 9.3, where we will also see (Proposition 9.11) that the bound given in Theorem 9.2 is essentially best possible for $m \geq 2n$.

Theorem 9.2 *Every labelled planar graph G with n vertices and m edges is contained in at least $\varepsilon 3^{(3n-m)/2}$ labelled triangulations on n vertices, where ε is an absolute constant.*

Combining our upper bounds with the result in [9] mentioned earlier, we will deduce the following result in the final section. Given a class $\mathcal{A}$ of graphs and a property $\mathcal{Q}$, we say that *almost all* graphs in $\mathcal{A}$ have $\mathcal{Q}$ if the proportion of graphs in $\mathcal{A}$ on n vertices which have $\mathcal{Q}$ tends to one as $n \to \infty$.

Theorem 9.3 *Almost all graphs in $\mathcal{P}_n$ have less than $2.56n$ edges.*

Very recently, Theorems 9.1 and 9.3 were improved by Bonichon, Gavoille and Hanusse [22]: they proved (also by using triangulations) that almost all lalled planar graphs on n vertices have at most $2.54n$ edges and that $|\mathcal{P}_n| \leq n!\,(32.2)^{n+o(n)}$.

9.2 Definitions and basic facts

For convenience, for $n \in \mathbb{N}$ we let $[n] := \{1, \ldots, n\}$. A *plane* graph is a planar graph together with an embedding into the plane. A planar graph G is called *rigid* if any two embeddings of G are equivalent. By the theorem of Whitney, every 3-connected planar graph is rigid (see e.g. [28]). Given a face f, we denote its boundary by $b(f)$. Moreover, if f is not the outer face, we say that the *bounding cycle* of f is the shortest cycle containing f in its interior and denote it by $bc(f)$. Note that $bc(f) \subset b(f)$ and that we have equality if G is 2-connected. We say that a vertex x *lies in* f if $x \in b(f)$.

Given a face f, we treat the boundary of f as an ordered sequence of not necessarily distinct vertices $v_1, \ldots, v_k$. If $i < j$ and $v_i \neq v_j$ are not adjacent, then the two sequences $v_{i+1}, \ldots, v_{j-1}$ and $v_{j+1}, \ldots, v_k, v_1, \ldots, v_{i-1}$ are non-empty and we call them the left and the right *v_i-v_j-side* of f, respectively.

For a set of *labelled* graphs $\mathcal{A}$, denote by $\mathcal{A}^u$ the set of distinct *unlabelled* graphs contained in $\mathcal{A}$, each representing an isomorphism class of $\mathcal{A}$. Denote by $\mathcal{A}_n$ the set of those graphs in $\mathcal{A}$ with vertex set $[n]$ and by $\mathcal{A}_{n,m}$ the set of those graphs in $\mathcal{A}_n$ with exactly m edges.

Denote by $\mathcal{P}$ the set of all labelled planar graphs. We shall need a few classes of special planar graphs. Define

$$\mathcal{T} \;:=\; \{G \in \mathcal{P} : \; |E(G)| = 3|V(G)| - 6\}\,.$$

Thus $\mathcal{T}$ is the set of all maximal planar graphs. It is well known that in every embedding of a graph $G \in \mathcal{T}$ all faces are bounded by triangles. Tutte [104] proved that

$$|\mathcal{T}_n^u| = (1+o(1))C\gamma^n n^{-5/2}, \qquad \text{where } \gamma = \frac{256}{27} = 9.48.. \tag{9.1}$$

and where C is some positive constant. It is well known (see e.g. [74] for a proof) that the number of triangulations of the interior of an ℓ-cycle is given by $\mathrm{Cat}(\ell)$, the ℓ-th Catalan number:

$$\mathrm{Cat}(\ell) = \frac{1}{\ell-1}\binom{2\ell-4}{\ell-2} = 4^{\ell+o(\ell)}. \tag{9.2}$$

9.3 Triangulating a planar graph

The aim of this section is to prove Theorem 9.2. The basic idea is as follows. Consider a plane graph $G \in \mathcal{P}_{n,m}$. We would like to generate as many triangulations containing G as possible, and the easiest way would be to simply triangulate each (non-triangular) face independently. This may of course not always be possible, because two non-adjacent vertices which are connected in order to triangulate one face can then not be connected in any of the neighbouring faces.

It turns out (see Proposition 9.5) that this approach does work for 3-connected planar graphs. However, in general, in order to generate many triangulations, we have to make use of the different embeddings that a planar graph may have. For instance, the graph in Figure 9.3 has only one triangulation when viewed as a plane graph but superexponentially many when viewed as a labelled planar graph.

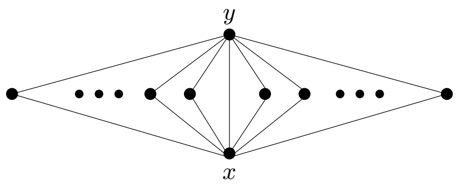


Figure 9.1: A plane graph with only one triangulation

Before dealing with this, let us first consider the 3-connected case. We say that a planar graph has the 1-*face property* if it has an embedding so that the intersection of the boundaries of two faces consists of either an edge, a vertex or is empty. Equivalently, if x and y lie on the boundary of some face f and are not adjacent, then they do not both also lie on the boundary of some other face f'.

Proposition 9.4 *A 3-connected plane graph G has the 1-face property.*

This seems to be a folklore result (see e.g. [84]). As the proof is short, we include it here for completeness. We mention (but will not make use of this) that the converse is also true: a 2-connected graph which has the 1-face property must be 3-connected.

Proof of Proposition 9.4. Suppose that G does not have the 1-face property. Then there exist vertices x and y which are contained in the boundaries of two faces f_1 and f_2 and where the edge xy (which may or may not be present in G) does not lie on the boundary of both faces. But this implies that there are two faces of $G - xy$ into which we can insert the edge xy and thus that $G + xy$ does not have a unique embedding. By Whitney's theorem, $G + xy$ cannot be 3-connected (and hence neither is G). □

The above proposition implies that it is easy to prove Theorem 9.2 for 3-connected graphs.

Proposition 9.5 *Let G be a 3-connected plane graph in $\mathcal{P}_{n,m}$. Then the number of triangulations on n vertices which contain G is at least 2^{3n-m-6}.*

Proof. Using Proposition 9.4, it is easy to see that the number of triangulations of G is equal to the product (over all faces) of the number of triangulations of each face. To calculate the product, first note that the boundary of each face of G is a cycle, and so the number of triangulations of such a face f is given by the Catalan number $\mathrm{Cat}(\ell)$, where ℓ is the length of the bounding cycle of f. Using (9.2), it is easy to show by induction that $\mathrm{Cat}(\ell) \geq 2^{\ell-3}$ for any $\ell \geq 3$. Now denote by ℓ_j the length of the boundary of the j-th face. Since every edge lies on the boundary of two faces, this gives us

$$\sum_j \ell_j = 2m, \tag{9.3}$$

where the sum is over all faces of G. Moreover, from Euler's formula we know that there are $2-n+m$ faces. Combining all this, the number of triangulations on n vertices of G is

$$\prod_j \mathrm{Cat}(\ell_j) \geq \prod_j 2^{\ell_j-3} = 2^{\sum_j(\ell_j-3)} = 2^{2m-3(2-n+m)} = 2^{3n-m-6}.$$

□

So we have proven that every 3-connected graph $G \in \mathcal{P}_{n,m}$ is contained in at least 2^{3n-6-m} triangulations. In other words, the deletion of an edge doubles the number of triangulations if the resulting graph is 3-connected. The above proof shows that this is tight if and only if all faces are triangles or quadrilaterals.

Before we move closer towards the proof of Theorem 9.2, we make some preliminary steps (Propositions 9.6 and 9.7). We say that two planar graphs G_0, $G_1 \in \mathcal{P}_n$ are *incomparable* if there is no planar graph $H \in \mathcal{P}_n$ containing both G_0 and G_1 as a subgraph. Obviously it is true that

$$\begin{aligned} &G_0 \subseteq H_0,\ G_1 \subseteq H_1,\ G_0 \text{ and } G_1 \text{ are incomparable} \\ &\quad \Rightarrow\ H_0 \text{ and } H_1 \text{ are incomparable.} \end{aligned} \tag{9.4}$$

(Here the H_i are allowed to have more vertices than the G_i). For a *planar* graph G denote by $\mathrm{Add}(G)$ the set of all edges $e \in E(\overline{G})$ so that $G+e$ is still planar. In contrast, for a *plane* graph G denote by $\mathrm{Ins}(G)$ the set of those edges $e \in \mathrm{Add}(G)$, so that e can be inserted into the current embedding of G. Suppose that G and H are planar graphs. Then

$$G \subseteq H \quad \Longrightarrow \quad E(\overline{G}) \cap E(H) \ \subseteq \ \mathrm{Add}(G). \tag{9.5}$$

Observe that for a rigid plane graph G we have that $\mathrm{Add}(G) = \mathrm{Ins}(G)$. Moreover note that rigidity is not necessarily preserved when adding vertices and/or edges to a graph. On the other hand, every subdivision of a rigid graph is rigid, as this means nothing else than replacing edges by paths. In particular, every subdivision of a 3-connected graph is rigid.

The following actually rather obvious proposition will turn out to be useful. Suppose we have two plane graphs and the first one has an edge which the second one does not have and cannot have – given its present embedding. If the second one is rigid, then this means that it will never be able to get it, and therefore there is no third planar graph containing both of the two graphs.

Proposition 9.6 *Let G_0 and G_1 be two plane graphs with vertex set $[n]$ where G_1 is a subdivision of a* 3*-connected graph. If there exists an edge $e_0 \in E(G_0) \cap E(\overline{G_1})$ such that $e_0 \notin \mathrm{Ins}(G_1)$, then G_0 and G_1 are incomparable.*

Proof. Suppose to the contrary that there is a planar graph $H \in \mathcal{P}_n$ containing G_0 and G_1 as subgraphs. Then $e_0 \in E(H)$, and applying (9.5) to G_1 and H shows that $e_0 \in \mathrm{Add}(G_1)$. On the other hand, G_1 is rigid because it is a subdivision of a 3-connected graph. Therefore $\mathrm{Add}(G_1) = \mathrm{Ins}(G_1)$. Hence $e_0 \in \mathrm{Ins}(G_1)$, contradicting the assumption. $\square$

Stepping back from planarity for a moment, consider a graph G containing a triangle T. Suppose that $\{x, y\}$ is a cut-set. Then there exists exactly one component H of $G - x - y$ which contains all vertices in $T \setminus \{x, y\}$ and we call H the *T-component* of $G - x - y$.

Proposition 9.7 *Let G be a graph which is* 2*-connected but not* 3*-connected and let T be a triangle in G. Let $\{x, y\}$ be a cut-set of G which minimizes the cardinality of the T-component H of $G - x - y$, and let H^+ be the subgraph of G induced by $V(H) \cup \{x, y\}$, where we only include the edge xy if $xy \in T$. Then H^+ is* 2*-connected.*

Proof. Suppose to the contrary that there exists a vertex z in H^+ such that $H^+ - z$ has two components H_1 and H_2. We can assume without loss of generality that $x \in H_1$ and $y \in H_2$, because if one of the H_i contains neither x nor y, then $G - z$ would not be connected. As there are no edges between H_1 and H_2, this implies that $xy \notin H^+$, and hence by definition of H^+, $xy \notin T$. But then T must lie either in $H_1 + z$ or in $H_2 + z$. Suppose (again without loss of generality) that $T \subseteq H_1 + z$. But then $H_1 - x$ is a component of $G - x - z$ which contains $T \setminus \{x, z\}$, and so it is a T-component which is smaller than H. $\square$

Returning to planar graphs, consider a plane 2-connected graph G whose outer face f is bounded by a triangle T. In this case we will refer to the T-component as the *external component.* Two paths are called *internally disjoint* if their intersection contains at most their endvertices. For two vertices u, w, a path denoted by $P(u, w)$ is a path connecting u and w.

Now reconsider our position with respect to proving Theorem 9.2. Recall that if G is not 3-connected, then we cannot apply Proposition 9.5. Instead, we consider a cut-set $\{x, y\}$ of G and consider the components $H_0, \ldots, H_k$ of $G-x-y$. In Lemma 9.8, which constitutes the core of the proof of Theorem 9.2, we fix the embedding of $H_0 + x + y$, and then embed all other components in several ways into that face of $H_0 + x + y$ which contains x and y . For every such embedding, we fix the positions of the H_i by inserting additional edges, so that the resulting graph is rigid as far as the relative positions of the H_i are concerned. Finally we make sure that the graphs obtained in this way are not only distinct but also incomparable – in other words, no matter how we later add more edges, the resulting graphs will continue to be distinct, so that in the end we really have the required number of distinct triangulations.

Lemma 9.8 *Let $G \in \mathcal{P}_{n,m}$ be a 2-connected plane graph whose outer face f is a triangle. Let $\{x, y\}$ be a cut-set which minimizes the cardinality of the external component H_0. Suppose that $G - x - y$ has $k + 1 \geq 2$ components. Set $t := k$ if $xy \in G$ and $t := k + 1$ if $xy \notin G$.*

Then there is a family of pairwise incomparable plane graphs $G_1, \ldots, G_s$ whose outer face is still f, such that for all $1 \leq i \leq s$, the embeddings of the G_i and G are the same when restricted to $V(H_0)$,

$$G_i \in \mathcal{P}_{n,m+t}, \quad G \subset G_i, \quad G_i - x - y \text{ is connected}$$

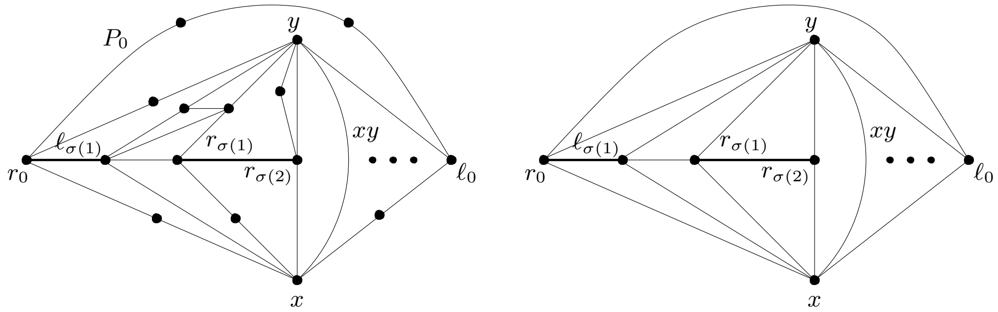
and $s \geq 3^{t/2}$. In the exceptional case where $xy \in f$, we only require that $s \geq 3^{(t-1)/2}$.

Proof. Denote by $H_0, \ldots, H_k$ the plane subgraphs of G which are induced by the connected components of $G - x - y$, where H_0 is the external component. Denote by H_i^+ the plane subgraph of G induced by $V(H_i) \cup \{x, y\}$, where we include the edge xy if and only if $i = 0$ and $xy \in f$. (By Proposition 9.7, this will imply that H_0^+ is 2-connected.) For each $i \geq 1$, it is clear that x and y lie in the outer face of H_i^+. As $xy \notin H_i^+$ for $i \geq 1$, following the notation in the beginning of Section 9.2, we choose a shortest path P_i in H_i connecting the left and the right x-y-side of the outer face of H_i^+, and let ℓ_i be the first and r_i be the last vertex of P_i. If $\ell_i \neq r_i$ then in H_i^+ there exist pairwise internally disjoint paths

$$P(\ell_i, x),\ P(r_i, x),\ P(\ell_i, y),\ P(r_i, y).$$

If $|P_i| = 1$ and $\ell_i = r_i$, then the respective paths above coincide.

Case 1*: $xy \notin G$.* Consider H_0^+. It is clear that there exists a face f' of H_0^+ such that x and y now lie on $b(f')$. By Proposition 9.7, H_0^+ is 2-connected, so both x and y must lie in $bc(f') = b(f')$. As xy is not contained in G and thus neither in H_0^+, this immediately implies that $b(f')$ has at least four vertices, and in particular $f' \neq f$. Moreover, let P_0 be a shortest path in H_0 connecting the



$G_{\sigma,h}$ for $\sigma(h) = 3$ $\qquad$ $A_{\sigma,h}$ is 3-connected

Figure 9.2: Illustrating the proof of Lemma 9.8. The thick edges are the edges which are added to G to form $G_{\sigma,h}$.

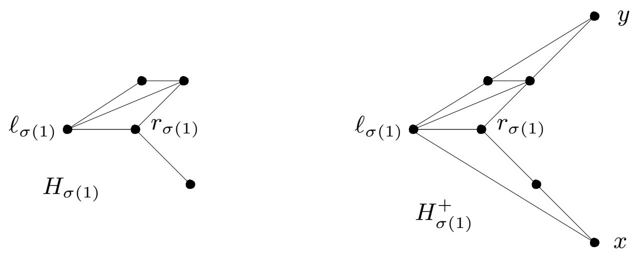


Figure 9.3: The graphs $H_{\sigma(1)}$ and $H^+_{\sigma(1)}$ corresponding to the example in the previous figure.

left and right x-y-side of $b(f')$. P_0 exists because H_0 is connected. Denote by $\ell_0 \neq r_0$ the two endvertices of P_0. Observe that $b(f') \cap P_0 = \{\ell_0, r_0\}$. Similarly to the paths in H_i^+, we divide $b(f')$ into four paths

$$P(\ell_0, x),\ P(r_0, x),\ P(\ell_0, y),\ P(r_0, y)$$

which are all pairwise internally disjoint.

Now choose a permutation σ on $[k]$ and an integer h from $\{0, \dots, k+1\}$. Given σ and h, construct the graph $G_{\sigma,h}$ as follows. For convenience, set $\sigma(0) := 0$ and $\sigma(k+1) := 0$. Successively, for every $i \in [k]$, embed $H^+_{\sigma(i)}$ into f' in such a way that the edge $r_{\sigma(i-1)}\ell_{\sigma(i)}$ can be added (and add it). Having done this for all i, add the edge $r_{\sigma(k)}\ell_0$. Finally, if $h > 0$, remove the edge $r_{\sigma(h-1)}\ell_{\sigma(h)}$ and insert the edge xy instead. Note that $G_{\sigma,h}$ has $m+t$ edges, where $t = k+1$ as required. Moreover, we have constructed

$$s := k!(k+2) \geq 3^{\frac{k+1}{2}} = 3^{\frac{t}{2}}$$

graphs $G_{\sigma,h}$. It is clear from the construction that each $G_{\sigma,h}$ is a plane graph whose outer face is identical to the outer face f of G and whose embedding when restricted to $V(H_0)$ is the same as that of G. Obviously, each $G_{\sigma,h}$ also contains G. As all components H_i of $G-x-y$ are now connected by new edges, we also know that $G_{\sigma,h} - x - y$ is connected.

It remains to prove that the $G_{\sigma,h}$ are pairwise incomparable. To this end, we define two auxiliary graphs which are obtained from $G_{\sigma,h}$ as follows. We first define $B_{\sigma,h} \subseteq G_{\sigma,h}$. For $h = 0$,

$$B_{\sigma,0} := \bigcup_{i=1}^{k+1} \left(r_{\sigma(i-1)}\ell_{\sigma(i)} \cup P_{\sigma(i)} \cup P(\ell_{\sigma(i)}, x) \cup P(r_{\sigma(i)}, x) \cup P(\ell_{\sigma(i)}, y) \cup P(r_{\sigma(i)}, y) \right)$$

and for $h > 0$, we obtain $B_{\sigma,h}$ from $B_{\sigma,0}$ by removing the edge $r_{\sigma(h-1)}\ell_{\sigma(h)}$ and inserting the edge xy instead.

Observe that the vertices in $B_{\sigma,h}$ which have degree at least 3 are exactly $x, y, \ell_0, r_0, \ldots, \ell_k, r_k$. Now obtain $A_{\sigma,h}$ by successively contracting all paths between these vertices into edges. Alternatively, $A_{\sigma,h}$ can also be obtained from a complete bipartite graph with classes $\{x, y\}$ and $\{\ell_0, r_0, \ldots, \ell_k, r_k\}$ by adding either a Hamilton cycle to the second class (if $h = 0$) or Hamilton paths to both classes (if $h > 0$). It is straightforward to check that $A_{\sigma,h}$ is 3-connected, and hence $B_{\sigma,h}$ is a subdivision of a 3-connected graph.

We now show that the $B_{\sigma,h}$ are pairwise incomparable. By the construction of $B_{\sigma,h}$ (in particular, since $B_{\sigma,h}$ is rigid), it is clear that there is no edge $r_i\ell_j \in \mathrm{Ins}(B_{\sigma,h})$ with $i \neq j$. Thus, when considering $B = B_{\sigma,h}$ and $B' = B_{\sigma',h'}$, it suffices to find an edge $r_i\ell_j \in E(B) \cap E(\overline{B'})$ with $i \neq j$ in order to apply Proposition 9.6.

If $\sigma \neq \sigma'$, then $k \geq 2$ and, recalling that $\sigma(0) = \sigma'(0) = 0$ and $\sigma(k+1) = \sigma'(k+1) = 0$, let $i \geq 1$ be the smallest integer so that $\sigma(i) \neq \sigma'(i)$ and let $j \leq k$ be the largest integer so that $\sigma(j) \neq \sigma'(j)$. Note that $i < j$. Then $r_{\sigma(i-1)}\ell_{\sigma(i)}, r_{\sigma(j)}\ell_{\sigma(j+1)} \in E(B) \cap E(\overline{B'})$, unless they are removed from B before adding xy. But as at most one of them will be removed, we are done. If $\sigma = \sigma'$, then $r_{\sigma(h'-1)}\ell_{\sigma(h')} \in E(B) \cap E(\overline{B'})$. Thus the $B_{\sigma,h}$ are pairwise incomparable, and hence by (9.4), so are the $G_{\sigma,h}$. This completes the proof of the lemma for Case 1.

Case 2a: $xy \in G$, $xy \notin f$. By definition, $xy \notin H_0^+$, so the only difference to Case 1 is that we need the edge xy to be present in all our graphs $G_{\sigma,h}$; in other words, we require that $1 \leq h \leq k+1$, so that the total number of graphs is $s := k!(k+1)$ when adding $t := k$ edges to G (one edge less as before, because xy already exists in G). Checking that

$$s = (k+1)! \geq 3^{\frac{k}{2}} = 3^{\frac{t}{2}}$$

completes this case.

Case 2b: $xy \in G$, $xy \in f$. As $xy \in H_0^+$ and H_0^+ is 2-connected, there are two faces in H_0^+ whose boundaries contain xy. Note that one of them is f, and denote the other one by f'. Furthermore the face f' into which we will be embedding the other H_i is bounded by a cycle $b(f')$ containing x and y. However, as $xy \subset b(f')$, we now cannot assume the existence of two vertices r_0 and ℓ_0 on $b(f')$ to which we can connect the H_i as before. So we still proceed as in Case 1, except that we fix $h := k+1$ and do not include ℓ_0 and P_0 in $B_{\sigma,k+1}$. (Nevertheless, it is easy to check that the resulting graphs $A_{\sigma,k+1}$ are still 3-connected.) Hence the total number of graphs is $s := k!$ when adding $t := k$ edges to G. By the statement of the lemma, this time we only need to check that

$$s = k! \geq 3^{\frac{k-1}{2}} = 3^{\frac{t-1}{2}}.$$

As our construction is a special case of the general construction for Case 1, it is clear that the graphs constructed in this way fulfil all the requirements of the lemma. $\square$

Observe that we can apply Lemma 9.8 iteratively, so that starting from a 2-connected graph G, we produce 3-connected graphs G_i satisfying the above requirements (once there are no more vertex pairs x and y to which we can apply the lemma, this means we have arrived at a graph which is 3-connected). We then apply Proposition 9.5 to the 3-connected graphs. So the only remaining problem is that we need to get started if G is not 2-connected. Roughly speaking, we solve this problem by simply embedding loose blocks into faces of 3-connected components and fixing them there.

Recall that a *block* of a graph G is a maximal 2-connected component. The following definitions will be convenient. Given a rooted tree, we say that the *children* of a vertex v are those vertices which are adjacent to v and whose distance to the root is greater than that of v. Given a plane graph G, a *triangulation tree of* G is a rooted tree whose vertices correspond to plane graphs (with vertex set $[n]$) containing G; whose root corresponds to G; whose leaves correspond to 3-connected graphs which are pairwise incomparable; for any vertex v of the tree, the graphs corresponding to the children of v all have the same number of edges and, finally, if a child has t more edges than its parent v, then v has at least $3^{t/2}$ children unless it is an *exceptional vertex*, in which case it still has at least $3^{(t-2)/2}$ children. Observe that in our definition we require the leaves of the tree to be pairwise incomparable, however this is guaranteed via (9.4) if the children of each vertex are pairwise incomparable. Our aim will be to construct a triangulation tree of G with few exceptional vertices on any path from the root to a leaf.

Lemma 9.9 *Consider a plane graph $G \in \mathcal{P}_{n,m}$ consisting of a 3-connected plane graph G' (whose outer face is a triangle) and a block H which has exactly one vertex x in common with G'. Then there is a triangulation tree of G with no exceptional vertices, unless x is contained in the boundary of the outer face of G and has degree at most six in G', in which case the root of the tree may be exceptional.*

Proof. Suppose first that x lies on the boundary of the outer face and has degree at most six. Let z be a vertex in H which is adjacent to x. Let y be a vertex on the boundary of the outer face which is adjacent to x and let v be a vertex which is not in the outer face but lies on the same face f as x and y in G' (such a v exists since x has degree at least three in G'). Let G'_1 be the graph obtained from G by inserting the edges zy and zv (thus H is embedded into f). Now apply Lemma 9.8 to G'_1. If the resulting graphs $G_1, \ldots, G_s$ are not 3-connected, then we repeatedly apply Lemma 9.8 to those graphs and

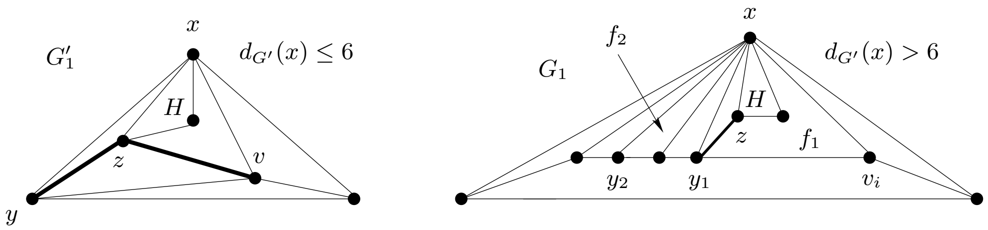


Figure 9.4: Illustrating the proof of Lemma 9.9 where the block H is a triangle. The thick edges are those which were added to G.

their children until we arrive at graphs which are 3-connected. We claim that these graphs then form the leaves of a triangulation tree of G_1' containing no exceptional vertices. In other words, we can never encounter the exceptional case Lemma 9.8 where the two cut-vertices lie in the outer face. This follows since G' is 3-connected and $V(H)\backslash\{x\}$ is connected to two distinct vertices of G' in G_1'. Thus we cannot separate the graph into several components by a cut consisting of exactly two vertices on the outer face. Since the vertices on the outer face are always the same, this proves the claim. We then obtain a triangulation tree of G by adding one more vertex (corresponding to G) and letting its only child be the vertex corresponding to G_1'. Since G_1' had two more edges than G, the root of this tree is an exceptional vertex.

Now suppose that the degree of x in G' is more than six. Let $v_1, \ldots, v_p$, where $p \geq 7$, be the neighbours of x in G', ordered clockwise when viewed from x. Then x has at least four neighbours $y_1', \ldots, y_4'$ (also ordered clockwise when viewed from x) which are not contained in the outer face. Let f_1 be the face of G' which contains x, y_1' and the neighbour v_i of x which lies to the left of y_1' when viewed from x. Then we may have $v_i = y_4'$ but we claim that f_1 cannot also contain y_2' or y_3'. (Indeed, if f_1 contains y_2' for instance, then one can easily verify that the removal of x and y_2' separates y_1' from v_i, contradicting the fact that G' is 3-connected.) Now let f_2 be the face of G' which contains x, y_3' and the neighbour v_j of x which lies to the left of y_3' when viewed from x. Then as above, we may have $v_j = y_2'$ but y_1' is not contained in f_2'. Similarly, it is easily checked that there is no face f which contains x, y_1' and y_3'. Now let $y_1 := y_1'$ and $y_2 := y_3'$. Let z be a vertex in H which is adjacent to x. For $i = 1$, 2, let G_i be the graph obtained from G' by inserting H and the edges xz and zy_i into f_i. Obviously, G_i contains G.

Hence an application of Lemma 9.8 to G_i gives us graphs $G_{i1}', \ldots, G_{is'}'$ and which contain G and which are incomparable. By repeatedly applying the Lemma 9.8 to the graphs G_{ij}' obtained and their children, we will eventually obtain graphs $G_{i1}, \ldots, G_{is}$ which have the additional property of being 3-connected and which form the leaves of a triangulation tree of G_i. As in the previous case, there are no exceptional vertices. (Consider two vertices w_1 and w_2 on the outer face of G_i and let $W := \{w_1, w_2\}$. Since f_i was not the outer face, we have $V(H) \cap W \subseteq \{x\}$. Since H is 2-connected, it follows that $H - W$ is connected. Also, since the y_i are not contained in the outer face, we have

$y_i \notin W$ and thus $G_i - W$ contains the edge zy_i. Finally, since G' is 3-connected, $G' - W$ is connected. Putting these three observations together, we see that $G_i - W$ is connected.)

We will now prove that for any j, G_{1j} is incomparable with G_2 (and thus with $G_{2j'}$, for any j'). For this, by applying Proposition 9.6 with $e_0 = zy_2$, it suffices to show that $zy_2 \in E(G_2) \cap E(\overline{G_{1j}})$ and $zy_2 \notin \mathrm{Ins}(G_{1j})$. To see this, first note that as G' is 3-connected and contains the outer face, in any of the above applications of Lemma 9.8, G' will be contained in the external component and thus the embeddings of G_{1j} when restricted to G' will be the same as that of G_2 when restricted to G'. Now consider the embedding of G_{1j} when restricted to $G' + z$. $G' + z$ contains both zy_1 and zx, so z must be embedded into a face of G' containing both x and y_1. Since G' contains no face which contains x and both of the y_i, this means that $zy_2 \notin E(G_{1j})$ and $zy_2 \notin \mathrm{Ins}(G_{1j})$ and thus the conditions of the proposition are satisfied.

Since $xz \in G$, the G_i have only one more edge than G and we thus obtain a triangulation tree of G (with no exceptional vertices) as follows. We form a single tree from the triangulation trees of the G_i by adding a root vertex corresponding to G and letting its two children correspond to the G_i. □

We can now prove Theorem 9.2 for connected graphs.

Theorem 9.10 *Every connected planar graph $G \in \mathcal{P}_{n,m}$ is contained in at least $\delta 3^{(3n-m)/2}$ labelled triangulations, where δ is an absolute constant.*

Proof. By inserting at most two edges into G, we can obtain a graph G' that has a block B which contains a triangle and which contains at least four vertices. Our aim is to construct a triangulation tree whose root corresponds to G' and whose vertices correspond to graphs obtained from repeated applications of Lemmas 9.8 and 9.9 (and so contain G').

Fix an embedding of B so that the outer face is a triangle. If B is not 3-connected, we apply Lemma 9.8 to obtain a set of children of B. We then apply Lemma 9.8 to those children which are not 3-connected and continue in this way until we have obtained a triangulation tree of B whose root corresponds to B, where the remaining vertices correspond to descendants of B and whose leaves are $B_1, \dots, B_s$, say. Note that we do not assume that the graphs corresponding to the leaves all have the same number of edges. Also on any path from the vertex to the root of this tree, the number of exceptional vertices is at most three – they can only appear if we apply the lemma to a pair of vertices on the outer face and the vertices on the outer face are always the same.

If $B = G'$ (i.e. G' was 2-connected), then we have a triangulation tree of G'. Now suppose that G' was not 2-connected and let H be a block of G' which has a vertex x in common with B (there will be exactly one such vertex in H). Now apply Lemma 9.9 to the graphs $B_i + H$ to obtain triangulation trees of $B_i + H$. We merge these into a single triangulation tree of $B \cup H$ as follows: we start with the triangulation tree of B except that a vertex of the tree which corresponded to a graph F in the triangulation tree of B now corresponds to $F + H$. We then identify the roots of the triangulation trees of $B_i + H$ with the leaves B_i of the tree of B. If $B + H = G'$, then again we have a triangulation tree of G'. If this is not the case, we take a new block H' and apply Lemma 9.9, until we have dealt with all the blocks of G' and thus obtained a triangulation tree of G'.

It is now easy to verify by backward induction on the distance from the root in the triangulation tree that each graph corresponding to a vertex of the tree is contained in the required number of triangulations. By Proposition 9.5, this is certainly true (with $\delta = 2^{-6}$) for the leaves of the tree because they are 3-connected. For the induction step, consider a vertex $F \in \mathcal{P}_{n,m}$ and suppose that its children all have $m+t$ edges and are all contained in at least q triangulations. Then Lemmas 9.8 and 9.9 imply that the number of triangulations of F is at least $q\,3^{t/2}$ (which is exactly what we are aiming for), unless we encountered the exceptional case of either of the lemmas, in which case we find at least $q\,3^{(t-2)/2}$ triangulations. But it is easily seen that on any path from a leaf to the root of the tree we can encounter the exceptional case of Lemma 9.8 at most three times when we build a triangulation tree of B. When we apply Lemma 9.9 to incorporate the other blocks, we can encounter the exceptional case (where x lies on the outer face and has degree at most six) at most 12 times altogether (since an application of the lemma increases the degree of the vertex x in the statement of the lemma by at least one and there are three vertices on the outer face, which are always the same ones). Since G' had at most two more edges than G, this proves the theorem (rather crudely) with $\delta = 3^{-(2/2+3+12)}\,2^{-6}$.

□

Finally, we are in a position to deal with the case where G is not necessarily connected.

Proof of Theorem 9.2. If G has a component K containing all but at most two vertices of G, then we are immediately done by applying Theorem 9.10 to C. So suppose that this is not the case. By inserting at most six edges into G, we can obtain a graph G' which contains two triangles t_a and t_b which are both contained in distinct components. Let v_a be a vertex on t_a and v_b one on t_b. Denote by $K_1, \dots, K_k$ the components of G' not containing t_a or t_b. We now construct 2^k connected plane graphs G_s which all contain G' as a subgraph. Consider some s with $0 \le s < 2^k$ and consider the binary expansion of s. If the ℓth entry is equal to 0, we place the K_ℓ into t_a and add an edge joining K_ℓ to v_a. If the ℓth entry is equal to 1, we do the same with t_b instead. Carrying this out for all ℓ with $1 \le \ell \le k$, this gives us a plane graph G_s.

Thus for each $0 \le i < 2^k$, the graph G_i has two components A_i and B_i with $n_{i,a}$ and $n_{i,b}$ vertices and $m_{i,a}$ and $m_{i,b}$ edges respectively (where $n_{i,a}+n_{i,b} = n$ and $m_{i,a} + m_{i,b} \le m + 6 + k$) and an embedding so that the outer face is bounded by t_a and t_b. Now we apply Theorem 9.10 to all the A_i and B_i. Clearly, we can turn any triangulation $T_{i,a}$ of A_i and $T_{i,b}$ of B_i into one of G_i by considering an embedding of $T_{i,a}$ and $T_{i,b}$ where t_a and t_b are on the outer face and triangulating the outer face of the resulting plane graph in an arbitrary way (we call such triangulation of G_i a *canonical* triangulation of G_i). Since in the last step we only added edges which have one endpoint in t_a and the other in t_b, the number of canonical triangulations of G_i is at least the product of the number of triangulations of A_i and B_i.

Finally, we claim that for any $0 \leq i \neq j < 2^k$, any canonical triangulation T of G_i is incomparable with G_j. Indeed, since $i \neq j$, there must be an ℓ so that K_ℓ was inserted into t_b (by adding an edge e_b between v_b and K_ℓ) in G_j but not in G_i. The claim now follows by applying Proposition 9.6 to e_b, G_j and T. Theorem 9.10 now shows that the number of triangulations of G' is at least

$$\sum_{i=0}^{2^k-1} \delta^2\, 3^{(3n_{i,a}-m_{i,a}+3n_{i,b}-m_{i,b})/2} \geq 2^k\, \delta^2\, 3^{(3n-m-6-k)/2} = \frac{\delta^2}{3^3} 3^{(3n-m)/2} \left(\frac{2}{\sqrt{3}}\right)^k$$

as required. □

We conclude this section by showing that Theorem 9.2 is essentially best possible for $m \geq 2n$.

Proposition 9.11 *For any c with $2 \leq c \leq 3$ and for all $n \in \mathbb{N}$, there is a graph G in $\mathcal{P}_{n',m}$, where $n' = n + o(n)$ and $m = cn + o(n)$, so that the number of triangulations on n' vertices which contain G is $3^{(3n'-m+o(n))/2}$.*

Proof. We construct G as follows. We begin with a square grid on $n/2 + o(n)$ vertices together with an arbitrary triangulation of the outer face of the grid (the number of edges needed to triangulate the outer face is $o(n)$) to obtain a plane graph D. Since we are considering labelled plane graphs, we can speak of the "top row" of the grid, etc. in what follows. We now "augment" each square inside the grid by adding a single new vertex into each face of the square grid and connecting it to both the bottom left and the top right vertex of the face (in other words, an augmented square is a four-cycle where two opposite vertices are connected by an additional path of length two). The resulting graph has $n + o(n)$ vertices, $2n + o(n)$ edges and $n/2 + o(n)$ augmented squares.

We obtain a graph G with $m = cn + o(n)$ edges by selecting $(c-2+o(1))n/2$ squares of the original grid and connecting the corresponding new vertices inside these squares to the remaining two vertices on the outside of this square. We call an augmented square not triangulated in this way *open*.

It remains to verify the claim about the number of triangulations of G. Since D is 3-connected, Proposition 9.4 implies that D has the 1-face property (which can of course also be easily verified directly). Hence, as already noted at the beginning of the proof of Proposition 9.5, the number of triangulations of D is the product of the number of triangulations of the squares in D. It is easily seen that this in turn implies that the number of triangulations of G is the product of the number of triangulations of the open augmented squares. But, clearly, the number of triangulations of each open augmented square is exactly three and there are

$$\frac{n}{2} - \frac{m-2n}{2} + o(n) = \frac{3n-m}{2} + o(n)$$

open squares. □

9.4 Upper bounds

In this section, we apply Theorem 9.2 to deduce upper bounds on the number of planar graphs (implying Theorem 9.1) and on their likely number of edges (Theorem 9.3). For $0 < x < 1$ let

$$H(x) := -x \log x - (1-x)\log(1-x) \tag{9.6}$$

be the *entropy function*, where log denotes the logarithm to base 2 and let $H(0) := H(1) := 0$ for convenience. It has the property that $\binom{n}{xn} = 2^{H(x)n+o(n)}$ for $0 \le x \le 1$.

For $G \in \mathcal{P}_n$ define

$$\beta(G) := |\{H \in \mathcal{T}_n \colon G \subseteq H\}|.$$

For $0 \le c \le 3$ we let

$$\beta(c,n) := \min_{G \in \mathcal{P}_{n,cn}} \beta(G).$$

By the result (9.1) of Tutte, the number $|\mathcal{T}_n|$ of labelled triangulations on n vertices is at most $n!\,\gamma^{n+o(n)}$. This implies that

$$|\mathcal{P}_{n,cn}| \le |\mathcal{T}_n| \frac{\binom{3n-6}{cn}}{\beta(c,n)} \le n!\,\gamma^{n+o(n)} \frac{2^{3H(c/3)n+o(n)}}{\beta(c,n)}. \tag{9.7}$$

Theorem 9.2 implies that $\beta(c,n) \ge 3^{(3-c)n/2+o(n)}$ and thus has the following immediate consequence, which is a generalization of Theorem 9.1.

Theorem 9.12 *For* $0 \le c \le 3$,

$$|\mathcal{P}_{n,cn}| \le n!\,\gamma^{n+o(n)}\, 2^{3H(c/3)n}\, 3^{-(3-c)n/2}.$$

In particular, $|\mathcal{P}_n| \le n!\,(37.3)^{n+o(n)}$ *(the maximum is attained at* $c \sim 1.902$*).*

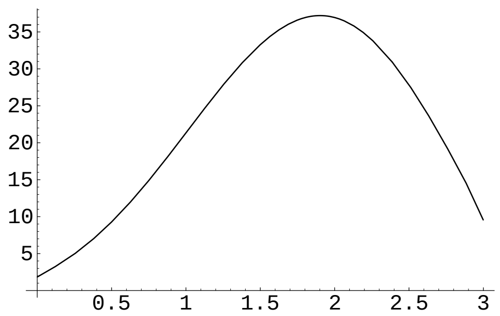


Figure 9.5: The upper bound on $|\mathcal{P}_{n,cn}|^{1/n}/n!$ in Theorem 9.12 versus c, where $0 \le c \le 3$.

Now we turn to the proof of Theorem 9.3, which stated that almost all graphs in $\mathcal{P}_n$ have at most $2.56n$ edges. Let $\mathcal{P}^2$ denote the set of 2-connected planar graphs. In the proof, we will employ the bound $|\mathcal{P}_n| \geq |\mathcal{P}_n^2| = n!\, \alpha^{n+o(n)}$, where $\alpha \sim 26.1876$, due to Bender, Gao and Wormald [9].

Proof of Theorem 9.3. Numerical calculation gives that $\gamma\, 2^{3H(c/3)}\, 3^{-(3-c)/2} < 26.18 - 0.1$ for $c = 2.56$. Note that this inequality also holds for $c \geq 2.56$ as $2^{3H(c/3)}\, 3^{c/2}$ is decreasing in this range. Thus if we apply Theorem 9.12 with $c \geq 2.56$, we have

$$\bigcup_{2.56n \leq m \leq 3n-6} |\mathcal{P}_{n,m}| \leq n!\, 0.44n\ (26.18 - 0.1)^{n+o(n)} = o(|\mathcal{P}_n|),$$

which implies the result. $\square$

Our upper bound in Theorem 9.12 applies only to labelled graphs. However, since any unlabelled graph on n vertices corresponds to at most $n!$ labelled graphs the above result of [9] immediately implies that $|\mathcal{P}_n^u| \geq |\mathcal{P}_n^{2,u}| \geq (26.18)^{n+o(n)}$. Similarly to the proof of Theorem 9.3, one can compare this with the upper bound

$$|\mathcal{P}_{n,cn}^u| \leq \gamma^{n+o(n)} \binom{3n-6}{cn} \tag{9.8}$$

(which follows from (9.1) again) to observe the following result.

Theorem 9.13 *Almost all graphs in $\mathcal{P}_n^u$ have at most $2.69n$ edges.*

Part II

The Regularity Lemma

Chapter 10

The extremal connectivity for topological cliques in bipartite graphs

10.1 Introduction

Given a natural number s, let $d(s)$ be the smallest number such that every graph of average degree $> d(s)$ contains a subdivision of the complete graph K_s of order s. The best asymptotic bounds for $d(s)$ known so far are

$$(1+o(1))\frac{9s^2}{64} \leq d(s) \leq (1+o(1))\frac{s^2}{2}. \tag{10.1}$$

The upper bound is due to Komlós and Szemerédi [58]. As observed by Łuczak, the lower bound is obtained by considering a random subgraph of a complete bipartite graph with edge probability 3/4 (see Proposition 10.12). It is widely believed that the lower bound gives the correct constant, i.e. that random graphs provide the extremal graphs. If true, this would mean that the situation is similar as for ordinary minors. Indeed, Thomason [100] was recently able to prove that random graphs are extremal for minors and Myers [85] showed that all extremal graphs are essentially disjoint unions of pseudo-random graphs.

In this chapter we show that the lower bound in (10.1) is correct if we restrict our attention to bipartite graphs whose connectivity is close to their average degree:

Theorem 10.1 *Given $s \in \mathbb{N}$, let $c_{bip}(s)$ denote the smallest number such that every $c_{bip}(s)$-connected bipartite graph contains a subdivision of K_s. Then*

$$c_{bip}(s) = (1+o(1))\frac{9s^2}{64}.$$

In Theorem 10.1 the condition of being bipartite can be weakened to being H-free for some arbitrary but fixed 3-chromatic graph H (Theorem 10.15). For arbitrary graphs, the best current upper bound on the extremal connectivity is the same as in (10.1). The proof of Theorem 10.1 yields the following improvement.

Theorem 10.2 *Given $s \in \mathbb{N}$, let $c(s)$ denote the smallest number such that every $c(s)$-connected graph contains a subdivision of K_s. Then*

$$(1+o(1))\frac{9s^2}{64} \leq c(s) \leq (1+o(1))\frac{s^2}{4}.$$

The lower bounds in Theorems 10.1 and 10.2 are provided by the random bipartite graphs mentioned above (since their connectivity is close to their average degree). Thus at least in the case of highly connected bipartite graphs they are indeed extremal.

By using methods as in the proof of Theorem 10.1, we also obtain a small improvement for the constant in the upper bound in (10.1).

Theorem 10.3 *Given $s \in \mathbb{N}$, let $d(s)$ denote the smallest number such that every graph of average degree $> d(s)$ contains a subdivision of K_s. Then*

$$(1+o(1))\frac{9s^2}{64} \leq d(s) \leq (1+o(1))\frac{10s^2}{23}.$$

The example of Łuczak mentioned above only gives us extremal graphs for Theorem 10.1 whose connectivity is about $3n/8$, i.e. whose connectivity is rather large compared to the order n of the graph. However, in Proposition 10.14 we show that there are also extremal graphs whose order is arbitrarily large compared to their connectivity. In contrast to this, the situation for ordinary minors seems to be quite different. In general a connectivity of order $s\sqrt{\log s}$ is needed to force a K_s minor, but the connectivity of the known extremal graphs is linear in their order. In fact, Thomason [101] even conjectured that there exists a constant c such that for all integers s there is an integer $n_0 = n_0(s)$ such that every graph of order at least n_0 and connectivity at least cs contains the complete graph K_s as minor. Thus the conjecture says that a linear connectivity suffices to force a K_s minor if we only consider sufficiently large graphs.

This chapter is organized as follows. In Section 10.2 we introduce the necessary notation and tools which we need in the proof of the upper bound of Theorem 10.1. In Section 10.3 we provide the (sparse) extremal examples for the lower bound. The proof of the upper bound of Theorems 10.1, 10.2 and 10.3 is contained in Section 10.4. It builds on results and methods of Komlós and Szemerédi [58]. Finally, in the last section we then briefly discuss the difficulties which arise if one tries to extend Theorem 10.1 to arbitrary graphs.

10.2 Notation and tools

Given constants $0 < \alpha, \beta < 1$, we write $\alpha \ll \beta$ if α is sufficiently small compared with β, i.e. there will always exist a positive $\alpha_0 = \alpha_0(\beta)$ such that the assertion in question holds for all $\alpha \leq \alpha_0$ and $\alpha \ll \beta$ means that $\alpha \leq \alpha_0$.

Our proof of Theorem 10.1 uses Szemerédi's Regularity lemma (Lemma 2.5). In Propositions 10.4 and 10.5 as well as in Lemmas 10.6 and 10.7 R will denote the reduced graph obtained by applying the Regularity lemma (Lemma 2.5) with parameters ε, d and $k_0 := 1/\varepsilon$ to the given graph G. G' and L will be as defined in this lemma. Given an edge $V_iV_j \in R$, for convenience we will refer to the density of $(V_i, V_j)_{G'}$ as the *density of the edge* V_iV_j.

Proposition 10.4 *Suppose that ε, c and d are positive numbers such that $2\varepsilon \leq d < c/2$ and suppose that G is a graph of minimum degree at least $c|G|$. Let μ be the maximum density of an edge in the reduced graph R. Then the minimum degree of R is at least $(c-2d)|R|/\mu$.*

Proof. Set $n := |G|$. Consider a cluster V and let $U \subseteq V(G)$ be the union of all those clusters which are neighbours of V in R. Then there exists a vertex $v \in V$ which has at most $\mu|U|$ neighbours in U (in the graph G'). Indeed, suppose not. Then there exists a cluster $W \in N_R(V)$ such that $e_{G'}(V, W) > \mu|V||W|$. This contradicts the fact that, by definition of μ, the density of $(V, W)_{G'}$ is at most μ. But for every vertex $v \in V$ with at most $\mu|U|$ neighbours in U, we have

$$\mu \cdot d_R(V)L = \mu|U| \geq d_{G'}(v) - |V_0| > d_G(v) - (d + 2\varepsilon)n.$$

Therefore

$$d_R(V) > \frac{(c - d - 2\varepsilon)n}{\mu L} \geq \frac{(c-2d)|R|}{\mu},$$

as required. □

Proposition 10.5 *Let V be a vertex of the reduced graph R and let A be a set of neighbours of V in R. Then, given $\ell \in \mathbb{N}$, there are at most $\ell\varepsilon L$ vertices $v \in V$ which have at most $(d-\varepsilon)L$ neighbours in at least $|A|/\ell$ clusters belonging to A (in the graph G').*

Proof. We say that a vertex $z \in V$ is *bad* for a cluster $W \in A$ if $|N_{G'}(z) \cap W| \leq (d-\varepsilon)L$. Let Z be the set of all those vertices $z \in V$ which are bad for at least $|A|/\ell$ clusters $W \in A$. By summing the number of vertices $z \in Z$ which are bad for W over the clusters $W \in A$, it is easy to see that for some cluster $W \in A$ there are at least $|Z|/\ell$ vertices in Z which are bad for W. Together with Proposition 2.3 this implies that $|Z|/\ell < \varepsilon L$. □

The following lemma is a special case of the well-known 'Embedding lemma' (see e.g. [14, Chapter IV, Thm. 31], [28, Lemma 7.3.2] or [56, Thm. 2.1] for a proof).

Lemma 10.6 *For every $0 < d \leq 1$ and every 3-chromatic graph H there exists a positive constant $\varepsilon_0 = \varepsilon_0(d, H)$ such that for each $0 < \varepsilon \leq \varepsilon_0$ there is an integer $n_0 = n_0(\varepsilon, d, H)$ for which the following holds. Let G be a graph of order at least n_0 and suppose that R contains a triangle. Then G contains a copy of H.*

The next simple observation will only be used in the proof of Theorem 10.3.

Lemma 10.7 *Given positive constants ε and d with $5\varepsilon \leq d < 1$, suppose that P is a path in R with endvertices U and W. Then G' contains at least $(1-2\varepsilon)L$ disjoint paths such that each of them starts in some vertex belonging to U, ends in some vertex belonging to W and contains only vertices belonging to clusters in $V(P)$ (precisely one vertex in each of these clusters).*

Proof. Suppose that $P = V_1 \dots V_r$. First apply Proposition 2.4 to P to obtain a $(1-2\varepsilon)L$-element subset V_i' of each cluster $V_i \in V(P)$ such that $(V_i', V_{i+1}')_{G'}$ is $(2\varepsilon, d-3\varepsilon)$-super-regular for all $1 \le i < r$. It is easily checked that the super-regularity of $(V_i', V_{i+1}')_{G'}$ implies that this graph satisfies Hall's matching condition and thus contains a perfect matching. The union of all these matchings forms a set of paths as required. $\square$

Given bipartite graph (U, W), a set $S \subseteq U$ and numbers $\ell \le |S|$ and $1/\ell \le \beta \le 1$, we say that S is (ℓ, β)-*dense for* W if for each ℓ-element subset S' of S there are at most $\beta|W|$ vertices in W which have less than $\beta|S'|$ neighbours in S'. If $U = S$, this notion can be viewed as a weakening of ε-regularity. Indeed, if (U, W) is ε-regular of density at least $\beta + \varepsilon$, $\beta \ge \varepsilon$ and $\ell \ge \varepsilon|U|$, then by Proposition 2.3 the set U is (ℓ, β)-dense for W. The following even weaker notion will also be convenient. S is called (ℓ, β)-*attached to* W if for each ℓ-element subset S' of S all but at most $\beta|W|$ vertices in W have a neighbour in S'.

Roughly speaking, the next lemma implies that with high probability the ε-regularity of a bipartite graph (U, W) is not lost completely when passing over to a subgraph (S, W) where S is a random subset of U. The point here is that $|S|$ need not be linear in $|U|$. A similar statement was proved independently of us by Gerke et al. [34] in the context of extremal subgraphs of random graphs.

Lemma 10.8 *Given constants $0 < \alpha, \beta, \varepsilon, d < 1$ with $\varepsilon \ll \alpha \le 1/2$, $\varepsilon \le \beta$ and $\beta \ll d$, there exists a natural number $s_0 = s_0(\varepsilon, \beta, \alpha, d)$ such that the following is true for all $s \ge s_0$. Set $\ell := \alpha s$ and suppose that $G = (U, W)$ is an ε-regular bipartite graph of density at least d such that $|U|, |W| \ge s$. Let S be a subset of U which is obtained by successively selecting s vertices in U uniformly at random without repetitions. Then with probability at least $1 - \mathrm{e}^{-s}$ the set S is (ℓ, β)-dense for W.*

Proof. Consider a subset S' of U which is obtained by successively selecting ℓ vertices in U uniformly at random without repetitions. We call S' a *failure* if there are at least $\beta|W|$ vertices in W which have less than $\beta|S'|$ neighbours in S'. We will first show that the probability that S' is a failure is very small. This will be done by grouping the vertices in S' into successive 'epochs' and by analyzing one such 'epoch' at a time. Set $r := d/(8\beta)$; r will be the number of such 'epochs' and so each 'epoch' will contain $\ell/r = 8\beta\ell/d$ vertices. We call the subset of S' which consists of the first ℓ/r vertices chosen for S' the *first epoch* of S' and denote it by S_1'. Similarly, given $2 \le i \le r$, we define the *ith epoch* S_i' of S'. Given $1 \le i \le r$, let W_i be the set of all those vertices $w \in W$ which have at least $\beta\ell$ neighbours in S_i'. For all $i \le r+1$ set $W_i' := W \setminus \bigcup_{j<i} W_j$. Thus W_i' contains all those vertices for which, after $i-1$ epochs, we cannot guarantee that they have enough neighbours in S'. We say that the ith epoch S_i' is *successful* if either $|W_i'| < \beta|W|$ or if at least half of the vertices in S_i' have at least $d|W_i'|/2$ neighbours in W_i'.

The aim now is to show that if the ith epoch is successful and W_i' is still large, then a significant proportion of the vertices in W_i' will belong to W_i. Since the probability that an epoch is successful will turn out to be quite large, this will then imply that with large probability the set W'_{r+1} is small and thus with large probability S' is not a failure. Set $\mu_i := |W'_i \cap W_i|/|W'_i|$ and suppose that the ith epoch S'_i is successful but $|W'_i| \geq \beta|W|$. By counting the edges between W'_i and S'_i and recalling that $|S'_i| = 8\beta\ell/d$, we get

$$\mu_i|W'_i| \cdot \frac{8\beta\ell}{d} + |W'_i| \cdot \beta\ell \geq e_G(W'_i, S'_i) \geq \frac{4\beta\ell}{d} \cdot \frac{d|W'_i|}{2}.$$

Hence

$$\mu_i \geq d/8.$$

We now bound the probability that an epoch is not successful. Since (U, W) is ε-regular and has density at least d, Proposition 2.3 implies that if $|W'_i| \geq \beta|W| \geq \varepsilon|W|$ then at most $\varepsilon|U|$ vertices in U have less than $d|W'_i|/2$ neighbours in W'_i. So in this case, for every $s \in S'_i$, the probability that s has less than $d|W'_i|/2$ neighbours in W'_i is at most $\varepsilon|U|/(|U| - \ell) \leq 2\varepsilon$. Thus for any event $\mathcal{A}_{i-1}$ depending only on the outcome of the first $i-1$ epochs, we have

$$\mathbb{P}(S'_i \text{ is not successful} \mid \mathcal{A}_{i-1}) \leq 2^{|S'_i|}(2\varepsilon)^{|S'_i|/2} = (8\varepsilon)^{4\beta\ell/d}.$$

Hence

$$\mathbb{P}(\text{at least } r/2 \text{ epochs are not successful}) \leq 2^r(8\varepsilon)^{2\beta\ell r/d} \leq (16\varepsilon)^{\ell/4}.$$

Let N denote the number of successful epochs. Then $|W'_{N+1}| \leq \max\{\beta|W|, (1-d/8)^N|W|\}$. But if $N \geq r/2$ we have

$$(1 - d/8)^N|W| \leq (1 - d/8)^{d/(16\beta)}|W| \leq \mathrm{e}^{-d^2/(128\beta)}|W| \leq \beta|W|.$$

This shows that with probability at most $(16\varepsilon)^{\ell/4}$, a random ℓ-set S' is a failure.

Now suppose that S is a random set as given by the lemma. Then, since every ℓ-element subset S' of S is again a random set whose distribution is uniform amongst all ℓ-element subsets of U,

$$\begin{aligned}
\mathbb{P}(S \text{ is not } (\ell,\beta)\text{-dense for } W) &\leq \sum_{S' \in S^{(\ell)}} \mathbb{P}(S' \text{ is a failure}) \\
&\leq \binom{s}{\ell}(16\varepsilon)^{\ell/4} \\
&\leq \left(\frac{\mathrm{e}s}{\ell}\right)^{\ell}(16\varepsilon)^{\ell/4} \\
&= \mathrm{e}^{\ell\log(\mathrm{e}/\alpha)}\mathrm{e}^{-(\ell/4)\log(1/(16\varepsilon))} \\
&\leq \mathrm{e}^{-(\ell/5)\log(1/\varepsilon)} \leq \mathrm{e}^{-s},
\end{aligned}$$

as required. (The third inequality is a weak form of Stirling's formula, see e.g. [16, p. 4].) □

The following special case of Lemma 10.8 was already proved by Komlós and Szemerédi [58]. A result which is slightly stronger than Corollary 10.9 was also proved earlier by Duke and Rödl [30].

Corollary 10.9 *Under the conditions of Lemma 10.8, the set S is (ℓ, β)-attached to W with probability at least $1 - \mathrm{e}^{-s}$.*

Given a positive number ε and sets $A, Q \subseteq T$, we say that A *is split ε-fairly by Q* if

$$\left| \frac{|A \cap Q|}{|Q|} - \frac{|A|}{|T|} \right| \leq \varepsilon.$$

Thus, if ε is small and A is split ε-fairly by Q, then the proportion of all those elements of T which lie in A is almost equal to the proportion of all those elements of Q which lie in A. We will use the following version of the well-known fact that if Q is random then it tends to split large sets ε-fairly.

Proposition 10.10 *For each $0 < \varepsilon < 1$ there exists an integer $q_0 = q_0(\varepsilon)$ such that the following holds. Given $t \geq q \geq q_0$ and a set T of size t, let Q be a subset of T which is obtained by successively selecting q elements uniformly at random without repetitions. Let $\mathcal{A}$ be a family of at most q^{10} subsets of T such that $|A| \geq \varepsilon t$ for each $A \in \mathcal{A}$. Then with probability at least $1/2$ every set in $\mathcal{A}$ is split ε-fairly by Q.*

To prove Proposition 10.10 we will use the following large deviation bound for the hypergeometric distribution (see e.g. [39, Thm. 2.10 and Cor. 2.3]).

Lemma 10.11 *Given $q \in \mathbb{N}$ and sets $A \subseteq T$ with $|T| \geq q$, let Q be a subset of T which is obtained by successively selecting q elements of T uniformly at random without repetitions. Let $X := |A \cap Q|$. Then for all $0 < \varepsilon < 1$ we have*

$$\mathbb{P}(|X - \mathbb{E}X| \geq \varepsilon \mathbb{E}X) \leq 2\mathrm{e}^{-\frac{\varepsilon^2}{3}\mathbb{E}X}.$$

Proof of Proposition 10.10. Given $A \in \mathcal{A}$, Lemma 10.11 implies that

$$\begin{aligned} \mathbb{P}(A \text{ is not split } \varepsilon\text{-fairly by } Q) &\leq \mathbb{P}(|\,|A \cap Q| - q|A|/t\,| \geq \varepsilon q|A|/t) \\ &\leq 2\mathrm{e}^{-\frac{\varepsilon^2}{3}\frac{q|A|}{t}} \leq 2\mathrm{e}^{\frac{-\varepsilon^3 q}{3}}. \end{aligned}$$

Hence, if q_0 is sufficiently large compared with ε, the probability that there is an $A \in \mathcal{A}$ which is not split ε-fairly is at most $2q^{10}\mathrm{e}^{-\varepsilon^3 q/3} < 1/2$, as required.

□

10.3 Proof of Theorem 10.1 – extremal graphs

As mentioned in [58], the following example of Łuczak shows that the function $c_{bip}(s)$ defined in Theorem 10.1 is at least $(1 + o(1))9s^2/64$ (and thus also the functions $c(s)$ and $d(s)$ defined in Section 10.1).

Proposition 10.12 *For every positive λ and each integer κ_0 there exists a bipartite graph G such that G is κ-connected for some $\kappa \geq \kappa_0$ and does not contain a subdivision of a clique of order at least $(1+\lambda)8\sqrt{\kappa}/3$.*

We include the proof of Proposition 10.12 here firstly for completeness and secondly because we will build on the argument in the proof of Proposition 10.14 below. In both proofs, the next simple and well-known fact (see e.g. [16, Ch. II, Thm. 2.1]) will be rather useful.

Theorem 10.13 *Let $n \in \mathbb{N}$ and let $0 < \varepsilon, p < 1$ be fixed. Let B_{np} be a bipartite random graph whose vertex classes A and B both have size n and where the edges between these classes are included with probability p independently of each other. Then, with probability tending to 1 as $n \to \infty$,*

$$(1-\varepsilon)p|U||W| \leq e_{B_{np}}(U,W) \leq (1+\varepsilon)p|U||W|$$

for all sets $U \subseteq A$ and $W \subseteq B$ with $|U|,|W| \geq (\log n)^2$.

Proof of Proposition 10.12. Throughout the proof we assume that λ is sufficiently small and n is sufficiently large for our estimates to hold. Let $\kappa := (1-\lambda/2)3n/4$ and $s := (1+\lambda)8\sqrt{\kappa}/3$. Put $p := 3/4$ and let B_{np} be a bipartite random graph as in Theorem 10.13. Using the lower bound in Theorem 10.13, one can easily show that B_{np} is κ-connected with probability tending to 1 as $n \to \infty$. We will show that, with probability tending to 1, there will not be any sets $U \subseteq A$ and $W \subseteq B$ such that $U \cup W$ can serve as the set of branch vertices of a subdivided K_s in B_{np}. Without loss of generality we assume that $|U| \geq |W|$. Clearly, if $|W| \leq (\log n)^2$, then B_{np} cannot contain a subdivided edge for all the pairs of vertices in U since each such edge must have an inner vertex in B and $|B| < \binom{s-(\log n)^2}{2}$. But Theorem 10.13 implies that with probability tending to 1 we have

$$e_{B_{np}}(U,W) \leq (1+\lambda/30)p|U||W| \tag{10.2}$$

for all U, W with $|U|,|W| \geq (\log n)^2$. However, if $U \cup W$ is the set of branch vertices of a TK_s, then B contains an inner vertex of each subdivided edge joining a pair of vertices in U as well as an inner vertex of each subdivided edge which joins some $a \in U$ to some $b \in W$ with $ab \notin B_{np}$. Thus, if (10.2) holds, then the number of all these subdivided edges is

$$\binom{|U|}{2} + |U||W| - e_{B_{np}}(U,W) > n.$$

This shows that with probability tending to 1 the graph B_{np} does not contain a subdivided K_s. Thus, with probability tending to 1, we can take B_{np} for the graph G in Proposition 10.12. $\square$

If we take a sequence of disjoint copies of the graph given by Proposition 10.12 and attach successive copies by inserting κ independent edges, then the next proposition shows that we obtain arbitrarily large κ-connected bipartite graphs which do not contain a subdivision of a large clique (and the density of these graphs is arbitrarily small).

Proposition 10.14 *For every positive λ and every integer κ_0 there exists an integer $\kappa \geq \kappa_0$ and arbitrarily large bipartite graphs G which are κ-connected and do not contain a subdivision of a clique of order at least $(1+\lambda)8\sqrt{\kappa}/3$.*

Proof. Throughout the proof we assume that λ is sufficiently small and n is sufficiently large for our estimates to hold. Let $G = (A, B)$ be the bipartite (random) graph given by the proof of Proposition 10.12. Thus $|A| = |B| = n$, G is κ-connected where $\kappa := (1-\lambda/2)3n/4$ and all sets $U \subseteq A$ and $W \subseteq B$ with $|U|, |W| \geq (\log n)^2$ satisfy

$$e(U, W) \leq (1+\lambda/30)3|U||W|/4. \tag{10.3}$$

Moreover, G does not contain a subdivision of K_s where $s := (1+\lambda)8\sqrt{\kappa}/3$. Given an integer k, let G^* denote the graph obtained from k disjoint copies $G_1 = (A_1, B_1), \dots, G_k = (A_k, B_k)$ of G by inserting κ independent edges between B_i and A_{i+1} (for all $1 \leq i < k$). Thus G^* is κ-connected and bipartite. We will show that G^* does not contain a subdivided K_s. So suppose not and choose a TK_s in G^*. For each $i \leq k$ let X_i be the set of all branch vertices of TK_s in $G_1 \cup \dots \cup G_i$ and let Y_i be the set of all branch vertices in $G_i \cup \dots \cup G_k$. Since each subdivided edge joining a branch vertex in X_i to a branch vertex in Y_{i+1} must contain one of the κ edges between B_i and A_{i+1}, we have $\kappa \geq |X_i||Y_{i+1}| = |X_i|(s - |X_i|)$. This implies that for each i either $|X_i| \leq 0.17s$ or $|X_i| \geq 0.83s$. Let i be the first index for which the latter holds. Thus $x := |X_{i-1}| \leq 0.17s$ and $y := |Y_{i+1}| \leq 0.17s$. Let S_A be the set of all branch vertices in A_i and let S_B be the set of all branch vertices in B_i. Put $X := X_{i-1}$, $Y := Y_{i+1}$, $s_A := |S_A|$ and $s_B := |S_B|$.

Let us now estimate the number of all those vertices in A_i which are contained in the TK_s. Firstly, since all the B_{i-1}–A_i edges are independent, A_i contains at least $x(s-x)$ vertices on subdivided edges joining a branch vertex in X to a branch vertex in $S_A \cup S_B \cup Y$. Secondly, at most $s_B/2$ subdivided edges joining two branch vertices in S_B begin and end with an S_B–A_{i+1} edge. But all the $\binom{s_B}{2} - s_B/2$ remaining such subdivided edges have an inner vertex in A_i. (Note that this also shows that $s_A \geq (\log n)^2$ since otherwise $\binom{s_B}{2} - s_B/2 > n$. Similarly, we have that $s_B \geq (\log n)^2$.) Thirdly, at most s_B subdivided edges joining some branch vertex $a \in S_A$ to some branch vertex $b \in S_B$ with $ab \notin G_i$ end with an S_B–A_{i+1} edge. Again all the at least $s_A s_B - e_{G_i}(S_A, S_B) - s_B$ remaining such subdivided edges must have an inner vertex in A_i. Since $s_A, s_B \geq (\log n)^2$, together with (10.3) this implies that

$$n = |A_i| \geq x(s-x) + \binom{s_B}{2} - \frac{s_B}{2} + \left(\frac{1}{4} - \frac{\lambda}{40}\right) s_A s_B - s_B. \tag{10.4}$$

Similarly, we arrive at an analogous inequality where A and B are interchanged and x is replaced by y. Adding (10.4) and this second inequality gives

$$x(s-x)+y(s-y)+\binom{s_A}{2}+\binom{s_B}{2}+\frac{s_A s_B}{2}-\frac{3}{2}(s_A+s_B)-\frac{\lambda s_A s_B}{20} \le 2n. \quad (10.5)$$

But $\binom{s_A}{2}+\binom{s_B}{2}+s_A s_B/2$ is minimized if $s_A = s_B$, i.e. if $s_A = s_B = (s-x-y)/2$. Thus (10.5) implies that

$$x(s-x)+y(s-y)+2\binom{\frac{s-x-y}{2}}{2}+\frac{1}{2}\left(\frac{s-x-y}{2}\right)^2-\frac{\lambda s^2}{16} \le 2n.$$

This shows that

$$2xs+2ys-5(x^2+y^2)+6xy+\lambda s^2 \le 0. \quad (10.6)$$

However, recall that $x, y \le 0.17s$. It is easy to check that (10.6) has no solution for such numbers x and y. □

10.4 Proof of Theorem 10.1 – upper bound

Clearly, it suffices to prove the following stronger statement. It implies that in Theorem 10.1 the condition of being bipartite can be weakened to being H-free where H is any fixed 3-chromatic graph.

Theorem 10.15 *For every $0 < \lambda < 1$ and every 3-chromatic graph H there exists $\kappa_0 = \kappa_0(\lambda, H)$ such that for every natural number $\kappa \ge \kappa_0$ each κ-connected graph G_0 contains a subdivision of a clique of order at least $8\sqrt{(1-\lambda)\kappa}/3$.*

For the proof of Theorem 10.15 we need the following consequence of Theorem 2.1 and Corollary 2.1 in Komlós and Szemerédi [58].

Theorem 10.16 *For all $\varepsilon^* > 0$ there are positive constants $c_0 = c_0(\varepsilon^*)$ and $d_0 = d_0(\varepsilon^*)$ such that every graph G^* of average degree at least $d^* \ge d_0$ either contains a subdivided clique of order at least $8\sqrt{d^*}/3$ or a subgraph G whose average degree d satisfies both $d \ge c_0|G|$ and $d \ge d^*/(1+\varepsilon^*)$.*

Very roughly, the strategy of the proof of Theorem 10.1 is as follows. By Theorem 10.16, we may assume that our given graph G_0 contains a dense subgraph G. We then apply the Regularity lemma to G to obtain a reduced graph R. If R contains a vertex X of rather large degree (Case 1), we choose the set of our branch vertices randomly inside X. In this case—similarly as in the proof of Komlós and Szemerédi [58]—most of the branch vertices can be joined by a path of length two whose midpoint lies in some cluster which is adjacent to X in R. The main difference is that here we need the more powerful Lemma 10.8 instead of Corollary 10.9 (which was sufficient in [58]). The left-over pairs of branch vertices are then joined by suitable paths of length four using special sets of vertices which we set aside earlier for this purpose.

If we cannot guarantee a vertex of large degree in R (Case 2) we proceed as follows. Let XY be an edge in R of maximum density. Proposition 10.4 implies that in Case 2 this density must be large. The branch vertices are now chosen within both X and Y. This has the advantage that many pairs of branch vertices can be connected directly by edges between them. The subdivided edges connecting two branch vertices in X (respectively two branch vertices in Y) are selected similarly as in Case 1. The main difficulty of the proof is that now we have to use the connectivity of G_0 in order to find subdivided edges joining every branch vertex $x \in X$ to all those branch vertices $y \in Y$ for which $xy \notin G_0$.

Proof of Theorem 10.15. Choose

$$\varepsilon^* \ll \lambda. \tag{10.7}$$

Let $c_0(\varepsilon^*)$ be as defined in Theorem 10.16 and choose constants

$$0 < \varepsilon \ll \beta \ll \alpha \ll \xi \ll \tau \ll d \ll \min\{c_0(\varepsilon^*), \varepsilon^*\}. \tag{10.8}$$

We will prove Theorem 10.15 for every κ_0 which is sufficiently large compared to each of $d_0(\varepsilon^*)$, $N(\varepsilon)$, $q_0(\varepsilon)$, $q_0((80N(\varepsilon))^{-1})$, $n_0(\varepsilon, d, H)$, $s_0(2\varepsilon, \beta, \alpha, d/2)$ and $s_0(2\varepsilon, \beta^2, \alpha, d/2)$, where d_0 is as defined in Theorem 10.16, $N(\varepsilon) := N(\varepsilon, 1/\varepsilon)$ is as defined in the Regularity lemma (Lemma 2.5), q_0 is as defined in Proposition 10.10, n_0 is as defined in Lemma 10.6 and s_0 is as defined in Lemma 10.8. Clearly, we may assume that the graph G_0 given in Theorem 10.15 does not contain a subgraph of connectivity greater than κ. By Theorem 10.16, we may assume that for some $c \geq c_0(\varepsilon^*)$ the graph G_0 contains a subgraph G of average degree

$$cn \geq \kappa/(1+\varepsilon^*), \tag{10.9}$$

where $n := |G|$. Then

$$4\kappa \geq cn \tag{10.10}$$

since otherwise, by Theorem 7.4, G would contain a subgraph whose connectivity is greater than κ. Set

$$s := 8\sqrt{(1-\lambda)\kappa}/3.$$

Apply the Regularity lemma (Lemma 2.5) with parameters ε, d and $k_0 := 1/\varepsilon$ to G to obtain a spanning subgraph G' of G and a reduced graph R. Let $N(\varepsilon) := N(\varepsilon, 1/\varepsilon)$. Throughout the proof, unless stated otherwise, we say that two vertices $x, y \in V(G) = V(G')$ are neighbours if they are neighbours in G'. Let L denote the size of the clusters given by the Regularity lemma and set $k := |R|$. Let μ denote the maximum density of an edge in R. Thus $\mu \geq d$ and Proposition 10.4 shows that

$$\delta(R) \geq (c - 2d)k/\mu =: \delta. \tag{10.11}$$

We will first deal with the case when $\mu \leq 9/32$.

Case 1. $\mu \leq 9/32$.

Let $X \in V(R)$ be any cluster. Choose disjoint sets N^1_X and N^2_X of neighbours of X in R such that $|N^1_X| = \tau k$ and $|N^2_X| = \delta - 10\tau k$. Next choose a set $\widetilde{N}^1_X$ consisting of τk vertices of R such that R contains a perfect matching between N^1_X and $\widetilde{N}^1_X$ and such that $\widetilde{N}^1_X$ is disjoint from each of $\{X\}$, N^1_X and N^2_X. We will fix such a perfect matching between N^1_X and $\widetilde{N}^1_X$. By Proposition 10.5, all but at most $3\varepsilon L$ vertices in X have at least $dL/2$ neighbours in at least $2/3$ of the clusters in N^1_X. Let $X' \subseteq X$ be the set of these vertices. Thus $|X'| \geq (1-3\varepsilon)L$. Together with Lemma 10.8 and Corollary 10.9 this implies that X' contains an s-element subset S which is $(\alpha s, \beta)$-dense for each cluster $W \in N^2_X$ and which in addition is $(\alpha s, \beta^2)$-attached to each cluster $W \in N^2_X$. (Indeed, since for each $W \in N^2_X$ the graph $(X', W)_{G'}$ is 2ε-regular of density at least $d/2$, Lemma 10.8 and Corollary 10.9 together imply that the probability that an s-element subset S of X' chosen uniformly at random without repetitions fails to satisfy the above conditions is at most $2|N^2_X|\mathrm{e}^{-s} < 1$.) S will be the set of branch vertices of our subdivided clique. Let Z be the set of all those vertices of G that lie in some cluster belonging to N^2_X.

To find the subdivided edges of our clique, for every pair of vertices $x, y \in S$ in turn, we select a vertex $z \in Z$ which is adjacent to both x and y and was not already chosen to connect a previous pair (provided that such a vertex exists). We call a vertex $x \in S$ *bad* if, after we have considered all pairs of vertices in S in this way, there are still at least αs vertices in S which are not yet joined to x. The following claim implies that we were able to join most of the pairs of branch vertices in the above way.

Claim A. *At most αs vertices in S are bad.*

Suppose not and let S' be an αs-element subset of S consisting of bad vertices. Let $Z' \subseteq Z$ be the set of all those vertices in Z which have not been selected to join some pair of vertices in S. Then, since $\mu \leq 9/32$,

$$\begin{aligned}
|Z'| \quad &> \quad |Z| - \binom{s}{2} \geq |N^2_X|L - \frac{s^2}{2} \\
&\geq \quad \frac{(c-2d)kL}{\mu} - 10\tau kL - \frac{32(1-\lambda)\kappa}{9} \\
&\overset{(10.9)}{\geq} \quad \frac{32(c-2d)(1-\varepsilon)n}{9} - 10\tau n - \frac{32(1-\lambda)(1+\varepsilon^*)cn}{9} \\
&\overset{(10.7),(10.8)}{\geq} \quad 2\beta n \geq 2\beta|Z|.
\end{aligned}$$

But since S was $(\alpha s, \beta)$-dense for each cluster belonging to N^2_X, it follows that at least half of the vertices in Z' have at least $\beta S'$ neighbours in the bad set S'. Thus there is a vertex $x \in S'$ with at least $\beta|Z'|/2 > \beta^2|Z|$ neighbours in Z'. Hence there exists a cluster $W \in N^2_X$ such that x has more than $\beta^2 L$ neighbours in $W \cap Z'$. Since S was $(\alpha s, \beta^2)$-attached to W, there must be an edge joining some neighbour z of x in $W \cap Z'$ to one of the at least αs vertices in S which are not yet joined to x, y say. But this means that when we considered the pair x, y we could have selected z in order to join them, a contradiction. This proves the claim.

Now we have to show that we can find a subdivided edge for each of the at most $2\alpha s^2$ left-over pairs of vertices in S. We will join up each such left-over pair greedily by a path of length 4. This 4-path will have its midpoint in some cluster $V \in \widetilde{N}^1_X$ and its other two inner vertices in the unique cluster in N^1_X that is matched to V. (Recall that we have fixed a perfect matching between N^1_X and $\widetilde{N}^1_X$.) We have to show that for all the left-over pairs in turn we can find (greedily) internally disjoint such paths. Suppose that we are about to join the left-over pair $x, y \in S$. Recall that, since $S \subseteq X'$, both x and y have at least $dL/2$ neighbours in at least $2/3$ of the clusters in N^1_X. Thus they have at least $1/3$ of these clusters in common. However, we may have used up some of the neighbours of x and y before to join up previous left-over pairs. But since the number of paths constructed previously is at most $2\alpha s^2$, we have used at most $6\alpha s^2 \leq 48\alpha n$ vertices for this. Thus at most

$$48\alpha n \cdot \frac{4}{dL} \overset{(10.8)}{<} \frac{\tau k}{3} = \frac{|N^1_X|}{3}$$

clusters in N^1_X contain at least $dL/4$ vertices which we have already used before. This shows that there is a cluster $U \in N^1_X$ in which both x and y still have at least $dL/4$ unused neighbours. Let $V \in \widetilde{N}^1_X$ be the cluster that is matched to U. Since by construction the number of used vertices in U is exactly twice the number of used vertices in V, there must be at least $dL/4$ vertices in V which we have not used already. Together with the ε-regularity of $(U, V)_{G'}$ this implies that V contains a vertex z which is joined to both some neighbour z_1 of x in U and some neighbour z_2 of y in U such that $z_1 \neq z_2$ and such that none of z_1, z_2, z_3 has been used to join previous left-over pairs. Thus xz_1zz_2y is a 4-path as required.

Case 2. $\mu > 9/32$.

The proof of this case is an extension of that of Case 1. Let XY be an edge in R of density μ. Since μ is large, the lower bound (10.11) on $\delta(R)$ is now weaker and so we cannot choose all the branch vertices in a single cluster, X say, as we did in Case 1. Indeed, the number of vertices lying in a neighbouring cluster of X could be smaller than $\binom{s}{2}$. So if we put all the branch vertices into X, there may not be enough room for all the subdivided edges of our topological clique. Therefore we split our branch vertices into two sets $S_X \subseteq X$ and $S_Y \subseteq Y$ such that the density of the bipartite subgraph between S_X and S_Y is about μ. Since μ is quite large, this has the advantage that we can join many pairs of branch vertices directly by these S_X–S_Y edges and so we need less vertices in the other neighbouring clusters of X (respectively of Y) for the remaining subdivided edges. However, we now face the additional difficulty that we also have to join each vertex in S_X to all those vertices in S_Y for which there is no S_X–S_Y edge. This is the point where we use the connectivity of the graph $G_0 \supseteq G$ we started with. (Note that in Case 1 we did not make any use of it.)

Select τk-element sets N^1_X and $\widetilde{N}^1_X$ and a $(\delta - 10\tau k)$-element set N^2_X similarly as in Case 1. But now we additionally require that Y does not belong to any of these sets. Next choose analogous sets N^1_Y, $\widetilde{N}^1_Y$ and N^2_Y. Since $G_0 \supseteq G$ is H-free, Lemma 10.6 implies that the neighbourhoods of both X and Y

are disjoint. Thus, we can choose all the sets $N_X^1, \widetilde{N}_X^1, N_X^2, N_Y^1, \widetilde{N}_Y^1, N_Y^2$ to be pairwise disjoint. (This is the only time we need the fact that G_0 is H-free.) The sets N_X^1 and $\widetilde{N}_X^1$ have the same purpose as in Case 1, namely to connect those left-over pairs $x, x' \in S_X$ of branch vertices by paths of length 4 which we were not able to link by paths of length 2. Each other path linking a pair of branch vertices will be routed through N_X^2 and/or N_Y^2 (see Fig. 10.1).

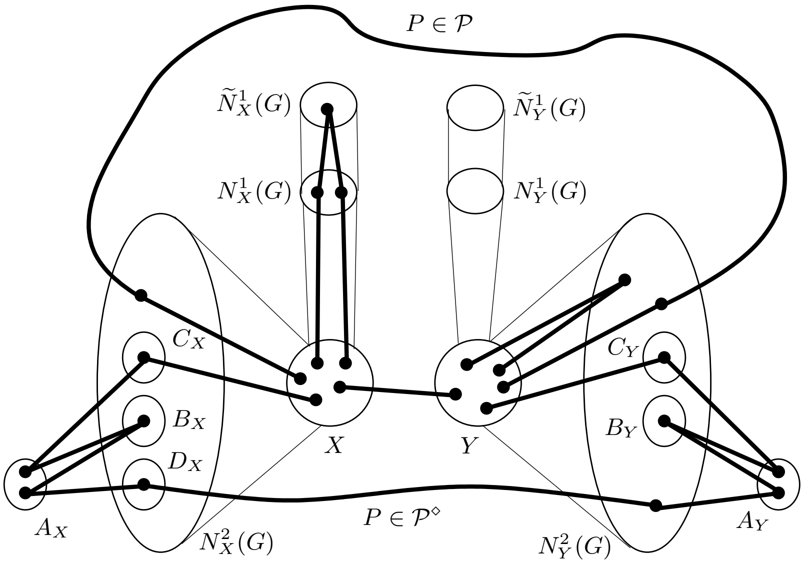


Figure 10.1: Five possible ways of connecting two branch vertices. The sets $N_X^1(G)$ etc. denote the subsets of $V(G)$ which correspond to the sets N_X^1 etc.

Let $1/2 \le \gamma \le 9/10$ be any number which satisfies the following two inequalities:

$$(1-\mu+10^6\varepsilon)\gamma s(1-\gamma)s \le (1-2\varepsilon^*)\kappa, \tag{10.12}$$

$$\binom{\gamma s}{2} + (1-\mu+10^6\varepsilon)\gamma s(1-\gamma)s + 10\alpha s^2 \le |N_X^2|L - \tau n. \tag{10.13}$$

We defer the technical proof of the existence of such a γ until the end of this section (Proposition 10.17). Inequality (10.12) will imply that the connectivity of G_0 is large enough to guarantee at least as many paths between the neighbourhood of S_X (inside clusters belonging to N_X^2) and the neighbourhood of S_Y as we will need to join all those pairs $x \in S_X$, $y \in S_Y$ of branch vertices for which $xy \notin G_0$. Inequality (10.13) will show that the neighbourhood of S_X is large enough to accommodate both an endvertex of each such path as well as a midpoint of each subdivided edge joining two branch vertices in S_X. (Similarly as in Case 1, we will join almost all pairs of branch vertices in S_X by paths of length 2.)

Set

$$s_X := \gamma s \quad \text{and} \quad s_Y := (1-\gamma)s.$$

We will now choose the set $S_X \cup S_Y =: S$ of branch vertices for our subdivided clique where $S_X \subseteq X$, $S_Y \subseteq Y$, $|S_X| = s_X$ and $|S_Y| = s_Y$. Note that, by

Propositions 2.3 and 10.5, all but at most $(10^5+4)\varepsilon L$ vertices $x \in X$ satisfy the following three properties:

(i) The proportion of clusters $U \in N_X^2$ for which x has at most $d|U|/2$ neighbours in U is at most 10^{-5}.

(ii) The proportion of the clusters $U \in N_X^1$ for which x has at most $d|U|/2$ neighbours in U is at most $1/3$.

(iii) The neighbourhood of x in Y has size at least $(\mu - \varepsilon)L$.

Let X' be the set of all those at least $(1-(10^5+4)\varepsilon)L$ vertices in X. Define $Y' \subseteq Y$ similarly. Just as in Case 1, one can apply Lemma 10.8 and Corollary 10.9 to obtain an s_X-element set $S_X \subseteq X'$ which is $(\alpha s_X, \beta)$-dense for each cluster $U \in N_X^2$ and which in addition is $(\alpha s_X, \beta^2)$-attached to each cluster $U \in N_X^2$. Similarly, using Lemma 10.8, Corollary 10.9 and Proposition 10.10, it is easy to see that there exists an s_Y-element subset $S_Y \subseteq Y'$ which is $(\alpha s_Y, \beta)$-dense for each cluster $V \in N_Y^2$, which in addition is $(\alpha s_Y, \beta^2)$-attached to each cluster $V \in N_Y^2$ and for which the bipartite graph $(S_X, S_Y)_{G'}$ has density at least $\mu - 10^6\varepsilon$. Indeed, to ensure that the latter property is also satisfied, let $\mathcal{A} := \{N_{G'}(x) \cap Y' \,|\, x \in S_X\}$. Since (iii) implies that $|A| \geq (\mu-\varepsilon)L - |Y \setminus Y'| \geq (\mu-(10^5+5)\varepsilon)|Y'|$ for all $A \in \mathcal{A}$, Proposition 10.10 (with $T := Y'$ and $Q := S_Y$) tells us that the probability that there exists a vertex $x \in S_X$ which has less than $(\mu - 10^6\varepsilon)s_Y$ neighbours in S_Y is at most $1/2$. This completes the choice of the branch vertices.

As indicated earlier, we will use the connectivity of G_0 to find a set $\mathcal{P}$ of almost κ disjoint paths whose first vertex lies in a cluster belonging to N_X^2 and whose last vertex lies in a cluster belonging to N_Y^2. Most of those pairs x, y of branch vertices for which $x \in S_X$, $y \in S_Y$ and $xy \notin G_0$ will be joined by a path of the form xPy where $P \in \mathcal{P}$. However, for some such pairs x, y this will not be possible. Each of those left-over pairs x, y will be joined by an extended path of the form $xu_1 \dots u_4 P v_4 \dots v_1 y$ where $P \in \mathcal{P}$. All these extension vertices $u_1, \dots, u_4$ and $v_1, \dots, v_4$ will lie in a relatively small set I' which we set aside (before determining $\mathcal{P}$) for this purpose and which will be avoided by the paths in $\mathcal{P}$. I' will be the union of six disjoint sets A_X, B_X, C_X, A_Y, B_Y and C_Y. All vertices of the form u_1 will lie in C_X, all vertices of the form u_2 and u_4 will lie in A_X and all vertices of the form u_3 will lie in B_X. The vertices of the form v_i will satisfy analogous properties for the sets A_Y, B_Y and C_Y (see Fig. 10.1).

Let us first choose the set A_X. For each cluster $U \in N_X^2$ we select a neighbour $W(U)$ in R such that all these $W(U)$ are distinct for different clusters U and such that none of them lies in

$$N_X^1 \cup \widetilde{N}_X^1 \cup N_Y^1 \cup \widetilde{N}_Y^1 \cup \{X, Y\} =: J. \tag{10.14}$$

Let U' be the set of all those vertices in U which have at least $dL/2$ neighbours in $W(U)$. Thus, by Proposition 2.3, $|U'| \geq (1-\varepsilon)L$. Apply Proposition 10.10 (with $T := W(U)$, $q := \tau L$ and $\mathcal{A} := \{N_{G'}(x) \cap W(U) \,|\, x \in U'\}$) to obtain a τL-element subset $A_X(U)$ of $W(U)$ such that every vertex in U' has at least $d|A_X(U)|/4$ neighbours in $A_X(U)$. Let $A_X := \bigcup_{U \in N_X^2} A_X(U)$. For all $U \in N_X^2$

choose any τL-element subset $B_X(U)$ of U'. Let $B_X := \bigcup_{U\in N_X^2} B_X(U)$. Thus A_X and B_X are disjoint. (This follows from the fact that G_0 is H-free and thus R is triangle-free, but here this fact is not necessary since we could simply choose each $B_X(U)$ in $U' \setminus A_X$.) Similarly, for each cluster $V \in N_Y^2$ we choose a neighbour $W(V)$ and define V' as well as τL-element sets $A_Y(V) \subseteq W(V)$ and $B_Y(V) \subseteq V'$ such that all the sets $A_Y(V)$ and $B_Y(V)$ are disjoint from $A_X \cup B_X$. Set $A_Y := \bigcup_{V\in N_Y^2} A_Y(V)$, $B_Y := \bigcup_{V\in N_Y^2} B_Y(V)$ and let

$$I := A_X \cup B_X \cup A_Y \cup B_Y.$$

Note that I meets each cluster in at most $4\tau L$ vertices. For every cluster $U \in N_X^2$, choose a ξL-element set $C_X(U) \subseteq U' \setminus I \subseteq U$ which contains at least $d|C_X(U)|/4$ neighbours of each vertex $x \in S_X$ that has at least $dL/2$ neighbours in U. (Indeed, to see that such a set $C_X(U)$ exists, observe that each vertex x with at least $dL/2$ neighbours in U has at least $dL/3$ neighbours in $U' \setminus I$ and apply Proposition 10.10 with $T := U' \setminus I$, $q = \xi L$ and $\mathcal{A} := \{N_{G'}(x) \cap (U' \setminus I) \,|\, x \in S_X\}$.) Thus condition (i) and the fact that $S_X \subseteq X'$ imply the following.

(iv) For each vertex $x \in S_X$ there are at least $(1-10^{-5})|N_X^2|$ clusters $U \in N_X^2$ such that x has at least $d|C_X(U)|/4$ neighbours in the set $C_X(U)$.

Set $C_X := \bigcup_{U\in N_X^2} C_X(U)$. For all $V \in N_Y^2$ define $C_Y(V) \subseteq V' \setminus I$ similarly and set $C_Y := \bigcup_{V\in N_Y^2} C_Y(V)$. Put

$$I' := I \cup C_X \cup C_Y$$

and

$$\kappa' := (1-\varepsilon^*)\kappa.$$

Note that $\kappa' \leq \kappa - 20\tau n$ by (10.8) and (10.10). Moreover,

$$\begin{aligned} \kappa' &\leq \frac{(1-\varepsilon^*)\kappa}{\mu} \overset{(10.9)}{\leq} \frac{(1-(\varepsilon^*)^2)cn}{\mu} \leq \frac{ckL + \varepsilon n - (\varepsilon^*)^2 cn}{\mu} \\ &\overset{(10.8)}{\leq} \frac{(c-2d-30\tau)kL}{\mu} \overset{(10.11)}{\leq} (\delta - 10\tau k)L - 20\tau kL \\ &= |N_X^2|L - 20\tau kL. \end{aligned} \tag{10.15}$$

Let $J(G)$ be the set of all those vertices in G which lie in a cluster belonging to J (which was defined in (10.14)). Since

$$\left| J(G) \cup I' \cup (\bigcup_{U\in N_X^2} U \setminus U') \cup (\bigcup_{V\in N_X^2} V \setminus V') \right| \leq 20\tau kL$$

and G_0 is κ-connected, Menger's theorem implies that we can choose a set $\mathcal{P}$ of κ' disjoint paths in the graph $G_0 \setminus (J(G) \cup I')$ such that each of these paths joins a vertex in $\bigcup_{U\in N_X^2} U'$ to a vertex in $\bigcup_{V\in N_Y^2} V'$ but has no other vertex in a cluster belonging to $N_X^2 \cup N_Y^2$.

Next we will choose a small set $\mathcal{P}^\diamond \subseteq \mathcal{P}$ which will be set aside to connect pairs $x \in S_X$, $y \in S_Y$ of branch vertices (with $xy \notin G_0$) for which we fail to find a path xPy with $P \in \mathcal{P}$. Each such pair x, y will be connected by a path of the form $xu_1 \dots u_4Pv_4 \dots v_1y$ with $P \in \mathcal{P}^\diamond$. For all pairs of clusters $U \in N_X^2$, $V \in N_Y^2$, the paths in $\mathcal{P}^\diamond$ will have the property that a significant proportion of paths in $\mathcal{P}$ joins U to V whenever a significant proportion of paths in $\mathcal{P}^\diamond$ joins U to V (see (10.17)). Roughly speaking, this property will enable us to deduce that every reasonably large set $\mathcal{P}_x \subseteq \mathcal{P}^\diamond$ of paths will have the property that the endvertices of these paths are distributed over a large number of clusters in N_Y^2. This in turn will enable us to find the path $v_1 \dots v_4$ joining y to some $P \in \mathcal{P}_x$. (The paths $\mathcal{P}_x$ will be defined in such a way that we can join their endvertices in the clusters belonging to N_X^2 to x via a suitable path $u_1 \dots u_4$.) For each cluster $U \in N_X^2$, let $\mathcal{P}(U)$ denote the set of all the paths in $\mathcal{P}$ that start in U (and thus in $U' \setminus I'$). Put

$$\eta := \frac{1}{80N(\varepsilon)}.$$

Let $\mathcal{P}'(U)$ denote the set of all those paths in $\mathcal{P}(U)$ which end in a cluster $V \in N_Y^2$ that meets (and thus contains the endvertices of) at least ηL paths in $\mathcal{P}(U)$. Note that

$$|\mathcal{P}(U) \setminus \mathcal{P}'(U)| \leq \eta L |N_Y^2| \leq \eta L k \leq L/80. \tag{10.16}$$

Thus Proposition 10.10 implies that for all $U \in N_X^2$ with $|\mathcal{P}(U)| \geq L/40$ we can choose a set $\mathcal{P}^\diamond(U)$ consisting of $\xi|\mathcal{P}'(U)|$ paths in $\mathcal{P}'(U)$ such that each cluster $V \in N_Y^2$ satisfies

$$\frac{\text{no. of paths in } \mathcal{P}'(U) \text{ ending in } V}{|\mathcal{P}'(U)|} \geq \frac{\text{no. of paths in } \mathcal{P}^\diamond(U) \text{ ending in } V}{2\xi|\mathcal{P}'(U)|}. \tag{10.17}$$

(To see this, apply Proposition 10.10 with parameters $\varepsilon := \eta$, $T := \mathcal{P}'(U)$, $q := \xi|\mathcal{P}'(U)|$ and let the set $\mathcal{A}$ consist of all the sets $\{P \in \mathcal{P}'(U) \,|\, P \text{ ends in } V\}$ for all those $V \in N_Y^2$ which meet more than ηL paths in $\mathcal{P}(U)$.) If $|\mathcal{P}(U)| < L/40$, set $\mathcal{P}^\diamond(U) := \emptyset$. Let $D_X(U) \subseteq U$ be the subset of all endvertices of paths in $\mathcal{P}^\diamond(U)$. Set $D_X := \bigcup_{U \in N_X^2} D_X(U)$, $\mathcal{P}^\diamond := \bigcup_{U \in N_X^2} \mathcal{P}^\diamond(U)$ and $\mathcal{P}^* := \mathcal{P} \setminus \mathcal{P}^\diamond$. Thus

$$|\mathcal{P}^*| \geq \kappa' - \xi n \overset{(10.8),(10.10)}{\geq} (1 - 2\varepsilon^*)\kappa + \tau n. \tag{10.18}$$

Since $\mu > 9/32$ we have

$$|\mathcal{P}| = (1 - \varepsilon^*)\kappa \overset{(10.10)}{\geq} \frac{ckL}{5} \overset{(10.11)}{\geq} \frac{\delta L}{20} \geq \frac{|N_X^2|L}{20}.$$

So on average at least $1/20$ of the vertices in a cluster $U \in N_X^2$ are endvertices of paths in $\mathcal{P}$. Hence the proportion of clusters $U \in N_X^2$ which satisfy $|\mathcal{P}(U)| \geq L/40$ is at least $1/40$, i.e. for at least $|N_X^2|/40$ clusters $U \in N_X^2$ the set $\mathcal{P}^\diamond(U)$ is non-empty and thus has size $\xi|\mathcal{P}'(U)|$. Together with (iv) this implies the following.

(v) For each vertex $x \in S_X$ there is a set $\mathcal{U}_x$ of at least $|N_X^2|/50$ clusters $U \in N_X^2$ such that for each $U \in \mathcal{U}_x$ the vertex x has at least $d|C_X(U)|/4$ neighbours in the set $C_X(U)$ and $|D_X(U)| = |\mathcal{P}^\diamond(U)| = \xi|\mathcal{P}'(U)| \geq \xi L/80$.

(The last inequality follows from (10.16).)

We will now choose the subdivided edges for all pairs x, y of branch vertices of the form $x \in S_X$, $y \in S_Y$. Clearly, we only have to consider pairs for which $xy \notin G_0$. For each such pair x, y in turn we first try to select a path $P \in \mathcal{P}^*$ whose first vertex is adjacent to x, whose last vertex is adjacent to y and such that P was not selected for a previous pair of branch vertices (if such a path P exists). We call a vertex $x \in S_X$ *useless* if after we have considered all such pairs of branch vertices there are still at least αs_Y vertices in S_Y which are not yet joined to x (neither by an edge $xy \in G_0$ nor by a path of the form xPy where $P \in \mathcal{P}^*$). The following claim implies that we were able to join most of these pairs of branch vertices in this way.

Claim B. *At most αs_X vertices in S_X are useless.*

Suppose not and let S'_X be an αs_X-element subset of S_X consisting of useless vertices. Let $\mathcal{P}'$ be the set of all those paths in $\mathcal{P}^*$ which we have not used to connect pairs x, y of branch vertices. Let Z' be the set of all those endvertices of paths in $\mathcal{P}'$ that lie in some cluster belonging to N_X^2. Recall that $d(S_X, S_Y)_{G'} \geq \mu - 10^6\varepsilon$. Together with inequalities (10.8), (10.12) and (10.18) this implies that $|Z'| = |\mathcal{P}'| \geq \tau n > 2\beta n$. But since S_X was $(\alpha s_X, \beta)$-dense for each cluster belonging to N_X^2, it follows that more than half of the vertices in Z' have at least $\beta|S'_X|$ neighbours in S'_X. Thus there is a vertex $x \in S'_X$ with more than $\beta|Z'|/2 > \beta^2 n$ neighbours in Z'. Let $\mathcal{P}''$ be the set of all those paths in $\mathcal{P}'$ that start in a neighbour of x in Z'. Thus $|\mathcal{P}''| > \beta^2 n$. Hence there must be a cluster $V \in N_Y^2$ which contains endvertices of more than $\beta^2 L$ paths in $\mathcal{P}''$. But since S_Y was $(\alpha s_Y, \beta^2)$-attached to each cluster belonging to N_Y^2 and thus also to V, there must be a path $P \in \mathcal{P}''$ whose endvertex in V is adjacent to one of the at least αs_Y vertices in S_Y that are not yet joined to x. Let $y \in S_Y$ be such a vertex. Then when considering the pair x, y we could have chosen P in order to connect it, a contradiction. This proves the claim.

Thus we are left with at most $2\alpha s_X s_Y \leq \alpha s^2$ pairs $x \in S_X$, $y \in S_Y$ of branch vertices for which we have not yet found a subdivided edge. As indicated before, for each such left-over pair x, y in turn, we will now select a subdivided edge P_{xy} which is of the form $xu_1 \dots u_4 P v_4 \dots v_1 y$ where P is some path in $\mathcal{P}^\diamond$. If U denotes the cluster in N_X^2 which contains an endvertex of P, then u_1 will be a neighbour of x in $C_X(U)$, both u_2 and u_4 will lie in $A_X(U)$ and u_3 will lie in $B_X(U)$. The path $v_1 \dots v_4$ will satisfy analogous properties.

We have to prove that for each left-over pair x, y in turn we can find such a path P_{xy} so that all these paths are internally disjoint. So suppose that we are about to consider the left-over pair x, y. Note that

$$\sum_{U \in \mathcal{U}_x} |C_X(U)| = \xi|\mathcal{U}_x|L \overset{\text{(v)}}{\geq} \frac{\xi|N_X^2|L}{50} \overset{(10.8)}{>} 2 \cdot 10^4 \alpha s^2,$$

where $\mathcal{U}_x$ was defined in (v). Thus, for at least half of the clusters $U \in \mathcal{U}_x$ at most $|C_X(U)|/10^4$ of the vertices in $C_X(U)$ have been used to join up previous (left-over) pairs. (Recall that each $C_X(U)$ is disjoint from all the paths in $\mathcal{P} \supseteq \mathcal{P}^\diamond$.) Let $\mathcal{U}'_x$ denote the set of all these clusters. So $|\mathcal{U}'_x| \geq |N_X^2|/100$. Consider the set $\mathcal{P}_x$ of all those paths in $\mathcal{P}^\diamond$ which we have not used for previous left-over pairs and whose first point lies in some set $D_X(U)$ with $U \in \mathcal{U}'_x$, i.e. $\mathcal{P}_x$ is obtained from $\bigcup_{U\in\mathcal{U}'_x} \mathcal{P}^\diamond(U)$ by deleting all the paths which we used before. Note that for each $U \in \mathcal{U}'_x$ the number of vertices in $D_X(U)$ which we used to join previous left-over pairs is precisely the number of vertices in $C_X(U)$ which we used to join previous left-over pairs. Thus for each $U \in \mathcal{U}'_x$ this number is at most $|C_X(U)|/10^4$. Hence

$$|\mathcal{P}_x| \geq \sum_{U\in\mathcal{U}'_x} (|\mathcal{P}^\diamond(U)| - |C_X(U)|/10^4) \overset{\text{(v)}}{\geq} |\mathcal{U}'_x| \left(\frac{\xi L}{80} - \frac{\xi L}{10^4}\right) \geq \frac{|N_X^2|\xi L}{10^4}. \quad (10.19)$$

We will now show that there exists a cluster $V \in N_Y^2$ which contains an endvertex of some path in $\mathcal{P}_x$ and for which the set $C_Y(V) \subseteq V$ contains a neighbour of y which is still unused. This neighbour will play the role of v_1. Let $\mathcal{V}_x$ denote the subset of all those clusters in N_Y^2 that contain an endvertex of some path in $\mathcal{P}_x$. Then (10.17) and (10.19) together imply that our original set of paths $\mathcal{P}$ contains at least $|\mathcal{P}_x|/2\xi \geq |N_X^2|L/(2\cdot 10^4)$ paths which end in a cluster belonging to $\mathcal{V}_x$ (and start in a cluster belonging to $\mathcal{U}'_x$). Thus $|\mathcal{V}_x| \geq |N_X^2|/(2\cdot 10^4) = |N_Y^2|/(2\cdot 10^4)$. But now the analogue of condition (iv) for vertices in S_Y shows that at least $|N_Y^2|(1/(2\cdot 10^4) - 1/10^5) \geq |N_Y^2|/10^5$ clusters $V \in N_Y^2$ contain an endvertex of some path in $\mathcal{P}_x$ and are such that y has at least $d|C_Y(V)|/4 = \xi dL/4$ neighbours in $C_Y(V)$. But since

$$\frac{|N_Y^2|}{10^5}\cdot\frac{\xi dL}{4} \overset{(10.8),(10.11)}{\geq} \frac{\xi cdkL}{10^6\mu} \geq \frac{\xi dcn}{10^7} \overset{(10.8)}{>} \alpha s^2,$$

there must be one such cluster V for which at least one of the neighbours of y in $C_Y(V)$ has not been used to connect previous left-over pairs. Let v_1 be such an unused neighbour, let P be any path in $\mathcal{P}_x$ that ends in V and let $v_5 \in V$ denote the endvertex of P. It remains to connect v_1 to v_5 via $A_Y(V)$ and $B_Y(V)$. Note that at most $2|C_Y(V)| = 2\xi L \leq \tau dL/8 = d|A_Y(V)|/8$ vertices in $A_Y(V)$ have been used for previous left-over pairs. Thus, since $v_1, v_5 \in V'$ and hence they have at least $d|A_Y(V)|/4$ neighbours in $A_Y(V)$, both v_1 and v_5 have at least $d|A_Y(V)|/8 > \varepsilon L$ unused neighbours in $A_Y(V)$. Since also a large proportion of the vertices in $B_Y(V)$ is still unused, we can use the fact that the graph $(V, W(V))_{G'} \supseteq (B_Y(V), A_Y(V))_{G'}$ is ε-regular of density at least d to find a neighbour v_2 of v_1 in $A_Y(V)$, a neighbour v_4 of v_5 in $A_Y(V)$ and a vertex $v_3 \in B_Y(V)$ adjacent to both v_2 and v_4 such that all these 3 vertices are still unused. Thus we have found a path $yv_1 \dots v_5$ connecting y to the endvertex of P in V. Similarly we can find a path $xu_1 \dots u_5$ connecting x to the other endvertex u_5 of P. This shows that we may join all the left-over pairs x, y of branch vertices by a path of the form $u_1 \dots u_4 P v_4 \dots v_1$.

Having joined all the pairs x, y of branch vertices with $x \in S_X$ and $y \in S_Y$ we now have to join the all the branch vertices in S_X to each other and also all the branch vertices in S_Y. We can do this in a similar way as in Case 1. Indeed, inequality (10.13) shows that the clusters in N_X^2 contain at least $\binom{s_X}{2} + \tau n$ vertices which we have not used before to connect a pair x, y of branch vertices with $x \in S_X$ and $y \in S_Y$. Thus exactly as in Case 1 one can show that all but at most $2\alpha s_X^2$ pairs $x_1, x_2 \in S_X$ can be joined by a path of length two whose midpoint lies in a cluster in N_X^2. Again, to join the remaining pairs we use the clusters in N_X^1 and in $\widetilde{N}_X^1$. The pairs $y_1, y_2 \in S_Y$ are then dealt with in a similar way. $\square$

Proposition 10.17 *For all* $9/32 \leq \mu \leq 1$, *there exists* γ *with* $1/2 \leq \gamma \leq 9/10$ *which satisfies inequalities* (10.12) *and* (10.13).

As one might expect, the only case for which (10.12) and (10.13) are sharp (if we ignore the error terms) is when the maximum density μ of the edges in the reduced graph is 3/4. This would be the case for the random graph considered in the proof of Proposition 10.12.

Proof of Proposition 10.17. Note that (10.15) implies that

$$|N_X^2|L - \tau n \geq \frac{\kappa}{\mu}(1 - \varepsilon^*). \tag{10.20}$$

We will now distinguish two cases.

Case 1. $\mu \geq 7/16$.

In this case we simply set $\gamma := 1/2$. Then (10.12) holds since

$$\begin{aligned}(1 - \mu + 10^6\varepsilon)\gamma s(1 - \gamma)s &\leq \left(\frac{9}{16} + 10^6\varepsilon\right)(1 - \lambda)\frac{16\kappa}{9} \\ &\leq \kappa + \frac{10^6 \cdot 16\varepsilon\kappa}{9} - \lambda\kappa \overset{(10.7),(10.8)}{\leq} (1 - 2\varepsilon^*)\kappa.\end{aligned}$$

Let us now show that (10.13) holds as well. If we multiply the left hand side of (10.13) with μ we obtain

$$\frac{s(s-2)\mu}{8} + \mu(1-\mu+10^6\varepsilon)\frac{s^2}{4} + 10\mu\alpha s^2 \leq \frac{8}{9}(1-\lambda)\kappa(\mu(3+2\cdot 10^6\varepsilon+80\alpha) - 2\mu^2).$$

Together with (10.20) this implies that in order to show that (10.13) holds, it suffices to check that $\mu a - 2\mu^2 \leq b$ where $a := 3 + 2 \cdot 10^6\varepsilon + 80\alpha$ and $b := \frac{9(1-\varepsilon^*)}{8(1-\lambda)}$. But $\mu a - 2\mu^2$ is maximized if $\mu = a/4$ and thus $\mu a - 2\mu^2 \leq b$ always holds since (10.7) and (10.8) imply that $a^2/8 < b$.

Case 2. $9/32 < \mu < 7/16$.

In this case we put

$$\gamma := \frac{3}{\sqrt{32}}\sqrt{\frac{1}{\mu} - 1}.$$

Since $9/32 < \mu < 7/16$ it follows that $6/10 < \gamma < 9/10$. We will first prove that γ satisfies the following 'pure versions' of inequalities (10.12) and (10.13):

$$(1 - \mu)(\gamma - \gamma^2)\frac{64\kappa}{9} \leq \kappa, \tag{10.21}$$

$$\frac{32\gamma^2\kappa}{9} + \kappa \leq \frac{\kappa}{\mu}. \tag{10.22}$$

Note that (10.22) is equivalent to

$$\mu \leq \frac{1}{\frac{32}{9}\gamma^2 + 1}. \tag{10.23}$$

But our definition of γ implies that (10.23) holds with equality. Therefore, to show that γ also satisfies (10.21), we may substitute (10.23) as an equality in (10.21) and thus it suffices to check that

$$\frac{\frac{32}{9}\gamma^2}{\frac{32}{9}\gamma^2 + 1}(\gamma - \gamma^2)\frac{64}{9} \leq 1,$$

i.e.

$$f(\gamma) := \gamma^4 - \gamma^3 + \frac{9}{64}\gamma^2 + \frac{3^4}{2^{11}} \geq 0.$$

To check this, we consider the roots of the derivative of $f(\gamma)$. But the only root of $f'(\gamma) = 4\gamma^3 - 3\gamma^2 + 9\gamma/32$ between $1/2$ and 1 is $3/8 + \sqrt{9/128} =: \gamma_0$. Since $f(\gamma_0) > 0$, $f(1/2) > 0$ and $f(1) > 0$, this shows that our γ satisfies (10.21).

It remains to show that γ also satisfies (10.12) and (10.13). But if we add $2\varepsilon^*\kappa$ to the left hand side of (10.12) we obtain

$$\begin{aligned}
&(1 - \mu + 10^6\varepsilon)(\gamma - \gamma^2)(1 - \lambda)\frac{64\kappa}{9} + 2\varepsilon^*\kappa \\
&\leq (1 - \mu)(\gamma - \gamma^2)\frac{64\kappa}{9} - \lambda(1 - \mu)(\gamma - \gamma^2)\frac{64\kappa}{9} + 10^6\varepsilon(\gamma - \gamma^2)\frac{64\kappa}{9} + 2\varepsilon^*\kappa.
\end{aligned}$$

Since γ satisfies (10.21), the first summand is at most κ. Moreover, $1-\mu \geq 9/16$ and $\gamma - \gamma^2 \geq 0.9 - 0.9^2$. Together with (10.7) and (10.8) this shows that the remaining sum is less than 0. Thus (10.12) holds. This implies that the left hand side of (10.13) is at most

$$\begin{aligned}
(1 - \lambda)\frac{32\gamma^2\kappa}{9} + \kappa - 2\varepsilon^*\kappa + \frac{640\alpha\kappa}{9} &\overset{(10.8),(10.22)}{\leq} \frac{\kappa}{\mu}\left(1 - \frac{32\lambda\gamma^2\mu}{9}\right) \\
&\overset{(10.7)}{\leq} \frac{\kappa}{\mu}(1 - \varepsilon^*) \overset{(10.20)}{\leq} |N_X^2|L - \tau n,
\end{aligned}$$

as desired. □

As the proof of Theorem 10.2 is similar to that of Theorem 10.15, we only sketch the argument.

Proof of Theorem 10.2 (Sketch). By Proposition 10.12, it suffices to prove the upper bound. Thus, given $0 < \lambda < 1$, we have to show that there exists $\kappa_0 = \kappa_0(\lambda)$ such that for every natural number $\kappa \geq \kappa_0$ each κ-connected graph G_0 contains a subdivision of a clique of order at least $2\sqrt{(1-\lambda)\kappa} =: s$. We start exactly as in the proof of Theorem 10.15. Since we are now only seeking a subdivision of a smaller clique, the calculation in Claim A shows that we can proceed as in Case 1 as long as $\mu \leq 1/2$. Thus we may assume that $\mu > 1/2$.

If the common neighbourhood of X and Y in R has size at most τk, we can discard it and proceed precisely as in the proof of Theorem 10.15 (Case 2). Otherwise we choose a τk-element set N^1_{XY} of common neighbours of X and Y and a τk-element set $\widetilde{N}^1_{XY}$ such that these sets are disjoint from each other and from $N^1_X, \widetilde{N}^1_X, N^2_X, N^1_Y, \widetilde{N}^1_Y, N^2_Y$ and such that R contains a perfect matching between N^1_{XY} and $\widetilde{N}^1_{XY}$. We set $\gamma := 1/2$ and choose the set $S_X \cup S_Y$ of branch vertices as in Case 2 of the proof of Theorem 10.15. (Note that when $\mu \geq 1/2$ the proof of Proposition 10.17 immediately shows that $\gamma = 1/2$ also works in the proof of Theorem 10.15.) The argument implies that we may additionally assume that each branch vertex has at least $d|U|/2$ neighbours in at least $2/3$ of the clusters $U \in N^1_{XY}$. Moreover, we may clearly assume that $T := N^2_X \cap N^2_Y$ is non-empty.

Next suppose that $|T|L \leq (1-\mu+10^6\varepsilon)s^2/4$. Thus the number of vertices lying in a cluster belonging to T is not larger than the required number of subdivided edges joining branch vertices in S_X to branch vertices in S_Y. We now join almost $|T|L$ pairs $x \in S_X$, $y \in S_Y$ of branch vertices (with $xy \notin G_0$) by a path of length two whose midpoint lies in a cluster belonging to T. (The existence of these paths follows similarly as in the proof of Claim A.) The set $\mathcal{P}$ of paths will now have size only $(1-\varepsilon^*)\kappa - |T|L$ and the paths in $\mathcal{P}$ will avoid all vertices lying in clusters belonging to T. As before (see Claim B), we can join most of the remaining pairs $x \in S_X$, $y \in S_Y$ of branch vertices by a path of the form xPy with $P \in \mathcal{P}$. As in the final part of the proof of Case 1, the sets N^1_{XY} and $\widetilde{N}^1_{XY}$ can then be used to join the small proportion of left-over pairs $x \in S_X$, $y \in S_Y$ by paths of length four. Since in total we have not used more vertices in N^2_X to join up the pairs $x \in S_X$, $y \in S_Y$ than in the proof of Case 2 in Theorem 10.15, all the pairs $x, x' \in S_X$ can be joined as before (and the same is true for the pairs $y, y' \in S_Y$).

Finally, suppose that $|T|L > (1-\mu+10^6\varepsilon)s^2/4$. In this case we again distribute the branch vertices evenly and proceed similarly as in the previous case except that this time we can find almost all of the subdivided edges joining pairs $x \in S_X$, $y \in S_Y$ (with $xy \notin G_0$) as paths of length two whose midpoint lies in a cluster belonging to T. Thus we do not have to use the connectivity of G_0 at all. Moreover, this time the number of all those vertices in clusters belonging to N^2_X which we have not used up so far is at least

$$\begin{aligned}
|N^2_X|L - (1-\mu+10^6\varepsilon)\frac{s^2}{4} &\overset{(10.15)}{\geq} (1-\varepsilon^*)\frac{\kappa}{\mu} + 2\tau n - (1-\mu+10^6\varepsilon)\frac{s^2}{4} \\
&\geq \frac{s^2}{4}\left(\frac{1}{\mu} - 1 + \mu\right) + 2\tau n \geq 2\binom{s/2}{2} + 2\tau n.
\end{aligned}$$

Thus there is still enough room to join up the pairs of the form $x, x' \in S_X$ and $y, y' \in S_Y$ as in the previous case. □

Roughly speaking, our aim in the proof of Theorem 10.3 is to find an edge XY in the reduced graph R whose density is large and which has the property that R contains many disjoint paths joining the neighbourhood of X to the neighbourhood of Y. Once we have found such an edge, we can proceed as in the proof of Theorem 10.2 since by Lemma 10.7 these paths correspond to many disjoint paths in the graph G_0 we started with. (Thus as before, the branch vertices are distributed within X and Y.) The following result of Mader [78] (see also [12]) implies that to find such an edge, it suffices to find a subgraph of R which has high minimum degree and in which every edge has large density.

Theorem 10.18 *In every graph G there exists an edge xy such that G contains $\delta(G)$ internally disjoint paths between x and y.*

Proof of Theorem 10.3 (Sketch). Let $\varrho := 1.15$ and $\sigma := 9/10$. Again, by Proposition 10.12, it suffices to show that for each $0 < \lambda < 1$ there exists $d^* = d^*(\lambda)$ such that for every $d_0 \geq d^*$ each graph G_0 of average degree d_0 contains a subdivision of a clique of order at least $\sqrt{2\varrho(1-\lambda)d_0} =: s$. We start by choosing constants as in (10.7) and (10.8) in the proof of Theorem 10.15. Similarly as there, we may assume that G_0 contains a subgraph G whose average degree is cn for some constant $c \geq c_0(\varepsilon^*)$ and such that $d_0/(1+\varepsilon^*) \leq cn \leq d_0$. (As before, n denotes the order of G.) By replacing G with a subgraph if necessary, we may assume that G contains no subgraph whose average degree is larger than cn and thus $\delta(G) \geq cn/2$. Next we apply the Regularity lemma to G. Proposition 10.4 implies that we obtain a reduced graph R which satisfies

$$\delta(R) \geq \left(\frac{c}{2} - 2d\right) k.$$

Put $c' := c - 2d$. Since we are now looking for a subdivision of a smaller clique, the calculation in Claim A in the proof of Theorem 10.15 shows that we can proceed as in Case 1 as long as $\Delta(R) \geq \varrho c' k$. (Indeed, take for X any vertex of maximum degree in R.) Thus we may assume that $\Delta(R) \leq \varrho c' k$.

Given a subgraph R' of R and a vertex $X \in V(R')$, we call

$$w_{R'}(X) := \sum_{Y \in N_{R'}(X)} \frac{e_{G'}(X,Y)}{L^2}$$

the *weight of X in R'*. Note that $d_{R'}(X) \geq w_{R'}(X)$. Moreover,

$$\sum_{X \in V(R)} w_R(X) L^2 = 2e(G' - V_0) \geq (c - (d+\varepsilon))n^2 - \varepsilon n^2 \geq c'(kL)^2.$$

Thus

$$\frac{\sum_{X \in V(R)} w_R(X)}{k} \geq c'k, \tag{10.24}$$

i.e. the average weight of the vertices in R is at least $c'k$. Let A be the set of all those vertices in R whose weight is less than $\sigma c' k$. Put $B := V(R) \setminus A$ and $b := |B|$. Let w_B be such that the average weight (in R) of the vertices in B is $w_B c' k$. Then (10.24) implies that $(k-b)\sigma c' k + b w_B c' k \geq c' k^2$ and hence

$$w_B \geq \frac{k}{b}(1-\sigma) + \sigma. \tag{10.25}$$

Let R_1 be the graph obtained from R by deleting all those edges which have both endvertices in A. Call an edge of R_1 *light* if its density is at most 1/2. For each $b \in B$, let ν_b be defined in such a way that $\nu_b c'k$ is the number of light edges of R_1 incident to b. Since $\Delta(R) \le \varrho c'k$, we have

$$\sum_{b \in B}(\nu_b c'k/2 + (\varrho - \nu_b)c'k) \ge \sum_{b \in B} w_{R_1}(b) = \sum_{b \in B} w_R(b) = w_B c'kb.$$

Thus, setting

$$\nu := 2(\varrho - w_B),$$

it follows that

$$\text{no. of light edges in } R_1 \le \sum_{b \in B} \nu_b c'k \le \nu c'kb. \tag{10.26}$$

Let R_2 be the graph obtained from R_1 by deleting all light edges. Then

$$\begin{aligned} d(R_2) &\overset{(10.26)}{\ge} d(R_1) - 2\nu c'b \ge \frac{b \cdot w_B c'k}{k} - 2\nu c'b = bc'(5w_B - 4\varrho) \\ &\overset{(10.25)}{\ge} c'k\left(5(1-\sigma) + \frac{b}{k}[5\sigma - 4\varrho]\right) \\ &\ge c'k(5 - 4\varrho) = 2c'k/5 =: 2\delta. \end{aligned}$$

(To see the last line, note that the square bracket is negative.) Finally, let R_3 be a subgraph of R_2 with minimum degree at least δ and set $\kappa_R := \delta - 1$. Apply Theorem 10.18 to find an edge $XY \in R_3$ such that R_3 contains a set $\mathcal{P}_R$ of κ_R disjoint paths between $N_R(X) \setminus \{Y\}$ and $N_R(Y) \setminus \{X\}$ which have no inner vertices in $N_R(X) \cup N_R(Y)$. We choose $\mathcal{P}_R$ in such a way that as many paths as possible are trivial. Since all edges in $E(R_3) \ni XY$ have at least one of their endvertices in B, we may assume that $X \in B$. Moreover, since no edge of $R_2 \supseteq R_3$ is light, the density of XY is at least 1/2.

Similarly as in the proof of Theorem 10.15 (Case 2), choose disjoint τk-element sets N_X^1, $\widetilde{N}_X^1$, N_Y^1 and $\widetilde{N}_Y^1$. If $|N_R(X) \cap N_R(Y)| \ge \kappa_R$, we also choose τk-element sets N_{XY}^1 and $\widetilde{N}_{XY}^1$ which are disjoint from each other and from the above four sets and such that R contains a perfect matching between N_{XY}^1 and $\widetilde{N}_{XY}^1$. Next choose a set N_X^2 of neighbours of X in $R - Y$ which is disjoint from the above sets and has size $(\sigma c' - 10\tau)k$. (This is possible since $X \in B$ and so $d_R(X) \ge w_R(X) \ge \sigma c'k$.) Also, choose a set N_Y^2 of $(\frac{c}{2} - 3d)k$ neighbours of Y which is disjoint from all the above sets except possibly from N_X^2. Moreover, we choose N_X^2 and N_Y^2 so that $\mathcal{P}_R$ contains at least $\kappa_R - 6\tau k$ paths which join N_X^2 to N_Y^2 and avoid each of N_X^1, $\widetilde{N}_X^1$, N_Y^1, $\widetilde{N}_Y^1$, N_{XY}^1 and $\widetilde{N}_{XY}^1$. Let $\mathcal{P}'_R \subseteq \mathcal{P}_R$ denote the set of all these paths.

We now proceed as in the proof of Theorem 10.2 except for two changes. Firstly, the set $\mathcal{P}$ of paths is now obtained by an application of Lemma 10.7 to all the paths in $\mathcal{P}'_R$. Thus $|\mathcal{P}| \ge (\kappa_R - 7\tau k)L$. Secondly, we have to check that we can distribute the branch vertices of our subdivided K_s among X and Y such that N_X^2, N_Y^2 and $\mathcal{P}$ are large enough to accommodate (almost) all the subdivided edges. For the latter, we distinguish two cases according to the size of $T := N_X^2 \cap N_Y^2$. Again, γ will denote the proportion of branch vertices which we choose in X.

Case 1. $|T| \leq \kappa_R$

In this case, we join all pairs x, y of branch vertices with $x \in X$, $y \in Y$ and $xy \notin G_0$ by paths of the form xPy with $P \in \mathcal{P}$. (Note that if $N_X^2 \cap N_Y^2 \neq \emptyset$, some or even all of these paths may be trivial.) This can be done as in the proof of Theorem 10.15 (Case 2) if the number of all these pairs x, y is a bit smaller than $|\mathcal{P}|$, i.e. if

$$\gamma(1-\gamma)s^2(1/2 + 10^6\varepsilon) \leq (\kappa_R - 20\tau k)L. \tag{10.27}$$

Almost all of the pairs x, x' of branch vertices with $x, x' \in X$ will be joined by a path of length two whose midpoint lies in a cluster belonging to $N_X^2 \setminus N_Y^2$ and was not used before to join some branch vertex in X to some branch vertex in Y. For this, we need that the number of all those unused vertices is a bit larger than the number of all the pairs x, x', i.e. that

$$\binom{\gamma s}{2} + \kappa_R L \leq (\sigma c' - 20\tau)kL. \tag{10.28}$$

The next inequality ensures that almost all pairs of branch vertices in Y can be joined in a similar way.

$$\binom{(1-\gamma)s}{2} + \kappa_R L \leq \left(\frac{c}{2} - 5d\right)kL \tag{10.29}$$

As before, the left-over pairs $x, x' \in X$ and $y, y' \in Y$ can be joined by using the sets N_X^1, $\widetilde{N}_X^1$ and N_Y^1, $\widetilde{N}_Y^1$ respectively. It is easy to check that (10.27), (10.28) and (10.29) hold if we set $\gamma := 0.78$.

Case 2. $|T| > \kappa_R$.

In this case, we join almost all of the pairs x, y of branch vertices with $x \in X$, $y \in Y$ and $xy \notin G_0$ by paths of length two whose midpoints lie in clusters belonging to T. (The left-over such pairs are then joined by paths of length four using the sets N_{XY}^1 and $\widetilde{N}_{XY}^1$ as in the proof of Theorem 10.2.) Thus, we need that the number of all these pairs x, y is at most $(|T| - \tau k)L$. Defining t by $|T| = tc'k$, this means that

$$\gamma(1-\gamma)s^2(1/2 + 10^6\varepsilon) \leq (tc' - \tau)kL. \tag{10.30}$$

Moreover, we will join almost all of the pairs x, x' of branch vertices with $x, x' \in X$ by paths of length two whose midpoints lie in a cluster belonging to $N_X^2 \setminus N_Y^2$. This will be possible if

$$\binom{\gamma s}{2} \leq (\sigma c' - tc' - 20\tau)kL. \tag{10.31}$$

Finally, we join almost all of the pairs y, y' of branch vertices with $y, y' \in Y$ by paths of length two whose midpoints lie in a cluster belonging to N_Y^2 but have not been used before to join some branch vertex in X to some branch vertex in Y. Thus we need that

$$\binom{(1-\gamma)s}{2} + \gamma(1-\gamma)s^2(1/2 + 10^6\varepsilon) \le \left(\frac{c}{2} - 5d\right) kL. \tag{10.32}$$

As before, all the left-over pairs $x, x' \in X$ and $y, y' \in Y$ of branch vertices will be joined by using the sets N_X^1, $\widetilde{N}_X^1$ and N_Y^1, $\widetilde{N}_Y^1$. It can be easily checked that inequalities (10.30), (10.31) and (10.32) hold if we put $\gamma := \sqrt{(\sigma - t)/\varrho}$. □

10.5 Concluding remarks

In this section, we briefly discuss the difficulties which arise if one tries to extend Theorem 10.15 to arbitrary graphs by removing the condition of H-freeness. The proof of Theorem 10.15 still works if the intersection of the neighbourhoods $N_R(X)$ and $N_R(Y)$ of X and Y in R is non-empty but not too large (here XY is an edge in R of maximum density). Indeed, as in the proof of Theorem 10.2 we can use this intersection to join a corresponding number of pairs $x \in X$, $y \in Y$ of branch vertices (with $xy \notin G_0$) by paths of length two whose midpoint belongs to a cluster in $N_R(X) \cap N_R(Y)$. The connectivity of G_0 is then only used to join the remaining such pairs.

However, the argument breaks down if $N_R(X) \cup N_R(Y)$ is too small, i.e. if the number of vertices belonging to a cluster in $N_R(X) \cup N_R(Y)$ is smaller than the required number of subdivided edges. In this case one is forced to distribute the branch vertices over more than two clusters. In fact, the following example shows that up to 9 clusters may be necessary in some cases. Suppose that G has a reduced graph R which consists of a large complete graph and whose edges all have density about 9/16. This will be the case (with high probability) if each subgraph of G corresponding to an edge of R is a bipartite random graph with edge probability 9/16 and G is empty otherwise. The connectivity of this graph is about $9n/16$ where $n := |G|$. Set $s := \frac{8}{3}\sqrt{\frac{9n}{16}}$. Then, if we distribute the branch vertices of a potential subdivision of K_s over t clusters, the number of subdivided edges one needs to find is at least about

$$t\binom{s/t}{2} + \frac{7}{16}\left(\frac{s}{t}\right)^2\binom{t}{2}$$

which is significantly larger than n unless $t \ge 9$. In this example, it is of course nevertheless easy to find a subdivision of K_s in G since the intersections of the the neighbourhoods of the clusters in R are identical (and so one can proceed as in the final case of the proof of Theorem 10.2). However, the example indicates that for arbitrary graphs a strategy similar to ours seems to lead to an enormous number of cases which need to be considered, as the case distinctions would not only depend on the sizes of the pairwise intersections but more generally on the sizes of the common neighbourhoods of each subset of the set of all those clusters which contain the branch vertices.

Chapter 11

Packings in dense regular graphs

11.1 Introduction, results and open problems

Given two graphs H and G, an *H-packing in G* is a collection of vertex-disjoint copies of H in G. It is *perfect* if all of the vertices of G are covered. Improving earlier bounds of Alon and Yuster [8], Komlós, Sárközy and Szemerédi [55] proved that given a graph H of chromatic number χ, there exists a constant c such that every sufficiently large graph G whose order n is divisible by $|H|$ and whose minimum degree is at least $(1 - 1/\chi)n + c$ has a perfect H-packing. This bound is best possible up to the value of c. However, Komlós [49] (see Theorem 12.10) showed that if we only want an H-packing covering almost all of the vertices of G, then in many cases a significantly smaller minimum degree suffices: given $\alpha > 0$ and a graph H of chromatic number χ, every sufficiently large graph G of minimum degree at least $(1 - 1/\chi')n$ has an H-packing covering all but at most αn vertices. Here χ' is the *critical chromatic number* of H, which is defined as $(\chi - 1)|H|/(|H| - \sigma)$, where σ denotes the minimum size of the smallest colour class in an optimal colouring of H. It is easy to see that $\chi - 1 < \chi' \leq \chi$ and that χ' is closer to $\chi - 1$ if there is an optimal colouring where one of the colours is used comparatively rarely. Again, the bound on the minimum degree is best possible (but Komlós conjectured that the error term αn can be reduced to a constant depending only on H).

Here we show that if we restrict our attention to packings of bipartite graphs H in (almost) regular graphs G then any linear bound on the minimum degree suffices. Given $a \geq b$, we say that a graph G is *$(a \pm b)$-regular* if its minimum degree is at least $a - b$ and its maximum degree is at most $a + b$.

Theorem 11.1 *Given a bipartite graph H and constants $0 < c, \alpha \leq 1$, there exist positive numbers $\gamma = \gamma(c, \alpha)$ and $n_0 = n_0(H, c, \alpha)$ such that every $(cn \pm \gamma n)$-regular graph G of order $n \geq n_0$ has an H-packing which covers all but at most αn vertices of G.*

The complete bipartite graph $G = K_{a,2a}$ (and $H := C_4$ say) shows that the restriction to almost regular graphs G is necessary. Also the restriction to bipartite graphs H is obviously necessary if $c \leq 1/2$. Our next result shows that if H has unequal vertex classes, then we can obtain an H-packing covering all but a constant number of vertices.

Theorem 11.2 *Given a bipartite graph H whose vertex classes have different size and a constant $0 < c \leq 1$, there exist $\gamma = \gamma(H, c) > 0$ and $C = C(H, c)$ such that every $(cn \pm \gamma n)$-regular graph G has an H-packing which covers all but at most C vertices of G.*

As the complete bipartite graph $G = K_{k,\ell}$ (with $k - \ell = \gamma n$ say) shows, we cannot hope for such a result if H has equal vertex classes. However, it could help to restrict one's attention to regular graphs G:

Question A *Is it true that for every $c > 0$ and every bipartite graph H there is a constant $C = C(c, H)$ such that every cn-regular graph G has an H-packing which covers all but at most C vertices of G?*

It is easy to see that in both Question A and Theorem 11.2 (even if we only consider regular graphs there too) the bound on the number of uncovered vertices must depend on H and c (consider the disjoint union G_r of r complete graphs whose order is $k|H| - 1$ for some positive integer k). Thus in contrast to the result of Komlós, Sárközy and Szemerédi mentioned above one cannot hope for a perfect packing when $|H|$ divides n. Moreover, the graphs G_r (with $r \to \infty$) show that in both Theorem 11.2 and Question A the requirement that the vertex degrees are linear is necessary. On the other hand, it may be true that the bound on the minimum degree in Theorem 11.1 can be improved. Since our proof of this theorem relies on Szemerédi's Regularity lemma, it seems that this would require a rather different approach from ours though. Note that if H contains a cycle, the bound on the minimum degree of G must of course depend on n. However, in the special case when H is a tree, it can be chosen to be independent of n: Kelmans, Mubayi and Sudakov [44] proved that for every $\alpha > 0$ and every tree T there exists a constant $d_0 = d_0(\alpha, T)$ such that for all $d \geq d_0$ every d-regular graph G has a T-packing which covers all but at most αn vertices of G. Triangle packings in sparse pseudo-random regular graphs were investigated by Krivelevich, Sudakov and Szabó [61].

Theorem 11.2 is related to a problem of Verstraëte [3, Conj. 3.4] (see also [106]) on packings of subdivisions in regular graphs. Given graphs H and G, a *TH-packing* in G is a collection of vertex-disjoint subdivisions of H in G (which are not required to be isomorphic).

Conjecture B (Verstraëte) *For every graph H and every positive ε, there exists an integer $r_0 = r_0(H, \varepsilon)$ such that for all $r \geq r_0$ every r-regular graph G contains a TH-packing which covers all but at most $\varepsilon|G|$ vertices of G.*

By the result in [44] Conjecture B holds for trees. Alon [3] proved it for cycles. Our results provide further support for this conjecture. In particular, Theorem 11.2 implies the following.

Corollary 11.3 *Given a graph H without isolated vertices which is not a union of cycles and a constant $0 < c \leq 1$, there exist $\gamma = \gamma(H, c) > 0$ and $C = C(H, c)$*

such that every $(cn \pm \gamma n)$-regular graph G has a TH-packing which covers all but at most C vertices of G.

Proof. Let H' be the graph obtained from H by subdividing each edge of H exactly once. Thus H' is a bipartite graph whose vertex classes have sizes $e(H)$ and $|H|$. Hence the corollary immediately follows from Theorem 11.2 if $e(H) \neq |H|$. So we may assume that $e(H) = |H|$. Since H is not a union of cycles, this implies that H (and thus also H') must contain at least one vertex of degree 1. But the graph H'' obtained from H' by deleting a vertex of degree 1 is still a subdivision of H. Since the vertex classes of H'' have different size, we are done by Theorem 11.2 again. □

Complete bipartite graphs $K_{k,\ell}$ with $k - \ell = \gamma n$ show that Corollary 11.3 is not true if H is a union of cycles. If $H = K_4$ or $H = K_5$, then we can even guarantee a perfect packing:

Theorem 11.4 *For all $0 < c \leq 1$ there exist positive numbers $\gamma = \gamma(c)$ and $n_0 = n_0(c)$ such that every $(cn \pm \gamma n)$-regular graph G of order $n \geq n_0$ has a perfect TK_r-packing for $r = 4$ and $r = 5$.*

It turns out that for $r \geq 6$ the vertex degrees in Corollary 11.3 have to be linear and that Theorem 11.4 does not extend to $r \geq 6$ (Proposition 11.10). However, as in Question A, it may help to consider only regular graphs G:

Question C *Given $r \geq 6$ and $0 < c \leq 1$, does every cn-regular graph of sufficiently large order n have a perfect TK_r-packing?*

As for Theorem 11.1, it is possible that in both Theorem 11.4 and Question C the condition on the minimum degree can be relaxed. On the other hand, at the end of Section 11.5 we give an example which shows that for all $r \geq 3$ we need a minimum degree of at least $\sqrt{n/2}$, even if G is regular.

A result of Jørgensen and Pyber [42] implies that Conjecture B holds if we do not require the subdivisions of H to be disjoint: Let $t(H)$ be so that every graph of average degree at least $t(H)$ contains a subdivision of H. Jørgensen and Pyber proved that the edges of every graph G can be covered by at most $10^4\, t(H)|G|$ subdivisions of H and edges. Hence if the average degree of G is large compared with $t(H)$, then almost all of its edges must lie in a subdivision of H (and thus, if G is regular, also almost all of its vertices). In the case when $H = K_r$, Lemma 2.3 of [42] implies that if G is 2-connected and has average degree at least $t(K_{r+1})$, then even all edges (and thus also all vertices) of G lie in a subdivided K_r. Using this, it is not hard to show that for all $\varepsilon > 0$ there exists an $r_0 = r_0(\varepsilon)$ such that for all $r \geq r_0$ every graph G of minimum degree at least $(1+\varepsilon)t(K_{r+1})$ has at most $\varepsilon|G|$ vertices and at most $\varepsilon e(G)$ edges that do not lie in a subdivision of K_r. (Indeed, consider the block tree T of G. All vertices and edges which lie in blocks of average degree $\geq t(K_{r+1})$ are contained in a subdivided K_r. For example, this is true for all blocks corresponding to leaves of T. But each block B with average degree $< t(K_{r+1}) < \delta(G)$ contains many vertices b which have at least one neighbour in another block. Since for different such b these neighbours must belong to distinct blocks, this implies that each such block B has many neighbours in T. The latter can be used to show that the blocks of average degree $< t(K_{r+1})$ contain only a small fraction of the vertices and edges of G.) In other words, to cover almost all of the vertices of G with (not necessarily disjoint) subdivisions of K_r we only need a minimum degree which is slightly larger than the average degree required for the mere existence of a subdivided K_{r+1}.

This chapter is organized as follows. In the next section we introduce some notation and the tools (bipartite Regularity lemma and Blow-up lemma) which we will need later on. In Section 11.3 we prove Theorem 11.1. In Section 11.4 we extend the argument to obtain Theorem 11.2. In the final section we then derive Theorem 11.4 from Theorem 11.1. Since our proofs of Theorems 11.1–11.4 can be derandomized by standard techniques (see e.g. [7]) and since both the Regularity lemma and the Blow-up lemma have algorithmic proofs (see [4] and [52]), it is easy to verify that the packings guaranteed by Theorems 11.1–11.4 can be found in polynomial time.

11.2 Notation and tools

We will use the following well-known bound on the number of edges in a bipartite $K_{s,t}$-free graph (it is a consequence of Lemma 4.6, see e.g. [12, Ch. VI, Thm. 2.2] for a proof).

Theorem 11.5 *For all $s \leq t$ there exists a constant $c_{s,t}$ such that every bipartite graph $G = (A, B)$ with at least $c_{s,t}|G|^{2-1/s}$ edges contains a copy of $K_{s,t}$ with s vertices in A and t vertices in B.*

Our proof of Theorem 11.1 is based on the Regularity lemma and that of Theorem 11.2 also employs the Blow-up lemma. We will use the following form of the Regularity lemma for bipartite graphs. The fact that the regularity partition can be required to refine the given bipartition (A, B) follows immediately from the proof of the Regularity lemma.

Lemma 11.6 (Regularity lemma) *For every $\varepsilon > 0$ there exists an $N = N(\varepsilon)$ such that for every number $d \in [0, 1]$ and for every bipartite graph $G = (A, B)$ with $|A| = |B|$ there exist partitions of A into $A_0, A_1, \dots, A_k$ and of B into $B_0, B_1, \dots, B_k$ and there is a spanning subgraph G' of G such that the following holds:*

- $k \leq N$,
- $|A_0 \cup B_0| \leq \varepsilon|G|$,
- $|A_1| = \cdots = |A_k| = |B_1| = \cdots = |B_k| =: L$,
- $d_{G'}(x) > d_G(x) - (d + \varepsilon)|G|$ *for all vertices* $x \in G$,
- *for all* $1 \leq i, j \leq k$ *the graph* $(A_i, B_j)_{G'}$ *is* ε*-regular and has density either* 0 *or* $> d$.

The sets A_i and B_i ($i \geq 0$) are called *clusters*. $A_0 \cup B_0$ is called the *exceptional set*. Given clusters and G' as in Lemma 11.6, the *reduced graph* R is the bipartite graph whose vertices are $A_1, \dots, A_k$ and $B_1, \dots, B_k$ (so we omit A_0 and B_0 here) and in which A_i is joined to B_j whenever $(A_i, B_j)_{G'}$ is ε-regular and has density $> d$. Thus A_iB_j is an edge of R if and only if G' has an edge between A_i and B_j.

In the proof of Theorem 11.2 we will use the following special case of the Blow-up lemma of Komlós, Sárközy and Szemerédi [51]. See [48] for a survey about this lemma.

Lemma 11.7 (Blow-up lemma) *For all $d > 0$ and all integers Δ there exists a positive number $\varepsilon_0 = \varepsilon_0(d, \Delta)$ such that for all $\varepsilon \leq \varepsilon_0$, all integers a, b and every subgraph H of $K_{a,b}$ with $\Delta(H) \leq \Delta$ each (ε, d)-super-regular bipartite graph $G = (A, B)$ with $|A| = a$ and $|B| = b$ contains H as a subgraph.*

11.3 H-packings covering all but a small fraction of vertices

The strategy of our proof of Theorem 11.1 is as follows. Given G, we will first find a spanning bipartite subgraph $G^* = (A, B)$ whose vertex classes have equal size and which is still 'almost regular' (see Proposition 11.8; we may assume that $|G|$ is even). We then apply the Regularity lemma to G^*. As G^* is 'almost regular', the reduced graph R has a matching M which misses only a small fraction of its vertices (Lemma 11.9). But as each edge $e \in M$ corresponds to an ε-regular graph G_e of sufficiently large density, we may successively apply Theorem 11.5 to pull out disjoint copies of H as long as an ε-fraction of the vertices remains uncovered in each of the vertex classes of G_e. Thus the copies of H can be chosen in such a way that we obtain an H-packing which covers a large fraction of the vertices of each G_e (and hence of G).

Proposition 11.8 *Given $\gamma, c > 0$ with $2\gamma < c \leq 1$, there exists $n_0 = n_0(\gamma, c)$ such that every $(cn \pm \gamma n)$-regular graph G of even order $n \geq n_0$ contains a spanning bipartite subgraph $G^* = (A, B)$ such that $|A| = |B|$ and $|d_{G^*}(x) - d_G(x)/2| \leq \gamma n$ for every vertex $x \in G$.*

Proof. Consider a random bipartition of $V(G)$ into sets A and $B = V(G) \setminus A$ which is obtained by including every vertex $x \in G$ into A with probability $1/2$ independently of all other vertices of G. Call a vertex $x \in G$ *bad* if $||N_G(x) \cap A| - d_G(x)/2| > \gamma n/2$. As $\mathbb{E}(|N_G(x) \cap A|) = d_G(x)/2$ and $\gamma n/2 \geq \gamma d_G(x)/2$, Lemma 3.6 implies that

$$\mathbb{P}(x \text{ is bad}) \leq 2\mathrm{e}^{-\beta(\gamma)d_G(x)/2} \leq 2\mathrm{e}^{-\beta(\gamma)cn/4}.$$

Hence if n is sufficiently large, then the expected number of bad vertices is at most $2n\mathrm{e}^{-\beta(\gamma)cn/4} < 1/2$. Thus Markov's inequality implies that

$$\mathbb{P}(\text{there is a bad vertex}) \leq 1/2.$$

Moreover, applying Lemma 3.6 again, we get

$$\mathbb{P}(|\,|A| - n/2| > \gamma n/2) \leq 2\mathrm{e}^{-\beta(\gamma)n/2} < 1/4.$$

Thus with probability at least $1/2 - 1/4 > 0$ there is an outcome A, B with $|\,|A| - n/2\,| \leq \gamma n/2$ for which no vertex of G is bad. We may assume that $|A| \leq |B|$. But then in the bipartite graph G^* obtained from $(A, B)_G$ by moving $(|B| - |A|)/2 \leq \gamma n/2$ vertices from B to A, every vertex $x \in G$ satisfies

$$|d_{G^*}(x) - d_G(x)/2| \leq \frac{\gamma n}{2} + \frac{\gamma n}{2} = \gamma n,$$

as required. $\square$

Lemma 11.9 *Given positive numbers c, d, ε and γ with $c > \gamma + d + 2\varepsilon$ and a $(cn \pm \gamma n)$-regular bipartite graph $G = (A, B)$ with $|A| = |B|$, let $A_0, A_1, \ldots, A_k$ and $B_0, B_1, \ldots, B_k$ be the clusters of a partition of $V(G)$ as given in Lemma 11.6. Let R be the corresponding reduced graph. Then R contains a matching of cardinality at least $k(1 - 2(\gamma + d + \varepsilon)/c)$.*

Proof. Given $I \subseteq \{1, \ldots, k\}$, let J be the set of all integers $j \geq 1$ such that in R the vertex B_j lies in the neighbourhood of $\{A_i \,|\, i \in I\}$. We will show that $|J| \geq |I|(1 - 2(\gamma + d + \varepsilon)/c))$. Then by the defect form of Hall's matching theorem (see e.g. [28, Cor. 2.1.3]) the bipartite graph R contains a matching of cardinality at least $k - 2k(\gamma + d + \varepsilon)/c$, as required.

Let $I' \subseteq A$ be the union of all A_i with $i \in I$ and define $J' \subseteq B$ to be the union of all B_j with $j \in J$. Let L and G' be as in Lemma 11.6. Recall that if $x \in A_i \subseteq I'$ and if y is a neighbour of x in G' lying in B_j for some $j \geq 1$ then A_iB_j is an edge of R and thus $j \in J$ and $y \in J'$. As

$$\delta(G') \geq \delta(G) - (d + \varepsilon)n \geq (c - \gamma - d - \varepsilon)n,$$

this shows that

$$e_{G'}(I', J') \geq \sum_{i \in I} |A_i|[(c - \gamma - d - \varepsilon)n - |B_0|] \geq |I|L(c - \gamma - d - 2\varepsilon)n.$$

But on the other hand,

$$e_{G'}(I', J') \leq e_{G'}(A, J') \leq \sum_{j \in J} |B_j|(c + \gamma)n = |J|L(c + \gamma)n$$

and so

$$\begin{aligned} |J| &\geq |I| \cdot \frac{1 - \frac{\gamma + d + 2\varepsilon}{c}}{1 + \frac{\gamma}{c}} \geq |I| \cdot \left(1 - \frac{\gamma + d + 2\varepsilon}{c}\right)\left(1 - \frac{\gamma}{c}\right) \\ &\geq |I| \cdot \left(1 - \frac{2(\gamma + d + \varepsilon)}{c}\right), \end{aligned}$$

as required. $\square$

Note that in the proof of Lemma 11.9 we cannot apply Hall's matching theorem directly to the reduced graph R since R is not necessarily almost regular.

Proof of Theorem 11.1. Clearly, it suffices to consider the case when $H = K_{s,t}$ where $s \le t$. By deleting one vertex of G if necessary, we may assume that the order n of G is even. We will prove the assertion for $\gamma := \alpha c/50$ and for sufficiently large n_0. Apply Proposition 11.8 to G to obtain a spanning bipartite $(cn/2 \pm 2\gamma n)$-regular subgraph $G^* = (A, B)$ with $|A| = |B|$. Set $\varepsilon := \gamma$, $d := 2\varepsilon$ and apply Lemma 11.6 to G^* to obtain clusters $A_0, A_1, \dots, A_k$ and $B_0, B_1, \dots, B_k$. Let L and G' be as in Lemma 11.6 and let R be the reduced graph. Then by Lemma 11.9, R has a matching M missing at most $8k(2\gamma + d + \varepsilon)/c = 40\varepsilon k/c$ vertices of R. Recall that for every edge $e = A_iB_j$ of M, the graph $(A_i, B_j)_{G'}$ is ε-regular and has density $> d = 2\varepsilon$. So if n (and thus L) is sufficiently large, then we may successively apply Theorem 11.5 to pull out disjoint copies of $K_{s,t}$ from $(A_i, B_j)_{G'}$ as long as there are at least εL vertices left in each of its vertex classes. Additionally, we may require that the vertex class of the $K_{s,t}$ of size s lies alternately in A_i and in B_j. In this way we get a $K_{s,t}$-packing in $(A_i, B_j)_{G'}$ that covers all but at most $2\varepsilon L + t$ of its vertices. Proceeding similarly for each edge of M, we have found a $K_{s,t}$-packing of G for which the number of uncovered vertices is at most

$$
\begin{aligned}
|A_0 \cup B_0| + L|V(R) \setminus V(M)| + (2\varepsilon L + t) \cdot e(M) &\le \varepsilon n + \frac{40\varepsilon k L}{c} + 3\varepsilon L k \\
&\le \frac{44\varepsilon n}{c} < \alpha n
\end{aligned}
$$

(since $t \le \varepsilon L$ if n is sufficiently large), as required. □

11.4 $K_{s,t}$-packings covering all but a constant number of vertices

Before proving Theorem 11.2, let us first sketch the idea. Similarly as in the proof of Theorem 11.1 we first choose an almost regular bipartite graph $G^* \subseteq G$, apply the Regularity lemma to G^* and choose a large matching M in the reduced graph R. Then we make every bipartite graph corresponding to some edge A_iB_j of M into an (ε', d')-super-regular graph $(A'_i, B'_j)_{G'}$ by adding a small fraction of vertices to the exceptional set $A_0 \cup B_0$ (see Proposition 2.4). We could now apply the Blow-up lemma to these super-regular graphs to obtain a $K_{s,t}$-packing that misses only a constant number of vertices in each of them. But then we would still be left with the exceptional vertices. So instead we proceed as follows. We first assign each exceptional vertex $x \in A_0 \cup B_0$ to one of the super-regular graphs in which x has many neighbours in such a way that to each super-regular graph we assign only a small number of exceptional vertices. If $(A'_i, B'_j)_{G'}$ is such a super-regular graph, then successively for each exceptional vertex x assigned to it we fix a $K_{s,t}$ which contains x but has all its other vertices in $(A'_i, B'_j)_{G'}$. Furthermore, all these $K_{s,t}$'s are chosen to be disjoint and so that the sizes of the remaining subsets of A'_i and B'_j differ at most by a constant. Having dealt with the exceptional vertices, our aim now is to apply the Blow-up lemma to the remainders of the super-regular graphs in order to find disjoint $K_{s,t}$'s that cover all but constantly many vertices. The only problem is that although the remainders of the super-regular graphs are still large, they may now contain vertices of small degree and thus need no longer be super-regular. But by being careful in the choice of the $K_{s,t}$'s containing the exceptional vertices, this can also be fixed. So let us now turn to the details.

Proof of Theorem 11.2. Clearly, it suffices to prove the theorem for the case when $H = K_{s,t}$ where $s < t$. Moreover, by deleting one vertex of G if necessary, we may assume that $|G| =: n$ is even. Set $\gamma := c^3/4000t$, $d := \gamma$ and $\varepsilon := \min\{\gamma/8, \varepsilon_0(d/4, t)/4\}$, where ε_0 is as defined in Lemma 11.7. By making C larger, we may assume that the order n of G is sufficiently large compared with c, s and t. We will show that G contains a $K_{s,t}$-packing which covers all but at most $2tN(\varepsilon)$ vertices, where $N(\varepsilon)$ is as defined in Lemma 11.6. As indicated before, we start similarly as in the proof of Theorem 11.1. First we apply Proposition 11.8 to G to obtain a spanning bipartite $(cn/2 \pm 2\gamma n)$-regular subgraph $G^* = (A, B)$ with $|A| = |B|$. Next we apply Lemma 11.6 to G^* to obtain clusters $A_0, A_1, \dots, A_k$ and $B_0, B_1, \dots, B_k$. Let L and G' be as in Lemma 11.6 and let R be the reduced graph. Then by Lemma 11.9, R has a matching M covering all but at most $8k(2\gamma + d + \varepsilon)/c \le 32\gamma k/c$ vertices of R. Let $L' := (1 - \varepsilon)L$. As for every edge A_iB_j of R the graph $(A_i, B_j)_{G'}$ is ε-regular of density $> d$, we may apply Proposition 2.4 to find sets $A'_i \subseteq A_i$ and $B'_j \subseteq B_j$ such that $|A'_i| = |B'_j| = L'$ and such that the graph $(A'_i, B'_j)_{G'}$ is $(2\varepsilon, d - 2\varepsilon)$-super-regular. Denote by M' the set of all pairs $A'_iB'_j$ for which A_iB_j is an edge of M. Let A'_0 be set of all those vertices in A that do not lie in some A'_i and define B'_0 similarly. Then

$$\begin{aligned} |B'_0| = |A'_0| &\le |A_0| + L \cdot \frac{|V(R) \setminus V(M)|}{2} + \varepsilon L \cdot |M'| \\ &\le \varepsilon n + \frac{16\gamma kL}{c} + \varepsilon kL \le \frac{18\gamma n}{c}. \end{aligned}$$

Given a vertex $x \in A'_0 \cup B'_0$ and a pair $A'_iB'_j \in M'$, we say that $A'_iB'_j$ is x-*friendly* if x has at least $cL'/4$ neighbours in $A'_i \cup B'_j$ (in the bipartite graph G^*). Denoting the number of x-friendly pairs in M' by N_x, we have

$$N_xL' + |M'| \cdot \frac{cL'}{4} \ge d_{G^*}(x) - |N_{G^*}(x) \cap (A'_0 \cup B'_0)| \ge \frac{cn}{2} - 2\gamma n - \frac{18\gamma n}{c} > \frac{cn}{4},$$

and therefore, as $|M'| \le n/2L'$,

$$N_x > \frac{cn}{4L'} - \frac{c|M'|}{4} \ge \frac{cn}{8L'}.$$

But setting $\alpha := 300\gamma/c^2$, this shows that

$$\frac{|A'_0 \cup B'_0|}{\alpha L'} \le \frac{36\gamma n}{c\alpha L'} < N_x$$

for every $x \in A'_0 \cup B'_0$. This implies that we can successively assign every vertex $x \in A'_0 \cup B'_0$ to an x-friendly pair in M' in such a way that to each pair there will be assigned at most $\alpha L'$ vertices.

Consider any pair $A'_iB'_j \in M'$ and let A''_i (respectively B''_j) denote the set of all those vertices in A'_0 (respectively B'_0) that are assigned to $A'_iB'_j$. So every vertex $x \in A''_i$ has at least $cL'/4$ neighbours in B'_j and similarly every vertex in B''_j has at least $cL'/4$ neighbours in A'_i. Let G_{ij} be the bipartite graph obtained from $(A'_i, B'_j)_{G'}$ by adding all vertices in $A''_i \cup B''_j$ and all edges of G^* between these vertices and $A'_i \cup B'_j$.

We now choose bipartitions S_A, T_A of A'_i and S_B, T_B of B'_j such that in the graph G_{ij} each vertex in $A'_i \cup A''_i$ has at least a third of its neighbours in each of S_B and T_B, and such that similarly each vertex of $B'_j \cup B''_j$ has at least a third of its neighbours in each of S_A and T_A. The following simple probabilistic argument shows the existence of such bipartitions. Consider a random bipartition S_A, T_A of A'_i which is obtained by including every vertex $x \in A'_i$ into S_A with probability $1/2$ independently of all other vertices of A'_i. Given a vertex $y \in B'_j \cup B''_j$, set $n_y := |N_{G_{ij}}(y) \cap A'_i|$. Thus, if $y \in B'_j$, then $n_y > (d - 2\varepsilon)L'$ since $(A'_i, B'_j)_{G'}$ is $(2\varepsilon, d - 2\varepsilon)$-super-regular. Also, if $y \in B''_j$, then $n_y \geq cL'/4 > (d - 2\varepsilon)L'$. Call $y \in B'_j \cup B''_j$ *bad* if either $|N_{G_{ij}}(y) \cap S_A| < n_y/3$ or $|N_{G_{ij}}(y) \cap S_A| > 2n_y/3$. Since the expected number of neighbours of y in S_A is $n_y/2$, Lemma 3.6 implies that

$$\mathbb{P}(y \text{ is bad}) \leq 2\mathrm{e}^{-\beta(1/6)n_y/2} \leq 2\mathrm{e}^{-\beta(1/6)(d-2\varepsilon)L'/2} \leq 2\mathrm{e}^{-\beta(1/6)\gamma L'/4}.$$

So if n is sufficiently large, then the expected number of bad vertices in $B'_j \cup B''_j$ is at most $2L' \cdot 2\mathrm{e}^{-\beta(1/6)\gamma L'/4} < 1$. Thus there exists an outcome S_A, T_A for which no vertex in $B'_j \cup B''_j$ is bad, i.e. a bipartition of A'_i having the required properties. The existence of S_B, T_B follows similarly.

Recall that every vertex $x \in A''_i$ (respectively $x \in B''_j$) has at least $n_x/3 \geq cL'/12$ neighbours in S_B (respectively S_A) and

$$\frac{cL'}{12} - |A''_i \cup B''_j|t \geq \frac{cL'}{12} - \alpha t L' = \frac{cL'}{12} - \frac{300\gamma t L'}{c^2} \geq \frac{cL'}{12} - \frac{cL'}{13} \geq 2\varepsilon L'.$$

Moreover, the graph $(A'_i, B'_j)_{G'} \supseteq (S_A, S_B)_{G'}$ is $(2\varepsilon, d - 2\varepsilon)$-super-regular, so we may apply Theorem 11.5 for each vertex in $x \in A''_i \cup B''_j$ in turn to find a $K_{s,t}$ in G_{ij} which contains x and has all its other vertices in $S_A \cup S_B$ and such that all these $K_{s,t}$'s are disjoint for different vertices $x \in A''_i \cup B''_j$. Denote the sets of those vertices in A'_i and B'_j that do not lie in such a $K_{s,t}$ by A^*_i and B^*_j. Note that we may assume that the $K_{s,t}$'s were chosen in such a way that $|\,|A^*_i| - |B^*_j|\,| \leq t - s$. Since firstly in the graph G' every vertex in A^*_i (respectively B^*_j) has more than $(d - 2\varepsilon)L'/3 \geq dL'/4$ neighbours in $T_B \subseteq B^*_j$ (respectively in $T_A \subseteq A^*_i$), since secondly the graph $(A'_i, B'_j)_{G'} \supseteq (A^*_i, B^*_j)_{G'}$ is $(2\varepsilon, d - 2\varepsilon)$-super-regular and since thirdly both A^*_i and B^*_j have size at least $L' - \alpha t L' \geq L'/2$, the graph $(A^*_i, B^*_j)_{G'}$ is still $(4\varepsilon, d/4)$-super-regular. So recalling the choice of ε, we may apply Lemma 11.7 to $(A^*_i, B^*_j)_{G'}$ to find disjoint copies of $K_{s,t}$ that cover all but at most $2(t-1)$ vertices of $(A^*_i, B^*_j)_{G'}$. Together with the $K_{s,t}$'s chosen earlier we have found a $K_{s,t}$-packing in G_{ij} that covers all but at most $2(t-1)$ of its vertices. Proceeding similarly for each pair in M', we have found a $K_{s,t}$-packing in G for which the number of uncovered vertices is at most $|M'|2(t-1) < N(\varepsilon)2t$. $\square$

11.5 Perfect TK_4- and TK_5-packings

Proof of Theorem 11.4. We only consider the case when $r = 4$ in detail, the proof for $r = 5$ works similarly. Set $s := \lceil 4/c \rceil$ and $\alpha := c/16s$. We will prove the theorem for $\gamma := \min\{c/2, \gamma(c, \alpha)\}$, where $\gamma(c, \alpha)$ is as defined in Theorem 11.1, and for sufficiently large n_0. First we apply Theorem 11.1 to obtain a $K_{s,s}$-packing M which covers all but at most αn vertices of G. Let X be the set of all remaining vertices.

We now successively assign each vertex $x \in X$ to a $K_{s,s}$ in M such that x has at least two neighbours in one of the colour classes of the $K_{s,s}$ and such that these $K_{s,s}$'s are distinct for different vertices in X. This can be done since for each $x \in X$

$$d_G(x) \geq cn - \gamma n \geq \frac{cn}{2} \geq 4s\alpha n + \frac{n}{s} > |N_G(x) \cap X| + 2s|X| + \frac{2n}{|K_{s,s}|}$$

and thus after each step for the vertex $x \in X$ to be considered next there must be a $K_{s,s}$ in M to which no vertex in X has been assigned yet and such that x has at least two neigbours in one of its colour classes.

As $s \geq 3$, it is easy to see that we can choose a spanning subdivision of K_4 in each $K_{s,s}$ of M to which we have assigned no vertex of X. Similarly, for every copy (A, B) of $K_{s,s}$ in M to which we have assigned a vertex $x \in X$ we can choose a subdivision of K_4 which contains x and all vertices in (A, B). Indeed, we may assume that x has two neighbours in A. These will be two of the branch vertices (which will be connected by the path of length 2 through x), we choose any other vertex in A as third branch vertex and any vertex in B as the fourth one (Fig. 11.1). In the case $r = 5$ we proceed similarly except that we now choose an additional branch vertex from B (Fig. 11.1). $\square$

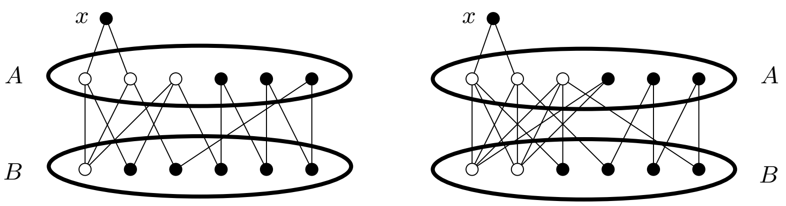


Figure 11.1: A spanning subdivision of K_4 (left) and K_5 (right). The white vertices are the branch vertices.

The graph $K_{n,n+1}$ immediately shows that we cannot hope for a perfect TK_r-packing if $r = 3$. The following proposition states that this is also the case for all $r \geq 6$ and thus the graph consisting of a disjoint union of $K_{n,n+1}$'s shows that for $r \geq 6$ the vertex degrees in Corollary 11.3 have to be linear.

Proposition 11.10 *For all integers n and $r \geq 6$ each TK_r-packing in $K_{n,n+1}$ misses at least one vertex.*

Proof. First consider the case when r is even, say $r = 2\ell$. Let A and B be the vertex classes of $K_{n,n+1}$. Consider any subdivision T of K_r in $K_{n,n+1}$. Let $k \in \mathbb{Z}$ be so that that $\ell + k$ branch vertices of T lie in A and $\ell - k$ branch vertices lie B. Then, no matter how the edges of K_r are subdivided,

$$|V(T) \cap B| - |V(T) \cap A| = \ell - k + \binom{\ell + k}{2} - (\ell + k) - \binom{\ell - k}{2} = (r-3)k.$$

But this implies that in every TK_r-packing in $K_{n,n+1}$ the difference between the number of covered vertices in A and the number of covered vertices in B is divisible by $r - 3 \geq 3$ and hence cannot be 1. So there exists no perfect TK_r-packing.

In the case when r is odd, $r = 2\ell + 1$ say, we let k be such that $\ell + k + 1$ branch vertices lie in A to obtain $|V(T) \cap B| - |V(T) \cap A| = (2k+1)(r-3)/2$ which gives us the same conclusion. $\square$

To conclude this section, we give an example of a regular graph G which shows that for a perfect TK_r-packing (where $r \geq 3$) a minimum degree of at least $\sqrt{n/2}$ is necessary. Let $k \geq 3$ be an odd integer and let H be the graph which is obtained from a $K_{k,k-1}$ by including $(k-1)/2$ independent edges in the larger vertex class. So all vertices of H have degree k except for one vertex of degree $k - 1$. Let G be the graph which consists of k disjoint copies of H together with a new vertex x that is joined to the vertices of degree $k - 1$ in all copies of H (Fig. 11.2). So x does not lie in a subdivision of K_r, G is k-regular and $|G| = k(2k-1) + 1$.

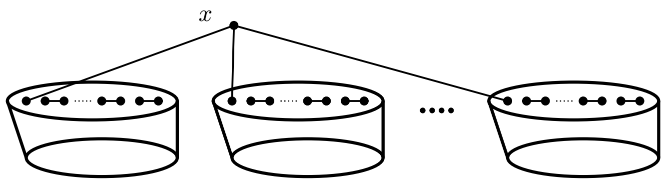


Figure 11.2: A regular graph of minimum degree at least $\sqrt{n/2}$ and no perfect TK_r-packing.

Chapter 12

Large planar subgraphs in dense graphs

12.1 Introduction

In this chapter we study the following extremal question: Given a function $m = m(n)$, how large does the minimum degree of a graph G of order n have to be in order to guarantee a planar subgraph with at least $m(n)$ edges?

If $m \leq n$, the answer is easy. Indeed, suppose that the minimum degree of G is at least one. Then every component C of G has a spanning tree with $|C| - 1 \geq |C|/2$ edges. So G has a (planar) spanning forest with at least $n/2$ edges, which is best possible if G consists of independent edges. Similarly, it is easy to see that if G has minimum degree at least two, then G contains a planar subgraph with n edges, which is best possible if G is a cycle.

On the other hand, if G is bipartite, then the facial cycles of any planar subgraph have length at least four and so Euler's formula implies that no planar subgraph of G has more than $2n - 4$ edges. So as long as the minimum degree is at most $n/2$, we cannot hope for a planar subgraph with more than $2n - 4$ edges. Our first theorem shows that a much smaller minimum degree already guarantees a planar subgraph with roughly $2n$ edges.

Theorem 12.1 *For every $0 < \varepsilon < 1$ there exists $n_0 = n_0(\varepsilon)$ such that every graph G of order $n \geq n_0$ and minimum degree $\delta \geq 1500\sqrt{n}/\varepsilon^2$ contains a planar subgraph with at least $(2 - \varepsilon)n$ edges.*

This is essentially best possible in two ways. Firstly, there are graphs with minimum degree $\sqrt{n}/2$ and girth at least 6 ([23], see also [12]). Hence Euler's formula shows that any planar subgraph of such a graph can have at most $\frac{3}{2}(n - 2)$ edges (as all of its facial cycles have length at least 6). Secondly, for $\delta \leq n/2$ consider the graph consisting of $n/2\delta$ disjoint copies of the complete bipartite graph $K_{\delta,\delta}$. It obviously has minimum degree δ, but again by Euler's formula it cannot contain a planar subgraph with more than $(2 \cdot 2\delta - 4)n/2\delta = 2n - 2n/\delta$ edges. This shows that as long as the minimum degree δ of G is $o(n)$, we cannot ask for a planar subgraph of G with $2n - C$ edges, where C does not depend on n. So if we want at least $2n - C$ edges in a planar subgraph, then a necessary condition is that $\delta \geq 2n/C$, i.e. δ must be linear in n. Our second theorem shows that the linearity of δ is also sufficient.

Theorem 12.2 *For every $\gamma > 0$ there exists $C = C(\gamma)$ such that every graph G of order n and minimum degree at least γn contains a planar subgraph with at least $2n - C$ edges.*

As we have already seen, this is best possible up to the value of the constant C as long as the minimum degree is at most $n/2$. If however the minimum degree is a little larger than this, we can already guarantee a planar subgraph which is a triangulation apart from a constant number of missing edges:

Theorem 12.3 *For every $\gamma > 0$ there exists $C = C(\gamma)$ such that every graph G of order n and minimum degree at least $(1/2 + \gamma)n$ contains a planar subgraph with at least $3n - C$ edges.*

Again, this is best possible in the sense that the constant C has to depend on γ and the additional term γn in the bound on the minimum degree cannot be replaced by a sublinear one (see Proposition 12.11).

Finally, we seek a spanning triangulation, i.e. a planar subgraph with $3n-6$ edges. As pointed out to us by Bollobás, the following 3-partite graph G shows that a minimum degree of $2n/3$ is necessary for this. G is obtained from two disjoint cliques C_1 and C_2 of order $n/3$ by adding an independent set X of $n/3$ new vertices and joining each of them to all the vertices in the two cliques. So G has minimum degree $2n/3 - 1$. Observe that any spanning triangulation in G would have two facial triangles T_1 and T_2 which share an edge and are such that T_i contains a vertex of C_i $(i = 1, 2)$. But this is impossible since every triangle of G containing a vertex of C_i can have at most one vertex outside C_i, namely in X. However, to guarantee a triangulation, it suffices to increase the minimum degree by a small amount:

Theorem 12.4 *For every $\gamma > 0$ there exists an integer $n_0 = n_0(\gamma)$ such that every graph G of order $n \geq n_0$ and minimum degree at least $(2/3+\gamma)n$ contains a triangulation as a spanning subgraph.*

In Chapter 13 we show that for sufficiently large graphs a minimum degree of $2n/3$ suffices. However, the proof of this is rather more involved than that of Theorem 12.4.

We also obtain an analogue of Theorem 12.4 for quadrangulations, i.e. plane subgraphs with $2n - 4$ edges in which every face is bounded by a 4-cycle.

Theorem 12.5 *For every $\gamma > 0$ there exists an integer $n_0 = n_0(\gamma)$ such that every graph G of order $n \geq n_0$ and minimum degree at least $(1/2+\gamma)n$ contains a quadrangulation as a spanning subgraph.*

The disjoint union of two cliques of order $n/2$ shows that apart from the error term γn, the minimum degree in Theorem 12.5 cannot be reduced.

We remark that Theorems 12.4 and 12.5 would immediately follow from the conjecture of Bollobás and Komlós [48] that for every $\gamma > 0$ and all $r, \Delta \in \mathbb{N}$ there are $\alpha > 0$ and $n_0 \in \mathbb{N}$ such that every graph G of order $n \geq n_0$ and minimum degree at least $(1 - \frac{1}{r} + \gamma)n$ contains a copy of every graph H of order n whose chromatic number is at most r, whose maximum degree is at most Δ and whose band-width is at most αn. (The *band-width* of a graph H is the smallest integer k for which there exists an enumeration $v_1, \dots, v_{|H|}$ of the vertices of H such that every edge $v_i v_j \in H$ satisfies $|i - j| \leq k$.) Indeed, to derive e.g. Theorem 12.4 from this conjecture it suffices to find for all $n \in \mathbb{N}$ a 3-partite triangulation of order n which has both bounded maximum degree and bounded band-width. It is easy to see that such triangulations exist (e.g. modify the graph H_1 in Fig. 12.3 below).

Theorems 12.1–12.5 give a fairly accurate picture of the maximum size of a planar subgraph when we consider graphs whose minimum degree δ is much larger than $\sqrt{n}$. However, we are not aware of any nontrivial lower bounds when δ lies between 2 and $\sqrt{n}$. An easy upper bound is obtained as follows. For $\ell \geq 3$ let $\delta_{2\ell} = \delta_{2\ell}(n)$ be the largest integer such that there are graphs G of order n, minimum degree at least $\delta_{2\ell}$ and girth at least 2ℓ. (The order of magnitude of $\delta_{2\ell}$ is only known for $\ell = 3, 4$ and 6, see e.g. [10, 23].) So all facial cycles in a planar subgraph of such a graph G have length at least 2ℓ and thus Euler's formula gives us an upper bound on the size of a planar subgraph of G. We believe that in general this upper bound is close to the truth (except maybe when the minimum degree is only a little larger than $\delta_{2\ell+2}$).

Our proofs immediately show that the planar subgraphs guaranteed by Theorems 12.1–12.5 can be found in polynomial time. For graphs with high minimum degree we therefore obtain improved approximation algorithms for the maximum planar subgraph problem which in a given graph G asks for a planar subgraph with the maximum number of edges. Călinescu et al. [24] showed that this problem is Max SNP-hard: there is a constant ε such that there cannot exist a polynomial time approximation algorithm with approximation ratio better than $1 - \varepsilon$, unless $P = NP$. The best known approximation algorithm has an approximation ratio of 4/9 [24]. (Note that a ratio of 1/3 is already achieved by producing spanning trees for all connected components.)

The problem of finding a large planar subgraph in a random graph was investigated by Schlatter [94], the case of triangulations was already considered earlier by Bollobás and Frieze [18].

The chapter is organized as follows. In Section 12.2 we give a brief sketch of the proofs of Theorems 12.1–12.5. In Section 12.3 we collect some notation and some information about the Regularity lemma and the Blow-up lemma we need for the proofs of Theorems 12.2–12.5. The proofs themselves are then given in the final section.

12.2 Sketch of proofs

The proof of Theorem 12.1 is rather different from those of the other results. In particular, it relies neither on the Regularity lemma nor on the Blow-up lemma. The strategy is to repeatedly find a suitable greedy covering of part of the vertices of the original graph G with disjoint complete bipartite graphs $K_{2,s}$, where s is large. (Note that if s is large then the planar graph $H := K_{2,s}$ has roughly $2|H|$ edges.) These partial coverings (which will overlap a little) are then combined into a single planar graph of the required size.

We now give a sketch of the proofs of Theorems 12.2–12.5. The structure of these proofs is similar: we first apply the Regularity lemma (Lemma 2.5) to obtain a partition of the vertices of G into a large but constant number of clusters. Since G has large minimum degree, this is also true for the 'reduced graph' R (whose vertices are the clusters and whose edges correspond to the pairs of clusters which are regular and have sufficient density). We will use this to cover almost all vertices of R by suitable disjoint graphs H of bounded size. Then we apply the Blow-up lemma (Lemma 12.7) to find spanning planar graphs P of the required density within the subgraphs H' of G corresponding to these graphs H. However, we also have to ensure that the exceptional vertices of G (i.e. the small proportion of those vertices of G which do not belong to some such H') can be incorporated into these planar graphs P without reducing their density. This also follows from the Blow-up lemma provided that we can assign each exceptional vertex v to some H which contains enough clusters with many neighbours of v in such a way that to each H we assign only a small number of exceptional vertices.

In the proof of Theorem 12.2 the graphs H will be stars of bounded size and the planar graphs P we seek within the graphs H' will be quadrangulations. For Theorem 12.3 we want the planar graphs P to be triangulations, which means that the graphs H can no longer be bipartite. Thus an obvious choice for H would be a triangle, but we cannot hope to cover almost all vertices of the reduced graph R by disjoint triangles since its minimum degree may be only a little larger than $|R|/2$. However, a recent result of Komlós (Theorem 12.10) implies that we can take H to be the complete 3-partite graph $K_{a,a,1}$ (where a is large) as it is in some sense close to being bipartite.

In the proof of Theorem 12.4 the minimum degree of the reduced graph R exceeds $2|R|/3$ and hence the Theorem of Corrádi and Hajnal [26] implies that R can be covered by disjoint triangles. However, this is not sufficient for our purposes as this time we seek a single triangulation containing all vertices of G (instead of a disjoint union of boundedly many triangulations as in the proof of Theorem 12.3). So we have to 'glue together' the different triangulations corresponding to the triangles covering R. For this we use suitable edges of R joining these triangles (as well as some additional vertices of G). Thus instead of merely covering R by disjoint triangles, we will start with the second power of a Hamilton path of R. The latter is guaranteed by a result of Fan and Kierstead [33].

The proof of Theorem 12.5 is similar to that of Theorem 12.4 but the gluing process is somewhat simpler. Instead of the second power, this time it suffices to work with an 'ordinary' Hamilton path.

12.3 Notation and tools

We often write n for the order of a graph G, if this is not ambiguous. Given a plane graph G, a *facial cycle* of G is a cycle in G which is the boundary of a face. G is a *triangulation* if all its faces are bounded by triangles and a *quadrangulation* if all faces are bounded by 4-cycles. So by Euler's formula a triangulation has $3n-6$ edges whereas a quadrangulation has $2n-4$ edges.

The proofs of Theorems 12.2–12.5 are based on the Regularity lemma. We will often use the following well-known and simple fact about the minimum degree of the reduced graph. Its proof is the only place in this chapter where the degree form of the Regularity lemma (Lemma 2.5) is more convenient than the classical form.

Proposition 12.6 *For every $\gamma > 0$ there exist $\varepsilon_0 = \varepsilon_0(\gamma)$ and $d_0 = d_0(\gamma)$ such that for all $\varepsilon \le \varepsilon_0$, $d \le d_0$ and every $c \ge 0$ an application of Lemma 2.5 to a graph G of minimum degree at least $(c+\gamma)|G|$ yields a reduced graph R of minimum degree at least $(c+\gamma/2)|R|$.*

Proof. Suppose that there is a vertex $V_i \in R$ whose degree in R is less than $(c+\gamma/2)k$. Let W denote the union of all those clusters V_j ($j \neq i$) for which $(V_i, V_j)_{G'}$ has density 0. Let u be any vertex in V_i. Then

$$\begin{aligned} d_{G'}(u) &\le |N_{G'}(u) \cap W| + d_R(V_i) \cdot L + |N_{G'}(u) \cap V_0| < 0 + (c+\gamma/2)kL + \varepsilon n \\ &\le (c + \gamma/2 + \varepsilon)n. \end{aligned}$$

But on the other hand, Lemma 2.5 states that $d_{G'}(u) > d_G(u) - (d+\varepsilon)n \ge (c+\gamma-d-\varepsilon)n$, a contradiction, provided that $\gamma \ge 2d + 4\varepsilon$. $\square$

We will also use the Blow-up lemma of Komlós, Sárközy and Szemerédi [51]. It implies that dense regular pairs behave like complete bipartite graphs with respect to containing bounded degree graphs as subgraphs.

Lemma 12.7 (Blow-up lemma) *Given a graph R on $\{1,\dots,r\}$ and numbers $d, c, \Delta > 0$, there are positive numbers $\varepsilon_0 = \varepsilon_0(d,\Delta,r,c)$ and $\alpha = \alpha(d,\Delta,r,c) \le 1/2$ such that the following holds. Given $L \in \mathbb{N}$ and $\varepsilon \le \varepsilon_0$, let $R(L)$ be the graph obtained from R by replacing each vertex $i \in R$ with a set V_i of L new vertices and joining all vertices in V_i to all vertices in V_j whenever ij is an edge of R. Let G be a spanning subgraph of $R(L)$ such that for every edge $ij \in R$ the graph $(V_i, V_j)_G$ is (ε, d)-super-regular. Then G contains a copy of every subgraph H of $R(L)$ with $\Delta(H) \le \Delta$. Furthermore, we can additionally require that for vertices $x \in H \subseteq R(L)$ lying in V_i their images in the copy of H in G are contained in (arbitrary) given sets $C_x \subseteq V_i$ provided that $|C_x| \ge cL$ for each such x and provided that in each V_i there are at most αL such vertices x.*

We say that the vertices x in Lemma 12.7 are *image restricted to C_x*.

12.4 Proofs

12.4.1 Planar subgraphs of size $2n - \varepsilon n$

In our proof of Theorem 12.1 we will use the following well-known upper bound on the number of edges of $K_{2,s}$-free graphs (see e.g. [12, Ch. VI, Thm. 2.2 and 2.3]).

Theorem 12.8 *Let $s \geq 2$ be an integer. Then every graph G with $e(G) \geq \sqrt{s}n^{3/2}$ contains a copy of $K_{2,s}$. Moreover, every bipartite graph $G = (A, B)$ with $e(G) \geq \sqrt{s}|A||B|^{1/2} + |B|$ contains a copy of $K_{2,s}$ with 2 vertices in A and s vertices in B.*

Proof of Theorem 12.1. Throughout the proof we assume that n is sufficiently large for our estimates to hold. For all $k \geq 1$ set $s_k := 2^{k^2+2}/\varepsilon^k$. We first greedily choose as many disjoint copies of K_{2,s_1} in G as possible. Let P_1 be the union of all these K_{2,s_1}'s, $X_1 := V(P_1)$ and let $Y_1 := V(G) \setminus X_1$. Thus $G[Y_1]$ is K_{2,s_1}-free and so Theorem 12.8 implies that $e(G[Y_1]) \leq \sqrt{s_1}|Y_1|^{3/2}$. Let Y_1' be the set of all those vertices in Y_1 which have at most $\delta/2$ neighbours in X_1. Then

$$\delta|Y_1'|/2 \leq 2e(G[Y_1]) \leq 2\sqrt{s_1}n^{3/2},$$

and thus

$$|Y_1'| \leq \frac{4\sqrt{s_1}n^{3/2}}{\delta}. \tag{12.1}$$

Let $Y_1^* := Y_1 \setminus Y_1'$. Next we greedily choose (as often as possible) disjoint copies of K_{2,s_2} in $(X_1, Y_1^*)_G$ having 2 vertices in X_1 and s_2 vertices in Y_1^*. Let P_2 be the union of all these K_{2,s_2}'s, $X_2 := V(P_2) \cap X_1$ and $Y_2 := V(P_2) \cap Y_1^*$. Let Y_2' be the set of all those vertices in $Y_1^* \setminus Y_2$ which have at most $\delta/2^2$ neighbours in X_2. Thus each vertex in Y_2' has at least $\delta/2^2$ neighbours in $X_1 \setminus X_2$ and so

$$e(X_1 \setminus X_2, Y_2') \geq \delta|Y_2'|/2^2.$$

On the other hand, $(X_1 \setminus X_2, Y_2')_G$ does not contain a K_{2,s_2} with 2 vertices in $X_1 \setminus X_2$ and s_2 vertices in Y_2'. Thus Theorem 12.8 implies

$$e(X_1 \setminus X_2, Y_2') \leq \sqrt{s_2} \cdot |X_1 \setminus X_2| \cdot |Y_2'|^{1/2} + |Y_2'| \leq \sqrt{s_2}n^{3/2} + |Y_2'|$$

and therefore

$$|Y_2'| \leq \frac{\sqrt{s_2}n^{3/2}}{\delta/2^2 - 1} \leq \frac{5\sqrt{s_2}n^{3/2}}{\delta}. \tag{12.2}$$

Let $Y_2^* := Y_1^* \setminus (Y_2 \cup Y_2')$ and greedily choose (again as often as possible) disjoint copies of K_{2,s_3} in $(X_2, Y_2^*)_G$ having 2 vertices in X_2 and s_3 vertices in Y_2^*. Let P_3 be the union of all these K_{2,s_3}'s, $X_3 := V(P_3) \cap X_2$ and $Y_3 := V(P_3) \cap Y_2^*$. Let Y_3' be the set of all those vertices in $Y_2^* \setminus Y_3$ which have at most $\delta/2^3$ neighbours in X_3. Let $Y_3^* := Y_2^* \setminus (Y_3 \cup Y_3')$ and continue in this fashion until $P_i = \emptyset$ (and thus $X_i = Y_i = \emptyset$ and $Y_i' = Y_{i-1}^*$). Let i be the smallest index such that $P_i = \emptyset$. Thus $i \leq \sqrt{\log n}$ since $s_{\sqrt{\log n}} > n$.

Using that $|X_{k-1} \setminus X_k| \leq |X_{k-1}| \leq 2n/s_{k-1}$ for all $3 \leq k \leq i$, a calculation similar to the case $k = 2$ shows that

$$|Y'_k| \leq \frac{5 \cdot 2^{k-1}\sqrt{s_k} n^{3/2}}{\delta s_{k-1}}. \tag{12.3}$$

Moreover, since $Y'_i = Y^*_{i-1}$,

$$|X_1| + \sum_{k=2}^{i-1} |Y_k| + \sum_{k=1}^{i} |Y'_k| = n. \tag{12.4}$$

Set

$$P := (P_1 - X_2) \cup (P_2 - X_3) \cup \cdots \cup (P_{i-2} - X_{i-1}) \cup P_{i-1}.$$

Clearly P is a planar subgraph of G. Notice that when removing X_k from P_{k-1}, we destroy at most $s_{k-1}|X_k|$ of its edges, but this is negligible compared to $e(P_k) = 2|Y_k|$, as s_k grows rather rapidly with k. Also, recall that $|X_k| \leq 2n/s_k$ for $k \geq 2$. Hence

$$\begin{aligned}
e(P) &\geq \sum_{k=1}^{i-1} e(P_k) - \sum_{k=2}^{i-1} s_{k-1}|X_k| \\
&\geq \frac{2s_1|X_1|}{s_1+2} + \sum_{k=2}^{i-1} 2|Y_k| - \sum_{k=2}^{i-1} \frac{2s_{k-1}n}{s_k} \\
&\overset{(12.4)}{=} 2(n - \sum_{k=1}^{i} |Y'_k|) - \frac{4|X_1|}{s_1+2} - \sum_{k=2}^{i-1} \frac{\varepsilon n}{2^{2k-2}} \\
&\overset{(12.1,12.2,12.3)}{\geq} 2n - \frac{32n^{3/2}}{\delta\sqrt{\varepsilon}} - \frac{80n^{3/2}}{\delta\varepsilon} - \sum_{k=3}^{i} \frac{5 \cdot 2^k \sqrt{s_k} n^{3/2}}{\delta s_{k-1}} - \frac{4n}{s_1} - \frac{\varepsilon n}{3} \\
&\geq 2n - \frac{\varepsilon n}{9} - \sum_{k=3}^{i} \frac{5\varepsilon^{k/2-1} n^{3/2}}{2^{k^2/2-3k+2}\delta} - \frac{\varepsilon n}{2} - \frac{\varepsilon n}{3} \\
&\geq 2n - \frac{80n^{3/2}}{\delta} - \frac{17\varepsilon n}{18} \\
&\geq (2-\varepsilon)n,
\end{aligned}$$

as required. □

We remark that the proof of Theorem 12.1 shows that we can let ε be any function of n with $\varepsilon(n) \leq 1$. Note that it does not make sense to take $\varepsilon(n) \leq n^{-1/4}$.

12.4.2 Planar subgraphs of size $2n - C$

For the proof of Theorem 12.2 we need the following simple proposition.

Proposition 12.9 *Given $0 < \gamma \leq 1/2$ and a graph G of minimum degree at least γn, there exists a set $\mathcal{S}$ of disjoint substars of G such that every vertex of G lies in some $S \in \mathcal{S}$ and such that each such S satisfies $1 \leq \Delta(S) \leq 1/\gamma$.*

Proof. Construct the stars in $\mathcal{S}$ greedily as follows. Suppose that we have already covered a set $X \subseteq V(G)$ with a set $\mathcal{S}'$ of disjoint substars of G such that $1 \leq \Delta(S) \leq 1/\gamma$ for every $S \in \mathcal{S}'$. Choose $x \in V(G) \setminus X$. If x has a neighbour y outside X, we may add the star consisting of the edge xy to $\mathcal{S}'$. So suppose that all neighbours of x lie in X. If x is joined to a leaf y of some star $S \in \mathcal{S}'$ then, if $|S| \geq 3$, we can replace S by $S - y$ and add the new star xy to $\mathcal{S}'$ or, if $|S| = 2$, we can replace S by $S \cup xy$. If x is only joined to midpoints of stars in $\mathcal{S}'$, then one such star must have at most $1/\gamma - 1$ leaves and so we can add x to this star. □

Proof of Theorem 12.2. Clearly, we may assume that $\gamma \leq 1/2$. Let $\varepsilon_0(\gamma)$ and $d_0(\gamma) =: d$ be as given in Proposition 12.6. Let $\varepsilon_0(d/2, 8/\gamma, 1+2/\gamma, \gamma/4) =: \varepsilon^*$ and $\alpha(d/2, 8/\gamma, 1+2/\gamma, \gamma/4) =: \alpha$ be as defined in the Blow-up lemma (Lemma 12.7). Put

$$\varepsilon := \min\left\{\varepsilon_0(\gamma), \frac{\varepsilon^*}{2}, \frac{\gamma^3\alpha}{72}, \frac{\gamma d}{6}\right\}.$$

Clearly, it suffices to show that every graph G whose order n is sufficiently large compared with γ contains a planar subgraph with at least $2n - 4N(\varepsilon, 2)$ vertices, where $N(\varepsilon, 2)$ is given by the Regularity lemma (Lemma 2.5). So throughout the proof we assume that n is sufficiently large.

We first apply the Regularity lemma to G to obtain an exceptional set V_0 and clusters $V_1, \dots, V_k$ where $2 \leq k \leq N(\varepsilon, 2)$. Let L and G' be as defined in the Regularity lemma and let R denote the reduced graph. Thus Proposition 12.6 implies that $\delta(R) \geq \gamma k/2$. So by Proposition 12.9 there exists a set $\mathcal{S}$ of disjoint substars of R such that every vertex of R lies in some star from $\mathcal{S}$ and such that $1 \leq \Delta(S) \leq 2/\gamma$ for each $S \in \mathcal{S}$.

Next we apply Proposition 2.6 to obtain sets $V_i' \subseteq V_i$ of size $(1 - 2\varepsilon/\gamma)L =: L'$ such that for all the edges V_iV_j of R lying in some star from $\mathcal{S}$ the graph $(V_i', V_j')_{G'}$ is $(2\varepsilon, d - (1 + 2/\gamma)\varepsilon)$-super-regular. Henceforth we will think of R and of the stars in $\mathcal{S}$ as graphs whose vertices are the new sets V_i'. Add all vertices of G which do not lie in some V_i' to the exceptional set V_0. By adding further vertices to V_0 if necessary, we may assume that L' is even. We still denote the enlarged exceptional set by V_0. Thus $|V_0| \leq \varepsilon n + 2\varepsilon kL/\gamma + k \leq 3\varepsilon n/\gamma$.

Given a vertex $v \in V_0$ and a star $S \in \mathcal{S}$, we say that S is *v-friendly* if there is a vertex $V_i' \in S$ such that v has at least $\gamma L'/4$ neighbours in V_i'. Let N_v denote the number of v-friendly stars $S \in \mathcal{S}$. Then

$$\gamma n/2 < (\gamma - 3\varepsilon/\gamma)n \leq d_G(v) - |V_0| \leq N_v \cdot (1 + 2/\gamma)L' + \sum_{S \in \mathcal{S}} |S| \cdot \gamma L'/4,$$

and therefore, since $\sum_{S \in \mathcal{S}} |S| = k$,

$$N_v > \frac{\gamma}{3L'} \cdot \left(\frac{\gamma n}{2} - \frac{\gamma k L'}{4}\right) \geq \frac{\gamma^2 n}{12L'}.$$

So

$$\frac{2|V_0|}{\alpha L'} \leq \frac{6\varepsilon n}{\gamma \alpha L'} < N_v$$

for every vertex $v \in V_0$. But this implies that we can greedily assign each vertex $v \in V_0$ to a v-friendly star $S \in \mathcal{S}$ in such a way that to every $S \in \mathcal{S}$ we assign at most $\alpha L'/2$ vertices from V_0.

Consider a fixed $S \in \mathcal{S}$ and let $X \subseteq V_0$ be the set of all vertices assigned to S. Let U_1 be the centre of S and let $U_2, \ldots, U_{|S|}$ be its other vertices. So each U_ℓ is a set of the form V_i'. Fix any bipartite quadrangulation P_S of maximum degree $4\Delta(S) \leq 8/\gamma$ whose vertex classes are U_1 and $U_2 \cup \cdots \cup U_{|S|}$ such that for each $\ell \leq |S|$ there is a set $\mathcal{C}_\ell$ of at least $L'/4 \geq |X|$ facial 4-cycles of P_S with the property that, firstly, each $C \in \mathcal{C}_\ell$ has two of its vertices in U_ℓ, secondly, these vertices are distinct for different $C \in \mathcal{C}_\ell$ and thirdly, each facial 4-cycle of P_S lies in at most one such $\mathcal{C}_\ell$. Recalling that L' is even, it is not difficult to see that such quadrangulations exist (see Fig. 12.1).

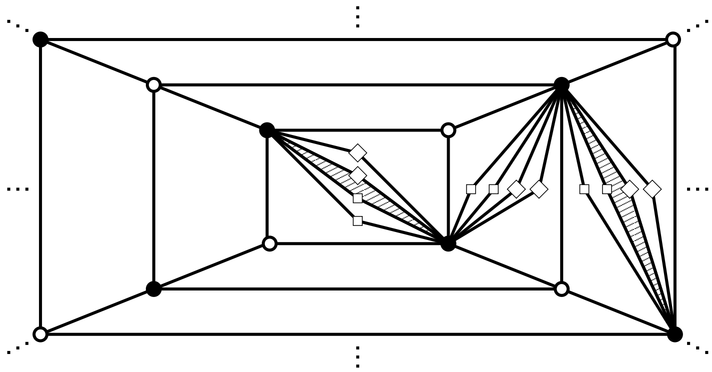


Figure 12.1: A quadrangulation P_S which corresponds to a star S with three leaves. The black vertices belong to U_1. The shaded faces indicate a possible choice for $\mathcal{C}_1$.

As each edge of S corresponds to a $(2\varepsilon, d/2)$-super-regular subgraph of G', the Blow-up lemma (Lemma 12.7) implies that the subgraph of G' corresponding to S (that is $G'[U_1 \cup \cdots \cup U_{|S|}]$) contains a spanning copy of P_S such that every vertex $v \in X$ is joined to two opposite vertices on some facial 4-cycle of P_S and such that these 4-cycles differ for distinct vertices $v \in X$. Indeed, this can be achieved as follows. By definition, each $v \in X$ has at least $\gamma L'/4$ neighbours in some U_ℓ $(1 \leq \ell \leq |S|)$. Assign v to a cycle $C_v \in \mathcal{C}_\ell$ such that these cycles C_v differ for distinct such v. When applying the Blow-up lemma, for each $v \in X$ the two vertices in $V(C_v) \cap U_\ell$ are image restricted to the neighbourhood of v in U_ℓ. (This can be done since the vertices in $V(C_v) \cap U_\ell$ are distinct for different v.)

The graph obtained from P_S by inserting all the vertices $v \in X$ in their facial 4-cycles C_v is still a quadrangulation. Hence G contains a planar subgraph which is a disjoint union of $|\mathcal{S}|$ quadrangulations and thus has $2n - 4|\mathcal{S}| \geq 2n - 4N(\varepsilon, 2)$ edges. $\square$

12.4.3 Planar subgraphs of size $3n - C$

The *critical chromatic number* $\chi_{cr}(H)$ of a graph H is defined as $(\chi(H) - 1)|H|/(|H| - \sigma)$, where σ denotes the minimum size of the smallest colour class in an optimal colouring of H. For the proof of Theorem 12.3 we need the following result of Komlós [49, Thm. 8].

Theorem 12.10 *For every $\varepsilon > 0$ and every graph H there exists an integer $k_0 = k_0(H, \varepsilon)$ such that all but at most εk vertices of every graph R of order $k \geq k_0$ and minimum degree $\delta(R) \geq (1 - 1/\chi_{cr}(H))k$ can be covered by disjoint copies of H.*

Note that Theorem 12.10 immediately implies that for all $\varepsilon, \gamma > 0$ there exists an integer $n_0 = n_0(\varepsilon, \gamma)$ such that every graph R of order $n \geq n_0$ and minimum degree at least γn contains a planar graph with at least $2n - \varepsilon n$ edges. Indeed, let $H := K_{2,s}$ in Theorem 12.10, where s is sufficiently large compared to ε and γ. Then the critical chromatic number of H is close to one and the disjoint union of all copies of H given by Theorem 12.10 is a planar subgraph of R of the required size. Similarly, as there exist large triangulations whose critical chromatic number is close to 2 (e.g. modify the graph in Fig. 12.2 below), Theorem 12.10 implies that Theorem 12.3 is true for large n if we only ask for a planar subgraph with $3n - \varepsilon n$ edges.

Proof of Theorem 12.3. By making γ smaller, we may assume that $1/\gamma$ is an integer divisible by 4. Let $\varepsilon_0(\gamma)$ and $d_0(\gamma) =: d$ be as given in Proposition 12.6. Set $a := 2/\gamma$ and $H := K_{a,a,1}$, the complete 3-partite graph with vertex classes of size a, a and 1. Let $\varepsilon_0(d/2, 8a, 2a+1, \gamma/4) =: \varepsilon^*$ and $\alpha(d/2, 8a, 2a+1, \gamma/4) =: \alpha$ be as defined in the Blow-up lemma (Lemma 12.7). Put

$$\varepsilon := \min\left\{\varepsilon_0(\gamma), \frac{\varepsilon^*}{2}, \frac{\gamma^3\alpha}{640}, \frac{\gamma d}{12}\right\}$$

and let $k_0 := k_0(H, \varepsilon)$ be defined as in Theorem 12.10. Clearly, it suffices to show that every graph G whose order n is sufficiently large compared with γ contains a planar subgraph with at least $3n - 6N(\varepsilon, k_0)$ vertices, where $N(\varepsilon, k_0)$ is given by the Regularity lemma (Lemma 2.5).

We first apply the Regularity lemma to G to obtain an exceptional set V_0 and clusters $V_1, \dots, V_k$ where $k_0 \leq k \leq N(\varepsilon, k_0)$. Let L and G' be as defined in the Regularity lemma and let R denote the reduced graph. Thus Proposition 12.6 implies that $\delta(R) \geq (1/2 + \gamma/2)k$. As $\chi_{cr}(H) = 2(2a+1)/2a = 2 + 1/a$ and therefore $\delta(R) \geq (1 - 1/\chi_{cr}(H))k$, we can apply Theorem 12.10 to obtain a set $\mathcal{H}$ of disjoint copies of H in R such that all but at most εk vertices of R lie in the union H' of all these copies. As $\Delta(H') = 2a$, we may apply Proposition 2.6 to find for every $V_i \in V(H')$ a set $V_i' \subseteq V_i$ of size $(1 - 2a\varepsilon)L =: L'$ such that for every edge $V_iV_j \in H'$ the graph $(V_i', V_j')_{G'}$ is $(2\varepsilon, d - (1 + 2a)\varepsilon)$-super-regular. We add all vertices of G which do not lie in some V_i' to the exceptional set V_0 and still denote this enlarged set by V_0. Thus

$$|V_0| \leq \varepsilon n + \varepsilon kL + 2a\varepsilon kL \leq 4a\varepsilon n.$$

Put $R' := R[V(H')]$. We will think of R' and of the graphs in $\mathcal{H}$ as graphs whose vertices are the new sets V_i'.

Given a vertex $v \in V_0$ and $S \in \mathcal{H}$, we say that S is *v-friendly* if there are vertices V_i' and V_j' lying in different classes of the $K_{a,a} \subseteq S$ such that v has at least $\gamma L'/4$ neighbours in both V_i' and V_j'. Let N_v denote the number of v-friendly $S \in \mathcal{H}$. Then

$$(1/2+\gamma/2)n < (1/2+\gamma-4a\varepsilon)n \le d_G(v)-|V_0| \le N_v(2a+1)L'+|\mathcal{H}|(a+1+\gamma a/4)L'$$

and therefore

$$\begin{aligned} N_v &> \frac{(1/2+\gamma/2)n}{(2a+1)L'} - \frac{k(a+1+\gamma a/4)L'}{(2a+1)^2L'} \\ &\ge \frac{n}{(2a+1)L'} \cdot \left(\frac{1}{2} + \frac{\gamma}{2} - \frac{a(1+1/a+\gamma/4)}{2a}\right) \ge \frac{n\gamma}{5L'} \cdot \frac{\gamma}{8} = \frac{\gamma^2 n}{40L'}. \end{aligned}$$

So

$$\frac{2|V_0|}{\alpha L'} \le \frac{8a\varepsilon n}{\alpha L'} < N_v$$

for every vertex $v \in V_0$. But this implies that we can successively assign each vertex $v \in V_0$ to a v-friendly $S \in \mathcal{H}$ in such a way that to every $S \in \mathcal{H}$ we assign at most $\alpha L'/2$ vertices from V_0.

Consider a fixed $S \in \mathcal{H}$ and the set $X \subseteq V_0$ of all vertices assigned to S. Let P_S be any 3-partite plane graph which satisfies the following three properties. Firstly, the classes of P_S have sizes aL', aL' and L' respectively. Secondly, $\Delta(P_S) \le 8a$ and, thirdly, P_S is a triangulation apart from $|X|$ disjoint facial 4-cycles and the vertices of each of these 4-cycles lie in the two larger vertex classes of P_S. Such plane graphs exist, see e.g. Fig. 12.2.

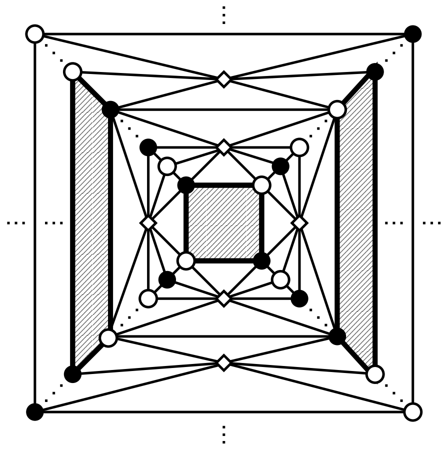


Figure 12.2: A triangulation apart from the shaded faces (into which the exceptional vertices will be inserted)

Since each edge of S corresponds to a $(2\varepsilon, d/2)$-super-regular subgraph of G', the Blow-up lemma (Lemma 12.7) implies that the subgraph of G' corresponding to S contains a spanning copy of P_S where every vertex $v \in X$ is joined to all vertices on one of the facial 4-cycles in P_S and these 4-cycles differ for distinct vertices from X. (The latter can be achieved in a similar way as in the proof of Theorem 12.2.) Thus by inserting the vertices from X into these facial 4-cycles of P_S we obtain a triangulation. Proceeding similarly for every element of $\mathcal{H}$, we obtain a spanning planar subgraph of G which is the disjoint union of $|\mathcal{H}|$ triangulations and thus has $3n - 6|\mathcal{H}| \geq 3n - 6N(\varepsilon, k_0)$ edges.

□

As a special case, the following proposition implies that the constant C in Theorem 12.3 must depend on γ and that the extra γn in the condition on the minimum degree cannot be replaced by a sublinear term.

Proposition 12.11 *For all positive integers k and n which satisfy $n/2 + k = r(2k+1)$ for some integer $r \geq 2$ there is a graph G of order n and minimum degree $n/2 + k$ which does not contain a planar subgraph with more than $3n - 6 - n/12k$ edges.*

Proof. Let G be the graph obtained from a disjoint union of r cliques $G_1, \dots, G_r$ of order $2k+1$ by adding a set Y of $n/2 - k$ new vertices and joining every vertex in Y to every vertex in $V(G_1) \cup \dots \cup V(G_r) =: X$. So G has order n and minimum degree $n/2 + k$. Consider a planar subgraph P of G with a maximum number of edges. Put $C := 3n - 6 - e(P)$. We will show that $C \geq n/12k$. Let E be a set of C edges such that $P + E$ is a triangulation, T say. Thus $E \cap E(G) = \emptyset$. Call an edge $e \in E$ *useful for G_i* if either

- e has an endvertex in G_i (and thus both endvertices of e lie in X) or
- e has both endvertices in Y and is an edge of a facial triangle of T which contains a vertex of G_i.

We claim that for every i there is an edge in E which is useful for G_i. Since a given edge from E lies in two faces of T and hence is useful for at most two cliques G_i, this would imply that

$$C = |E| \geq \frac{r}{2} = \frac{n/2 + k}{4k + 2} \geq \frac{n}{8k + 4} \geq \frac{n}{12k},$$

as desired. So fix $i \leq r$ and let us now show that there is an edge in E which is useful for G_i. Suppose not. Then every vertex of G_i lies in a facial triangle of T which is contained in G. So each such triangle contains at least one edge of G_i. We say that all these facial triangles of T are of *type I* and all other facial triangles (i.e. those which do not contain an edge of G_i) are of *type II.* So no vertex of $X - V(G_i)$ lies in a facial triangle of type I and thus there are facial triangles of type II. Since T is a triangulation, there is a path in the dual graph from a triangle of type I to a triangle of type II. Hence there is a triangle of type I which shares an edge with some triangle D of type II. But D cannot be contained in G, and so it contains an edge e from E. It is now easy to check that e is useful for G_i, a contradiction. □

12.4.4 Triangulations and Quadrangulations

The *square* G^2 of a graph G is the graph obtained from G by adding an edge between every two vertices of distance two in G. For the proof of Theorem 12.4 we will use the following result of Fan and Kierstead [33]. (It was extended to arbitrary powers of Hamilton cycles by Komlós, Sárközy and Szemerédi [54], see also [53].)

Theorem 12.12 *Every graph of minimum degree at least* $2|G|/3$ *contains the square of a Hamilton path.*

Proof of Theorem 12.4. Clearly, we may assume that $\gamma < 1/3$. Apply Proposition 12.6 to obtain $\varepsilon_0(\gamma)$ and $d_0(\gamma)$. Put $d := \min\{\gamma, d_0(\gamma)\}$. Let $\varepsilon_0(d/2, 8, 3, (d/2)^4) =: \varepsilon^*$ and $\alpha(d/2, 8, 3, (d/2)^4) =: \alpha$ be as given in the Blow-up lemma (Lemma 12.7). Set

$$\varepsilon := \min\left\{\varepsilon_0(\gamma), \frac{\varepsilon^*}{3}, \frac{\gamma\alpha}{252}, \frac{d^3}{16}\right\}.$$

and $k_0 := \max\{2/\varepsilon, 20/\gamma\}$. Throughout the proof we assume that n is sufficiently large for our estimates to hold.

Apply the Regularity lemma (Lemma 2.5) to G to obtain an exceptional set V_0 and clusters $V_1, \dots, V_k$ where $k_0 \le k \le N(\varepsilon, k_0)$. Let L and G' be as defined in the Regularity lemma. By adding at most 2 of the V_i to the exceptional set V_0 if necessary, we may assume that 3 divides k. We still denote the enlarged exceptional set by V_0. Thus $|V_0| \le \varepsilon n + 2L \le \varepsilon n + 2n/k_0 \le 2\varepsilon n$. Let R denote the reduced graph. By Proposition 12.6 we have $\delta(R) \ge (2/3 + \gamma/2)k - 2$. So Theorem 12.12 implies that R contains the square of a Hamilton path P. As $\Delta(P^2) = 4$, we may apply Proposition 2.6 to obtain adjusted clusters $V_i' \subseteq V_i$ ($i \ge 1$) of size $(1 - 4\varepsilon)L =: L'$ such that every edge of P^2 corresponds to a $(2\varepsilon, d - 5\varepsilon)$-super-regular subgraph of G'. We add all vertices that do not lie in some V_i' to the exceptional set V_0. Thus $|V_0| \le 2\varepsilon n + 4\varepsilon kL \le 6\varepsilon n$. Given a vertex $x \in R$, we will write $V'(x)$ for the adjusted cluster corresponding to x. Since $|V'(x)|, |V'(y)| \ge L/2$ for every edge $xy \in R$, it follows from the ε-regularity of the original pair that the graph $(V'(x), V'(y))_{G'}$ corresponding to xy is 2ε-regular and has density $> d - \varepsilon$.

Partition the vertices of P^2 into $k' := k/3$ disjoint sets $D_1, \dots, D_{k'}$, each containing 3 consecutive vertices of P. So the vertices in each D_i induce a triangle of P^2. For all $1 \le i < k'$ let N_i be the number of vertices of R which are joined to at least five of the six vertices in $D_i \cup D_{i+1}$. Then

$$6\delta(R) - 2e(R[D_i \cup D_{i+1}]) \le e_R(D_i \cup D_{i+1}, V(R) \setminus (D_i \cup D_{i+1})) \le 6N_i + 4|R| \tag{12.5}$$

and thus

$$N_i \ge \delta(R) - 2|R|/3 - e(R[D_i \cup D_{i+1}])/3 \ge \gamma k/2 - 2 - 5 > 0.$$

So for each $1 \le i < k'$ we can find a vertex $a_i \in R$ as well as vertices $s_i, t_i \in D_i$ and $u_{i+1}, w_{i+1} \in D_{i+1}$ with $s_i u_{i+1} \in P^2$ and such that in R each of

$s_i, t_i, u_{i+1}, w_{i+1}$ is joined to a_i. (Here the vertices a_i need not be distinct for different i.) As each edge of R corresponds to a 2ε-regular subgraph of G' of density $> d - \varepsilon$, it easily follows from repeated applications of Proposition 2.3 that there are vertices $x_i \neq y_i$ in $V'(a_i)$ such that in the graph G' their common neighbourhood in each of $V'(s_i), V'(t_i), V'(u_{i+1}), V'(w_{i+1})$ has size at least $(d-3\varepsilon)^2 L'$. Moreover, all these vertices x_i and y_i can be chosen to be distinct. Roughly speaking, the proof now proceeds as follows. We apply the Blow-up lemma to obtain for all i an (almost-) triangulation which is a spanning subgraph of the subgraph of G' corresponding to D_i. (Each exceptional vertex will also be added to one of these triangulations.) The vertices x_i and y_i will be used to 'glue together' all these triangulations into a single triangulation containing all vertices of G. In this gluing process we will also use two edges between $V'(s_i)$ and $V'(u_{i+1})$.

So let $S_i \subseteq V'(s_i)$ be any set consisting of $(d-3\varepsilon)^3 L'$ vertices which lie in the common neighbourhood of x_i and y_i but are not of the form x_j or y_j $(1 \leq j < k')$. Note that this is possible since $(d-3\varepsilon)^3 L' \leq (d-3\varepsilon)^2 L' - 2k'$. Define T_i, U_{i+1} and W_{i+1} similarly. Since we still have $|U_{i+1}| \geq 2\varepsilon L'$, we can apply Proposition 2.3 again to find a set $S'_i \subseteq S_i$ of size $(d-3\varepsilon)^4 L' \leq |S_i| - 2\varepsilon L'$ such that in G' each vertex from S'_i has at least $(d-3\varepsilon)|U_{i+1}| \geq (d-3\varepsilon)^4 L'$ neighbours in U_{i+1}.

Remove all x_i and y_i from the adjusted clusters to which they belong (but do not add them to V_0). Then the sizes of the clusters thus obtained lie between $L' - 2k'$ and L'. Set $\ell := \lfloor (L' - 2k')/4 \rfloor$. By moving a constant number of vertices into V_0 if necessary, we may assume that for all $1 \leq i \leq k'$ every cluster belonging to D_i has size $4\ell =: L''$. We still denote by $V'(x)$ the (re)-adjusted cluster corresponding to a vertex $x \in R$ and by V_0 the enlarged exceptional set. Thus $|V_0| \leq 7\varepsilon n$ and each pair of clusters in D_i still corresponds to a $(3\varepsilon, d/2)$-super-regular subgraph of G'. Furthermore, we can easily ensure that each newly adjusted cluster of the form $V'(s_i)$, $V'(t_i)$, $V'(u_i)$ or $V'(w_i)$ still contains S'_i, T_i, U_i or W_i respectively.

Let H_1, H_2 and H_3 be the 3-partite plane graphs of order $3L''$ given in Fig 12.3. So each H_i has maximum degree 8 and all of its vertex classes have size $L'' = 4\ell$. Moreover, both H_1 and H_2 are triangulations apart from two disjoint facial 4-cycles. In H_1 the vertices on these 4-cycles lie in the same two vertex classes while in H_2 one of the 4-cycles has its vertices in the first and second vertex class and the other one in the second and third vertex class. H_3 is a triangulation apart from one facial 4-cycle.

The Blow-up lemma implies that for all $1 \leq i \leq k'$ the subgraph of G' corresponding to $R[D_i]$ contains a spanning copy of each of H_1, H_2 and H_3. However, before we apply the Blow-up lemma we also have to take care of the exceptional vertices. So given a vertex $v \in V_0$ and $1 \leq i \leq k'$, we say that D_i is *v-friendly* if each of the three newly adjusted clusters in D_i contains at least $\gamma L''$ neighbours of v. Let N_v denote the number of v-friendly D_i's. Then

$$(2/3 + \gamma/2)n < d_G(v) - |V_0| \leq N_v \cdot 3L'' + k' \cdot (2+\gamma)L''.$$

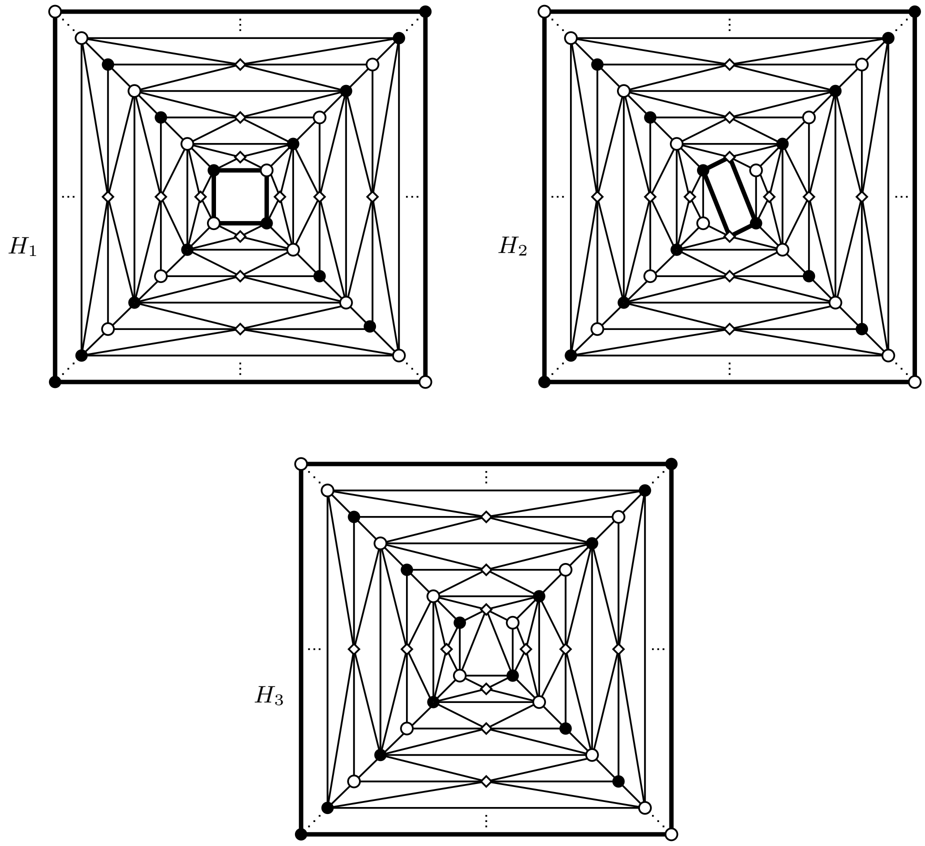


Figure 12.3: The graphs H_1, H_2 and H_3, the only non-triangular facial cycles are indicated with thick lines

Thus

$$N_v > \frac{n}{3L''}\left(\frac{2}{3}+\frac{\gamma}{2}-\frac{k(2+\gamma)L''}{3n}\right) \geq \frac{n}{3L''}\left(\frac{\gamma}{2}-\frac{\gamma}{3}\right) = \frac{\gamma n}{18L''}$$

and hence

$$\frac{2|V_0|}{\alpha L''} \leq \frac{14\varepsilon n}{\alpha L''} < N_v$$

for every $v \in V_0$. This shows that we can successively assign each exceptional vertex $v \in V_0$ to some v-friendly D_i in such a way that to each D_i we assign at most $\alpha L''/2$ vertices.

We are now ready to construct our spanning triangulation of G. We first apply the Blow-up lemma to find a spanning copy P_1 of H_3 in the subgraph of G' corresponding to $R[D_1]$ so that the vertices of the unique facial 4-cycle in P_1 lie alternately in S_1' and T_1 and so that every exceptional vertex v assigned to D_1 is joined to all vertices on some facial triangle of P_1 where these facial triangles are disjoint for distinct such vertices $v \in V_0$. (This can be done in a similar way as in the proof of Theorem 12.2 since H_3 contains at least $\alpha L''/2$ disjoint facial triangles which are also disjoint from the unique facial 4-cycle of

H_3.) Let $x_S^1, y_S^1 \in S_1'$ and $x_T^1, y_T^1 \in T_1$ be the vertices of the facial 4-cycle of P_1 and call this cycle C_{ST}^1.

For $1 < i < k'$, we now say that D_i is of *type I* if the unordered pairs s_i, t_i and u_i, w_i coincide and of *type II* if they differ. The pair s_i, t_i will be used to 'glue' the (almost-) triangulation P_i corresponding to D_i to that corresponding to D_{i+1}, whereas the pair u_i, w_i will be used to 'glue' P_i to the (almost-) triangulation corresponding to D_{i-1}. As the next step, we apply the Blow-up lemma to find a spanning copy P_2 of H_1 if D_2 is of type I, or of H_2 if it is of type II, in the subgraph of G' corresponding to $R[D_2]$ such that the vertices of one facial 4-cycle lie alternately in S_2' and T_2, the vertices of the other facial 4-cycle lie alternately in U_2 and W_2 and such that every exceptional vertex v assigned to D_2 is joined to all vertices on some facial triangle of P_2. (Again, these facial triangles are disjoint for distinct such vertices v.) Let $x_S^2, y_S^2 \in S_2'$ and $x_T^2, y_T^2 \in T_2$ be the vertices of the first facial 4-cycle C_{ST}^2 and let $x_U^2, y_U^2 \in U_2$ and $x_W^2, y_W^2 \in W_2$ be the vertices of the other facial 4-cycle C_{UW}^2. As, by definition of S_1', each of x_S^1, y_S^1 has at least $(d - 3\varepsilon)^4 L'$ neighbours in U_2, we may also require that x_U^2 is joined to x_S^1 and y_U^2 is joined to y_S^1. (To achieve this, we restrict the image of x_U^2 to the neighbourhood of x_S^1 in U_2 and the image of y_U^2 to the neighbourhood of y_S^1 in U_2.) Furthermore, by definition of S_1', T_1, U_2 and W_2, both x_1 and y_1 are joined to all vertices of C_{ST}^1 and C_{UW}^2. Thus x_1 and y_1 may be used to 'glue' P_1 and P_2 together in order to obtain a planar graph which is a triangulation apart from one facial 4-cycle, namely C_{ST}^2 (Fig. 12.4).

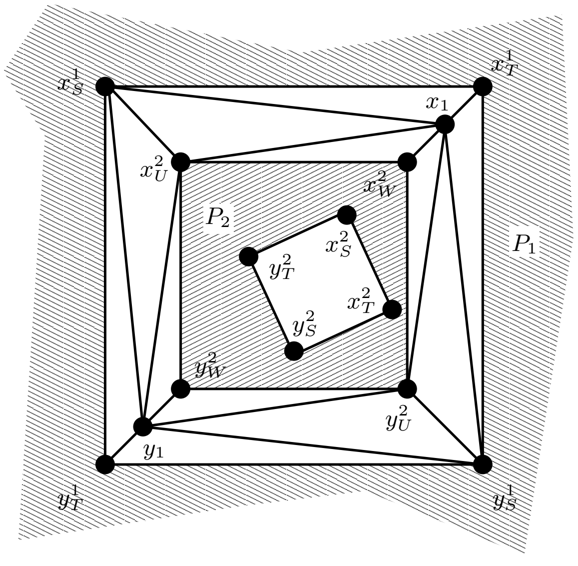


Figure 12.4: Gluing two almost-triangulations P_1 and P_2

We may continue in this fashion to obtain a spanning triangulation. Indeed, for $P_{k'}$ we again choose a copy of H_3 such that the vertices on the unique facial 4-cycle $C_{UW}^{k'}$ of $P_{k'}$ lie alternately in $U_{k'}$ and $W_{k'}$ and such that one of the two vertices from $U_{k'}$ on $C_{UW}^{k'}$ is joined to $x_S^{k'-1}$ while the other one is joined to $y_S^{k'-1}$. Thus if we glue $P_{k'}$ into the planar graph constructed in the previous step, we obtain a triangulation T. As each exceptional vertex v is joined to all vertices on some facial triangle of T and all these are distinct, we can add the exceptional vertices to T to obtain a triangulation containing all vertices of G. $\square$

Proof of Theorem 12.5 (sketch). The proof proceeds in a similar way as that of Theorem 12.4 except for a few modifications (and simplifications) which we describe below. We may now assume that the reduced graph R has even order and contains a Hamilton path P (instead of the square of a Hamilton path). We partition P into $|P|/2 := k'$ independent edges $D_1, \ldots D_{k'}$. We then adjust the clusters such that each edge D_i corresponds to a $(2\varepsilon, d-2\varepsilon)$-super-regular subgraph of G'. A calculation similar to (12.5) shows that for every pair D_i, D_{i+1} there is a vertex $a_i \in R$ which is joined to both a vertex $s_i \in D_i$ and a vertex $u_{i+1} \in D_{i+1}$. We choose two vertices $x_i, y_i \in V'(a_i)$ which have many common neighbours in both $V'(s_i)$ and $V'(u_{i+1})$. Finally, we apply the Blow-up lemma to obtain spanning quadrangulations P_i of the subgraphs of G' corresponding to the D_i which are 'glued together' into a single quadrangulation P using the vertices x_i and y_i (Fig. 12.5). These quadrangulations are chosen

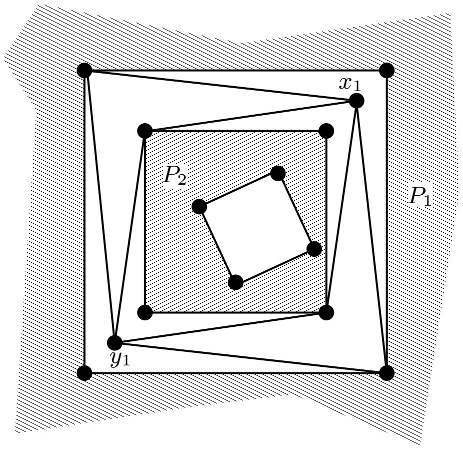


Figure 12.5: Gluing two quadrangulations P_1 and P_2

so that every exceptional vertex v is joined to two opposite vertices on some facial 4-cycle where these 4-cycles are disjoint for distinct exceptional vertices v. So all the exceptional vertices can be added to P to obtain a spanning quadrangulation of G. $\square$

As remarked towards the end of Section 12.1, the planar graphs guaranteed by Theorems 12.2–12.5 can be constructed in polynomial time: both the Regularity lemma and the Blow-up lemma can be implemented in polynomial time (see [4] and [52]). As the order of the reduced graph is constant, the remaining steps can also be carried out in polynomial time.

Chapter 13

Spanning triangulations in graphs

13.1 Introduction

The aim of this chapter is to prove that the error term γn in Theorem 12.4 can be omitted:

Theorem 13.1 *There exists an integer n_0 such that every graph G of order $n \geq n_0$ and minimum degree at least $2n/3$ contains a triangulation as a spanning subgraph.*

Our proof of Theorem 13.1 can easily be extended to obtain a spanning triangulation of an arbitrary surface (see Section 13.6).

The example before Theorem 12.4 shows that Theorem 13.1 is best possible for all integers n which are divisible by 3. For convenience of the reader, we repeat it here. Consider the graph G^* which is obtained from two disjoint cliques A and B of order $n/3$ by adding an independent set C of $n/3$ new vertices and joining each of them to all the vertices in the two cliques. So G^* has minimum degree $2n/3 - 1$. Observe that any spanning triangulation in G^* would have two facial triangles T_1 and T_2 which share an edge and are such that T_1 contains a vertex of A and T_2 contains a vertex of B. But this is impossible since every triangle of G^* containing a vertex of A (respectively B) can have at most one vertex outside A (respectively B), namely in C. In Proposition 13.17 we will extend this example slightly to show that for all n a minimum degree of $\lceil 2n/3 \rceil - 1$ does not ensure a spanning triangulation.

The spanning triangulation guaranteed by Theorem 13.1 can be found in polynomial time (see Section 13.6). In other words, the maximum planar subgraph problem (which in a given graph G asks for a planar subgraph with the maximum number of edges) can be solved in polynomial time for graphs G of minimum degree at least $2n/3$.

Our proof of Theorem 13.1 relies on Szemerédi's Regularity lemma, the Blow-up lemma of Komlós, Sárközy and Szemerédi [51] and several ideas which were introduced in [50] by the same authors. (In [50] they proved the related result that every graph of sufficiently large order n and minimum degree at least $2n/3$ contains the square of a Hamilton cycle.)

This chapter is organized as follows. In the next section we introduce the necessary definitions and tools, mainly concerning the Blow-up lemma. In Section 13.3 we give an overview of the proof. In Section 13.4 we then prove Theorem 13.1 for the case when G is rather similar to the graph G^* described in the introduction. Based on this, the general case is then dealt with in the Section 13.5. In the final section, we prove Proposition 13.17, describe how Theorem 13.1 can be extended to obtain triangulations of arbitrary surfaces and discuss some open questions.

13.2 Notation and tools

Given a vertex $x \in G$, we denote the set of its neighbours in some set $A \subseteq V(G)$ by $N_A(x)$. Given two disjoint subgraphs H and H' of G, we write $e_G(H, H')$ for $e_G(V(H), V(H'))$, i.e. for the number of all those edges of G which have one endvertex in H and the other in H'. The *common neighbourhood of* a subgraph H of G is the set of all those vertices of G which are joined to all vertices of H. Given a plane graph P, a *facial cycle* of P is a cycle which is the boundary ∂F of some face F of P. P is a *triangulation* if all its faces are bounded by triangles. So by Euler's formula a triangulation has $3n - 6$ edges. A 4-*face* of a plane graph is a face bounded by a cycle of length 4.

In our proof, we will build up the spanning triangulation from smaller fragments. For this, we introduce the following notation. Given an integer $k \geq 1$, we call a plane graph P an *almost-triangulation (with k attachment faces)* if precisely k faces of P are bounded by cycles of length at least 4 and all other faces are bounded by triangles. Each face which is not bounded by a triangle is an *attachment face*. Given almost-triangulations P and P' and attachment faces F of P and F' of P' where F' is the outer face of P', we say that a graph H is obtained by *attaching P and P'* (*via F and F'*) if H is (isomorphic to the graph) obtained by inserting P' into F and adding enough edges between ∂F and $\partial F'$ to triangulate the resulting face whose boundary is $\partial F \cup \partial F'$. Note that H is a triangulation if both P and P' had only one attachment face. (We do not allow triangular faces as attachment faces as this turns out to be inconvenient.)

One of our main tools will be the following theorem of Corrádi and Hajnal [26].

Theorem 13.2 *Every graph R whose minimum degree is at least $2|R|/3$ contains $\lfloor |R|/3 \rfloor$ disjoint triangles.*

Given $0 < \varepsilon < 1$, a graph G of order n and sets $X \subseteq A \subseteq V(G)$, we say that a bipartition A_1, A_2 of A *splits X ε-fairly* if

$$\big|\,|X \cap A_1| - |X \cap A_2|\,\big| \leq \varepsilon n.$$

We say that a bipartition A_1, A_2 of A is *ε-fair* if A as well as the sets $N_A(x)$ are split ε-fairly for every vertex $x \in G$.

Proposition 13.3 *For each $0 < \varepsilon < 1$ there exists an integer $n_0 = n_0(\varepsilon)$ such that for every graph G of order $n \geq n_0$, for every set $A \subseteq V(G)$ and for every family $\mathcal{X}$ of at most n^2 subsets of A, there exists an ε-fair bipartition of A which splits each set in $\mathcal{X}$ ε-fairly.*

Proof. Consider a random bipartition A_1, A_2 of A which is obtained by including each vertex of A into A_1 with probability $1/2$ independently of all other vertices in A. Let $\mathcal{X}'$ be the family which consists of A, all the sets in $\mathcal{X}$ as well as all sets of the form $N_A(x)$ ($x \in G$). We say that a set $X \in \mathcal{X}'$ is *bad* if it is not split ε-fairly. Let $\mathcal{X}^* \subseteq \mathcal{X}'$ be the family of all $X \in \mathcal{X}'$ with $|X| \geq \varepsilon n$. Thus no $X \in \mathcal{X}' \setminus \mathcal{X}^*$ will be bad and for each $X \in \mathcal{X}^*$ we have $n/2 \geq \mathbb{E}(|A_1 \cap X|) = |A \cap X|/2 \geq \varepsilon n/2$. Hence Lemma 3.6 implies that for each $X \in \mathcal{X}^*$

$$\mathbb{P}(X \text{ is bad}) = \mathbb{P}(|\,|A_1 \cap X| - \mathbb{E}(|A_1 \cap X|)\,| > \varepsilon n/2) \leq 2\mathrm{e}^{-\beta(\varepsilon)\varepsilon n/2}.$$

Therefore, if n is sufficiently large, the expected number of bad sets $X \in \mathcal{X}'$ is at most $|\mathcal{X}'| \cdot \mathrm{e}^{-\beta(\varepsilon)\varepsilon n/2} < 1$. So there is an outcome A_1, A_2 for which no set $X \in \mathcal{X}'$ is bad, i.e. a bipartition A_1, A_2 of A as desired. $\square$

Suppose that R is the reduced graph obtained by an application of the Regularity lemma (Lemma 2.5) to some graph G. Given a graph $H \subseteq R$, we denote by $G(H)$ the subgraph of G which is induced by all those vertices of G that lie in some cluster $V_i \in V(H)$. $G'(H)$ is defined similarly (where G' is as defined in Lemma 2.5). Given a set $X \subseteq V(R)$, we denote by $G(X)$ the set of all those vertices of G which lie in some cluster $V_i \in X$.

We will use the next well-known fact. Its simple proof is similar to that of Proposition 12.6.

Proposition 13.4 *Suppose that c, d and ε are positive numbers such that $\varepsilon \leq d/2$. Then an application of Lemma 2.5 to a graph G of minimum degree at least $c|G|$ yields a reduced graph R of minimum degree at least $(c - 2d)|R|$.*

A key ingredient of our proof will be the following special case of the Blow-up lemma of Komlós, Sárközy and Szemerédi [51]. It implies that dense 3-partite graphs G behave like complete 3-partite graphs with respect to containing bounded degree graphs as subgraphs if all the 3 bipartite graphs between the vertex classes of G are super-regular.

Lemma 13.5 (Blow-up lemma) *For all positive d, c and Δ, there are positive numbers $\varepsilon_0 = \varepsilon_0(d, \Delta, c)$ and $\alpha = \alpha(d, \Delta, c) \leq 1/2$ such that the following holds for all $0 < \varepsilon \leq \varepsilon_0$. Given $L \in \mathbb{N}$, let $K_{L,L,L}$ denote the complete 3-partite graph with vertex classes V_1, V_2 and V_3 of size L. Let G be a spanning subgraph of $K_{L,L,L}$ such that for all $1 \leq i < j \leq 3$ the graph $(V_i, V_j)_G$ is (ε, d)-super-regular. Then G contains a copy of every subgraph H of $K_{L,L,L}$ with $\Delta(H) \leq \Delta$. Furthermore, we can additionally require that for vertices $x \in H \subseteq K_{L,L,L}$ lying in V_i their images in the copy of H in G are contained in (arbitrary) given sets $C_x \subseteq V_i$ provided that $|C_x| \geq cL$ for each such x and provided that there are at most αL such vertices x.*

The vertices x in Lemma 13.5 are said to be *image restricted to C_x*.

Suppose that G is a graph as described in the Blow-up lemma. Then this lemma implies that G contains a triangulation as spanning subgraph. Indeed, all we have to show is that for each $L \in \mathbb{N}$ there is a 3-partite triangulation of bounded maximum degree with vertex classes of size L. But such triangulations can easily be constructed (see e.g. Figure 13.1 for the case when L is divisible by 4). The following simple corollary shows that much more is true.

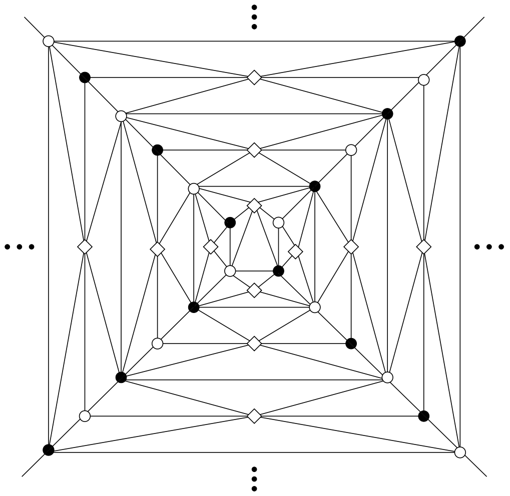


Figure 13.1: A 3-partite triangulation of maximum degree 8 with vertex classes of equal size

Corollary 13.6 *Given positive numbers d and c, there is a positive number $\varepsilon_0 = \varepsilon_0(d, c)$ and an integer $n_0 = n_0(d, c)$ such that for all $0 < \varepsilon \leq \varepsilon_0$ the following holds. Suppose that G is a 3-partite graph of order $n \geq n_0$ whose vertex classes A, B and C have the same size and such that each of the bipartite graphs $(A, B)_G$, $(A, C)_G$ and $(B, C)_G$ is (ε, d)-super-regular. Furthermore, suppose that X is a set of additional vertices (i.e. $X \cap V(G) = \emptyset$) of size at most $100\varepsilon n$ such that each vertex $x \in X$ has at least cn neighbours in at least two of the vertex classes of G. Finally, suppose that $A' \subseteq A$, $B' \subseteq B$ and $C' \subseteq C$ are sets of size at least cn. Then there exists a triangulation whose vertex set is $V(G) \cup X$ and which has a facial triangle containing a vertex from each of A', B' and C'.*

Proof. Set $\varepsilon_0 := \min\{\varepsilon_0(d, 8, c), \alpha(d, 8, c)/1300\}$, where $\varepsilon_0(d, 8, c)$ and $\alpha(d, 8, c)$ are as defined in the Blow-up lemma. Throughout the proof, we suppose that n is sufficiently large compared with ε for our estimates to hold. Let X_{AB} be the set of all those vertices in X which have at least cn neighbours in both A and B. Let $X_{BC} \subseteq X \setminus X_{AB}$ be the set of all vertices with at least cn neighbours in both B and C. Set $X_{AC} := X \setminus (X_{AB} \cup X_{BC})$. Thus each $x \in X_{AC}$ has at least cn neighbours in A and at least cn neighbours in C.

Consider any 3-partite almost-triangulation P whose vertex classes are A, B and C and which satisfies the following properties:

- $\Delta(P) \leq 8$.
- P has precisely $|X|$ attachment faces, all of which are bounded by 4-cycles. All these 4-cycles are disjoint. For $|X_{AB}|$ (respectively $|X_{BC}|$; $|X_{AC}|$) attachment faces F of P, the vertices on the boundary ∂F of F lie alternately in A and B (respectively in B and C; in A and C).
- P has a facial triangle D which is disjoint from all the facial 4-cycles of P.

Such almost-triangulations can easily be constructed, for example modify the graph in Figure 13.1. Assign each vertex $x \in X_{AB}$ to an attachment face F_x of P such that the vertices of ∂F_x lie alternately in A and B and such that the F_x are different for distinct x. Proceed similarly for all $x \in X_{BC} \cup X_{AC}$. The Blow-up lemma implies that G contains a copy of P such that the copy of D in G meets each of A', B' and C' and such that each $x \in X$ is joined to all vertices in the copy of ∂F_x in G. (To achieve the former, when applying the Blow-up lemma we only have to restrict the image of the unique vertex in $V(D) \cap A$ to A', that of the vertex in $V(D) \cap B$ to B' and that of the remaining vertex of D to C'. Similarly, for each $x \in X_{AB}$ (say) we have to restrict the images of the vertices on ∂F_x in A and B to $N_A(x)$ and $N_B(x)$ respectively.) The graph obtained from the copy of P in G by including each vertex x into F_x and joining it to all vertices on ∂F_x is a triangulation as required in the corollary. □

13.3 Overview of the proof

13.3.1 Applying the Regularity lemma and covering the reduced graph with triangles

Fix positive constants ε and d with $\varepsilon \ll d \ll 1$. We first apply the Regularity lemma (Lemma 2.5) to G to obtain an exceptional set V_0 and clusters $V_1, \ldots, V_k$ where $1/\varepsilon \leq k \leq N(\varepsilon, 1/\varepsilon)$. Let L and G' be as defined in the Regularity lemma and let R denote the reduced graph. Proposition 13.4 shows that $\delta(R) \geq (2/3 - 2d)k$. Now Theorem 13.2 implies that there exists a family $\mathcal{T}$ of disjoint triangles of R such that all but at most $13dk$ vertices of R lie in some triangle from $\mathcal{T}$. (To see this, add $6dk$ new vertices to R and join them to all other vertices. The graph R' thus obtained has minimum degree at least $2|R'|/3$. Theorem 13.2 shows that R' contains $\lfloor |R'|/3 \rfloor$ disjoint triangles. Then all but at most $12dk + 2$ vertices of R are contained in a triangle lying entirely in R.)

Delete all those vertices from R which do not lie in a triangle from $\mathcal{T}$ and still denote this graph by R. Let $T_1, \ldots, T_{k'}$ be the triangles in $\mathcal{T}$. So $|R| = 3k'$ and

$$\delta(R) \geq (2/3 - 2d)k - 13dk \geq (2/3 - 15d)|R|. \tag{13.1}$$

Clearly, we may assume that $V_1, \ldots, V_{3k'}$ are the vertices of R. We add all vertices of G which do not lie in some V_i with $1 \leq i \leq 3k'$ to the exceptional set and still denote this set by V_0. Thus $|V_0| \leq \varepsilon n + 13dkL$. Next we apply Proposition 2.4 to the union $R^* \subseteq R$ of all the triangles in $\mathcal{T}$ to obtain for each cluster V_i a subset of size $(1 - 2\varepsilon)L =: L'$ such that for all the edges

$V_iV_j \in R^*$ the bipartite subgraphs of G' between the corresponding subsets are $(2\varepsilon, d-3\varepsilon)$-super-regular. For each vertex $V_i \in R$, we add all those vertices in V_i which do not lie in the selected subset of V_i to V_0 and still denote this selected subset by V_i and the enlarged exceptional set by V_0. Thus

$$|V_0| \leq \varepsilon n + 13dkL + 2\varepsilon kL \leq 14dn. \tag{13.2}$$

Note that the bipartite subgraph of G' corresponding to an edge $V_iV_j \in R$ is still 2ε-regular of density $> d-\varepsilon$.

13.3.2 Gluing together the almost-triangulations corresponding to the triangles T_i

The Blow-up lemma immediately implies that for each triangle $T_i \in \mathcal{T}$ the corresponding subgraph $G(T_i)$ of G contains a spanning triangulation P_i. However, in order to obtain a spanning triangulation in G, we need to 'glue' all these together into a single one. As already indicated in Section 13.2, to make this 'gluing process' work, some faces of P_i (the attachment faces) will be 4-faces, i.e. P_i will only be an almost-triangulation. We will 'glue' P_i to P_{i+1} by using a very small almost-triangulation $G_i \subseteq G$ which has two attachment 4-faces F_i' and F_i''. F_i' is used to attach G_i to P_i while F_i'' is used to attach it to P_{i+1}. For each $1 \leq i < k'$ we set aside a suitable such graph G_i in advance, i.e. we remove its vertices from the clusters they belong to before applying the Blow-up lemma. In order to attach G_i to P_i, we need to make sure that there are enough edges between the boundary of one of the attachment faces of P_i, F_i say, and the boundary $\partial F_i'$ of F_i'. But if the neighbourhood of $\partial F_i'$ in $G(T_i)$ is large in some sense, then when applying the Blow-up lemma P_i can be chosen so that these ∂F_i–$\partial F_i'$ edges exist. Depending on the (common) neighbourhood of T_i and T_{i+1} in R, there are three ways in which this 'gluing process' will be done (Fig. 13.5). In Figure 13.5(a) and (c) the graph G_i is just a 4-cycle, in Figure 13.5(b) it is the graph obtained from a $K_{2,4}$ by adding two independent edges in the larger vertex class.

13.3.3 Incorporating the exceptional vertices

The preceding argument gives us a triangulation which contains all vertices of G apart from the exceptional vertices in V_0. These will be incorporated as follows. We assign each exceptional vertex $x \in V_0$ to a triangle T_i such that x has many neighbours in at least two clusters of T_i and such that not too many exceptional vertices are assigned to the same T_i.

For every $i \leq k'$ we now proceed as follows. For the first exceptional vertex x assigned to T_i, we choose a suitable almost-triangulation $Q_x \subseteq G(T_i)$ of bounded order which has two attachment 4-faces F_x and F_x'. All the vertices on the boundary of F_x will be adjacent to x, so F_x can be triangulated by inserting x. For the next exceptional vertex y assigned to T_i we choose a suitable almost-triangulation $Q_y \subseteq G(T_i) - Q_x$ with three attachment 4-faces F_y, F_y' and F_y''. Again, y will be inserted in F_y and we attach Q_y to Q_x via F_y'' and F_x'. We continue in this way until we have dealt with all the exceptional vertices assigned to T_i. In this way we obtain an almost-triangulation Q_i with one attachment 4-face.

Our aim is to apply the Blow-up lemma to find a spanning almost-triangulation in the remainder of $G(T_i)$ which has three suitable attachment 4-faces: one to attach Q_i, one to attach the 'gluing graph' G_i and one to attach G_{i-1}. (If $i = 1$ or $i = k'$ this almost-triangulation will have only two attachment 4-faces. So the graph obtained by carrying out all these attachments will be a spanning triangulation in G.) The only difficulty is that for this the remainder of $G(T_i)$ should remain super-regular. But by choosing the graphs Q_x carefully enough (using a simple probabilistic argument) the super-regularity can also be maintained.

13.3.4 Extremal case

The strategy described in Section 13.3.2 breaks down if G is very similar to the graph G^* given in Section 13.1 which had minimum degree $\lceil 2n/3 \rceil - 1$ but no spanning triangulation. In this case, the vertices of our given graph G can be partitioned into three sets A, B and C of roughly the same size such that the density of the bipartite graph $(A, B)_G$ between A and B is very small. This implies that $G[A]$ and $G[B]$ as well as the bipartite graphs $(A, C)_G$ and $(B, C)_G$ are nearly complete. (If also the maximum degree of G is not much larger than $2n/3$, then the graph $G[C]$ will be nearly empty.) In this extremal case we will find our spanning triangulation in G directly. The structure of G makes it rather easy to find a spanning subgraph consisting of two triangulations (one containing A and part of C and the other B and the remainder of C). The main difficulty is that we may have rather few edges which can be used to glue together these two triangulations into a single one.

13.4 The extremal case

Given $\varepsilon > 0$ and a partition A, B, C of the vertex set of a graph G, we say that a vertex $x \in A$ (respectively $x \in B$) is *typical* if x has at least $(1-\varepsilon)|A|$ neighbours in A (respectively at least $(1-\varepsilon)|B|$ neighbours in B) and at least $(1-\varepsilon)|C|$ neighbours in C. Similarly, we call a vertex $x \in C$ *typical* if it has at least $(1-\varepsilon)|A|$ neighbours in A and at least $(1-\varepsilon)|B|$ neighbours in B.

Definition 13.7 Given $\varepsilon > 0$, a partition A, B, C of the vertex set of a graph G is *ε-good* if it satisfies the following properties.

(i) Each vertex of C has at least $3|A|/4$ neighbours in A and at least $3|B|/4$ neighbours in B.

(ii) In each of the classes A, B and C all but at most an ε-fraction of the vertices are typical.

(iii) No vertex in $A \cup B$ has both at least $7|A|/8$ neighbours in A and at least $7|B|/8$ neighbours in B.

(iv) For every vertex $a \in A$ we have $|N_B(a)| \leq |N_A(a)| + \varepsilon n$. Similarly, every vertex $b \in B$ satisfies $|N_A(b)| \leq |N_B(b)| + \varepsilon n$.

(v) $(1-\varepsilon)n/3 \leq |A|, |B|, |C| \leq (1+\varepsilon)n/3$.

Note that if $\delta(G) \geq 2n/3$ and ε is sufficiently small, then conditions (iv) and (v) imply that each vertex $x \in A$ (respectively $x \in B$) has at least $|A|/4$ neighbours in A (respectively at least $|B|/4$ neighbours in B) and (iii) together with (v) imply that x has at least $|C|/10$ neighbours in C.

Throughout the proofs in this section we assume that ε is sufficiently small and that n is sufficiently large for our estimates to hold.

Lemma 13.8 *For each sufficiently small positive ε there exists a positive $\gamma = \gamma(\varepsilon)$ and an integer $n_0 = n_0(\varepsilon)$ such that the following holds. Suppose that G is a graph of order $n \geq n_0$ and minimum degree at least $2n/3$ whose vertex set can be partitioned into A', B' and C' such that $(1-\gamma)n/3 \leq |A'|, |B'| \leq n/3$ and such that the density of $(A', B')_G$ is at most γ. Then G has an ε-good partition.*

Proof. Set $D := 64/\varepsilon^2$ and $\gamma := 1/D^2$. We will show that the partition A', B', C' can be made into an ε-good partition by moving only a small fraction of the vertices of G into different partition classes. Let A^* be the set of all those vertices in A' which have at least $|B'|/D$ neighbours in B' and define $B^* \subseteq B'$ similarly. Since $|A^*| \cdot |B'|/D \leq e_G(A', B') \leq \gamma |A'||B'|$ and the same is true for B^*, we have

$$|A^*| \leq \frac{|A'|}{D} < \frac{\varepsilon n}{500} \quad \text{and} \quad |B^*| \leq \frac{|B'|}{D} < \frac{\varepsilon n}{500}. \tag{13.3}$$

Note that every vertex in $A' \setminus A^*$ has at least $2n/3 - n/3D$ neighbours in $A' \cup C'$ and thus at least $n/3 - n/3D \geq (1-\varepsilon/2)|C'|$ neighbours in C' and at least $n/3 - 2\gamma n/3 - n/3D \geq (1-\varepsilon/2)|A'|$ neighbours in A'. Thus in A' (respectively in B') only a small fraction of the vertices might become non-typical if we move a few of the vertices of G into different partition classes. We will now show that the same is true for C'. Let C'_A be the set of those vertices in C' which have at most $(1-\varepsilon/2)|A'|$ neighbours in A' and define $C'_B \subseteq C'$ similarly. Then

$$\begin{aligned} |C'_A| \cdot (1-\varepsilon/2)|A'| + |C' \setminus C'_A| \cdot |A'| \quad &\geq \quad e_G(A', C') \geq \left(\frac{n}{3} - \frac{n}{3D}\right) \cdot |A' \setminus A^*| \\ &\overset{(13.3)}{\geq} \quad \frac{n}{3} \cdot \left(1 - \frac{1}{D}\right)^2 |A'| \end{aligned}$$

and therefore

$$\begin{aligned} |C'_A| &\leq \frac{2}{\varepsilon}\left(|C'| - \frac{n}{3} \cdot \left(1 - \frac{2}{D}\right)\right) \leq \frac{2}{\varepsilon} \cdot \frac{n}{3}\left(2\gamma + \frac{2}{D}\right) \\ &\leq \frac{2}{\varepsilon} \cdot \frac{n}{3} \cdot \frac{2}{D} \cdot \frac{3}{2} \leq \frac{\varepsilon n}{32}. \end{aligned} \tag{13.4}$$

Similarly, $|C'_B| \leq \varepsilon n/32$. Let $C^* \subseteq C'$ be the set of all those vertices which either have less than $6|A'|/7$ neighbours in A' or less than $6|B'|/7$ neighbours in

B' (or both). (So C^* contains all vertices in C' which, when we move a few of the vertices of G into different partition classes, might come into conflict with condition (i) of Definition 13.7. Thus we will move each vertex in C^* to either A' or B'.) Note that $C^* \subseteq C'_A \cup C'_B$. Let $C^*_A \subseteq C^*$ be the set of all vertices which have more neighbours in A' than in B' and set $C^*_B := C^* \setminus C^*_A$.

Let $A^*_C \subseteq A'$ be the set of all vertices which have at least $6|A'|/7$ neighbours in A' and at least $6|B'|/7$ neighbours in B'. Note that $A^*_C \subseteq A^*$. Define $B^*_C \subseteq B^*$ similarly. (So $A^*_C \cup B^*_C$ contains all vertices which might come into conflict with condition (iii) of Definition 13.7 and thus we will move these vertices to C'.) Moreover, note that each vertex in A' which has more neighbours in B' than in A' lies in A^*. Let A^*_B be the set of all those vertices in $A^* \setminus A^*_C$ which have more neighbours in B' than in A' and define $B^*_A \subseteq B^* \setminus B^*_C$ similarly. Finally, set $A := (A' \setminus (A^*_C \cup A^*_B)) \cup C^*_A \cup B^*_A$, $B := (B' \setminus (B^*_C \cup B^*_A)) \cup C^*_B \cup A^*_B$ and $C := (C' \setminus C^*) \cup A^*_C \cup B^*_C$.

Using that all the vertices which we moved into another partition class lie in $A^* \cup B^* \cup C^*$ and that this set has size at most $\varepsilon n/15$, it is not hard to check that the partition A, B, C is ε-good. For example,

$$|C'_A \cup C'_B \cup A^*_C \cup B^*_C| \overset{(13.3),(13.4)}{\leq} \frac{\varepsilon n}{16} + \frac{\varepsilon n}{250} \leq \varepsilon |C|$$

and it is easy to see that all vertices in $C \setminus (C'_A \cup C'_B \cup A^*_C \cup B^*_C)$ have at least $(1-\varepsilon)|A|$ neighbours in A and at least $(1-\varepsilon)|B|$ neighbours in B. So all but at most an ε-fraction of the vertices in C are typical. $\square$

Lemma 13.9 *For each sufficiently small positive ε there exists an integer $n_0 = n_0(\varepsilon)$ such that the following holds. Suppose that G is a graph of order $n \geq n_0$ and minimum degree at least $2n/3$ which has an ε-good partition A, B, C for which there are two independent A–B edges a_1b_1 and a_2b_2 such that each a_i is typical and such that each b_i has at most $|A|/3$ neighbours in A. Then G contains a triangulation as spanning subgraph.*

Proof. The fact that $\delta(G) \geq 2n/3$ and condition (v) of Definition 13.7 imply that each b_i has at least $7|C|/12$ neighbours in C. Since both a_1 and a_2 are typical, this implies that there are two typical vertices $c_1, c_2 \in C$ which are both joined to each of a_1, a_2, b_1, b_2. Since each of a_1, a_2, c_1, c_2 is typical, their common neighbourhood in A has size at least $(1-4\varepsilon)|A|$ and thus contains two typical vertices a' and a''. Similarly, since each b_i has at least $7|B|/12$ neighbours in B, the common neighbourhood of b_1, b_2, c_1, c_2 contains two typical vertices b' and b''. Set $A' := A \setminus \{a_1, a_2, a', a''\}$, $B' := B \setminus \{b_1, b_2, b', b''\}$ and $C' := C \setminus \{c_1, c_2\}$. Apply Proposition 13.3 to obtain an ε-fair partition of A' into A'_1 and A'_2, an ε-fair partition of B' into B'_1 and B'_2 and an ε-fair partition of C' into C'_A and C'_B such that firstly, the common neighbourhood of the 4-cycle $a_1a'a_2a'' =: D_a$ in A' as well as that in C' are split ε-fairly and secondly, such that the common neighbourhood of $b_1b'b_2b'' =: D_b$ in B' as well as that in C' are split ε-fairly.

By removing a set X_A of at most $10\varepsilon n$ vertices from $A_1 \cup A_2 \cup C_A$ and a set X_B of at most $10\varepsilon n$ vertices from $B_1 \cup B_2 \cup C_B$, we may assume that

$|A_1| = |A_2| = |C_A|$, $|B_1| = |B_2| = |C_B|$ and that all these six sets consist only of typical vertices. Using this, it is easy to see that each of the bipartite graphs $(A_1, A_2)_G$, $(A_1, C_A)_G$ and $(A_2, C_A)_G$ is $(10\varepsilon, 1/2)$-super-regular. Condition (i) of Definition 13.7 together with the remark after Definition 13.7 imply that for each $x \in X_A$ its neighbourhood $N_{A'}(x)$ in A' has size at least $|A'|/5$. Since $N_{A'}(x)$ was split ε-fairly, it follows that each vertex $x \in X_A$ has at least $|A_i|/10$ neighbours in A_i $(i = 1, 2)$. Moreover, also the common neighbourhoods of D_a in A_1, A_2 and C_A have sizes at least $|A_1|/10$, $|A_2|/10$ and $|C_A|/10$ respectively. Thus if $\varepsilon \leq \varepsilon_0(1/2, 1/30)/10$ and $n \geq 4n_0(1/2, 1/30)$, where ε_0 and n_0 are as defined in Corollary 13.6, then Corollary 13.6 implies that $G[A_1 \cup A_2 \cup C_A \cup X_A]$ contains a spanning triangulation P_A which has a facial triangle that lies in the common neighbourhood of D_a. Similarly, it follows that $G[B_1 \cup B_2 \cup C_B \cup X_B]$ contains a spanning triangulation with one facial triangle lying in the common neighbourhood of D_b. P_A and P_B can be 'glued' together into a spanning triangulation of G by using c_1 and c_2 and the 4-cycles D_a and D_b as well as suitable edges (Fig. 13.2). $\square$

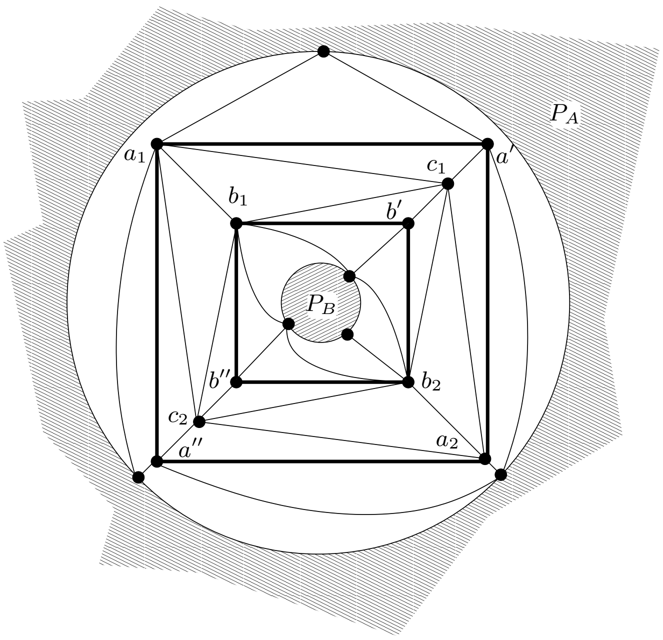


Figure 13.2: Gluing P_A and P_B together into a spanning triangulation in the proof of Lemma 13.9

Lemma 13.10 *For each sufficiently small positive ε there exists an integer $n_0 = n_0(\varepsilon)$ such that the following holds. Suppose that G is a graph of order $n \geq n_0$ and minimum degree at least $2n/3$ which has an ε-good partition A, B, C for which there are vertices $b_1, b_2 \in B$ and typical vertices $c_1, c_2 \in C$ such that each c_i is joined to both b_1 and b_2 and each b_i has at least $4\varepsilon|A|$ neighbours in A. Then G contains a triangulation as spanning subgraph.*

Proof. Since both c_1 and c_2 are typical, their common neighbourhood in A has size at least $(1-2\varepsilon)|A|$. Thus there are typical vertices a_1 and a_2 such that a_1 is joined to each of b_1, c_1, c_2 while a_2 is joined to each of b_2, c_1, c_2. Now one can proceed in a similar way as in the proof of Lemma 13.9. First find two typical vertices $a', a'' \in A$ and six typical vertices $b_1', b', b_2', b_1'', b'', b_2'' \in B$ such that $a_1 a' a_2 a'' =: D_a$ is a 4-cycle lying in the common neighbourhood of c_1 and c_2 and such that $b_1 b_1' b' b_2' b_2 b_2'' b'' b_1'' =: D_b$ is an 8-cycle in the common neighbourhood of c_1 and c_2 (since b_1 and b_2 might have no common neighbours in B, we cannot take D_b to be a 4-cycle anymore). Similarly as in the proof of Lemma 13.9, choose disjoint triangulations P_A and P_B such that, firstly, $V(P_A) \cup V(P_B) = V(G) \setminus (V(D_a \cup D_b) \cup \{c_1, c_2\})$, secondly, one of the facial triangles of P_A lies in the common neighbourhood of D_a and thirdly, P_B has a facial triangle which lies in the common neighbourhood of $D_b - \{b_1, b_2\}$ and meets both $N_B(b_1)$ and $N_B(b_2)$. (Since b_1 and b_2 might have no common neighbours in B, this time we cannot guarantee a facial triangle of P_B that lies in the common neighbourhood of the entire cycle D_b.) As c_1 and c_2 are joined to all vertices on D_a and D_b, they can be used to 'glue' P_A and P_B together to obtain a spanning triangulation in G (Fig. 13.3). $\square$

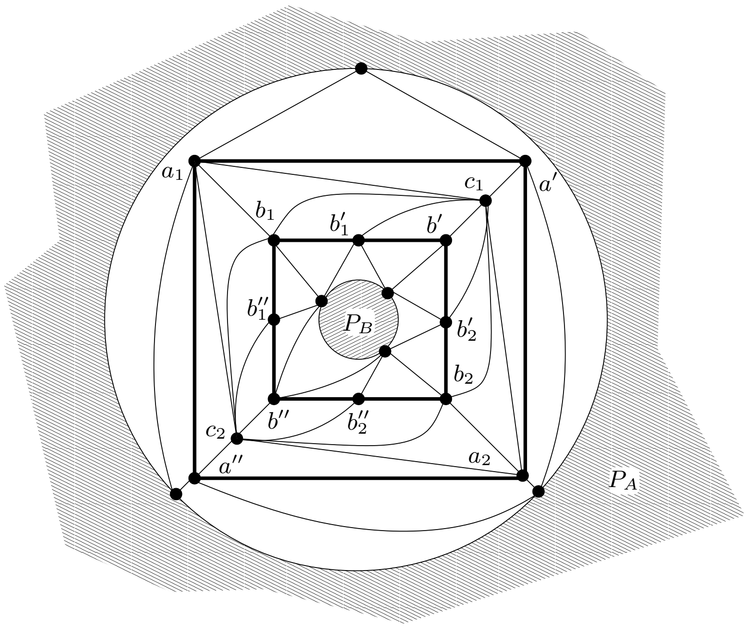


Figure 13.3: Gluing P_A and P_B together into a spanning triangulation in the proof of Lemma 13.10

Corollary 13.11 *For each sufficiently small positive ε there exists an integer $n_0 = n_0(\varepsilon)$ such that the following holds. Suppose that G is a graph of order $n \geq n_0$ and minimum degree at least $2n/3$ which has an ε-good partition A, B, C for which there are distinct vertices $b_1, b_2, \ldots, b_{11} \in B$ such that each b_i has* *at least $4\varepsilon|A|$ neighbours in A. Then G contains a triangulation as spanning subgraph.*

Proof. As remarked after Definition 13.7, each b_i has at least $|C|/10$ neighbours in C. Hence an easy calculation shows that the must be two of the b_i whose common neighbourhood in C contains at least two typical vertices. So we may assume that there are two typical vertices $c_1, c_2 \in C$ which are joined to both b_1 and b_2. Now the corollary follows immediately from Lemma 13.10. $\square$

Proposition 13.12 *Given $0 < \varepsilon \leq 1$, every graph H of average degree at least $2 + \varepsilon$ contains a cycle of length at most $6(\log_2 |H|)/\varepsilon$.*

Proof. Note that the average degree of H will not decrease below $2 + \varepsilon$ by deleting a vertex of degree one or by deleting all the vertices on a path of length at least $\lceil 2/\varepsilon \rceil - 1 =: \ell$ whose vertices all have degree two in H. So by considering a subgraph of H if necessary, we may assume that the minimum degree of H is at least 2 and that H does not contain paths of length at least ℓ all of whose vertices have degree 2 in H.

Since we may assume that H does not contain a cycle of length at most $\ell + 2$, this implies that each vertex x of degree 2 in H lies on a unique path P_x of length at most $\ell + 1$ such that all the inner vertices of P_x, but not its endvertices, have degree 2 in H and such that the endvertices of P_x are not joined by an edge. (So P_x is induced.) Moreover, since we may assume that H does not contain a cycle of length at most $2\ell + 2$, two such paths P_x and P_y either coincide or meet at most in one of their endpoints. Thus the graph H' obtained from H by replacing each such P_x by an edge has minimum degree at least 3 (and no multiple edges).

It is easy to see that H' contains a cycle C of length at most $2\log_2 |H'| \leq 2\log_2 |H|$. As each edge of C either lies in H or corresponds to some P_x, C corresponds to a cycle in H whose length is at most $(\ell + 1)|C|$. $\square$

Lemma 13.13 *For each sufficiently small positive ε there exists an integer $n_0 = n_0(\varepsilon)$ such that the following holds. Suppose that G is a graph of order $n \geq n_0$ and minimum degree at least $2n/3$ which has an ε-good partition A, B, C for which $G[C]$ contains a cycle D of length at most $(\log_2 n)^2$. Then G contains a triangulation as spanning subgraph.*

Proof. Recall that by condition (i) in Definition 13.7 the common neighbourhood in A of up to 2 vertices on D has size at least $|A|/2$. Construct a cycle $D_1 = a_1 a_2 \dots a_{|D|}$ of length $|D|$ which consists of typical vertices from A as follows. Let a_1 be a typical vertex in A which is joined to both d_1 and $d_{|D|}$. For a_2 take a typical vertex in A which is joined to each of a_1, d_1 and d_2 and for a_3 take one which is joined to each of a_2, d_2 and d_3. The last vertex $a_{|D|}$ of D_1 will be a typical vertex in A which is joined to $a_1, a_{|D|-1}, d_{|D|-1}$ and $d_{|D|}$. So the subgraph of G induced by $V(D \cup D_1)$ contains an almost-triangulation with precisely two attachment faces, bounded by D and D_1. Similarly, as now all vertices on D_1 are typical and so every 10 (say) consecutive vertices on D_1

have a large common neighbourhood in A, there is a cycle D_2 of length at most $|D_1|/2$ consisting of typical vertices in A such that the subgraph of G induced by $V(D_1 \cup D_2)$ contains an almost-triangulation with precisely two attachment faces, bounded by D_1 and D_2. Continuing in this fashion, we may assume that for some $i \le \log_2((\log_2 n)^2)$ the cycle D_i is a triangle. Thus G contains an almost-triangulation P'_A whose vertex set is $V(D) \cup V(D_1) \ldots V(D_i)$ and which has only one attachment face, bounded by D. Note that $|P'_A| \le (\log_2 n)^3$. Similarly, G contains an almost-triangulation P'_B of order at most $(\log_2 n)^3$ such that all vertices of P'_B either lie on D or are typical vertices from B and such that P'_B has only one attachment face, bounded by D. So $P' := P'_A \cup P'_B$ is a triangulation.

Choose a facial triangle D'_A of P' consisting of typical vertices from A and a facial triangle D'_B consisting of typical vertices from B. Similarly as in the proof of Lemma 13.9, one can now find disjoint triangulations P_A and P_B with $V(P_A) \cup V(P_B) = V(G - P')$ and such that one facial triangle D_A of P_A lies in the common neighbourhood of the triangle D'_A while one facial triangle D_B of P_B lies in the common neighbourhood of D'_B. Thus the graph obtained from $P_A \cup P_B \cup P'$ by adding suitable D_A–D'_A edges and suitable D_B–D'_B edges is a spanning triangulation in G. □

The next corollary is an immediate consequence of Proposition 13.12 and Lemma 13.13.

Corollary 13.14 *For each sufficiently small positive ε there exists an integer $n_0 = n_0(\varepsilon)$ such that the following holds. Suppose that G is a graph of order $n \ge n_0$ and minimum degree at least $2n/3$ which has an ε-good partition A, B, C for which the average degree of $G[C]$ is at least $2 + \varepsilon$. Then G contains a triangulation as spanning subgraph.* □

Lemma 13.15 *For each sufficiently small positive ε there exists an integer $n_0 = n_0(\varepsilon)$ such that the following holds. Suppose that G is a graph of order $n \ge n_0$ and minimum degree at least $2n/3$ which has an ε-good partition A, B, C for which the minimum degree $G[C]$ is at least 2 and for which at least $4\varepsilon|A|$ vertices in A and at least $4\varepsilon|B|$ vertices in B are joined to all vertices in C. Then G contains a triangulation as spanning subgraph.*

Proof. Condition (ii) of Definition 13.7 implies that there is a triangle D'_A that consists of typical vertices in A which are joined to all vertices in C. Similarly, there is a triangle D'_B that consists of typical vertices in B which are joined to all vertices in C. Let D'_C be a cycle in $G[C]$. Choose disjoint triangulations P_A and P_B such that $V(P_A) \cup V(P_B) = V(G - D'_A - D'_B - D'_C)$ and such that one facial triangle D_A of P_A lies in the common neighbourhood of D'_A while one facial triangle D_B of P_B lies in the common neighbourhood of D'_B. The existence of such triangulations can be shown similarly as in the proof of Lemma 13.9. The only difference is that D'_C might no longer be small compared to C. So if $|D'_C| \ge 10\varepsilon n$ (say), we first choose an ε-fair partition of $C \setminus V(D'_C)$ into C_A and C_B. But instead of doing the same for A and B (as we did in the proof of Lemma 13.9), we now proceed a little differently: the proofs of Proposition 13.3 and Lemma 13.9 immediately show that we can partition $A \setminus V(D'_A)$ into three sets A_0, A_1 and A_2 such that $A_0 \cup C_A$, A_1 and A_2 have roughly the same size and each of the three bipartite graphs between these sets can be made super-regular by removing all vertices which are not typical. Similarly, we split $B \setminus V(D'_B)$ into three sets B_0, B_1 and B_2. Having found P_A and P_B, we can extend $P_A \cup P_B \cup D'_A \cup D'_B \cup D'_C$ to a spanning triangulation of G by adding suitable edges since all vertices of D'_A and D'_B are joined to every vertex on D'_C. □

Lemma 13.16 *For each sufficiently small positive γ there exists an integer $n_0 = n_0(\gamma)$ such that the following holds. Suppose that G is a graph of order $n \geq n_0$ and minimum degree at least $2n/3$ whose vertex set can be partitioned into A', B', C' such that $(1-\gamma)n/3 \leq |A'|, |B'| \leq n/3$ and such that the density of $(A', B')_G$ is at most γ. Then G contains a triangulation as spanning subgraph.*

Proof. Let ε be sufficiently small to satisfy the requirements of all the previous results in this section. We may assume that $\gamma \leq \gamma(\varepsilon)$ where $\gamma(\varepsilon)$ is as defined in Lemma 13.8. Apply this lemma to obtain an ε-good partition A, B, C of $V(G)$.

Case 1 $|A| \geq \lfloor n/3 \rfloor$ or $|B| \geq \lfloor n/3 \rfloor$.

We only consider the case when $|B| \geq \lfloor n/3 \rfloor$, the other case is similar. So set $x := |A| - \lfloor n/3 \rfloor$ and $y := |B| - \lfloor n/3 \rfloor$. Thus, by our assumption, $y \geq 0$. Since $\delta(G) \geq 2n/3$, it follows that each vertex in A sends at least

$$\delta_A := \left\lceil \frac{2n}{3} \right\rceil - (|A| - 1 + |C|) = |B| - \lfloor n/3 \rfloor + 1 = y + 1 \geq 1$$

edges to B and that each vertex in B sends at least $\max\{x+1, 0\} =: \delta_B$ edges to A. We call a vertex $a \in A$ *rich* if it has at least $|B|/10$ neighbours in B. Similarly, a vertex $b \in B$ is *rich* if it has at least $|A|/10$ neighbours in A. Let $R_A \subseteq A$ and $R_B \subseteq B$ be the sets of rich vertices in A and B.

Assume first that $|R_B| \leq \delta_A$. For every vertex $a \in A$ choose δ_A edges joining a to B. Note that there are at least $|A|/9$ vertices in A which send at least one of their chosen edges to $B \setminus R_B$. Indeed, suppose that this is not the case. Then $|R_B| = \delta_A$ and each vertex in R_B is joined to all but at most $|A|/9$ vertices in A, and thus, by condition (iv) in Definition 13.7, to at least $8|B|/9 - 2\varepsilon n > 7|B|/8$ vertices of B, which contradicts condition (iii). So there are at least $|A|/10$ typical vertices in A which send (at least) one of their chosen edges to $B \setminus R_B$. These chosen edges cannot all have the same endvertex in B, since this would be rich. Therefore, there are two independent A–B edges whose endvertices in A are typical and whose endvertices in B are not rich and thus have at most $|A|/10$ neighbours in A. So Lemma 13.9 implies that G contains a spanning triangulation.

Thus we may assume that $|R_B| > \delta_A$. On the other hand, Corollary 13.11 shows that we may assume that $|R_B| \leq 10$. Moreover, Lemma 13.10 implies that we may assume that every two typical vertices in C have at most one common neighbour in R_B. Thus all but at most $\binom{|R_B|}{2} \leq 45$ typical vertices in C have at most one neighbour in R_B. Similarly, if $x \geq 0$ and thus $\delta_B \geq 1$, we may assume that $\delta_B < |R_A| \leq 10$ and that all but at most 45 typical vertices of C have at most one neighbour in R_A.

Suppose that $x \geq 0$ and consider a (typical) vertex $c \in C$ which is joined to at most one vertex in R_B and to at most one vertex in R_A. Then

$$\left\lceil \frac{2n}{3} \right\rceil \leq d_G(c) \leq |N_C(c)| + (|A| - |R_A| + 1) + (|B| - |R_B| + 1)$$
$$\leq |N_C(c)| + \left\lfloor \frac{n}{3} \right\rfloor + x - (x+1) + \left\lfloor \frac{n}{3} \right\rfloor + y - (y+1) \tag{13.5}$$

and therefore $|N_C(c)| \geq 2$. The same calculation shows that if $x < 0$ then every (typical) vertex $c \in C$ which is joined to at most one vertex in R_B must have degree at least two in $G[C]$. Thus in both cases all but at most 90 typical vertices in C have degree at least two in $G[C]$ and therefore at most $\varepsilon|C| + 90$ vertices in C have degree less than two in $G[C]$. Since by Corollary 13.14 we may assume that the average degree of $G[C]$ is less than $2 + \varepsilon$, this implies that at most $3\varepsilon|C| + 180$ vertices in C have degree at least 3 in $G[C]$. (Note that if n is not divisible by 3, then the above calculation already shows that all but at most $\varepsilon|C| + 90$ vertices in C have degree at least 3 in $G[C]$; and thus we are done by Corollary 13.14 in that case. However, in what follows we will not make use of this fact.)

Call a typical vertex $c \in C$ *useful* if its degree in $G[C]$ is two and if c has at most one neighbour in R_B and, in the case when $x \geq 0$, if in addition c has at most one neighbour in R_A. Thus all but at most $4\varepsilon|C| + 270 \leq 5\varepsilon|C|$ vertices in C are useful. Furthermore, note that inequality (13.5) (or its analogue for the case when $x < 0$) shows that a useful vertex c must be joined to exactly one vertex in R_B.

Let us now show that $G[C]$ either contains a cycle of length at most 20 (in which case we are done by Lemma 13.13) or $G[C]$ contains a nontrivial path $Q = c_1 \dots c_2$ of length less than 20 whose endvertices c_1 and c_2 are useful and are joined to the same vertex $b \in R_B$. Indeed, consider a family $\mathcal{P}$ of disjoint subpaths of $G[C]$ which is obtained as follows. For the first path $P_1 \in \mathcal{P}$ take a longest path in $G[C]$. For the second path $P_2 \in \mathcal{P}$ take a longest path in $G[C] - P_1$. The third path $P_3 \in \mathcal{P}$ is a longest path in $G[C] - P_1 - P_2$. Continue in this fashion until all vertices of $G[C]$ lie on some path in $\mathcal{P}$. Consider first a path $P \in \mathcal{P}$ of length at least 20 (if such a path exists). If P does not contain a subpath Q with the required properties, then at most $2|P|/3$ vertices on P can be useful. (To see this, use that $|R_B| \leq 10$ and divide P into consecutive subpaths on 20 vertices and a leftover path.) As at most $5\varepsilon|C|$ vertices in C are not useful, this implies that at most $10\varepsilon|C|$ useful vertices from C are covered by paths in $\mathcal{P}$ of length at least 20, i.e. at least $(1 - 15\varepsilon)|C|$ useful vertices from C are covered by a path in $\mathcal{P}$ of length less than 20. If no such path can be taken for Q, then all these paths contain at most 10 useful vertices. Thus there are at least $(1-15\varepsilon)|C|/10$ paths of length less than 20 in $\mathcal{P}$. Let $\mathcal{P}'$ denote the set of all these paths. By construction of $\mathcal{P}$, each endvertex of a path $P \in \mathcal{P}'$ either has degree one in $G[C]$ or has a second neighbour on P (in which case $G[C]$ contains a cycle of length at most 20, as required) or it is joined to a vertex in C of degree at least 3 (this vertex then lies in the interior of a path from $\mathcal{P}$ that was chosen before P). Since at most $\varepsilon|C| + 90$ vertices in C have degree at most one, we may assume that at least $|C|/20$ paths $P \in \mathcal{P}'$ have an endvertex which is joined to some vertex in C of degree at least 3. Thus at least $|C|/20$ edges of $G[C]$ are incident to some vertex in C of degree at least 3, which contradicts the fact that

$$\begin{aligned}\sum_{c\in C:\,|N_C(c)|\geq 3} |N_C(c)| &\leq 2e(G[C]) - \sum_{c\in C:\,|N_C(c)|=2} 2 \\ &\leq (2+\varepsilon)|C| - 2(1-5\varepsilon)|C| = 11\varepsilon|C|.\end{aligned}$$

So we may assume that there exists a nontrivial path $Q = c_1 \dots c_2$ of length less than 20 whose endvertices c_1 and c_2 are useful and are joined to the same vertex $b \in R_B$. As $|N_A(b)| \geq |A|/10$ and both c_1 and c_2 are useful (and thus typical), there is a typical vertex $a \in A$ which is joined to each of b, c_1 and c_2. Similarly as in the proof of Lemma 13.13, one can now find planar graphs P_A and P_B satisfying the following properties. Both P_A and P_B are almost-triangulations having only one attachment face. The attachment face of P_A is bounded by the cycle ac_1Qc_2 and that of P_B is bounded by bc_1Qc_2. The vertex set of P_A consists of A, the vertices on Q and of about half of the vertices in $C \setminus V(Q)$. The vertex set of B consists of B, the vertices on Q and of the remainder of $C \setminus V(Q)$. Thus adding the edge ab to $P_A \cup P_B$ yields a spanning triangulation in G (Fig. 13.4).

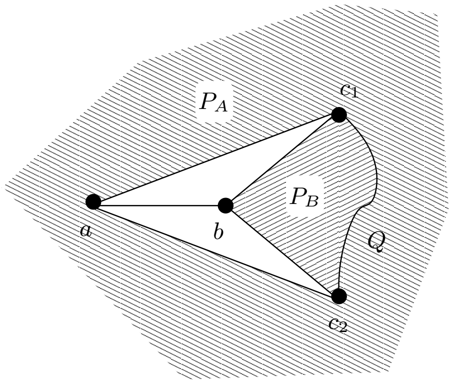


Figure 13.4: A spanning triangulation obtained by adding the edge ab

Case 2 $|A| < \lfloor n/3 \rfloor$ and $|B| < \lfloor n/3 \rfloor$.

Set $x := \lfloor n/3 \rfloor - |A|$ and $y := \lfloor n/3 \rfloor - |B|$. Thus $x, y \geq 1$. Since $\delta(G) \geq 2n/3$, the minimum degree of $G[C]$ is at least $x + y \geq 2$. Thus Corollary 13.14 shows that we may assume that $x = y = 1$.

Case 2.1 All but at most $4\varepsilon|A|$ vertices in A have at least one neighbour in B or all but at most $4\varepsilon|B|$ vertices in B have at least one neighbour in A.

Suppose that the former holds, the other case is similar. As in Case 1, call a vertex $b \in B$ *rich* if it has at least $|A|/10$ neighbours in A and denote the set of all rich vertices in B by R_B. Again, it follows that we may assume that

$2 \leq |R_B| \leq 10$ and that all but at most 45 typical vertices $c \in C$ have at most one neighbour in R_B. But for each such vertex c we have

$$\left\lceil \frac{2n}{3} \right\rceil \leq d_G(c) \leq |N_C(c)| + |A| + |B| - 1 = |N_C(c)| + 2\left\lfloor \frac{n}{3} \right\rfloor - x - y - 1$$

and so c has at least 3 neighbours in C. This shows that all but at most $\varepsilon|C| + 45$ vertices of C have degree at least 3 in $G[C]$ and so we are done by Corollary 13.14.

Case 2.2 At least $4\varepsilon|A|$ vertices in A have no neighbours in B and at least $4\varepsilon|B|$ vertices in B have no neighbours in A.

Since $\delta(G) \geq 2n/3$ and $x = y = 1$, each of these vertices is joined to all vertices in C. So we are done by Lemma 13.15. □

13.5 Proof of Theorem 13.1

13.5.1 Applying the Regularity lemma and covering the reduced graph with triangles

Throughout the proof we assume that the order n of G is sufficiently large for our estimates to hold. Moreover, we will assume that ε, c and d are sufficiently small positive constants such that $\varepsilon \ll d \ll c$. With these parameters ε and d, we proceed exactly as described in Section 13.3.1 to obtain a set $\mathcal{T} = \{T_1, \dots, T_{k'}\}$ of triangles covering the (modified) reduced graph R. Thus every edge $V_iV_j \in R$ corresponds to a 2ε-regular bipartite subgraph of G' which has density $> d - \varepsilon$. Moreover, if V_iV_j lies in some triangle from $\mathcal{T}$, then this subgraph is also $(2\varepsilon, d - 3\varepsilon)$-super-regular.

13.5.2 Choosing the gluing graphs G_i

As outlined in Section 13.3.2, for all $1 \leq i < k'$ in turn we now choose the graphs G_i that will be used later to glue the almost-triangulation corresponding to the triangle $T_i \in \mathcal{T}$ to that corresponding to T_{i+1}. Given $1 \leq i < k'$, we assign to every vertex $W \in R$ the label $L_i(W) := (x, y)$ where $x := |N_{V(T_i)}(W)|$ and $y := |N_{V(T_{i+1})}(W)|$. So $0 \leq x, y \leq 3$. We need to distinguish the following cases.

Case 1 There exists a vertex $W \in R - V(T_i \cup T_{i+1})$ with label $L_i(W) = (x, y)$ such that $x, y \geq 2$.

Choose two neighbours A_i and B_i of W in T_i and two neighbours I_{i+1} and J_{i+1} of W in T_{i+1}. Let X_i denote the common neighbourhood of A_i and W in R. So $|X_i| \geq (1/3 - 30d)|R|$ by (13.1). By making X_i smaller if necessary, we may assume that $|X_i| \leq |R|/3$. Similarly, we choose a set Y_{i+1} in the common neighbourhood of I_{i+1} and W with $(1/3 - 30d)|R| \leq Y_{i+1} \leq |R|/3$.

Case 1.1 X_i and Y_{i+1} are not disjoint.

In this case G_i will be a 4-cycle. Choose a vertex W' in $X_i \cap Y_{i+1}$. Take G_i to be any 4-cycle $v_iv_i'w_iw_i'$ in G with $v_i, w_i \in W$ and $v_i', w_i' \in W'$ and such that the vertices of G_i have at least $(d - 3\varepsilon)^4 L'$ common neighbours in both A_i and I_{i+1} and such that both v_i and w_i have at least $(d - 3\varepsilon)L'$ neighbours in B_i and at least $(d - 3\varepsilon)L'$ neighbours in J_{i+1}. (The existence of such vertices follows from repeated applications of Proposition 2.3 since each edge of R corresponds to a 2ε-regular subgraph of G' of density $> d - \varepsilon$.) So G_i is a (rather degenerate) almost-triangulation whose attachment faces F_i' and F_i'' are the interior and the exterior of G_i. Recall from Section 13.3.2 that we wish to apply the Blow-up lemma to find an almost-triangulation P_i in $G(T_i)$ which can be attached to G_i via an attachment 4-face F_i of P_i and F_i'. The above conditions on the neighbourhood of G_i ensure that this can be done as indicated in Figure 13.5(a) if we restrict the images of two opposite vertices of ∂F_i to the common neighbourhood of G_i in A_i and the images of its remaining two opposite vertices to the neighbourhoods of v_i and w_i in B_i. When applying the

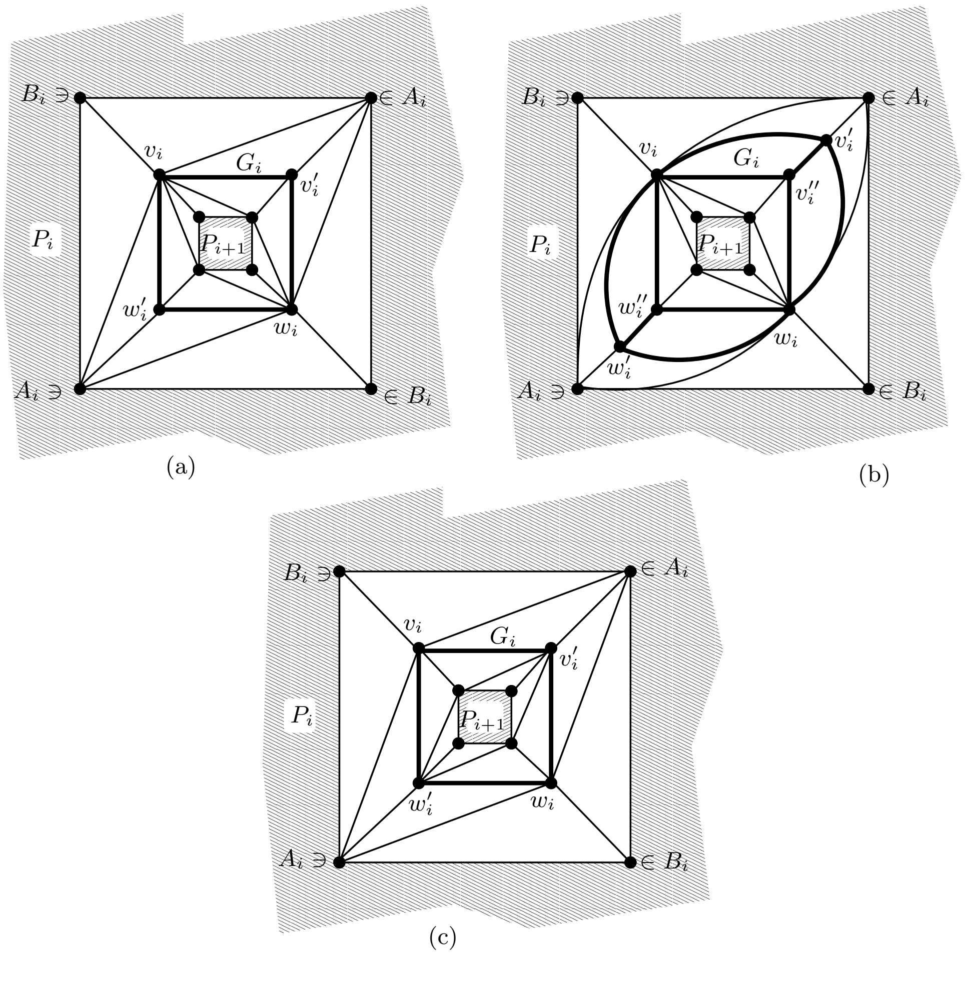


Figure 13.5: Three ways of gluing together the triangulations P_i and P_{i+1} by using a gluing graph G_i. The graphs G_i are indicated by thick lines.

Blow-up lemma, the latter can always be guaranteed since by our choice of G_i all these neighbourhoods are sufficiently large.

Case 1.2 X_i and Y_{i+1} are disjoint.

In this case G_i will be the graph obtained from a $K_{2,4}$ by adding two independent edges in the larger vertex class. Note that, since $X_i \cap Y_{i+1} = \emptyset$, Lemma 13.16 implies that we may assume that there exists an X_i–Y_{i+1} edge $W'W''$ in R. Indeed, if there is no such edge, then the bipartite subgraph of G' between the sets $G(X_i)$ and $G(Y_{i+1})$ is empty. Let γ be the density of the bipartite subgraph of G between $G(X_i)$ and $G(Y_{i+1})$. Then it is easily checked that $G(X_i)$ contains a vertex x with $d_G(x) \geq \gamma n/4$. Since $0 = d_{G'}(x) > d_G(x) - (d+\varepsilon)n$, this implies that $\gamma \leq 4(d+\varepsilon)$. Moreover, the sizes of both $G(X_i)$ and $G(Y_{i+1})$ lie between $(1-150d)n/3$ and $n/3$ and so if d is sufficiently small we are done by Lemma 13.16.

Choose vertices $v_i, w_i \in W$, $v'_i, w'_i \in W'$ and $v''_i, w''_i \in W''$ such that both v_i and w_i are joined to each of v'_i, v''_i, w'_i, w''_i and such that v'_i is joined to v''_i and w'_i is joined to w''_i. Furthermore, we choose these vertices so that v_i and w_i have at least $(d-3\varepsilon)L'$ neighbours in both B_i and J_{i+1}, v_i, w_i, v'_i, w'_i have at least $(d-3\varepsilon)^4 L'$ common neighbours in A_i and so that v_i, w_i, v''_i, w''_i have at least $(d-3\varepsilon)^4 L'$ common neighbours in I_{i+1}. (Again, the existence of such vertices follows from repeated applications of Proposition 2.3.) So v_i and w_j form the vertex class of size two of a $K_{2,4}$ in G_i and the independent edges added to the larger vertex class of this $K_{2,4}$ are $v'_i v''_i$ and $w'_i w''_i$. Thus G_i can be viewed as an almost-triangulation whose two attachment faces F'_i and F''_i are bounded by the 4-cycles $v_i v'_i w_i w'_i$ and $v_i v''_i w_i w''_i$ respectively. Figure 13.5(b) indicates how G_i will be used in Case 1.2 to glue together the almost triangulations P_i and P_{i+1} corresponding to T_i and T_{i+1} respectively.

Case 2 All vertices W in $R - V(T_i \cup T_{i+1})$ have a label $L_i(W) = (x, y)$ for which either $x \leq 1$ or $y \leq 1$ (or both).

In this case G_i will again be a 4-cycle. Let X_i be the set of all those vertices in $R - V(T_i \cup T_{i+1})$ whose label L_i is $(1,3)$ and let Y_{i+1} be the set of all those vertices in $R - V(T_i \cup T_{i+1})$ whose label L_i is $(3,1)$. Then

$$\begin{aligned}4|R| - 90d|R| - 30 \overset{(13.1)}{\leq} 6 \cdot \delta(R) - 2\binom{6}{2} &\leq e_R(T_i \cup T_{i+1}, R - T_i - T_{i+1}) \\ &\leq 4(|X_i| + |Y_{i+1}|) + 3(|R| - |X_i| - |Y_{i+1}|)\end{aligned}$$

and therefore

$$|X_i| + |Y_{i+1}| \geq |R| - 100d|R|. \tag{13.6}$$

On the other hand,

$$\begin{aligned}2|R| - 45d|R| - 15 \overset{(13.1)}{\leq} 3 \cdot \delta(R) - 15 &\leq e_R(T_i, R - T_i - T_{i+1}) \\ &\leq |X_i| + 3(|R| - |X_i|) = 3|R| - 2|X_i|\end{aligned}$$

and so $|X_i| \leq |R|/2 + 25d|R|$. Combining this with (13.6) gives $|Y_{i+1}| \geq |R|/2 - 125d|R|$. Similarly it follows that $|X_i| \geq |R|/2 - 125d|R|$. Since (13.1) implies that every vertex $x \in X_i$ has more than $|R| - |Y_{i+1}|$ neighbours, this means that R contains an X_i–Y_{i+1} edge $W'W$. Let A_i be the unique neighbour of $W' \in X_i$

in T_i and let B_i be any other vertex of T_i. Let I_{i+1} be the unique neighbour of $W \in Y_{i+1}$ in T_{i+1} and let J_{i+1} be any other vertex of T_{i+1}. Take G_i to be any 4-cycle $v_i v_i' w_i w_i'$ in G with $v_i, w_i \in W$ and $v_i', w_i' \in W'$ and such that the vertices on G_i have at least $(d-3\varepsilon)^4 L'$ common neighbours in both A_i and I_{i+1} and such that furthermore both v_i and w_i have at least $(d-3\varepsilon)L'$ neighbours in B_i while both v_i' and w_i' have at least $(d-3\varepsilon)L'$ neighbours in J_{i+1}. (The existence of such vertices follows from repeated applications of Proposition 2.3 again.) So G_i is an almost-triangulation whose attachment faces F_i' and F_i'' are the interior and the exterior of G_i. In Case 2 we will use G_i to glue P_i to P_{i+1} as indicated in Figure 13.5(c).

Since k' does not depend on n, it is not hard to show that for all $1 \leq i < k'$ we can choose the graphs G_i to be disjoint from each other. We then remove all the vertices of G lying in the G_i's from the clusters they belong to. So the size of each cluster now lies between $L' - 2k'$ and L'. We remove further vertices if necessary and add them to the exceptional set to ensure that all the clusters have the same size $L'' \geq L' - 2k'$. We still denote by V_0 the enlarged exceptional set. Thus

$$|V_0| \overset{(13.2)}{\leq} 14dn + 2k'|R| \leq 15dn. \tag{13.7}$$

13.5.3 Incorporating the exceptional vertices

Given an exceptional vertex $x \in V_0$, we say that a triangle T_i is *good for x* if x has at least cL'' neighbours in at least two of the (modified) clusters of T_i. Thus the number N of good triangles for x satisfies

$$3NL'' + (1+2c)L''k' \geq |N_G(x) \setminus V_0| \overset{(13.7)}{\geq} \frac{2n}{3} - 15dn.$$

Hence

$$N \geq \frac{1}{3L''}\left(\frac{2n}{3} - 15dn - (1+2c)\frac{n}{3}\right) > \frac{n}{12L''}$$

and therefore

$$\frac{|V_0|}{180dL''} \overset{(13.7)}{<} N.$$

But this shows that we can assign each exceptional vertex $x \in V_0$ greedily to a triangle $T_i \in \mathcal{T}$ such that T_i is good for x and such that to each triangle we assign at most $180dL''$ exceptional vertices. Let E_i be the set consisting of all exceptional vertices assigned to T_i.

Recall that A_i and B_i are two of the vertices of T_i and denote its third vertex by C_i. Moreover, recall that F_i' and F_i'' are the two attachment faces of the gluing graph G_i. For each $1 \leq i \leq k'$, we will now incorporate the exceptional vertices in E_i as indicated in Section 13.3.3. However, to ensure that after this process each of the 3 bipartite subgraphs between the vertex classes of the remainder of $G'(T_i)$ is still super-regular, we first apply Proposition 13.3 to the graph $G'(T_i)$ to obtain a partition A_i', A_i'' of A_i such that firstly $|A_i'| \geq |A_i|/4$ and $|A_i''| \geq |A_i|/4$, secondly, for each vertex $v \in G'(T_i)$ the set $N_{G'}(v) \cap A_i$ is split ε-fairly and, thirdly, for each $v \in E_i \cup V(\partial F_i' \cup \partial F_{i-1}'')$ the set $N_G(v) \cap A_i$

is split ε-fairly. We then apply Proposition 13.3 to obtain analogous partitions B_i', B_i'' of B_i and C_i', C_i'' of C_i. So in particular, each of the bipartite graphs $(A_i', B_i')_{G'}$, $(A_i', C_i')_{G'}$ and $(B_i', C_i')_{G'}$ is still 10ε-regular and has density $> d-3\varepsilon$. Let $G_i' := G'[A_i' \cup B_i' \cup C_i']$. Note that for each exceptional vertex $v \in E_i$ at least two of the vertex classes of G_i' contain at least $(cL'' - \varepsilon|G'(T_i)|)/2 \geq cL''/4$ neighbours of v.

Fix an enumeration of the vertices in E_i. Let x be the first vertex in E_i. Applying Proposition 2.3 repeatedly, it is not hard to find in G_i' a copy Q_x of the graph in Figure 13.6 such that x is joined to all the vertices on the boundary of the face F_x of Q_x. (The latter can be guaranteed since x has at least $cL''/4$ neighbours in two of the vertex classes of G_i'.) So Q_x is a 3-partite

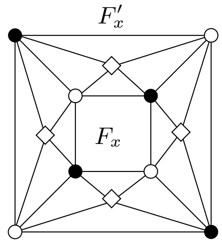


Figure 13.6: The graph Q_x used to incorporate the exceptional vertex x

almost-triangulation with two attachment 4-faces F_x and F_x'. Moreover, the boundary $\partial F_x'$ of F_x' contains only vertices from two of the vertex classes of G_i'. F_x' will be used to attach Q_x to the graph Q_y which we will choose below to incorporate the second vertex $y \in E_i$. To ensure that this can be done, we choose Q_x such that $\partial F_x'$ has at least $(d-13\varepsilon)^4 L''/4$ common neighbours in the unique vertex class of G_i' that avoids $\partial F_x'$ and such that every vertex $v \in \partial F_x'$ has at least $(d-13\varepsilon)L''/4$ neighbours in each of the two vertex classes of G_i' not containing v.

Now we apply Proposition 2.3 to choose a subgraph Q_y of $G_i' - Q_x$ for the next vertex y in our enumeration of E_i. Q_y will be one of the 3-partite almost-triangulations in Figure 13.7. Again, Q_y is chosen in such a way that y is joined to all the vertices on ∂F_y. (To ensure this, we take Q_y to be the graph in Figure 13.7(a) when y has many neighbours in each of the two clusters meeting $\partial F_x'$, and take the graph in Figure 13.7(b) for Q_y—possibly with the black and the white circular vertices interchanged—otherwise.) Moreover, $\partial F_y'$ will satisfy conditions analogous to $\partial F_x'$. Additionally, this time, Q_y has a third attachment 4-face F_y'' which will be used to attach Q_y to Q_x via F_y'' and F_x' as shown in Figure 13.8. (Our choice of $\partial F_x'$ implies that all the necessary $\partial F_y''$–$\partial F_x'$ edges can be guaranteed.)

We continue in this fashion until we have dealt with all the vertices in E_i. By attaching the graphs Q_z ($z \in E_i$) to each other, inserting each z into the attachment face F_z of Q_z and joining z to all vertices on ∂F_z we obtain an almost-triangulation Q_i in G_i' which has only one attachment face, namely F_v', where v is the last vertex in the enumeration of E_i. Note that the definition of G_i' implies that the graph $G_i'' := G'(T_i) - Q_i$ is still $(10\varepsilon, d/4)$-super-regular

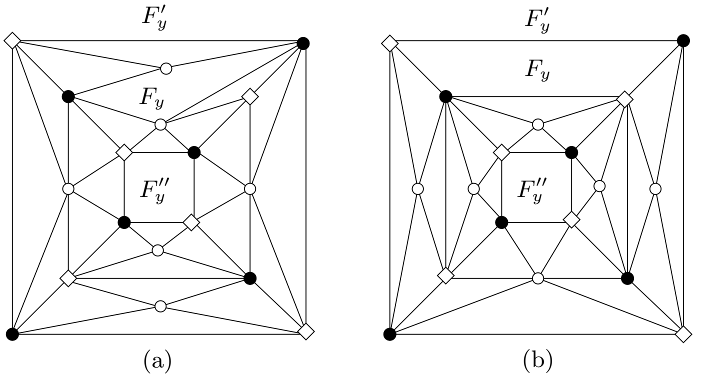


Figure 13.7: The two possibilities for the graph Q_y which is used to incorporate the exceptional vertex y

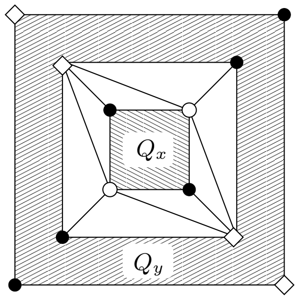


Figure 13.8: Attaching Q_y to Q_x via F''_y and F'_x

and that the vertices on the 4-cycle $v_i v'_i w_i w'_i$ bounding the attachment face F'_i of the gluing graph G_i still have at least $(d-3\varepsilon)^4 L''/4$ common neighbours in $A_i \cap V(G''_i)$ and that both v_i and w_i have at least $(d-3\varepsilon)L''/4$ neighbours in $B_i \cap V(G''_i)$. The 4-cycle bounding the attachment face F''_{i-1} of G_{i-1} satisfies analogous conditions. Moreover, note that all the graphs Q_z contain the same number of vertices in each cluster of T_i and hence the same applies to G''_i.

13.5.4 Applying the Blow-up lemma

Having incorporated the exceptional vertices, we can now apply the Blow-up lemma to obtain a spanning almost-triangulation P_1 in G''_1 which has two attachment 4-faces if $E_1 \neq \emptyset$ and only one attachment 4-face otherwise. The boundary of the first attachment 4-face F_1 will have two opposite vertices in the common neighbourhood of the boundary $\partial F'_1$ of the attachment face F'_1 of G_1 in $A_1 \cap V(G''_1)$ while its other two vertices lie the neighbourhoods of v_1 and w_1 in $B_1 \cap V(G''_1)$ respectively. Thus P_1 can be attached to G_1 via F_1 and F'_1 (see Fig. 13.5). Similarly, if $E_1 \neq \emptyset$, then the second attachment face of P_1 will be chosen in such a way that it can be used to attach P_1 to Q_1. Again, this is possible since the boundary of the unique attachment face of Q_1 has a large (common) neighbourhood in G''_1.

Next we apply the Blow-up lemma to obtain a spanning almost-triangulation P_2 in G_2'' which has three attachment faces if $E_2 \neq \emptyset$ and only two attachment faces otherwise. Similarly as for P_1, each of of these will be a 4-face, one will be chosen in such a way that it can be used to attach P_2 to G_2 and one to attach P_2 to Q_2 (if $E_2 \neq \emptyset$). The third attachment face will be chosen so that P_2 can be attached to G_1 via this face and F_1''. Continuing in this way, we obtain a spanning triangulation of G. ($P_{k'}$ has two attachment faces if $E_{k'} \neq \emptyset$ and only one attachment face otherwise.) This completes the proof of Theorem 13.1.

13.6 Concluding remarks

The example G^* in Section 13.1 showed that the bound in Theorem 13.1 is sharp if the order n of the graph G is divisible by 3. The following proposition states that this is in fact the case for all n.

Proposition 13.17 *For all $n \in \mathbb{N}$ there is a graph G of order n and minimum degree $\lceil 2n/3 \rceil - 1$ which contains no triangulation as a spanning subgraph.*

Proof. Clearly, we may assume that $n \geq 3$. Write $n = 3k + \ell$ where $\ell \in \{0, 1, 2\}$. If $\ell = 0$ then we can take for G the graph G^* described in Section 13.1. If $\ell = 1$, we take a similar graph: this time the cliques A and B both have order k and the independent set C has order $k + 1$. So we may assume that $\ell = 2$. Also in this case our graph G will again look rather similar to G^*. The cliques A and B both have order k, the set C has size $k + 2$, but to ensure that $\delta(G) = 2k + 1 = \lceil 2n/3 \rceil - 1$ we now additionally insert a maximal set of independent edges into C and, if k is odd, one extra edge joining the unique vertex in C that is not covered by these independent edges to some other vertex in C. So at most one vertex in C has degree two in $G[C]$ and all other vertices have degree one.

We have to show that G does not contain a spanning triangulation. Suppose on the contrary that P is such a triangulation. Then P must have two facial triangles T_1 and T_2 which share an edge $xy \in G[C]$ and are such that T_1 contains a vertex of A while T_2 contains a vertex of B. We may assume that x has degree 1 in $G[C]$. On the other hand, in P the vertex x has neighbours in both A and B. Since P is a triangulation, this implies that P contains an A–B edge joining two neighbours of x, a contradiction. □

Note that the graph G in the proof of the $\ell = 2$ case of the proposition already contains the square of a Hamilton cycle if n is large. On the other hand, the proof of that case can be extended to show that one can take for $G[C]$ even a Hamilton path without creating a spanning triangulation. As soon as $G[C]$ contains a Hamilton cycle, a spanning triangulation obviously exists though.

Theorem 13.1 can easily be extended to arbitrary surfaces:

Theorem 13.18 *For each surface S there exists an integer $n_0 = n_0(S)$ such that every graph of order $n \geq n_0$ and minimum degree at least $2n/3$ contains some triangulation of S as a spanning subgraph.*

To see this, first note that in every case of the proof of Theorem 13.1 we apply the Blow-up lemma at least once to find a large planar (almost-) triangulation P similar to the one in Figure 13.1. (In particular, P is 3-chromatic and its colour classes have equal size.) For an arbitrary surface S, all we have to change is that in one of these applications we choose a 3-chromatic (almost-) triangulation of S with colour classes of equal size instead of P. (Such triangulations of S are easily seen to exist. For example, they too can be obtained by a slight modification of the graph in Figure 13.1.)

As remarked in the introduction, the proof of Theorem 13.1 shows that the spanning triangulation can be found in polynomial time. Indeed, both the Regularity lemma and the Blow-up lemma can be implemented in polynomial time (see [4] and [52] respectively). Furthermore, the probabilistic splitting argument in Proposition 13.3 can be derandomized using standard techniques (see e.g. [7]) and it is easy to see that the cycle of logarithmic length in Proposition 13.12 can be found in polynomial time using a shortest path algorithm. As the order of the reduced graph R is constant, the remaining steps can also be done in polynomial time.

In the remainder of this section we discuss how Theorem 13.1 might perhaps be strengthened. Obviously a minimum degree of $2n/3$ will not force every given triangulation P of order n as a subgraph. For example, G might be 3-partite, which implies that we can only hope for triangulations P with chromatic number 3. Of course, we cannot guarantee all of these either, as there are triangulations whose chromatic number is 3 and whose maximum degree is $n-2$. However, in view of our proof of Theorem 13.1, it might be helpful to restrict one's attention to triangulations P of bounded band-width, as this imposes a linear structure on P. (The *band-width* of a graph H is the smallest integer k for which there exists an enumeration $v_1, \dots, v_{|H|}$ of the vertices of H such that every edge $v_iv_j \in H$ satisfies $|i-j| \le k$.) Bollobás and Komlós [48] conjectured that for every $\gamma > 0$ and all $r, \Delta \in \mathbb{N}$ there are $\alpha > 0$ and $n_0 \in \mathbb{N}$ such that every graph G of order $n \ge n_0$ and minimum degree at least $(1 - \frac{1}{r} + \gamma)n$ contains a copy of every graph H of order n whose chromatic number is at most r, whose maximum degree is at most Δ and whose band-width is at most αn.

This would imply that every sufficiently large graph of minimum degree at least $(2/3 + \gamma)n$ contains every 3-chromatic triangulation of bounded band-width. Even in this special case the error term γn cannot be omitted completely: there are 3-chromatic triangulations whose colour classes have different sizes (modify the graph in Figure 13.1). These obviously do not embed into the complete 3-partite graph whose vertex classes have equal size. However, it might be true that for all integers b there exists a constant $C = C(b)$ such that every graph of order n and minimum degree at least $2n/3 + C$ contains every 3-chromatic triangulation of order n and band-width at most b as a subgraph.

Also, we do not know whether one can strengthen Theorem 13.1 in the following way. Given n, is there a triangulation P_n of order n which is contained in every graph G of order n and minimum degree at least $2n/3$? When n is divisible by 3, then the preceding arguments show that P_n would have to be 3-chromatic with equal size colour classes. Moreover, P_n would have to contain induced cycles of many different lengths. To see the latter, consider a graph G which is similar to the graph G^* from Section 13.1. This time the cliques have order $n/3 - 1$, the independent set C has order $n/3 + 2$ and we insert a 2-factor into C. Then the proof of Proposition 13.17 shows that every spanning triangulation of G must contain one of the cycles in $G[C]$.

Bibliography

[1] M. Ajtai, J. Komlós, and E. Szemerédi, A note on Ramsey numbers, *J. Combin. Theory Ser. A* **29** (1980), 354–360.

[2] M. Ajtai, J. Komlós and E. Szemerédi, A dense infinite Sidon sequence, *Europ. J. Combin.* **2** (1981), 1–11.

[3] N. Alon, Problems and results in extremal combinatorics, Part I, *Discrete Math.*, to appear.

[4] N. Alon, R.A. Duke, H. Lefmann, V. Rödl and R. Yuster, The algorithmic aspects of the Regularity lemma, *J. Algorithms* **16** (1994), 80–109.

[5] N. Alon, M. Krivelevich and B. Sudakov, Coloring graphs with sparse neighbourhoods, *J. Combin. Theory B* **77** (1999), 73–82.

[6] N. Alon, L. Rónyai and T. Szabó, Norm-graphs: variations and applications, *J. Combin. Theory B* **76** (1999), 280–290.

[7] N. Alon and J. Spencer, *The Probabilistic Method* (2nd edition), Wiley-Interscience 2000.

[8] N. Alon and R. Yuster, H-factors in dense graphs, *J. Combin. Theory B* **66** (1996), 269–282.

[9] A. Bender, Z. Gao and N.C. Wormald, The number of labeled 2-connected planar graphs, *Electronic J. Combin.* **9** (2002), R43.

[10] C.T. Benson, Minimal regular graphs of girths eight and twelve, *Canad. J. Math.* **26** (1966), 1091–1094.

[11] N. Biggs and M.J. Hoare, The sextet construction for cubic graphs, *Combinatorica* **3** (1983), 153–165.

[12] B. Bollobás, *Extremal Graph Theory*, Academic Press 1978.

[13] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.

[14] B. Bollobás, *Modern Graph Theory*, Graduate Texts in Mathematics 184, Springer-Verlag 1998.

[15] B. Bollobás, Extremal graph theory, in: R.L. Graham et al. (ed.), *Handbook of Combinatorics*, Vol. 2, Elsevier 1995, 1231–1292.

[16] B. Bollobás, *Random Graphs* (2nd edition), Cambridge studies in Advanced Mathematics 73, Cambridge University Press 2001.

[17] B. Bollobás, P.A. Catlin and P. Erdős, Hadwiger's conjecture is true for almost every graph, *Europ. J. Combin.* **1** (1980), 195–199.

[18] B. Bollobás and A. Frieze, Spanning maximal subgraphs of random graphs, *Random Struct. Alg.* **2** (1991), 225-231.

[19] B. Bollobás and A. Thomason, Highly linked graphs, *Combinatorica* **16** (1996), 313–320.

[20] B. Bollobás and A. Thomason, Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs, *Europ. J. Combin.* **19** (1998), 883–887.

[21] J.A. Bondy and M. Simonovits, Cycles of even length in graphs, *J. Combin. Theory Ser. B* **16** (1974), 97–105.

[22] N. Bonichon, C. Gavoille and N. Hanusse, An information-theoretic upper bound of planar graphs using triangulations, H. Alt and M. Habib (eds.) STACS 2003, *Springer LNCS* **2607** (2003) 499–510.

[23] W.G. Brown, On hamiltonian regular graphs of girth six, *J. London Math. Soc.* **42** (1967), 514–520.

[24] G. Călinescu, C.G. Fernandes, U. Finkler and H. Karloff, A better approximation algorithm for finding planar subgraphs, *J. Algorithms* **28** (1998), 105-124.

[25] F. Chung and R. Graham, *Erdős on Graphs. His legacy of unsolved problems*, A K Peters 1998.

[26] K. Corrádi and A. Hajnal, On the maximal number of independent circuits in a graph, *Acta Math. Acad. Sci. Hungar.* **14** (1963), 423–439.

[27] A. Denise, M. Vasconcellos and D. Welsh, The random planar graph, *Congressus numerantium* **113** (1996), 61–79.

[28] R. Diestel, *Graph Theory*, Graduate Texts in Mathematics 173, Springer-Verlag 1997.

[29] R. Diestel and C. Rempel, Dense minors in graphs of large girth, *Combinatorica*, to appear.

[30] R.A. Duke and V. Rödl, On graphs with small subgraphs of large chromatic number, *Graphs and Combinatorics* **1** (1985), 91–96.

[31] P. Erdős, Some remarks on the theory of graphs, *Bulletin Amer. Math. Society* **53**, 292–294.

[32] P. Erdős, Graph theory and probability, *Canad. J. Mathematics* **11** (1959), 34–38.

[33] G. Fan and H.A. Kierstead, Hamilton square paths, *J. Combin. Theory B* **67** (1996), 167–182.

[34] S. Gerke, Y. Kohayakawa, V. Rödl and A. Steger, in preparation.

[35] S. Gerke and C. McDiarmid, On the number of edges in a random planar graph, *Combin. Probab. Comput.*, to appear.

[36] A. Gyárfás, On Ramsey covering-numbers, *Coll. Math. Soc. János Bolyai*, in: *Infinite and Finite Sets*, North Holland/American Elsevier, New York 1975.

[37] E. Györi, C_6-free bipartite graphs and product representation of squares, *Discrete Mathematics* **165/166** (1997), 371–375.

[38] R. Häggkvist and A. Scott, Arithmetic progressions of cycles, preprint.

[39] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience 2000.

[40] T.R. Jensen and B. Toft, *Graph Coloring Problems*, Wiley-Interscience 1995.

[41] A.R. Johansson, Asymptotic choice number for triangle free graphs, preprint.

[42] L.K. Jørgensen and L. Pyber, Covering a graph by topological complete subgraphs, *Graphs and Combinatorics* **6** (1990), 161–171.

[43] H.A. Jung, Eine Verallgemeinerung des n-fachen Zusammenhangs für Graphen, *Math. Ann.* **187** (1970), 95–103.

[44] A. Kelmans, D. Mubayi and B. Sudakov, Asymptotically optimal tree-packings in regular graphs, *Electronic J. Combin.* **8** (2001), #R38.

[45] H.A. Kierstead and S.G. Penrice, Radius two trees specify χ-bounded classes, *J. Graph Theory* **18** (1994), 119–129.

[46] J.H. Kim, The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$, *Random Struct. Alg.* **7** (1995), 173–207.

[47] J. Kollár, L. Rónyai and T. Szabó, Norm graphs and bipartite Turán numbers, *Combinatorica* **16** (1996), 399–406.

[48] J. Komlós, The Blow-up lemma, *Combin. Probab. Comput.* **8** (1999), 161–176.

[49] J. Komlós, Tiling Turán theorems, *Combinatorica* **20** (2000), 203–218.

[50] J. Komlós, G. N. Sárközy and E. Szemerédi, On the square of a Hamilton cycle in dense graphs, *Random Struct. Alg.* **9** (1996), 193–211.

[51] J. Komlós, G. N. Sárközy and E. Szemerédi, Blow-up lemma, *Combinatorica* **17** (1997), 109–123.

[52] J. Komlós, G. N. Sárközy and E. Szemerédi, An algorithmic version of the blow-up lemma, *Random Struct. Alg.* **12** (1998), 297–312.

[53] J. Komlós, G. N. Sárközy and E. Szemerédi, On the Pósa-Seymour conjecture, *J. Graph Theory* **29** (1998), 167–176.

[54] J. Komlós, G. N. Sárközy and E. Szemerédi, Proof of the Seymour conjecture for large graphs, *Ann. Combin.* **2** (1998), 43–60.

[55] J. Komlós, G. N. Sárközy and E. Szemerédi, Proof of the Alon-Yuster conjecture, *Discrete Math.* **235** (2001), 255–269.

[56] J. Komlós and M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, *Bolyai Society Mathematical Studies 2, Combinatorics, Paul Erdős is Eighty (Vol. 2)* (D. Miklós, V. T. Sós and T. Szőnyi eds.), Budapest (1996), 295–352.

[57] J. Komlós and E. Szemerédi, Topological cliques in graphs, *Comb. Probab. Comput.* **3** (1994), 247–256.

[58] J. Komlós and E. Szemerédi, Topological cliques in graphs II, *Combin. Probab. Comput.* **5** (1996), 79–90.

[59] A. V. Kostochka, A lower bound for the Hadwiger number of graphs by the average degree, *Combinatorica* **4** (1984), 307–316.

[60] M. Krivelevich, On the minimal number of edges in color-critical graphs, *Combinatorica* **17** (1997), 401–426.

[61] M. Krivelevich, B. Sudakov and T. Szabó, Triangle factors in pseudo-random graphs, preprint.

[62] D. Kühn and D. Osthus, Topological minors in graphs of large girth, *J. Combin. Theory B* **86** (2002), 364–380.

[63] D. Kühn and D. Osthus, Minors in graphs of large girth, *Random Struct. Alg.* **22** (2003), 213–225.

[64] D. Kühn and D. Osthus, Induced subdivisions in $K_{s,s}$-free graphs of large average degree, *Combinatorica*, to appear.

[65] D. Kühn and D. Osthus, Complete minors in $K_{s,s}$-free graphs, *Combinatorica*, to appear.

[66] D. Kühn and D. Osthus, Large topological cliques in graphs without a 4-cycle, *Combin. Probab. Comput.*, to appear.

[67] D. Kühn and D. Osthus, Subdivisions of K_{r+2} in graphs of average degree at least $r + \varepsilon$ and large but constant girth, *Combin. Probab. Comput.*, to appear.

[68] D. Kühn and D. Osthus, Packings in dense regular graphs, *Combin. Probab. Comput.*, to appear.

[69] D. Kühn and D. Osthus, Forcing unbalanced complete bipartite minors, submitted.

[70] D. Kühn and D. Osthus, Extremal connectivity for topological cliques in bipartite graphs, submitted.

[71] D. Kühn and D. Osthus, Spanning triangulations in graphs, submitted.

[72] D. Kühn and D. Osthus, 4-cycles in graphs without a forbidden even cycle, submitted.

[73] D. Kühn, D. Osthus and A. Taraz, Large planar subgraphs in dense graphs, submitted.

[74] L. Lovász, Combinatorial Problems and Exercises (2nd edition), North-Holland, 1993.

[75] W. Mader, Homomorphieeigenschaften und mittlere Kantendichte von Graphen, *Math. Annalen* **174** (1967), 265–268.

[76] W. Mader, Homomorphiesätze für Graphen, *Math. Annalen* **178** (1968), 154–168.

[77] W. Mader, Existenz n-fach zusammenhängender Teilgraphen in Graphen genügend großer Kantendichte, *Abh. Math. Sem. Univ. Hamburg* **37** (1972), 86–97.

[78] W. Mader, Grad und lokaler Zusammenhang in endlichen Graphen, *Math. Ann.* **205** (1973), 9–11.

[79] W. Mader, Topological subgraphs in graphs of large girth, *Combinatorica* **18** (1998), 405–412.

[80] W. Mader, An extremal problem for subdivisions of K_5^-, *J. Graph Theory* **30** (1999), 261–276.

[81] W. Mader, Subdivisions of a graph of maximal degree $n+1$ in graphs of average degree $n+\varepsilon$ and large girth, *Combinatorica* **21** (2001), 251–265.

[82] C. McDiarmid, A. Steger and D. Welsh, Random planar graphs, preprint.

[83] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, 1995.

[84] R.C. Mullin and P.J. Schellenberg, The enumeration of c-nets via quadrangulations, *J. Combin. Theory* **4** 259–276, 1968.

[85] J.S. Myers, Graphs without large complete minors are quasi-random, *Combin. Probab. Comput.* **11** (2002), 571–585.

[86] J.S. Myers, The extremal function for unbalanced bipartite minors, *Discrete Math.* **271** (2003), 209–222.

[87] J.S. Myers and A. Thomason, The extremal function for noncomplete minors, preprint.

[88] D. Osthus, H.J. Prömel and A. Taraz, Almost all graphs of high girth and suitable density have high chromatic number, *J. Graph Theory* **37** (2001), 220–226.

[89] D. Osthus, H.J. Prömel and A. Taraz, On random planar graphs, the number of planar graphs and their triangulations, *J. Combin. Theory B* **88** (2003), 119–134.

[90] D. Osthus, H.J. Prömel and A. Taraz, For which densities are random triangle-free graphs almost surely bipartite?, *Combinatorica* **23** (2003), 105–150.

[91] D. Osthus and A. Taraz, Random maximal H-free graphs, *Random Struct. Alg.* **18** (2001), 61–82.

[92] H.J. Prömel and A. Steger, Counting H-free graphs, *Discrete Math.* **154** (1996), 311–315.

[93] L. Pyber, V. Rödl and E. Szemerédi, Dense graphs without 3-regular subgraphs, *J. Combin. Theory B* **63** (1995), 41–54.

[94] D. Schlatter, *On the Evolution of Random Discrete Structures: Poset Dimension and Maximum Planar Subgraphs*, Diploma thesis, Humboldt-Universität zu Berlin, 2002.

[95] A. Scott, Induced trees in graphs of large chromatic number, *J. Graph Theory* **24** (1997), 297–311.

[96] D.P. Sumner, Subtrees of a graph and chromatic number, in: *The Theory and Applications of Graphs*, G. Chartrand (ed.), John Wiley, New York 1981.

[97] E. Szemerédi, Regular partitions of graphs, in *Colloques Internationaux C.N.R.S. No. 260 – Problèmes Combinatoires et Théorie des Graphes*, Orsay 1976, pp. 148–192.

[98] R. Thomas and P. Wollan, An improved linear edge bound for graph linkages, preprint.

[99] A. Thomason, An extremal function for contractions of graphs, *Math. Proc. Cambridge Philos. Soc.* **95** (1984), 261–265.

[100] A. Thomason, The extremal function for complete minors, *J. Combin. Theory B* **81** (2001), 318–338.

[101] A. Thomason, Extremal function for graph minors, preprint.

[102] C. Thomassen, Girth in graphs, *J. Combin. Theory B* **35** (1983), 129–141.

[103] C. Thomassen, Subdivisions of graphs with large minimum degree, *J. Graph Theory* **8** (1984), 23–28.

[104] W.T. Tutte, A census of planar triangulations, *Canadian J. Mathematics* **14**, 21–38, 1962.

[105] J. Verstraëte, On arithmetic progressions of cycle lengths in graphs, *Combin. Probab. Comput.* **9** (2000), 369–373.

[106] J. Verstraëte, A note on vertex-disjoint cycles, *Combin. Probab. Comput.* **11** (2002), 97–102.

[107] A. Weiss, Girth of bipartite sextet graphs, *Combinatorica* **4** (1984), 241–245.