BOLYAI SOCIETY
MATHEMATICAL STUDIES, 15

Conference on Finite and Infinite Sets
Budapest, pp. 359–380.

Extremal Functions for Graph Minors

A. THOMASON

The extremal problem for graph minors is to determine, given a fixed graph H, how many edges a graph G can have if it does not have H as a minor. It turns out that the extremal graphs are pseudo-random; the sense of this has best been expressed by Vera T. Sós in a question answered by Joseph Myers.

This survey describes what is known about the extremal function and discusses some related matters.

1. Introduction

We say that the graph H is a *minor* or *subcontraction* of the graph G, written $G \succ H$, if H can be obtained from G by deleting some vertices and edges and by contracting some other edges. This is equivalent to the statement that $V(G)$ contains disjoint subsets W_u, $u \in V(H)$, such that the subgraph $G[W_u]$ induced by W_u is connected for each $u \in V(H)$ and there is an edge in G between W_u and W_v whenever $uv \in E(H)$.

This survey describes what is currently known about the fundamental extremal question regarding graph minors, namely, how many edges are needed in G to ensure that $G \succ H$? It is now possible to give a fairly full answer to this question. In the first place, it turns out that there is a close connection with the theory of random graphs and with the theory of pseudo-random graphs. This connection is expressed best by a question of Vera T. Sós; her question, and the answer subsequently given by Joseph Myers, are discussed in §5. Secondly, the variation of the extremal function with H can be described in terms of a structural property of H, reminiscent of the way in which, in classical extremal graph theory, the extremal function depends on the chromatic number. In the present case, the relevant structural property is again a kind of partition of H, by means of weights, that is defined in §1.2 and discussed in detail in §6.

We also describe briefly (in §8 and §9) some other extremal problems for minors, such as what connectivity or girth forces a graph to have a given minor. This area has enjoyed some substantial recent advances, but there remain significant open questions about which little, as yet, is known.

1.1. Background

The source of the basic extremal problem for minors is, arguably, the remarkable paper of Wagner [38], in which he proved that the Four Colour Theorem is equivalent to the assertion that $G \succ K_5$ for every graph G that needs five colours to colour it. Hadwiger [10] in 1943 famously conjectured that $G \succ K_t$ for every graph G that needs t colours to colour it. This assertion is trivial for $t \leq 3$, and Hadwiger proved it for $t = 4$. Much more recently, Robertson, Seymour and Thomas [31] have proved the conjecture for $t = 6$ by showing that it follows from the Four Colour Theorem. For a good survey of Hadwiger's conjecture see Toft [37].

In 1964 Wagner [39] proved that $G \succ K_t$ provided the chromatic number of G is sufficiently large (2^{t-3} will do). Mader [22] then developed the idea that the chromatic number might not be the significant parameter; he showed that $G \succ K_t$ provided merely that the *average degree* of G is sufficently large. He therefore introduced the function

$$c(t) = \min\left\{ c : e(G) \geq c|G| \text{ implies } G \succ K_t \right\},$$

proving that $c(t) \leq 2^{t-3}$ (see Lemma 2.1) and later [23] that $c(t) \leq 8\lceil t \log_2 t \rceil$. Thus we are led to the extremal problem for complete graph minors.

In fact, for small t, much more precise information is available. Write $F+G$ for the join of two graphs F and G, meaning their disjoint union with all edges added between. Observe that the graph $K_{t-2} + \overline{K}_{n-t+2}$ does not have a K_t minor, and neither does the graph $K_{t-5} + P$ if P is a maximal planar graph. These graphs all have $(t-2)|G| - \binom{t-1}{2}$ edges. Dirac [7] demonstrated that if $t \leq 5$ then this is the exact maximum number of edges in G if $G \not\succ K_t$, and Mader [23] extended this to $t \leq 7$. But the seductive pattern stops here; as Mader pointed out, the complete 5-partite graph with two vertices in each class has $40 = 6|G| - 20$ edges and no K_8 minor. (Jørgensen [12] later proved that this is the maximum size of graphs with no K_8 minor, and characterized the extremal graphs. He could thereby (see [11]) extend to $t \le 8$ the cases in which the following conjecture is known to hold: that if G has a partition into $V_1, \dots, V_t$ such that $G[V_i \cup V_j]$ is connected for $i \ne j$, then $G \succ K_t$. This conjecture is one of several, related to Hadwiger's conjecture, made by Las Vergnas and Meyneil [21].)

For larger values of t the divergence of the extremal function from the simple pattern just described is much greater. Random graphs provide examples showing that $c(t)$ is of order at least $t\sqrt{\log t}$. This was noticed by several people at about the same time (for example Kostochka [15, 16], and also Fernandez de la Vega [9] based on Bollobás, Catlin and Erdős [2]). Kostochka [15, 16] proved that the correct order of growth for $c(t)$ is indeed $t\sqrt{\log t}$ (see also [32]).

1.2. Recent developments

Recently, the asymptotic value of $c(t)$ was determined.

Theorem 1.1 ([34]). *There exists a constant $\alpha = 0.3190863\dots$ such that*

$$c(t) = \big(\alpha + o(1)\big)\, t\sqrt{\log t}.$$

The constant α can be explicitly described (see §3); it is simply the best constant that can be obtained from randomly generated lower bounds (note that logarithms are natural unless stated otherwise).

It is evident from Theorem 1.1 that there is a connection between random graphs and extremal functions for minors, though the connection is still closer than first appears. The extremal graphs must be pseudo-random graphs of specified order and density, or else a more-or-less disjoint union of such graphs ([34, 27]). The connection has been captured best by Vera T. Sós in a question which, loosely speaking, is this: if a graph of positive density has no minor bigger than what might be found in a random graph of the same density, must the graph itself be pseudo-random? Myers [26] has given a positive answer to this question. We explain this question more precisely, together with its answer, in §5.

Even more recently, the asymptotic value of the average degree that implies a general H minor has been determined, and the strong connection with pseudo-random graphs persists. Let

$$c(H) = \min\left\{c : e(G) \geq c|G| \text{ implies } G \succ H\right\},$$

so that $c(t) = c(K_t)$. The results about $c(H)$ are expressed in terms of a parameter $\gamma(H)$ of the graph H, defined as the minimum average vertex weight amongst weightings satisfying a certain condition.

Definition 1.2. Let H be a graph of order t. We define

$$\gamma(H) = \min_{w} \frac{1}{t} \sum_{u \in H} w(u) \qquad \text{such that} \qquad \sum_{uv \in E(H)} t^{-w(u)w(v)} \leq t,$$

where the minimum is over all assignments $w : V(H) \to \mathbf{R}^+$ of non-negative weights to the vertices of H.

A uniform weighting w shows that $0 \leq \gamma(H) \leq 1$ for all H and, more generally, $\gamma(H) \leq \sqrt{\tau}$ if H has at most $|H|^{1+\tau}$ edges. In §6.2 we shall describe ways of estimating $\gamma(H)$ fairly precisely, but it is worth pointing out here that, amongst H with $|H|^{1+\tau}$ edges, almost all H and all regular H satisfy $\gamma(H) \approx \sqrt{\tau}$; indeed, $\gamma(H)$ will not be significantly smaller than this unless H has some very restrictive structure.

The extremal result for H, if H has t vertices, is then this.

Theorem 1.3 ([28]). *There exists a constant $\alpha = 0.3190863\ldots$ such that*

$$c(H) = \big(\gamma(H)\alpha + o(1)\big)\, t\sqrt{\log t}$$

for every graph H of order t, where the $o(1)$ term is a term tending to zero as $t \to \infty$.

1.3. Contents of this article

We begin in §2 with some preliminary remarks about the extremal function; in particular, it is seen why only dense graphs are of importance in the study of the extremal problem. There follows in §3 a discussion of minors of random graphs and in §4 an explanation of what lies behind Theorem 1.1. The discussion of Sós's question in §5 should nevertheless be comprehensible without first reading the earlier parts.

After that, we go on in §6 to consider the general extremal problem for contractions to a fixed graph H (not necessarily complete). In §7 we comment on an application of the extremal problem to linking in graphs. We finish with some remarks about other conditions on a graph that imply it has large minors; in §8 it is seen how large girth can replace large minimal degree as such a condition, and lastly in §9 we look at how large connectivity might do the same.

2. Initial observations

Here is a simple lemma that implies the existence of the function $c(t)$.

Lemma 2.1. *Let d be an integer and let G be minimal, with respect to taking minors, in the class*

$$\{G : e(G) \geq d|G|\}.$$

Then every edge of G is in at least d triangles; in particular, if H is the neighbourhood subgraph of some vertex, then $e(H) \geq \frac{d}{2}|H|$.

Proof. If G is minimal then G is non-empty and, for every edge uv, the graph G/uv obtained by contracting uv satisfies $e(G/uv) < d(|G|-1)$. Thus more than d edges are lost by contracting uv, meaning that uv is in at least d triangles. So, if H is the neighbourhood graph of u, then $\delta(H) \geq d$. ■

The bound $c(t) \leq 2^{t-3}$ follows at once from Lemma 2.1 by induction on t, because a graph G with $e(G) \geq 2^{t-3}|G|$ contracts to a graph containing a vertex u joined to a graph H with $H \succ K_{t-1}$.

Now if G is minimal in $\{G : e(G) \geq d|G|\}$ then $e(G) = d|G|$ (else just remove an edge), so if u is a vertex of minimal degree then $|H| = \delta(G) \leq 2d$. Thus, if we can find a large complete minor in any graph H with $\delta(H) \geq |H|/2$, we can find a large complete minor in any graph at all. In fact, the function $c(t)$ is completely determined by minors of dense graphs, as we explain in §4.

The simple idea of Lemma 2.1 can be exploited further by considering graphs minimal in the class $\{G : e(G) \geq f(|G|),\ |G| \geq m\}$ where $f(n)$ is an integer-valued function chosen so that $f(m) > \binom{m}{2}$ for some m. Then the class contains no graph of order m so a minimal graph must, by the argument above, satisfy $G \succ H$ where $|H| \leq 2f(|G|)$ and $\delta(H) \geq f(|G|) - f(|G|-1)$. A couple of choices that are helpful in different contexts, both essentially due to Mader [23], are these.

First, let $f(|G|) = d|G| - kd$. Provided $k \leq d/2$ we can take $m = d$. This choice gives the same conclusion as Lemma 2.1 but with the extra property that $\kappa(G) \geq k+1$, as can easily be shown. This choice is useful when determining the extremal function $c(t)$.

Secondly, with the choice $f(G) = \lceil \beta d|G|(1 + \log(|G|/\beta d))/2 \rceil$, where β satisfies $1 = \beta(1 + \log(2/\beta))$, we can take $m = \lceil \beta d \rceil$. The function is chosen both so that $f(|G|) - f(|G|-1)$ is large for $|G| \leq 2d$ and also so that the graph H from Lemma 2.1, with $|H| \leq 2d$ and $\delta(H) \geq d$, lies in the class. Applying the above arguments to this H produces, after a little calculation, the following result.

Lemma 2.2. *Let $\beta = 0.37\ldots$ be as above. Let G be a graph with $e(G) \geq d|G|$. Then $G \succ H$, where $|H| \leq d+2$ and $2\delta(H) \geq |H| + \lfloor \beta d \rfloor - 1$.*

The main point of this lemma is that the minimum degree is bounded below away from $|H|/2$. This has useful consequences, as we describe in §7.

3. Random graphs

Let $G(n,p)$ denote a graph of order n whose edges are chosen independently and at random with probability p.

Theorem 3.1. *Given $\varepsilon > 0$ there exists $T = T(\varepsilon)$ with the following property. Let $t > T$, let $\varepsilon < p < 1-\varepsilon$, let $q = 1-p$ and let $n = \lfloor (1-\varepsilon)\, t\sqrt{\log_{1/q} t} \rfloor$. Then $G(n,p) \succ K_t$ with probability less than ε.*

By choosing $q = \lambda$ where $\lambda = 0.284668\ldots$ is the root of the equation $1 - \lambda + 2\lambda \log \lambda = 0$, we obtain from Theorem 3.1 graphs that have no K_t minor and that have average degree $pn \sim \alpha t\sqrt{\log t}$ where $\alpha = (1-\lambda)/2\sqrt{\log(1/\lambda)}$. This straightaway gives half of Theorem 1.1, namely $c(t) \geq (\alpha + o(1))\, t\sqrt{\log t}$.

Theorem 3.1 is best possible, as shown by Bollobás, Catlin and Erdős [2], in the sense that if $n = (1+\varepsilon)\, t\sqrt{\log_{1/q} t}$ then $G(n,p)$ almost surely has a

K_t minor, but this follows in any case from the stronger Theorem 4.1 in §4. For our purposes, random graphs are needed only as a supply of graphs without H minors, for any specified H.

It is worth seeing what determines whether or not $G(n,p) \succ H$ with high probability. Let the vertices of $G(n,p)$ be partitioned into sets W_u, $u \in V(H)$. We need $G[W_u]$ to be connected and we need an edge between W_u and W_v whenever $uv \in E(H)$. The first of these is, in practice, easily arranged — it is the second condition that is the harder to satisfy. The probability that it is satisfied for a particular partition is

$$\prod_{uv \in E(H)} \left(1 - q^{|W_u||W_v|}\right) \approx \exp\left\{ - \sum_{uv \in E(H)} q^{|W_u||W_v|} \right\}.$$

So the partitions most likely to work are those where $\sum_{uv \in E(H)} q^{|W_u||W_v|}$ is minimized, and it is the way in which this sum minimizes, for a particular H, that decides which random graphs have H minors and so, in turn, decides the value of $c(H)$.

By far the most common case is that where, in the minimizing choice, all $|W_u|$ are equal; that is, $|W_u| = n/t$ where $t = |H|$. The expected number of successful partitions is then around $t^n \exp\left\{ - e(H) q^{n^2/t^2} \right\}$, there being about t^n possible partitions. For a graph with $e(H) = t^{1+\tau}$ edges this expected value is small or large according to whether n is less than, or greater than, $\sqrt{\tau}\, t \sqrt{\log_{1/q} t}$, so this is the threshold value of n at which H minors appear.

For general H, put $w(u) = |W_u| / \sqrt{\log_{1/q} t}$, and write $\overline{w} = n/t\sqrt{\log_{1/q} t}$ for the average value of w. Choosing $|W_u|$ to minimize the sum above is the same as choosing w to minimize $\sum_{uv \in E(H)} t^{-w(u)w(v)}$. Writing M for this minimum value, the expectation becomes $t^n \exp(-M)$; since $n = \overline{w}\, t \sqrt{\log_{1/q} t}$, the threshold region for n is when M is approximately t. It can now be seen that the quantity $\overline{w}$ determining this threshold is precisely the parameter $\gamma(H)$ defined in §1.2.

4. Complete minors of dense graphs

The main theorem relevant to the extremal properties of complete minors is the following one, a slightly weakened version of that appearing in [34].

Theorem 4.1 ([34]). *Given $\varepsilon > 0$ there exists $T = T(\varepsilon)$ with the following property. Let $t > T$, let $\varepsilon < p < 1 - \varepsilon$, let $q = 1 - p$ and let $n = \left\lceil (1+\varepsilon)\, t\sqrt{\log_{1/q} t}\right\rceil$. Then every graph G of order n and connectivity $\kappa(G) \geq n(\log\log\log n)/(\log\log n)$ has a K_t minor.*

Thus, every graph of positive density (except those which are nearly disconnected) has complete minors at least as large as those in random graphs of the same density. Some kind of connectivity requirement is obviously required since, for example, the minors of a union of two disjoint graphs of order $n/2$ and density $1/2$ are the minors in the individual components, and they would not be expected to correspond to the minors in a typical graph of order n and density $1/4$.

To prove Theorem 4.1 we must find a partition of $V(G)$ into sets W_u, $u \in V(K_t)$, such that each $G[W_u]$ is connected and such that there is an edge between W_u and W_v whenever $uv \in E(K_t)$. Just as in §3, the first requirement can be arranged fairly straightforwardly, and it is the second that needs care. A natural approach would be to take a random partition of the n vertices into t parts of size n/t each, in the hope that, even if not all the required edges materialize, at most $o(t)$ of them fail, and by dropping any vertex of K_t that is incident with one of these failed edges, we are still left with a complete minor on $t - o(t)$ vertices, which is good enough.

The reason this approach does not succeed directly is because the degrees in the graph G may vary wildly. In order for the argument to work it is necessary that a randomly chosen part of size $l = \sqrt{\log t}$ be joined to all but not much more than nq^l vertices; a second random part would then fail to have an edge to the first random part with probability around $q^{l\times l}$, so behaving much as if the graph were itself random. However, the expected number of vertices not joined to our first random part is $\sum_{x\in G} q(x)^l$, where x has $q(x)n$ non-neighbours, and this expected value can be much larger than nq^l if the degrees differ.

It transpires that two properties of a randomly chosen part are needed to make things work: both the part itself, and its set of non-neighbours, must be spread uniformly throughout the vertices of different degrees; that is, these sets must contain their fair share of the vertices of each degree, in a sense that can be made precise. All but $o(t)$ of the parts, which can be discarded, have both these two properties, and between the remaining parts, all but $o(t)$ of the desired edges materialize, and so we can proceed according to our initial strategy. (In the proof given in [34], the parts are in fact chosen at random only from those that are spread uniformly through the vertices, and so only the spread of the non-neighbours is an issue. On the other hand, in the proof given in [28] of Theorem 6.2 below, which extends Theorem 4.1 to general H, the parts are chosen entirely at random.)

4.1. The extremal function $c(t)$

The remaining half of Theorem 1.1, that is, the upper bound on $c(t)$, can be derived from Theorem 4.1 in this way. Writing $d = \alpha t\sqrt{\log t}$, it is enough to show that if G is minimal in the class $\{ G : e(G) \geq d|G| \}$ then $G \succ K_t$. This minimal graph G is either small and dense, or sparse but large. In the first case, Theorem 4.1 implies straightaway that $G \succ K_t$. In the second case, we can assume by the arguments of §2 that G is reasonably well connected and that each edge is in at least d triangles. A few judicious applications of Theorem 4.1 then produce a large number of small minors that can be combined to form a K_t minor. In fact, a minor much larger than K_t can be formed, and from this it follows that extremal graphs arise only from the first case, and they are therefore essentially disjoint unions of small dense pieces.

4.2. Directed graphs

All the above arguments can be made to work for directed graphs, where the minor being sought is DK_t, the complete directed graph of order t with an edge in each direction between each pair of vertices. The extremal digraphs turn out just to be those obtained from the undirected case by replacing each edge by a double edge — details are in [34].

5. Pseudo-randomness and Sós's question

As indicated in the §4.1, the extremal graphs for the function $c(t)$ are formed by first taking random-like graphs of the appropriate order and density, and then forming as large a graph as desired by taking (almost) disjoint unions of the random-like pieces. Thus extremal graphs must be looked for in the class of pseudo-random, or quasi-random, graphs as discussed by Chung, Graham and Wilson [4] or in [33].

Now it is not true that all pseudo-random graphs behave as well as random graphs in terms of not having large minors. In fact, in [35] it is shown that most of the standard examples of pseudo-random graphs with n vertices have complete minors with $\Theta(n)$ vertices, compared with only $\Theta\big(n/\sqrt{\log n}\,\big)$ for random graphs. Indeed, Mader's request [25] for an explicit graph whose largest complete minor has $o(n)$ vertices remains unanswered; in general it seems hard to find a graph G whose largest minor has $o\big(\delta(G)\big)$ vertices. Alon [1] has nevertheless shown that random Cayley graphs have minors no larger than $\Theta\big(n/\sqrt{\log n}\,\big)$.

Sós has expressed the connection between the extremal theorems and quasi-randomness in the most succinct way. Although quasi-randomness does not preclude the presence of large minors, she asked whether quasi-randomness is necessary for the absence of large minors. To be precise, she asked whether a graph of density p and order $t\sqrt{\log_{1/q} t}$, and having no K_t minor, must necessarily be quasi-random.

The standard arguments about quasi-random graphs, even when properly quantified, are not quite strong enough to answer Sós's question. The issue has been settled by Myers [26] in the following way (at the same time giving a more precise description of the extremal graphs for the function $c(t)$.)

To understand Myers' theorem, consider a graph G whose vertex set is partitioned into two sets, X and Y, and define the three densities

$$p_X = \frac{e(X)}{\binom{|X|}{2}}, \qquad p_{XY} = \frac{e(X,Y)}{|X|\,|Y|}, \qquad p_Y = \frac{e(Y)}{\binom{|Y|}{2}}$$

where $e(X)$, $e(Y)$ and $e(X,Y)$ are the numbers of edges of G spanned by X, spanned by Y and joining X to Y. Likewise define $q_X = 1 - p_X$, $q_{XY} = 1 - p_{XY}$ and $q_Y = 1 - p_Y$. It is the principal feature of quasi-random graphs that G is quasi-random if and only if $p_{X'}$ differs little from p_X for every X' with $|X'| = |X|$, which of course implies that each of p_X, p_{XY} and p_Y are close to p, the density of G. Note that, whether or not G is quasi-random, the density of G satisfies

$$q = x^2 q_X + 2x(1-x)q_{XY} + (1-x)^2 q_Y$$

if G is large, where $q = 1 - p$ and $x = |X|/|G|$.

Consider now a randomly generated graph $G(n, x, p_X, p_{XY}, p_Y)$, having n vertices partitioned into two sets X and Y, where $|X| = xn$; the edges are chosen independently, with probability p_X inside X, p_{XY} between X and Y and p_Y inside Y. The proof of Theorem 3.1 is readily modified to show that the threshold value of n at which a K_t minor almost surely appears in $G(n, x, p_X, p_{XY}, p_Y)$ is

$$n = \big(1 + o(1)\big)\, t\sqrt{\log_{1/q^*} t} \qquad \text{where} \qquad q^* = q_X^{x^2}\, q_{XY}^{2x(1-x)}\, q_Y^{y^2}.$$

By taking logarithms and applying Jensen's inequality it can be seen that

$$q \geq q^*$$

with equality if and only if $q_X = q_{XY} = q_Y = q$.

Thus, so far as the sizes of complete minors are concerned, the constrained random graph $G(n, x, p_X, p_{XY}, p_Y)$ of density $1 - q$ behaves like the ordinary but denser random graph $G(n, 1 - q^*)$.

We can now state Myers' generalization of Theorem 4.1.

Theorem 5.1 (Myers [26]). *Given $\varepsilon > 0$ there exists $T = T(\varepsilon)$ with the following property. Let $t > T$, let $\varepsilon < p < 1 - \varepsilon$, let $q = 1 - p$ and let $n = \left\lceil (1 + \varepsilon)\, t\sqrt{\log_{1/q} t} \right\rceil$. Let G be a graph of order n and connectivity $\kappa(G) \geq n(\log\log\log n)/(\log\log n)$, having a vertex partition into X and Y as described above, where $\varepsilon < q_X, q_{XY}, q_Y \leq 1$ and $q^* < 1-\varepsilon$. Then $G \succ K_s$ where*

$$s = \left\lceil \sqrt{\frac{\log(1/q^*)}{\log(1/q)}}\; t \right\rceil .$$

In other words, a graph G with a partition as described will have complete minors at least as large as those found in $G(n, 1 - q^*)$. It follows immediately that if a graph as described in Theorem 4.1 has no minor significantly larger than K_t then q_X is approximately equal to q for every subset X of size xn, implying that G is quasi-random.

The proof of Theorem 5.1 is similar to that of Theorem 4.1, except that the vertices of X and Y are ordered separately, and the parts W_u are chosen so that each is sure to contain a representative sample of both X and Y. The principal difficulty is that the ordering of X, say, must respect the number of neighbours a vertex has both in X and in Y; however, by ordering with respect to a certain subtle parameter, a suitable linear order can be effected.

6. The extremal problem for general H

In this section we describe what is known about the function $c(H)$ for general H. Up until recently nothing was known, but although the situation at the time of writing is still a little fluid, the following description should be fairly accurate. Throughout this section t will stand for the number of vertices of H.

We would like to answer the following questions: (a) how does the function $c(H)$ behave, (b) is there some reasonable structural property that determines its value and (c) do the extremal graphs continue to be pseudo-random?

The answer to these questions appears to be that the function $c(H)$ behaves very similarly to $c(t)$ (indeed, for most graphs H, $c(H)$ is indistinguishable from $c(t)$) and that, at least for graphs with more than $t^{1+\varepsilon}$ edges, the extremal graphs behave in much the same way as before. When asking for a structural property that determines $c(H)$ we have in mind the classical situation of the Erdős-Stone-Simonovits theorem [8], in which the extremal function (for whether H must appear as an ordinary subgraph) is determined by the chromatic number of H.

The fact that the extremal graphs here are pseudo-random, however, makes the situation more complicated than the classical case, for two reasons. First of all, the results must necessarily be of an asymptotic kind (that is, as $|H| \to \infty$, as opposed to the classical case where perhaps $n \to \infty$ but H is allowed to be fixed). Secondly, the extremal function will be insensitive to small changes in the structure of H, such as the addition of an edge, or a handful of edges. This is because such a change in H will have a negligible effect on whether H appears as a minor of a random graph, and random graphs are the extremal graphs. This insensitivity to change is in marked contrast to the classical case, where of course the addition of a single edge can increase the chromatic number and so dramatically affect the extremal function.

As evidenced by Theorem 1.3, $c(H)$ can be described in terms of the parameter $\gamma(H)$ defined in §1.2. The implication of the previous remarks is that some leeway is possible in the definition; if $\gamma'(H)$ were another parameter with $\gamma'(H) = \gamma(H) + o(1)$, where $o(1)$ denotes something tending to zero as $t \to \infty$, then $\gamma'(H)$ could be used just as well as $\gamma(H)$ in all the results. The definition given is chosen because it seems to be the cleanest one that works, and its form is easily related to the appearance of H as a minor in $G(n,p)$, as we noted in §3.

6.1. General H minors

Here are two theorems that generalize Theorems 3.1 and 4.1 to general H. The way we state them, though, is slightly different to before.

Theorem 6.1 ([28]). *Given $\varepsilon > 0$ there exists $T = T(\varepsilon)$ with the following property.*

Let H be a graph with $t > T$ vertices and with $\gamma(H) \geq \varepsilon$. Let $\varepsilon \leq p \leq 1-\varepsilon$, let $q = 1-p$ and let $n = \left\lfloor \gamma(H)\, t \sqrt{\log_{1/q} t} \right\rfloor$. Then H is a minor of a random graph $G(n, p-\varepsilon)$ with probability less than ε.

The essence of the proof of this theorem has already been given in §3. More work is needed to prove the next theorem, in which the density of G, as usual, means $|E(G)|/\binom{n}{2}$.

Theorem 6.2 ([28]). *Given $\varepsilon > 0$ there exists $T = T(\varepsilon)$ with the following property.*

Let H be a graph with $t > T$ vertices and with $\gamma(H) \geq \varepsilon$. Let $\varepsilon \leq p \leq 1-\varepsilon$, let $q = 1-p$ and let $n = \left\lfloor \gamma(H)\, t \sqrt{\log_{1/q} t} \right\rfloor$. Let G be a graph of order n, density $p+\varepsilon$ and connectivity $\kappa(G) \geq n(\log\log\log n)/(\log\log n)$. Then H is a minor of G.

Theorems 3.1 and 4.1 show that the threshold probability p at which an H minor appears in $G(n,p)$ is the threshold density at which H minors appear in every reasonably connected graph of density p. This fact is at the heart of why Theorem 1.3 is true.

The modification to the proof of Theorem 4.1 needed to prove Theorem 6.2 is that the size of the parts W_u varies, being in fact proportional to the optimal weight $w(u)$ that determines $\gamma(H)$. This is the reason behind the change of approach remarked upon in §4.

Arguments similar to those in §4.1, in particular the separate treatment of dense and sparse minimal graphs and the finding of large complete minors in sparse minimal graphs, can be used to derive the extremal function $c(H)$ from Theorem 6.2, so proving Theorem 1.3. The discussion in §5 can also be carried over to general H minors, showing that, apart from a change in constants, the extremal graphs have the same pseudo-random structure as they do when H is complete.

6.2. Estimating $\gamma(H)$

It is straightforward to evaluate $\gamma(H)$ when H is complete or complete bipartite, but otherwise it appears to be difficult. We know, though, that if H has $t^{1+\tau}$ edges then assigning weight $\sqrt{\tau}$ to every vertex shows that $\gamma(H) \le \sqrt{\tau}$. Suppose that w is an optimal weighting of $V(H)$ that realizes $\gamma(H)$. Then there cannot be a significant proportion of edges uv such that $w(u)w(v) < \tau$. So, if we group together vertices of roughly equal weight, there will be almost no edges between the class containing u and the class containing v if $w(u)w(v) < \tau$. This leads us to approximate H as a subgraph of a blowup of a small graph, in the following way.

A *shape* is defined to be a pair (F, f), where F is a graph (in which loops, but not multiple edges, are allowed) and $f : V(F) \to \mathbf{R}^+$ is a function assigning non-negative numbers to the vertices such that $\sum_{a \in V(F)} f(a) = 1$. We say that the graph H of order t is an *ε-fit* to shape (F, f) if there is a partition of $V(H)$ into sets V_a, $a \in V(F)$, such that $\lfloor f(a)t \rfloor \le |V_a| \le \lceil f(a)t \rceil$, and

$$\left|\left\{ uv \in E(H) \,:\, u \in V_a,\ v \in V_b \text{ and } ab \notin E(F) \right\}\right| \le t^{-\varepsilon} \left|E(H)\right|.$$

So H is an ε-fit to (F, f) if there is a partition of H into classes indexed by $V(F)$ and of sizes proportional to f, so that all but a tiny fraction of the edges of H lie between classes corresponding to edges of F. The fact that F might have loops allows H to have edges within the corresponding classes; in particular, every H fits the shape consisting of a single vertex with a loop.

The parameter of the shape (F, f) that is related to $\gamma(H)$ is the parameter $m(F, f)$, given by

$$m(F, f) = \max_{x \cdot f = 1} \min_{ab \in E(F)} x(a)x(b).$$

Here the maximum is over all functions $x \in [0, \infty)^{V(F)}$ of $V(F)$, and $x \cdot f$ stands for the standard inner product $\sum_{a \in F} x(a)f(a)$. This definition allows $x(a) > 1$ even though we always have $f(a) \le 1$. The constant function $x(a) = 1$ satisfies $x \cdot f = 1$ and so $m(F, f) \ge 1$ always holds. Also, if F has a single vertex a with a loop then $f(a) = 1$ and $m(F, f) = 1$.

Some calculation then supplies the crucial fact that, if H has $t^{1+\tau}$ edges, then H is an ε-fit to some shape (F, f) with $|F| \le (1/\varepsilon)$ and $\gamma(H) \ge \sqrt{\tau/m(F,f)} - 4\sqrt{\varepsilon}$. So a lower bound on $\gamma(H)$ can be given by checking that H is not an ε-fit to any small shape (F, f) with $m(F, f)$ large. In so doing it is necessary only to check *critical* shapes: these are shapes (F, f) for which $m(F', f') < m(F, f)$ for any F' resulting from F either by the addition of an edge or by the *merger* of two vertices of F. (The merger of $a, b \in F$ is the replacement of a and b by a single vertex c joined to every vertex previously joined to either a or b, with $f'(c) = f(a)+f(b)$ and $f' = f$ on the other vertices of F'.) This is because if H is an ε-fit to (F, f) then it is also an ε-fit to (F', f').

What makes these observations useful is that the check required is quite short; there are very few critical shapes, and we can describe them explicitly.

Theorem 6.3 ([28]). *A shape (F, f) with $|F| = k+1$ is critical if and only if F is the* half-graph *of order $k+1$ that is,*

$$V(F) = \{0, 1, \ldots, k\} \qquad \text{and} \qquad E(F) = \{ij : i + j \ge k\},$$

and moreover f satisfies

$$\frac{f(k)}{f(0)} < \frac{f(k-1)}{f(1)} < \frac{f\big(k - \lfloor (k-1)/2 \rfloor\big)}{f\big(\lfloor (k-1)/2 \rfloor\big)} < 1.$$

For these shapes,

$$m(F, f) = \left\{ \sum_{i=0}^{k} \sqrt{f(i) f(k-i)} \right\}^{-2}.$$

So, if we know the structure of H, it is fairly easy to check whether H is an ε-fit to a small critical shape, and hence to get a lower bound on $\gamma(H)$. The simplest, and commonest, case is where H fails to fit any shape apart from the shape with one vertex and a loop. This case can be reformulated in the statement that H has a *tail*, which is a large subset T whose neighbours lie almost entirely inside a smaller subset S; here is a precise version.

Theorem 6.4 ([28]). *Let $\varepsilon > 0$ and let H be a graph of order $t \ge 1/\varepsilon^2$ with $t^{1+\tau}$ edges such that $\gamma(H) \le \sqrt{\tau} - 5\sqrt{\varepsilon}$. Then H has an ε-*tail *— that is, $V(H)$ has a partition $R \cup S \cup T$, with $|T| > |S| + \varepsilon t$, such that $\big|E(T, T \cup R)\big| \le t^{1+\tau-\varepsilon}$.*

Now regular graphs cannot have a tail, nor indeed can graphs that are almost regular, and this includes almost all graphs. We have the following conclusion.

Corollary 6.5. *All regular graphs and almost all graphs H of order t with $t^{1+\tau}$ edges have $\gamma(H) = \sqrt{\tau} + o(1)$.*

As a further corollary we can evaluate $\gamma(H)$ for, for example, complete multipartite graphs; these all have $\gamma(H) \approx 1$ unless the largest part has size βt with $\beta > 1/2$, in which case $\gamma(H) \approx \sqrt{4\beta(1-\beta)}$.

It should be pointed out, however, that this method for approximating $\gamma(H)$ can sometimes give a bound much less than the correct value. This is because the property of being an ε-fit to a shape is insensitive to the introduction of a very sparse subgraph H^*, though this subgraph might be what actually determines $\gamma(H)$. The situation is analogous to that in the classical extremal theory where the chromatic number of H might be determined by $\chi(H^*)$ and not just by the chromatic number of some dense subgraph. An example is when H is the union of $K_{t/8,7t/8}$ with a $t^{1/2}$-regular graph H^* on the same vertex set. We know that $\gamma(K_{t/8,7t/8}) = \sqrt{7}/4 + o(1)$ whereas $\gamma(H^*) = 1/\sqrt{2} + o(1)$. So $\gamma(H) \geq \max(\sqrt{7}/4, 1/\sqrt{2}) + o(1) = 1/\sqrt{2} + o(1)$. But, for every $\varepsilon > 0$, if t is large this graph is an ε-fit to a two vertex shape with $f = (1/8, 7/8)$ and $m(F, f) = 16/7$, so our lower bound method gives only $\gamma(H) \geq \sqrt{7}/4 + o(1)$.

We conclude this section with another lower bound on $\gamma(H)$ based just on the density of the graph. This shows that $\gamma(H)$ can never be close to zero for graphs of positive density.

Theorem 6.6 ([28]). *Let H be a graph of order $t \geq (1/\varepsilon)^{1/\varepsilon}$ and density p. Then $\gamma(H) \geq p - 5\sqrt{\varepsilon}$.*

7. Linking

A graph G is said to be *k-linked* if, for any sequence $s_1, \ldots, s_k, t_1, \ldots, t_k$ of distinct vertices, we can find s_i–t_i paths P_i that are disjoint, $1 \leq i \leq k$. Larman and Mani [20] and Jung [13] noticed that if $\kappa(G) \geq 2k$ and G contains a subdivided complete graph of order $3k$ then G is k-linked. Mader [22] proved that if the average degree of a graph exceeds $2^{\binom{k}{2}}$ then it contains a subdivided K_k, and so, if $\kappa(G)$ is sufficiently large, G is k-linked. (For a survey of subdivisions of graphs, see Mader [24].)

Robertson and Seymour [30], as part of their deep study of graph minors, established a connection between linking and graph minors; they strengthened the above remarks by showing that G is k-linked if $\kappa(G) \geq 2k$ and $G \succ K_{3k}$. It follows from Theorem 1.1 that the connectivity required to force k-linking is only $O(k\sqrt{\log k})$.

Bollobás and Thomason [3] weakened the condition $G \succ K_{3k}$ still further to $G \succ H$ where H is *any* graph such that $2\delta(H) \geq |H| + 4k - 2$. In consequence of Lemma 2.2 they could then show that G is k-linked provided $\kappa(G) \geq 22k$.

The reason we point this out in this survey is to contrast the average degree required to obtain some *specific* H with $2\delta(H) \geq |H| + 4k - 2$, with that needed to achieve just *some* H. By Theorem 1.3 and Theorem 6.6 the former would still require average degree $\Theta(k\sqrt{\log k})$, whereas Lemma 2.2 shows the latter to hold given average degree only $\Theta(k)$.

Added in proof. Thomas and Wollan have recently shown that G is k-linked if $\kappa(G) \geq 10k$.

8. Minors and Girth

The simple fact underlying the observations in §2 is that contracting an edge of a graph tends to increase the average degree unless the edge lies in many triangles. In particular, if a graph has large girth then many edges can be contracted, each contraction increasing the average degree.

Thomassen [36] made a systematic study of this phenomenon — his aim was to show that many consequences of a graph having large average degree could be derived also for graphs having minimum degree only three but having large girth. His fundamental tool was the following theorem, whose simple and elegant proof we include here. We use $g(G)$ to denote the girth of G.

Theorem 8.1 (Thomassen [36]). *If $\delta(G) \geq 3$ and $g(G) \geq 4k - 5$ then $G \succ H$ where $\delta(H) \geq k$.*

Proof. We may assume $k \geq 4$. Take a partition $A_1, \ldots, A_t$ of $V(G)$ with t maximal such that $G[A_i]$ is connected and $|A_i| \geq 2k-3$ for $1 \leq i \leq t$. If $G[A_i]$ contains a cycle C, then $|C| \geq 4k-5$, so by splitting C into two paths we can partition A_i into A_i^1 and A_i^2, with $G[A_i^l]$ is connected and $|A_i^l| \geq 2k-3$ for $l = 1, 2$; the maximality of t thus implies $G[A_i]$ must in fact be a tree. Suppose now we could find A_i and A_j with $1 \leq i < j \leq t$ for which there were three edges between A_i and A_j. Then we could find vertices $u \in A_i$ and $v \in A_j$ together with three disjoint u–v paths P_1, P_2, P_3 in $G[A_i \cup A_j]$. Any two of these paths have at least $4k-5$ edges between them and so in particular two of them, say P_1 and P_2, must have length at least $2k-2$. So we could partition $A_i \cup A_j$ into three sets A^1, A^2, A^3, with A^l containing $2k-3$ vertices from $P_l - \{u, v\}$, $l = 1, 2$, and A^3 containing the rest of $P_1 \cup P_2 \cup P_3$, such that $G[A^l]$ is connected and $|A^l| \geq 2k-3$ for $l = 1, 2, 3$. Hence the maximality of t implies that there are at most two edges between A_i and A_j for $1 \leq i < j \leq t$.

Now, of course, we contract each A_i to a single vertex a_i. In the resultant multigraph H^*, every pair of vertices is joined by at most two edges; throw away one edge from each double edge to obtain a graph H. The degree of a vertex a_i in H^* is at least $3|A_i| - 2(|A_i| - 1) \geq 2k-1$, and so its degree in H is at least $\lceil (2k-1)/2 \rceil = k$, as desired. ■

Diestel and Rempel [5] have reduced the girth required here to $6 \log_2 k + 4$. More recently, Kühn and Osthus [18] reduced it to $4 \log_2 k + 27$. They obtained results close to best possible for minors with specified minimum degree and girth; an example is this.

Theorem 8.2 (Kühn and Osthus [18]). *Let $k \geq 1$ and $d \geq 3$ be integers, and let $g = 4k+3$. If $g(G) \geq g$ and $\delta(G) \geq d$ then $G \succ H$ where $\delta(H) \geq (d-1)^{(g+1)/4}/48$.*

As a further consequence of their methods they also show that Hadwiger's conjecture holds for graphs of girth at least 19 (Kawarabayashi [14] also found this result).

One natural way of weakening the constraint of large girth is to forbid $K_{s,s}$ as a subgraph, in the hope that this constraint still yields complete minors in graphs of low average degree. (Note that forbidding a non-bipartite subgraph will not help, since the extremal graphs for complete minors contain bipartite subgraphs with at least half as many edges.) Kühn and Osthus [19] have investigated this condition, obtaining the following result, which is again close to best possible provided a standard conjecture about the extremal function for $K_{s,s}$ is true..

Theorem 8.3 (Kühn and Osthus [19]). *Given $s \geq 2$ there exists a constant $c = c(s)$, such that every $K_{s,s}$-free graph of average degree at least r has a K_t minor for $t = \lfloor cr^{1+2/(s-1)}(\log r)^{-3} \rfloor$.*

As might be expected, the proofs of these results are much more substantial than the proof of Theorem 8.1.

9. Minors and connectivity

Large average degree is the simplest property forcing a graph to have a K_t minor. Robertson and Seymour, in their series of papers on Graph Minors, have investigated more complex structural properties that give rise to minors; one of their fundamental results [29] is that a graph has large tree-width if and only if it contains a large grid minor. Diestel, Jensen, Gorbunov and Thomassen [6] gave a short proof of this result, and introduced the notion of *external connectivity*: a set $X \subset V(G)$ is *externally k-connected* if $|X| \geq k$ and for all subsets $Y, Z \subset X$ with $|Y| = |Z| = k$ there are $|Y|$ disjoint Y–Z paths in G without inner vertices or edges inside X. A large grid that has high external connectivity yields a large complete minor; Kühn [17] has shown that the same conclusion holds even if the large grid is replaced by a large number of large disjoint binary trees, each having an extra vertex joined to its leaves.

There is a simple, and as yet unsolved, problem relating (ordinary) connectivity to complete minors. What connectivity is needed to force a K_t minor? Since $\kappa(G) \leq \delta(G)$ with equality for random graphs, the answer to this question is $\big(2\alpha + o(1)\big)t\sqrt{\log t}$, by Theorem 1.1. But the only examples achieving this are pseudo-random graphs of bounded (in t) order; the extremal graphs of larger order for Theorem 1.1 have very low connectivity. It might well be that, for graphs of large order, a lower connectivity will suffice for a K_t minor. We therefore make the following conjecture.

Conjecture 9.1. There is an absolute constant C and a function $n(t)$ such that if $|G| \geq n(t)$ and $\kappa(G) \geq Ct$ then $G \succ K_t$.

Perhaps even $\kappa(G) \geq t+1$ is enough (though $\kappa(G) = t$ is not, as a 5-connected planar graph joined to K_{t-5} shows). For $t = 6$ Jørgensen [12] (see also [31]) has a related conjecture, that every 6-connected graph with no K_6 minor has a vertex joined to all the others.

Myers [27] has a partial result in this area; if t is odd, a $(t+1)$-connected graph G, with a long sequence of cutsets $S_1, S_2, \ldots$ of size $t+1$ such that S_j separates $S_1, \ldots, S_{j-1}$ from $S_{j+1}, S_{j+2}, \ldots$, has a K_{t-3} minor if the $G[S_j]$'s are 2-edge-connected.

Added in proof. Böhme, Kawarayabashi, Maharry and Mohar have recently shown that every large $23t$-connected graph has a K_t minor.

References

[1] N. Alon, personal communication.

[2] B. Bollobás, P. Catlin and P. Erdős, Hadwiger's conjecture is true for almost every graph, *Europ. J. Combin. Theory,* **1** (1980), 195–199.

[3] B. Bollobás and A. Thomason, Highly linked graphs, *Combinatorica* **16** (1996), 313–320.

[4] F. R. K. Chung, R. L. Graham and R. M. Wilson, Quasi-random graphs, *Combinatorica,* **9** (1989), 345–362.

[5] R. Diestel and C. Rempel, *Dense minors in graphs of large girth,* (preprint).

[6] R. Diestel, T. Jensen, K. Gorbunov and C. Thomassen, Highly connected sets and the excluded grid theorem, *J. Combinatorial Theory Ser. B,* **75** (1999), 61–73.

[7] G. A. Dirac, Homomorphism theorems for graphs, *Math. Ann.,* **153** (1964), 69–80.

[8] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.,* **1** (1966), 51–57.

[9] W. Fernandez de la Vega, On the maximum density of graphs which have no subcontraction to K^s, *Discrete Math.,* **46** (1983), 109–110.

[10] H. Hadwiger, Über eine Klassifikation der Streckenkomplexe, *Vierteljahresschr. Naturforsch. Ges. Zürich,* **88** (1943), 133–142.

[11] L. Jørgensen, Contractions to complete graphs, *Annals of Discrete Maths.,* **41** (1989), 307–310. **18** (1994), 431–448.

[12] L. Jørgensen, Contractions to K_8, *J. Graph Theory,* **18** (1994), 431–448.

[13] H. A. Jung, Verallgemeinerung des n-fachen zusammenhangs für Graphen, *Math. Ann.,* **187** (1970), 95–103.

[14] K. Kawarabayashi, *Hadwiger's conjecture on girth condition,* (pre-print).

[15] A. V. Kostochka, The minimum Hadwiger number for graphs with a given mean degree of vertices (in Russian), *Metody Diskret. Analiz.*, **38** (1982), 37–58.

[16] A. V. Kostochka, A lower bound for the Hadwiger number of graphs by their average degree, *Combinatorica*, **4** (1984), 307–316.

[17] D. Kühn, Forcing a K_r minor by high external connectivity, *J. Graph Theory*, **39** (2002), 241–264.

[18] D. Kühn and D. Osthus, *Minors in graphs of large girth* (preprint).

[19] D. Kühn and D. Osthus, *Complete minors in $K_{s,s}$-free graphs* (preprint).

[20] D.G. Larman and P. Mani, On the existence of certain configurations within graphs and the 1-skeletons of polytopes, *Proc. London Math. Soc.*, **20** (1970), 144–160.

[21] M. Las Vergnas and H. Meyniel, Kempe classes and the Hadwiger conjecture, *J. Combin. Theory Ser. B*, **31** (1981), 95–104.

[22] W. Mader, Homomorphieeigenschaften und mittlere Kantendichte von Graphen, *Math. Ann.*, **174** (1967), 265–268.

[23] W. Mader, Homomorphiesätze für Graphen, *Math. Ann.*, **178** (1968), 154–168.

[24] W. Mader, Topological minors in graphs of minimum degree n, Contemporary trends in discrete mathematics (Štiřín Castle, 1997), *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, **49**, Amer. Math. Soc., Providence, RI, (1999), 199–211.

[25] W. Mader, personal communication.

[26] J. S. Myers, Graphs without large complete minors are quasi-random, (to appear in *Combinatorics, Probability and Computing*).

[27] J. S. Myers, Extremal theory of graph minors and directed graphs, Ph.D. thesis, University of Cambridge (2002).

[28] J. S. Myers and A. Thomason, *The extremal function for non-complete minors* (in preparation).

[29] N. Robertson and P. D. Seymour, Graph Minors. V. Excluding a planar graph, *J. Combin. Theory Ser. B*, **41** (1986), 92–114.

[30] N. Robertson and P. D. Seymour, Graph Minors. XIII. The disjoint paths problem, *J. Combin. Theory Ser. B*, **63** (1995), 65–110.

[31] N. Robertson, P. Seymour and R. Thomas, Hadwiger's conjecture for K_6-free graphs, *Combinatorica*, **13** (1993), 279–361.

[32] A. Thomason, An extremal function for contractions of graphs, *Math. Proc. Cambridge Phil. Soc.*, **95** (1984), 261–265.

[33] A. Thomason, Pseudo-random graphs, in: *Random Graphs '85* (M. Karonski and Z. Palka, Eds), Annals of Discrete Math., **33** (1987), 307–331.

[34] A. Thomason, The extremal function for complete minors, *J. Combinatorial Theory Ser. B*, **81** (2001), 318–338.

[35] A. Thomason, Complete minors in pseudo-random graphs, *Random Structures and Algorithms*, **17** (2000), 26–28.

[36] C. Thomassen, Girth in graphs, *J. Combinatorial Theory Ser. B,* **35** (1983), 129–141.

[37] B. Toft, A survey of Hadwiger's conjecture, Surveys in graph theory (San Francisco, CA, 1995), *Congr. Numer.,* **115** (1996), 249–283.

[38] K. Wagner, Über eine Eigenschaft der ebenen Komplexe, *Math. Ann.,* **114** (1937), 570–590.

[39] K. Wagner, Beweis einer Abschwächung der Hadwiger-Vermutung, *Math. Ann.,* **153** (1964), 139–141.

Andrew Thomason

Department of Pure Mathematics and
Mathematical Statistics
Centre for Mathematical Science
Wilberforce Road
Cambridge CB3 0WB, U.K.

A.G.Thomason@dpmms.cam.ac.uk