A Separator Theorem for String Graphs and its Applications

JACOB FOX[1†] and JÁNOS PACH[2‡]

[1]Department of Mathematics, Princeton, NJ 08544, USA
(e-mail: jacobfox@math.princeton.edu)

[2]EPFL Lausanne, IMB-DCG, Station 8, CH-1015, Switzerland
and
Rényi Institute, Budapest, H-1364, Hungary
(e-mail: pach@cims.nyu.edu)

A *string graph* is the intersection graph of a collection of continuous arcs in the plane. We show that any string graph with m edges can be separated into two parts of roughly equal size by the removal of $O(m^{3/4}\sqrt{\log m})$ vertices. This result is then used to deduce that every string graph with n vertices and no complete bipartite subgraph $K_{t,t}$ has at most $c_t n$ edges, where c_t is a constant depending only on t. Another application shows that locally tree-like string graphs are globally tree-like: for any $\epsilon > 0$, there is an integer $g(\epsilon)$ such that every string graph with n vertices and girth at least $g(\epsilon)$ has at most $(1 + \epsilon)n$ edges. Furthermore, the number of such labelled graphs is at most $(1 + \epsilon)^n T(n)$, where $T(n) = n^{n-2}$ is the number of labelled trees on n vertices.

1. Introduction

A large part of computational geometry deals with representation and manipulation of various geometric objects. Special attention is paid to pairs of objects that are in contact with each other: detecting intersections among line segments, for example, belongs to the oldest and best-studied chapter of computational geometry, already addressed in the first monograph devoted to the subject [37]. Yet, even in the special case of segments, little is known about elementary structural properties of the arising intersection patterns. The recognition of such intersection patterns (intersection graphs) is known to be NP-hard [21, 22].

Given a collection $C = \{\gamma_1, \ldots, \gamma_n\}$ of arcwise connected sets in the plane, their *intersection graph* $G = G(C)$ is a graph on the vertex set C, where γ_i and γ_j ($i \neq j$) are connected by an edge if and only if $\gamma_i \cap \gamma_j \neq \emptyset$. It is easy to show that every such intersection graph can be obtained as an intersection graph of a collection of (simple) continuous curves in the plane. Therefore, the intersection graphs of arcwise connected sets in the plane are often called *string graphs*.

Given a graph $G = (V, E)$ with vertex set V and edge set E, a *weight function* $w : V \to \mathbb{R}_{\geq 0}$ is a non-negative function on the vertex set such that the sum of the weights is at most 1. The *weight of a subset* $S \subseteq V$, denoted by $w(S)$, is defined as $\sum_{v \in S} w(v)$.

A *separator* in a graph $G = (V, E)$ with respect to a weight function w is a subset $S \subseteq V$ for which there is a partition $V = S \cup V_1 \cup V_2$ such that $w(V_1), w(V_2) \leq 2/3$ and there is no edge between V_1 and V_2. If the weight function is not specified, it is assumed that $w(v) = 1/|V|$ for every vertex $v \in V$.

The Lipton–Tarjan separator theorem [26] states that, for every planar graph G with n vertices and for every weight function w for G, there is a separator of size $O(n^{1/2})$. This has been generalized in various directions: to graphs embedded in a surface of bounded genus [16], graphs with a forbidden minor [1], intersection graphs of balls in $\mathbb{R}^d$ [29], intersection graphs of Jordan regions [11], and intersection graphs of convex sets in the plane [11]. Our main result is a separator theorem for string graphs.

Theorem 1.1. *For every string graph G with m edges and for every weight function w for G, there is a separator of size $O\left(m^{3/4}\sqrt{\log m}\right)$ with respect to w.*

We do not believe that the bound on the separator size in Theorem 1.1 is tight. In fact, as in [14], we make the following conjecture.

Conjecture 1.2. *Every string graph with m edges has a separator of size $O(\sqrt{m})$.*

This conjecture is known to be true in several special cases: (1) for intersection graphs of convex sets in the plane with bounded clique number [11], (2) for intersection graphs of curves, any pair of which has a bounded number of intersection points [11], and (3) for *outerstring graphs*, that is, intersection graphs of collections C of curves with the property that there is a suitable curve γ such that each member of C has one endpoint on γ, but is otherwise disjoint from it [12].

Separator theorems have many important applications (see, *e.g.*, [25] and [27]). Despite the apparent weakness of the bound in Theorem 1.1, it is still strong enough to yield some interesting corollaries.

For any graph H, a graph G is called *H-free* if it does not have a (not necessarily induced) subgraph isomorphic to H. Given H and a positive integer n, the *extremal number* $ex(H, n)$ is defined as the maximum number of edges over all H-free graphs on n vertices. The study of this parameter is a classical area of Turán-type *extremal graph theory*: see [3]. The problem of investigating the same maximum restricted to intersection graphs of arcwise connected sets, convex bodies, segments, *etc.*, was initiated in [34]. For partial results in this directions, see [34], [38], [11].

In the present paper, we use Theorem 1.1 to prove that for any bipartite graph H, there is a constant c_H such that every H-free intersection graph of n arcwise connected sets in the plane has at most $c_H n$ edges. Clearly, it is sufficient to prove this statement for *balanced complete* bipartite graphs $H = K_{t,t}$, as every bipartite graph with t vertices is a subgraph of $K_{t,t}$.

Theorem 1.3. *For any positive integer t, every $K_{t,t}$-free string graph with n vertices has at most $t^{c \log \log t} n$ edges, where c is an absolute constant.*

A graph G is called *d-degenerate* if every subgraph of G has a vertex of degree at most d. Every d-degenerate graph has chromatic number at most $d + 1$. Theorem 1.3 implies that every $K_{t,t}$-free intersection graph of arcwise connected sets in the plane is $2t^{c \log \log t}$-degenerate. Thus, we obtain the following result.

Corollary 1.4. *For any positive integer t, the chromatic number of every $K_{t,t}$-free intersection graph of n arcwise connected sets in the plane is at most $2t^{c \log \log t} + 1$.*

In [11], it was shown that every $K_{t,t}$-free intersection graphs of n curves, no pair of which has more than a fixed constant number of points in common, has at most $c_t n$ edges, where the dependence on t is exponential. In this case, our separator-based approach gives a tight bound. In Section 6, we establish the following result.

Theorem 1.5. *Let k and t be positive integers. There exists a constant C_k depending only on k, such that the maximum number of edges of any $K_{t,t}$-free intersection graph G of n curves in the plane, no pair of which has more than k points in common, is at most $C_k t n$. Apart from the value of the constant C_k, this bound cannot be improved.*

A collection of curves in the plane is called a collection of *pseudo-segments* if no two of them has more than one point in common. The *girth* of a graph is the length of its shortest cycle. Kostochka and Nešetřil [19] proved that for any $\epsilon > 0$, there is a positive integer $g(\epsilon)$ such that the intersection graph of any collection of pseudo-segments with girth at least $g(\epsilon)$ has at most $(1 + \epsilon)n$ edges. Using our separator theorem, Theorem 1.1, this statement can be extended to all string graphs.

Theorem 1.6. *For any $\epsilon > 0$, there is a positive integer $g(\epsilon)$ such that every string graph on n vertices with girth at least $g(\epsilon)$ has at most $(1 + \epsilon)n$ edges.*

In particular, this theorem implies that there exists a positive integer g_0 such that every string graph with girth at least g_0 has chromatic number at most 3. It would be interesting to determine the smallest such integer g_0.

We mention another application of Theorem 1.1. The *bandwidth* of a graph G with n vertices is the minimum b for which there is a labelling of the vertices of G by distinct integers such that the labels of adjacent vertices differ by at most b. Chung [6] showed that every tree with n vertices and maximum degree Δ has bandwidth $O(n/\log_\Delta n)$.

Böttcher, Pruessmann, Taraz and Würfl [4] used the separator theorem for planar graphs to extend this result to show that every planar graph with n vertices and maximum degree Δ has bandwidth $O(n/\log_\Delta n)$. Replacing the separator theorem for planar graphs by Theorem 1.1 in the proof of this result and using the bound in Theorem 1.3, we obtain the following extension to all string graphs with a forbidden bipartite subgraph.

Corollary 1.7. *Every $K_{t,t}$-free string graph with n vertices and maximum degree Δ has bandwidth at most $c_t n/\log_\Delta n$, where c_t only depends on t.*

In the next section, we prove an inequality bounding the pair-crossing number of a string graph by the number of short paths in the graph. We use this to deduce an upper bound on the bisection width of a string graph, and to obtain a proof of our separator theorem, Theorem 1.1, for the uniform weight function $w \equiv 1/|V|$. We then give a different proof for the full version of the separator theorem in Section 3. In Section 4, we apply the (weak) separator theorem to prove a qualitative version of Theorem 1.3, which states that $K_{t,t}$-free string graphs with n vertices have at most $c_t n$ edges. The proof of Theorem 1.3 is given in Section 5. In Section 6, we prove Theorem 1.5 and a similar result for intersection graphs of convex sets in the plane. In Section 7, we deduce Theorem 1.6 and two other results that can be obtained similarly. In the concluding remarks, we discuss the strength of the constant factor dependence on t in Theorem 1.3 and the asymptotic number of string graphs with a forbidden bipartite subgraph. Throughout the paper, we systematically omit floor and ceiling signs, whenever they are not crucial for the sake of clarity of the presentation. All logarithms in this paper are base 2 unless otherwise noted.

2. Crossing number, bisection width, and separators for string graphs

A *topological graph* is a graph drawn in the plane with vertices as points and edges as curves connecting its vertices. These curves are disjoint from the vertices except for their endpoints. The *pair-crossing number* $\mathrm{pcr}(G)$ of a graph G is the minimum number of pairs of edges that intersect in a drawing of G. The length of a path in a graph is the number of its edges. We prove the following upper bound on the pair-crossing number of string graphs.

Lemma 2.1. *If G is a string graph, then $\mathrm{pcr}(G)$ is at most the number of paths of length 2 or 3 in G.*

Proof. Let C be a collection of curves whose intersection graph is G. For each curve $\gamma \in C$, let $p(\gamma)$ be an arbitrary point on γ. For each pair $q = \{\gamma, \gamma'\}$ of distinct intersecting curves in C, let $\alpha(q)$ denote a curve which starts at $p(\gamma)$, goes along γ until it comes to an intersection point of γ and γ', and then continues along γ' until it ends at $p(\gamma')$. Note that this provides a drawing of G in the plane, in which the vertices are the points $p(\gamma)$ and the edges are the curves $\alpha(q)$. Suppose that two edges $\alpha(q_1)$ and $\alpha(q_2)$ in this drawing intersect. Since $\alpha(q_i)$ lies along the union of the two curves of which q_i is composed, one of the curves in q_1 intersects one of the curves in q_2. If q_1 and q_2 have a curve γ in

common, then we obtain a path of length two in G with middle vertex γ. Otherwise, q_1 and q_2 consist of distinct curves, and one of the curves in q_1 intersects one of the curves in q_2, which gives rise to a path of length three in G with these intersecting curves as the two middle vertices. □

The *crossing number* $\mathrm{cr}(G)$ of a graph G is the minimum number of edge crossings in any drawing of G. It is a challenging open problem [35] to determine whether $\mathrm{pcr}(G) = \mathrm{cr}(G)$ holds for every graph G. We can prove an inequality similar to Lemma 2.1 for intersection graphs of convex sets in the plane, replacing pair-crossing number by crossing number.

Lemma 2.2. *If G is an intersection graph of convex sets in the plane, then $\mathrm{cr}(G)$ is at most four times the number of paths of length 2 or 3 in G.*

Proof. Let C be a collection of convex sets in the plane whose intersection graph is G. For each convex set $\gamma \in C$, let $p(\gamma)$ be an arbitrary point in γ. For each pair $q = \{\gamma, \gamma'\}$ of convex sets in C that intersects, let $\alpha(q)$ be a curve which is a polygonal path consisting of two segments, the first segment starts at $p(\gamma)$ and ends at an intersection point of γ and γ', and the second segment starts at this intersection point and ends at $p(\gamma')$. Note that this provides a drawing of G in the plane: the vertices are the points of the form $p(\gamma)$ and the edges are the curves $\alpha(q)$. Just as in Lemma 2.1, the number of pairs of crossing edges in this drawing is at most the number of paths of length 2 or 3 in G. Since each edge in this drawing is a union of two segments, and pairs of segments can cross at most once, each pair of crossing edges has at most four crossings. □

The *bisection width* $b(G)$ of a graph $G = (V, E)$ is the least integer for which there is a partition $V = V_1 \cup V_2$ such that $|V_1|, |V_2| \leqslant 2|V|/3$ and the number of edges between V_1 and V_2 is $b(G)$. For any graph G, let $\mathrm{ssqd}(G) = \sum_{v \in V(G)} (\deg(v))^2$. Pach, Shahrokhi and Szegedy [33] used the Lipton–Tarjan separator theorem to show that

$$b(G)^2 = O\big(\mathrm{cr}(G) + \mathrm{ssqd}(G)\big).$$

Kolman and Matoušek [18] use a result of Leighton and Rao [24] on multicommodity flows to obtain the following analogous result for pair-crossing number.

Lemma 2.3 (Kolman and Matoušek [18]). *Every graph G on n vertices satisfies*

$$b(G) \leqslant c \log n \big(\sqrt{\mathrm{pcr}(G)} + \sqrt{\mathrm{ssqd}(G)}\big),$$

where c is an absolute constant.

Noting that $\mathrm{ssqd}(G)$ is twice the number of paths of length 1 or 2 in G, we have the following corollary of Lemmas 2.1 and 2.3.

Corollary 2.4. *Let G be a string graph on n vertices and let p denote the number of paths of length at most 3 in G. Then the bisection width of G satisfies*

$$b(G) = O(p^{1/2} \log n).$$

Replacing Lemma 2.1 by Lemma 2.2 and Lemma 2.3 by the result of Pach, Shahrokhi and Szegedy, we have that for G an intersection graph of convex sets in the plane, $b(G) = O(p^{1/2})$. This bound and Corollary 2.4 (up to the logarithmic factor) cannot be improved. Indeed, for the complete graph on n vertices, which is an intersection graph of segments, the number of paths of length at most three is $\Theta(n^4)$ while the bisection width is $\Theta(n^2)$.

Every graph $G = (V, E)$ has a separator with at most $b(G)$ vertices. Indeed, let $V = V_1 \cup V_2$ be a vertex partition with $|V_1|, |V_2| \leqslant 2|V|/3$ and $b(G)$ edges between V_1 and V_2. Let V_0 denote those vertices in V_1 which are adjacent to a vertex in V_2, so $|V_0| \leqslant b(G)$. Since $V = V_0 \cup (V_1 \setminus V_0) \cup V_2$ is a vertex partition, $|V_1 \setminus V_0|, |V_2| \leqslant 2|V|/3$, and there are no edges between $V_1 \setminus V_0$ and V_2, then V_0 is a separator for G.

Let m denote the number of edges of G and let $\Delta \geqslant 1$ denote the maximum degree of G. The number of paths of length 2 in G is

$$\sum_{v \in V(G)} \binom{\deg(v)}{2} \leqslant \frac{\Delta - 1}{2} \sum_{v \in V(G)} \deg(v) = m(\Delta - 1).$$

Each edge of G is the middle edge of at most $(\Delta - 1)^2$ paths of length 3 in G, so the number of paths of length 3 in G is at most $m(\Delta - 1)^2$. Putting these bounds together, the number of paths of length at most 3 is at most $m + m(\Delta - 1) + m(\Delta - 1)^2 \leqslant m\Delta^2$.

From Corollary 2.4 and the discussion in the previous two paragraphs, we get the following separator theorem for string graphs. We may assume that the string graph has no isolated vertices, so $\log n$ and $\log m$ are within a constant factor of each other.

Theorem 2.5. *Every string graph with $m \geqslant 2$ edges and maximum degree Δ has a separator of size*

$$O\left(\Delta m^{1/2} \log m\right).$$

We can quickly deduce an unweighted version of Theorem 1.1 from this result. Indeed, let G be a string graph with m edges and $d = m^{1/4}/\sqrt{\log m}$. Let V_0 consist of those vertices of degree at least d together with the vertices of a smallest separator in the remaining induced subgraph. The number of vertices of degree at least d is at most $2m/d = 2m^{3/4}\sqrt{\log m}$. The remaining induced subgraph has at most m edges and maximum degree at most d, so Theorem 2.5 implies it has a separator of size $O\left(dm^{1/2} \log m\right) = O(m^{3/4}\sqrt{\log m})$. The set V_0 is a separator for G of size $O(m^{3/4}\sqrt{\log m})$. $\quad\square$

3. Proof of Theorem 1.1

The *bisection width* $b_w(G)$ of a graph $G = (V, E)$ with respect to a weight function w is the least integer for which there is a partition $V = V_1 \cup V_2$ such that $w(V_1), w(V_2) \leqslant 2/3$ and the number of edges between V_1 and V_2 is $b_w(G)$. Note that $b(G) = b_w(G)$ if w is the uniform weight function defined by $w(v) = 1/|V|$ for all $v \in V$.

By iterating Lemma 2.3, we obtain the following result.

Theorem 3.1. *Let G be a topological graph with n vertices and maximum degree d, and assume that every edge of G intersects at most D other edges. For any weight function w, we have*

$$b_w(G) = O\left(\left(\sqrt{dD} + d\right)\sqrt{n}\log n\right).$$

Proof. The maximum degree is d, so the number of edges of G is at most $dn/2$. Since each edge of G intersects at most D other edges, the pair-crossing number of G is at most $\frac{dn}{2}\frac{D}{2} = dDn/4$.

Let A_0 denote the vertex set of G. By Lemma 2.3, there is a partition $A_0 = A_1 \cup B_1$ such that $|A_1|, |B_1| \leqslant \frac{2}{3}n$, and the number of edges with one vertex in A_1 and the other in B_1 is at most

$$c\log n\left(\sqrt{\operatorname{pcr}(G)} + \sqrt{\operatorname{ssqd}(G)}\right) \leqslant c\log n\left(\sqrt{dDn/4} + \sqrt{d^2 n}\right).$$

Without loss of generality, we may assume that $w(A_1) \geqslant w(B_1)$.

After i iterations, we have a vertex subset A_i with at most $\left(\frac{2}{3}\right)^i n$ vertices. By Lemma 2.3 applied to the subgraph $G[A_i]$ of G induced by A_i, there is a partition $A_i = A_{i+1} \cup B_{i+1}$ such that $|A_{i+1}|, |B_{i+1}| \leqslant \frac{2}{3}|A_i| \leqslant \left(\frac{2}{3}\right)^{i+1}n$, and the number of edges with one vertex in A_{i+1} and the other in B_{i+1} is at most

$$c\log n\left(\sqrt{\operatorname{pcr}(G[A_i])} + \sqrt{\operatorname{ssqd}(G[A_i])}\right) \leqslant c\log\left(\left(\frac{2}{3}\right)^i n\right)\left(\sqrt{dD\left(\frac{2}{3}\right)^i n/4} + \sqrt{d^2\left(\frac{2}{3}\right)^i n}\right)$$

$$\leqslant \left(\frac{2}{3}\right)^{i/2} c\left(\sqrt{dD} + d\right)\sqrt{n}\log n.$$

Without loss of generality, we may assume that $w(A_{i+1}) \geqslant w(B_{i+1})$.

We stop the iterative process with i_0 if $w(A_{i_0}) \leqslant \frac{2}{3}$. Since $w(A_{i_0}) + w(B_{i_0}) = w(A_{i_0-1}) > 2/3$, we have $1/3 < w(A_{i_0}) \leqslant 2/3$. Let $X = A_{i_0}$ and $Y = A_0 \setminus A_{i_0} = B_1 \cup \cdots \cup B_{i_0}$. By construction, the number of edges of G with one vertex in X and the other vertex in Y is less than

$$\sum_{i=0}^{\infty}\left(\frac{2}{3}\right)^{i/2} c\left(\sqrt{dD} + d\right)\sqrt{n}\log n \leqslant 6c\left(\sqrt{dD} + d\right)\sqrt{n}\log n.$$

Thus, $A_0 = X \cup Y$ is a partition of the vertex set demonstrating that the bisection width of G with respect to w is $O\left(\left(\sqrt{dD} + d\right)\sqrt{n}\log n\right)$. $\qquad\square$

We next prove a separator theorem for string graphs of maximum degree Δ.

Theorem 3.2. *Let C be a collection of curves in the plane whose intersection graph G has m edges and maximum degree Δ, and let w be a weight function on G. Then G has a separator of size $O\left(\Delta m^{1/2}\log m\right)$ with respect to w.*

Proof. By slightly perturbing the curves in C, if necessary, we can assume that no *three* curves in C share a point in common. We may also assume without loss of generality that every element of C intersects at least one other element.

For each pair of intersecting curves, pick an arbitrary point of intersection, and let P be the set of these m points. Define the topological graph T on the vertex set P by connecting a pair of points of P with an edge if and only if they are consecutive points of P along a curve in C. The number of vertices of T is m. Since no *three* curves in C have a point in common, the maximum degree of the vertices of T is at most *four*. Each curve in C gives rise to a path in the topological graph T with at most Δ vertices and at most $\Delta - 1$ edges. Since each curve in C intersects at most Δ other curves, each edge of T crosses at most Δ curves, besides the one in which it is contained. Each of these at most Δ curves contains at most $\Delta - 1$ edges of T. Therefore, each edge of T intersects altogether at most $\Delta(\Delta - 1) < \Delta^2$ other edges.

For any $\gamma \in C$, let $d(\gamma)$ denote the number of points of P on γ, *i.e.*, the number of curves in C that intersect γ. To each vertex v of T that is the intersection of two elements $\gamma_1, \gamma_2 \in C$, assign the weight

$$w'(v) = \frac{w(\gamma_1)}{d(\gamma_1)} + \frac{w(\gamma_2)}{d(\gamma_2)}.$$

Notice that $w'(P) = w(C) = 1$.

We now apply Theorem 3.1 to the topological graph T and to the weight function w'. Recall that T has m vertices, maximum degree at most *four*, and every edge intersects at most Δ^2 other edges. So there is a partition $P = P_1 \cup P_2$ with $w'(P_1), w'(P_2) \leqslant 2/3$ and the number of edges with one vertex in P_1 and the other in P_2 is

$$O\left((\Delta^2 m)^{1/2} \log m\right) = O\left(\Delta m^{1/2} \log m\right).$$

Let C_0 consist of those curves in C that contain an edge of the topological graph T with one vertex in P_1 and the other in P_2. There are $O\left(\Delta m^{1/2} \log m\right)$ such edges, therefore we have $|C_0| = O\left(\Delta m^{1/2} \log m\right)$.

For $i \in \{1, 2\}$, let C_i consist of those curves of C all of whose intersection points in P belong to P_i. Note that, by construction, $w(C_i) \leqslant w'(P_i) \leqslant 2/3$ and the sets C_0, C_1, C_2 are pairwise disjoint.

No curve in C_1 intersects a curve in C_2 as otherwise they have an intersection point in P, and this point must be in both P_1 and P_2, a contradiction. To show that every curve in C belongs to exactly one of the sets C_0, C_1, C_2, it is enough to notice that any curve $\gamma \in C$ which contains a point in P_1 and one in P_2 must belong to C_0. Indeed, such a curve gives rise to a path in T, and hence contains an edge from P_1 to P_2. Therefore, C_0 is a separator with respect to w of the desired size. $\square$

Just as we deduced the unweighted version of Theorem 1.1 from Theorem 2.5, Theorem 1.1 follows from Theorem 3.2.

4. *H*-free string graphs have linearly many edges

In this section, we show how to deduce quickly from our separator theorem a qualitative version of Theorem 1.3.

A weaker version of Theorem 1.3, established in [34], states that every $K_{t,t}$-free string graph on n vertices has at most $n \log^{c'_t} n$ edges. Combining this theorem with Theorem 1.1, we obtain the following corollary.

Corollary 4.1. *For every $K_{t,t}$-free string graph G on n vertices and for every weight function w for G, there is a separator of size $n^{3/4} \log^{c''_t} n$ with respect to w, where c_t is a constant depending only on t.*

A family of graphs is *hereditary* if it is closed under taking induced subgraphs. The following lemma of Lipton, Rose, and Tarjan [25] shows that if all members of a hereditary family of graphs have small separators, then the number of edges of these graphs is at most linear in the number of vertices. Another proof with a slightly better bound can be found in [11].

Lemma 4.2 (Lipton, Rose and Tarjan [25]). *Let $\epsilon > 0$, and let F be a hereditary family of graphs such that every member of F with n vertices has a separator of size $O(n/(\log n)^{1+\epsilon})$. Then every graph in F on n vertices has at most $c_F n$ edges, where c_F is a suitable constant.*

Clearly, the family of $K_{t,t}$-free string graphs is hereditary. Therefore, Corollary 4.1 combined with Lemma 4.2 immediately implies that every $K_{t,t}$-free string graph on n vertices has at most $c_t n$ edges, where c_t only depends on t. $\qquad\square$

5. Proof of Theorem 1.3

The aim of this section is to prove Theorem 1.3. (A proof disregarding the dependence of the constant on t was given in Section 4.)

The first ingredient of the proof of Theorem 1.3 is a weaker upper bound on the number of edges of a $K_{t,t}$-free string graph on n vertices. Pach and Sharir [34] proved that every $K_{t,t}$-free string graph on n vertices has at most $n \log^{c_t} n$ edges. Their proof shows that we may take $c_t = 2^{ct}$ for some absolute constant c. We first show how to modify their proof technique, in combination with other extremal results on string graphs, to show that the result also holds with $c_t = c \log t$.

Lemma 5.1. *Every string graph G with n vertices and more than $n \log^{c_1 \log t} n$ edges has $K_{t,t}$ as a subgraph, where c_1 is an absolute constant.*

To prove this lemma, we need the following two auxiliary results. The first of these results, from [13], shows that every n-vertex string graph with positive constant edge density contains a balanced complete bipartite graph with $\Omega(n/\log n)$ vertices.

Lemma 5.2 ([13]). *Every string graph with n vertices and ϵn^2 edges has $K_{t,t}$ as a subgraph with $t = \epsilon^{c_3} n/\log n$ for some absolute constant c_3.*

The following lemma guarantees that topological graphs on n vertices with sufficiently many edges contain s pairwise crossing edges with distinct vertices. The same result was proved in [12], except that the s pairwise crossing edges were allowed to share endpoints. As we will need the slightly stronger version for the proof of Theorem 1.3, we include its proof here.

Lemma 5.3. *There is an absolute constant c_2 such that every topological graph with $n \geqslant 2$ vertices and more than $n(\log n)^{c_2 \log s}$ edges has s pairwise crossing edges with distinct vertices.*

We will use the following lemma, which shows that for every graph G with n vertices and $m \gg n$ edges, almost all induced subgraphs of G have roughly $m/4$ edges.

Lemma 5.4. *Let G be a graph with n vertices and $m \geqslant n$ edges. Let H be an induced subgraph of G taken uniformly at random and let X be the random variable denoting number of edges of H. For every $\lambda > 0$,*

$$\mathbb{P}[|X - m/4| \geqslant \lambda\sqrt{mn}/2] \leqslant 1/\lambda^2.$$

Proof. The proof uses the *second moment method* (see, *e.g.*, Section 4 of [2]).

We first show that the expected value of the random variable X is $m/4$. We pick a vertex to be in H with probability $1/2$ independently of the other vertices. For each edge e of G, let X_e be the indicator random variable of the event that e is an edge of H. That is, $X_e = 1$ if e is an edge of H and $X_e = 0$ otherwise. We have $\mathbb{E}[X_e] = 1/4$, and by linearity of expectation, $\mathbb{E}[X] = \sum_{e \in E(G)} \mathbb{E}[X_e] = m/4$.

We next compute the variance of the random variable X. Since $X = \sum_{e \in E(G)} X_e$, we have

$$\mathrm{Var}[X] = \sum_{e \in E(G)} \mathrm{Var}[X_e] + \sum_{e \neq e'} \mathrm{Cov}[X_e, X_{e'}],$$

where the variance is defined by $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ and the covariance is defined by $\mathrm{Cov}[X_e, X_{e'}] = \mathbb{E}[X_e X_{e'}] - \mathbb{E}[X_e]\mathbb{E}[X_{e'}]$. Since X_e is an indicator random variable, we have $\mathrm{Var}[X_e] = \mathbb{E}[X_e^2] - \mathbb{E}[X_e]^2 = \mathbb{E}[X_e] - \mathbb{E}[X_e]^2 = 3/16$. The covariance of independent variables is 0. In particular, if e and e' do not share a vertex, then $\mathrm{Cov}[X_e, X_{e'}] = 0$. If e and e' share a vertex, then $X_e X_{e'} = 1$ if and only if the three vertices of e or e' are all vertices of H. Hence, in this case, $\mathbb{E}[X_e X_{e'}] = 1/8$ and $\mathrm{Cov}[X_e, X_{e'}] = \mathbb{E}[X_e X_{e'}] - \mathbb{E}[X_e]\mathbb{E}[X_{e'}] = 1/8 - 1/16 = 1/16$. Let $d_1, \ldots, d_n$ be the degree sequence of G. The number of pairs of distinct edges that share an edge, by counting over the vertex in common of the two edges, is precisely $\sum_{i=1}^{n} \binom{d_i}{2}$. Putting this together and using linearity of expectation,

$$\mathrm{Var}[X] = \sum_{e \in E(G)} \mathrm{Var}[X_e] + \sum_{e \neq e'} \mathrm{Cov}[X_e, X_{e'}] = \frac{3}{16}m + \frac{2}{16}\sum_{i=1}^{n}\binom{d_i}{2}$$

$$= \frac{1}{16}\left(m + \sum_{i=1}^{n} d_i^2\right) \leqslant \frac{1}{16}\left(m + \frac{2m}{n}n^2\right) = \frac{1}{16}(m + 2mn) \leqslant mn/4,$$

where the first inequality uses the convexity of the function $f(y) = y^2$ together with the inequalities $0 \leqslant d_i \leqslant n$ and the equation $\sum_{i=1}^{n} d_i = 2m$. The desired inequality is just Chebyshev's inequality substituting in the above upper bound on the variance of X. $\quad\square$

The next statement is an easy consequence of Lemma 5.4.

Lemma 5.5. *Let G_1 and G_2 be graphs on the same vertex set V of cardinality n and denote the number of edges of G_i by m_i. If $m_i \geqslant 64n$ for $i = 1, 2$, then there is a partition $V = V_1 \cup V_2$ such that the subgraph $G_i[V_i]$ of G_i induced by V_i has at least $m_i/8$ edges, for $i = 1, 2$.*

Proof. Let $\lambda = 2$. Pick $V_1 \subset V$ uniformly at random, and let $V_2 = V \setminus V_1$. For each $i = 1, 2$, since $\lambda \sqrt{m_i n}/2 \leqslant m_i/8$, Lemma 5.4 implies that the probability that the number of edges of $G_i[V_i]$ is less than $m_i/8$ is at most $1/4$. Hence, with probability at least $1/2$, for each $i = 1, 2$, the number of edges of $G_i[V_i]$ is at least $m_i/8$. Since this event occurs with positive probability, there is a partition $V = V_1 \cup V_2$ such that the number of edges of $G_i[V_i]$ is at least $m_i/8$, for $i = 1, 2$. $\quad\square$

We have now established the necessary lemmas to present the proof of Lemma 5.3. Our proof is similar to the proof of Theorem 11 from [12], but it also guarantees that the s pairwise crossing edges have distinct vertices.

Proof of Lemma 5.3. Let $P(n, s)$ denote the maximum number of edges of a topological graph on n vertices with no s pairwise crossing edges with distinct vertices. We will prove by induction on n and s the upper bound

$$P(n, s) \leqslant n(\log n)^{c_2 \log s},$$

which implies Lemma 5.3. For $n \leqslant 2$, the inequality follows from $P(n, s) \leqslant \binom{n}{2}$ and for $s = 1$ from $P(n, 1) = 0$. These are our base cases. The induction hypothesis is that if $s' \leqslant s$ and $n' \leqslant n$ are positive integers and $(n', s') \neq (n, s)$, then $P(n', s') \leqslant n'(\log n')^{c_2 \log s'}$. Let $G = (V, E)$ be a topological graph with n vertices, $m = P(n, s)$ edges, and no s pairwise crossing edges with distinct vertices. Let F be the intersection graph of the edges of G, and x denote the number of edges of F, *i.e.*, G has x pairs of crossing edges. Let $y = 100c^2 \log^4 n$, where c is the absolute constant from Lemma 2.3.

Case 1: $x < m^2/y$. Note that x is an upper bound on the pair-crossing number of G. By Lemma 2.3, there is a partition $V = V_1 \cup V_2$ into non-empty subsets such that $|V_1|, |V_2| \leqslant \frac{2}{3}|V|$ and the number of edges with one vertex in V_1 and the other in V_2 satisfies

$$e(V_1, V_2) = b(G) \leqslant c \log n \left(\sqrt{\mathrm{pcr}(G)} + \sqrt{\mathrm{ssqd}(G)} \right).$$

Note that $\mathrm{ssqd}(G) \leqslant \frac{2m}{n} n^2 = 2mn$ because the function $f(z) = z^2$ is convex, the degrees of the vertices in G lie between 0 and n, and the sum of the degrees of the vertices in G is $2m$. If $m < 2ny = 200c^2 n \log^4 n$, then we are done. Thus, we may assume that $m \geqslant 2ny$

and it follows that $\sqrt{x} + \sqrt{\mathrm{ssqd}(G)} \leqslant 2my^{-1/2}$. Hence, $e(V_1, V_2) \leqslant c \log n \cdot 2my^{-1/2} = \frac{m}{5 \log n}$. For $i = 1, 2$, the subgraph of G induced by V_i also has no s pairwise crossing edges with distinct vertices. Hence,

$$m \leqslant P(|V_1|, s) + P(|V_2|, s) + \frac{m}{5 \log n}.$$

Using the induction hypothesis, the inequality $|V_1|, |V_2| \leqslant 2n/3$, and that c_2 is a sufficiently large constant, we have

$$\begin{aligned}
P(n, s) = m &\leqslant \left(1 - \frac{1}{5 \log n}\right)^{-1} \big(P(|V_1|, s) + P(|V_2|, s)\big) \\
&\leqslant \left(1 - \frac{1}{5 \log n}\right)^{-1} \big(|V_1|(\log |V_1|)^{c_2 \log s} + |V_2|(\log |V_2|)^{c_2 \log s}\big) \\
&\leqslant \left(1 - \frac{1}{5 \log n}\right)^{-1} n\big(\log(2n/3)\big)^{c_2 \log s} < n(\log n)^{c_2 \log s},
\end{aligned}$$

which completes this case.

Case 2: $x \geqslant m^2/y$. So F, the intersection graph of the edges of G, has at least m^2/y edges. Since F is a string graph, Lemma 5.2 implies that there is an absolute constant c_3 such that F contains $K_{t,t}$ as a subgraph with

$$t = y^{-c_3} m/\log m = 100^{-c_3} c^{-2c_3} (\log n)^{-4c_3} m/\log m \geqslant m(\log n)^{-c'},$$

for some absolute constant c'. Hence, there are two edge subsets E_1, E_2 of G, each of size at least t, such that every edge in E_1 crosses every edge in E_2. Applying Lemma 5.5, there are edge subsets $E_1' \subset E_1$ and $E_2' \subset E_2$, each of cardinality at least $t/8$, such that the vertices of the edges in E_1' are distinct from the vertices in E_2'. Since G has no s pairwise crossing edges with distinct vertices, there exists $i \in \{1, 2\}$ such that E_i' does not contain $s/2$ pairwise crossing edges with distinct vertices. Hence,

$$m(\log n)^{-c'}/8 \leqslant t/8 \leqslant |E_i'| \leqslant P(n, \lceil s/2 \rceil) \leqslant n(\log n)^{c_2 \log \lceil s/2 \rceil},$$

which implies $m \leqslant n(\log n)^{c_2 \log s}$ since c_2 was chosen to be a sufficiently large absolute constant. This completes the proof. $\qquad\square$

Having gathered the required lemmas, we now prove Lemma 5.1. A *Jordan region* is a closed region of the plane, bounded by a simple closed Jordan curve. In other words, a Jordan region is homeomorphic to the closed unit disk.

Proof of Lemma 5.1. Let s be the smallest positive integer such that $(1/16)^{c_3}(2s)/\log 2s \geqslant t$, where c_3 is the absolute constant from Lemma 5.2. Let $c_t = c_2 \log s = \Theta(\log t)$, where c_2 is the absolute constant from Lemma 5.3.

Let G be a string graph with n vertices and more than $n(\log n)^{c_t}$ edges. Let $\mathcal{C} = \{C_1, \ldots, C_n\}$ be a collection of n Jordan regions whose intersection graph is the string graph G and which has the property that any two intersecting Jordan regions in $\mathcal{C}$ intersect in their interiors (it is easy to see that, by slightly fattening compact connected sets, every

string graph is the intersection graph of such a collection of Jordan regions). Fix distinct points p_i in the interior of C_i for $i = 1, \ldots, n$. For each intersecting pair $C_i, C_j \in \mathcal{C}$ with $i < j$, let p_{ij} be a point in $C_i \cap C_j$ such that all the points in $\{p_1, \ldots, p_n\} \cup \{p_{ij} : C_i \cup C_j \neq \emptyset\}$ are distinct, and let γ_{ij} be a simple (non-intersecting) curve contained in $C_i \cup C_j$ such that:

(1) γ_{ij} has endpoints p_i and p_j,
(2) γ_{ij} does not contain any other p_ℓ,
(3) γ_{ij} can be split into two subcurves γ_{ij}^0 and γ_{ij}^1 such that γ_{ij}^0 is contained in C_i and has endpoints p_i and p_{ij} and γ_{ij}^1 is contained in C_j and has endpoints p_{ij} and p_j.

The points $\{p_1, \ldots, p_n\}$ are the vertex set and curves $\{\gamma_{ij} : C_i \cap C_j \neq \emptyset\}$ are the edge set of a topological graph T with n vertices and more than $n(\log n)^{c_t}$ edges. Since $c_t = c_2 \log s$, by Lemma 5.3, there are at least s pairwise intersecting edges in T with distinct vertices. Each edge consists of two subcurves and these $2s$ subcurves have at least $\binom{s}{2} \geqslant \frac{1}{16}(2s)^2$ intersecting pairs. By Lemma 5.2, the intersection graph of these $2s$ subcurves contains $K_{h,h}$ with $h = (1/16)^{c_3}(2s)/\log 2s \geqslant t$. It follows from the construction that G contains $K_{t,t}$. $\qquad\square$

The second ingredient of the proof of Theorem 1.3 is our separator theorem, Theorem 1.1, which, together with Lemma 5.1, implies that every $K_{t,t}$-free string graph has $m < n(\log n)^{c_t}$ edges and hence a separator of size $O(m^{3/4}\sqrt{\log m}) < O(n^{3/4}\log^{3c_t/4+1/2} n)$. Thus, for $n_0 = 2^{O(c_t \log c_t)} \leqslant t^{c' \log \log t}$ (where c' is an absolute constant), every $K_{t,t}$-free string graph with $n \geqslant n_0$ vertices has a separator of size $n^{7/8}$. This fact, together with the following lemma from [11] (which is a more precise version of Lemma 4.2), immediately implies Theorem 1.3.

Lemma 5.6 ([11]). *Let $\phi(n)$ be a monotone decreasing non-negative function defined on the set of positive integers, and let n_0 and C be positive integers such that*

$$\phi(n_0) \leqslant \frac{1}{12} \quad and \quad \prod_{i=0}^{\infty} \left(1 + \phi\left(\lceil (4/3)^i n_0 \rceil\right)\right) \leqslant C.$$

If F is an $n\phi(n)$-separable hereditary family of graphs, then every graph in F on $n \geqslant n_0$ vertices has fewer than $\frac{Cn_0}{2}n$ edges.

6. Stronger versions of Theorem 1.3

First we establish Theorem 1.5, that is, we show how to improve Theorem 1.3 considerably for intersection graphs of collections of curves in which every pair of curves intersects in at most a fixed constant number k of points.

In [11], we proved the following lemma.

Lemma 6.1. *The intersection graph of a collection of curves with x crossings has a separator of size $O(\sqrt{x})$.*

If each pair in a collection of curves intersects in at most k points, then the number m of edges of the intersection graph is at least x/k, and we obtain a separator of size

$O(\sqrt{km})$. In [14], the following result was established, which is an analogue of Lemma 5.2 for families of curves in which each pair intersects in at most a constant number k of points.

Lemma 6.2. *Let G be the intersection graph of a collection of n curves in the plane, any pair of which intersects in at most k points. If G has at least ϵn^2 edges, then it contains a complete bipartite subgraph $K_{t,t}$ with $t \geqslant c_k \epsilon^c n$, where c is an absolute constant and c_k is a constant that only depends on k.*

We now have the necessary tools to prove Theorem 1.5. This theorem states that for positive integers k and t, there exists a constant C_k depending only on k, such that any $K_{t,t}$-free intersection graph G of n curves in the plane, no pair of which has more than k points in common, has at most $C_k t n$ edges.

Proof of Theorem 1.5. Suppose that G has ϵn^2 edges. By Lemma 6.2, we have $t \geqslant c_k \epsilon^c n$, that is, $\epsilon \leqslant \left(c_k^{-1} \frac{t}{n}\right)^{1/c}$. Thus, according to Lemma 6.1, G has a separator of size

$$O(\sqrt{km}) < O\left(\sqrt{k\left(c_k^{-1}\frac{t}{n}\right)^{1/c} \cdot n^2}\right) < c_k'(t/n)^{c_1} n,$$

where $c_1 = 1/(2c) > 0$ and c_k' only depends on k.

Letting $\phi(n) = c_k'(t/n)^{c_1}$ and $n_0 = (12c_k')^{1/c_1} t$, Lemma 5.6 implies that G has at most $C_k t n$ edges for some constant C_k only depending on k. $\square$

We similarly prove the following result.

Theorem 6.3. *Every $K_{t,t}$-free intersection graph G of n convex sets in the plane has $O(t^3 n)$ edges.*

Proof. Suppose that G has ϵn^2 edges. In [15], it was shown that every intersection graph of n convex sets in the plane with ϵn^2 edges contains a complete bipartite subgraph $K_{t,t}$ with $t \geqslant c\epsilon^2 n$ for some absolute constant $c > 0$. Hence, $\epsilon \leqslant \left(\frac{t}{cn}\right)^{1/2}$. A separator lemma from [11] states that every K_s-free intersection graph of convex sets in the plane with m edges has a separator of size at most $c'\sqrt{sm}$ for some absolute constant c'. Hence, G has a separator of size $c'\sqrt{2tm} \leqslant 2c'\sqrt{t}\epsilon^{1/2} n \leqslant 2c'c^{-1/4}t^{3/4}n^{3/4}$. Letting $\phi(n) = 2c'c^{-1/4}t^{3/4}n^{-1/4}$ and $n_0 = 24^4 c'^4 c^{-1} t^3$, Lemma 5.6 implies that G has $O(t^3 n)$ edges. $\square$

7. Proof of Theorem 1.6 and related results

Theorem 1.6 is a direct corollary of Theorem 1.1 and the following lemma, which shows that all graphs of large girth that belong to a hereditary family of graphs with small separators are quite sparse.

Lemma 7.1. *Let $\alpha > 0$, and let F be a hereditary family of graphs such that every member of F with n vertices has a separator of size $O(n/(\log n)^{1+\alpha})$. Then for each $\epsilon > 0$ there is a positive integer $g = g_F(\epsilon)$ such that every graph in F on n vertices and girth at least g has at most $(1+\epsilon)n$ edges.*

The aim of this section is to prove Lemma 7.1 and to discuss some of its consequences. The similarity between Lemma 7.1 and 4.2 is no coincidence; their proofs are very similar.

Before turning to the proof, we briefly outline its main idea. Consider a hereditary family F of graphs in which every graph has a small separator. We show that every graph G in F with n vertices has an induced subgraph with at most $\frac{3}{4}n$ vertices, whose average degree is not much smaller than the average degree of G. We repeatedly use this fact until we find an induced subgraph of G with fewer than g vertices, whose average degree is not much smaller than that of G. But if the girth of G is at least g, then this induced subgraph of G with fewer than g vertices is a forest and so has average degree less than 2. If g is chosen sufficiently large, we conclude that G has average degree at most $2 + 2\epsilon$ and hence at most $(1+\epsilon)n$ edges.

Now we work out the details of the proof of Lemma 7.1. Given a non-negative function f defined on the set of positive integers, we say that a family F of graphs is *f-separable* if every graph in F with n vertices has a separator of size at most $f(n)$.

Lemma 7.2. *Let $\phi(n)$ be a monotone decreasing non-negative function defined on the set of positive integers, let g be a positive integer, and let $\epsilon > 0$ be such that*

$$\phi(g) \leqslant \frac{1}{12} \quad and \quad \prod_{i=0}^{\infty} \left(1 + \phi\left(\lceil (4/3)^i g\rceil\right)\right) \leqslant 1 + \epsilon.$$

If F is an $n\phi(n)$-separable hereditary family of graphs, then every graph in F on n vertices with girth at least g has fewer than $(1+\epsilon)n$ edges.

Proof. Let $G_0 = (V, E)$ be a member of the family F with n vertices, girth at least g, and average degree d. If $n < g$, then G_0 is a forest and hence has at most $n - 1$ edges. We may therefore assume that $n \geqslant g$. By definition, there is a partition $V = V_0 \cup V_1 \cup V_2$ with $|V_0| \leqslant n\phi(n)$, $|V_1|, |V_2| \leqslant \frac{2}{3}n$, such that no vertex in V_1 is adjacent to any vertex in V_2.

Let d' and d'' denote the average degree of the vertices in the subgraphs of G_0 induced by $V_0 \cup V_1$ and $V_0 \cup V_2$, respectively. Every edge of G_0 is contained in at least one of these two induced subgraphs. Hence,

$$d'(|V_0| + |V_1|) + d''(|V_0| + |V_2|) \geqslant 2|E| = d|V|,$$

so

$$d'\frac{|V_0| + |V_1|}{|V| + |V_0|} + d''\frac{|V_0| + |V_2|}{|V| + |V_0|} \geqslant d\frac{|V|}{|V| + |V_0|}.$$

Since $|V| = |V_0| + |V_1| + |V_2|$, then

$$\frac{|V_0| + |V_1|}{|V| + |V_0|} + \frac{|V_0| + |V_2|}{|V| + |V_0|} = 1.$$

and the left-hand side of the above inequality is a weighted mean of d' and d''. Consequently, d' or d'' is at least

$$d\frac{|V|}{|V|+|V_0|} \geqslant d\frac{1}{1+\phi(n)}.$$

Suppose without loss of generality that d' is at least as large as this number, and let G_1 denote the subgraph of G induced by $V_0 \cup V_1$. By assumption, we have that $\phi(n) \leqslant \frac{1}{12}$ and $|V_0| \leqslant n\phi(n)$. Therefore, G_1 has $|V_0| + |V_1| \leqslant \frac{1}{12}n + \frac{2}{3}n = \frac{3}{4}n$ vertices.

Proceeding like this, we find a sequence of induced subgraphs $G_0 \supset G_1 \supset G_2 \supset \cdots$ with the property that, if G_i has $n_i \geqslant g$ vertices and average degree d_i, then G_{i+1} has at most $\frac{3}{4}n_i$ vertices and average degree at least $\frac{1}{1+\phi(n_i)}d_i$. We stop with G_j if the number of vertices of G_j is less than g.

Since G_j is an induced subgraph of G, it also has girth at least g. The number of vertices of G_j is less than g, so G_j must be a forest and therefore has average degree less than 2. The above argument also shows that the average degree of G_j is at least $\frac{1}{1+\epsilon}d$, so $d < 2(1+\epsilon)$, and the number of edges of G is $dn/2 < (1+\epsilon)n$, completing the proof. □

Taking logarithms and approximating $\ln(1+x)$ by x, we obtain that

$$\prod_{i=0}^{\infty}(1 + \phi(\lceil (4/3)^i g\rceil)) \neq \infty$$

if and only if $\sum_{i=0}^{\infty} \phi(\lceil (4/3)^i\rceil) \neq \infty$ if and only if $\sum_{i=0}^{\infty} \phi(2^i) \neq \infty$. (For a formal proof of the elementary fact that $\prod_{i=1}^{\infty}(1 + a_i)$ with each $a_i > 0$ converges if and only if $\sum_{i=1}^{\infty} a_i$ converges, see, *e.g.*, Theorem 3 of Section 3.7 in [20].) Therefore, Lemma 7.2 has the following corollary.

Corollary 7.3. *Let F be an $n\phi(n)$-separable hereditary family of graphs, where $\phi(n)$ is a monotone decreasing non-negative function such that $\sum_{i=0}^{\infty} \phi(2^i) \neq \infty$. Then, for each $\epsilon > 0$ there exists $g_F(\epsilon)$ such that every graph in F on n vertices and girth at least g has at most $(1+\epsilon)n$ edges.*

Since $\sum_{i=1}^{\infty} 1/i^{1+\alpha}$ converges for all $\alpha > 0$, Lemma 7.1 is an immediate consequence of Corollary 7.3.

The condition that a connected graph has large girth means that the graph is locally 'tree-like'. In general, this local condition does not imply that the graph also has some global tree-like properties. For instance, in 1959 Erdős [9] proved the existence of graphs with arbitrarily large girth and chromatic number. However, according to Lemma 7.1, if every member of a hereditary family F of graphs has a small separator, then the condition that a connected graph in F has large girth does imply that the graph is *globally tree-like*. Indeed, if $\epsilon < 1/2$, then every graph in F with girth at least $g_F(\epsilon)$ is 2-degenerate and hence has chromatic number at most 3.

Furthermore, any graph G in F on n vertices with girth at least $\max(g_F(\epsilon/3), 3/\epsilon)$ can be turned into a forest by the removal of at most ϵn edges. Indeed, we may assume that G has minimum degree at least 2 since we can remove vertices of degree less than 2

as they are in no cycles. Also, G has $e(G) \leqslant (1 + \epsilon/3)n$ edges. Let $d_1, \ldots, d_n$ denote the degree sequence of G. We have $\sum_i d_i - 2 = 2e(G) - 2n \leqslant \frac{2}{3}\epsilon n$. Delete $d_i - 2$ edges from each vertex i, so the remaining subgraph G_0 has vertices of degree at most 2. The number of edges deleted so far is at most $\frac{2}{3}\epsilon n$. The connected components of G_0 are trees and cycles of length at least the girth of G, which is at least $3/\epsilon$. Hence, the number of cycles of G_0 is at most $\frac{\epsilon}{3}n$. Delete one edge from each of these cycles. The total number of edges deleted is at most ϵn, and the remaining subgraph is a forest.

We end this section by presenting two other corollaries of Lemma 7.1. The separator theorem for graphs with an excluded minor [1], together with Lemma 7.1, imply the following.

Corollary 7.4. *For any $\epsilon > 0$ and positive integer t, there exists a positive integer $g(\epsilon, t)$ such that every K_t-minor-free graph on n vertices with girth at least $g(\epsilon, t)$ has at most $(1 + \epsilon)n$ edges.*

A well-known result of Thomassen [39] (see also Chapter 8.2 in [7]) states that for any positive integer t, there exists another integer $g(t)$ such that every graph with minimum degree at least 3 and girth at least $g(t)$ contains K_t as a minor. Obviously, Corollary 7.4 implies Thomassen's result. In fact, it can be shown by a simple argument that the two statements are equivalent. In the special case of planar graphs and, more generally, for graphs with bounded genus, the statement easily follows from Euler's polyhedral formula.

The separator theorem for intersection graphs of balls in $\mathbb{R}^d$ [29] together with Lemma 7.1 imply the following result.

Corollary 7.5. *For any $\epsilon > 0$ and positive integer d, there exists a positive integer $g(\epsilon, d)$ such that every intersection graph of balls in $\mathbb{R}^d$ with girth at least $g(\epsilon, d)$ has at most $(1 + \epsilon)n$ edges.*

8. Concluding remarks: a conjecture and counting string graphs

Theorem 1.3, with a much worse dependence of the coefficient of n on t, can also be deduced from the following result of Kuhn and Osthus [23]. For any graph H and any positive integer t, there is a constant $c(H, t)$ such that every graph with n vertices and at least $c(H, t)n$ edges contains an induced subdivision of H or $K_{t,t}$ as a subgraph. Let H_0 be the graph obtained from the complete graph K_5 by replacing each edge by a path of length two. Using the non-planarity of K_5, it is easy to see that no subdivision of H_0 is a string graph. Since the family of string graphs is closed under taking induced subgraphs, no string graph contains an induced subdivision of H_0. Thus, the result of Kuhn and Osthus implies that any $K_{t,t}$-free string graph on n vertices has at most $c(H_0, t)n$ edges. However, this proof only shows that

$$c(H_0, t) < 2^{2^{2^{2^{2^{ct \log t}}}}},$$

for some absolute constant c.

The dependence of the coefficient of n on t in Theorem 1.3 could be further improved if we could prove Conjecture 1.2. Indeed, Conjecture 1.2 combined with Lemmas 5.2 and 5.6 would imply the following.

Conjecture 8.1. *Every $K_{t,t}$-free string graph with n vertices has $O((t \log t)n)$ edges.*

Conjecture 8.1, if true, would be tight up to the constant factor. According to a construction in [10] and [36], there are string graphs with n vertices and $(1 - o(1))n^2/2$ edges, in which the size of the largest balanced bipartite subgraph is $O(n/\log n)$.

Another consequence of Conjecture 1.2 would be that, together with Lemma 5.2, it would imply that every K_t-free string graph with n vertices has chromatic number at most $(\log n)^{c \log t}$ for some absolute constant c. This was shown in [12] for intersection graphs of curves in which each pair of curves intersects in at most a fixed constant number of points. It is not even known if every triangle-free string graph with n vertices has chromatic number $n^{o(1)}$.

A family of graphs is *small* if it contains at most $n! \alpha^n$ labelled graphs on n vertices, for some constant α. For example, a classical result of Cayley asserts that the number of labelled trees on n vertices is n^{n-2}, so the family of trees is small. A family F of graphs is *addable* if $G \in F$ if and only if all connected components of G are in F, and if $G_1, G_2 \in F$ and v_i is a vertex of G_i for $i \in \{1, 2\}$ implies that the graph obtained from the disjoint union of G_1 and G_2 by adding the edge $\{v_1, v_2\}$ is also in F. It was shown by McDiarmid, Steger and Welsh [28] that if F is small and addable, then there is a constant $\alpha = \alpha(F)$ such that the number of graphs in F on n vertices is $n! \alpha^{(1+o(1))n}$.

Norine, Seymour, Thomas and Wollan [30] showed that all proper minor closed graphs are small, which answered a question of Welsh. Norine and Dvořák [8] recently found a much simpler proof with a divide-and-conquer approach using the separator theorem [1] for graphs with a forbidden minor. They show that if F is an $f(n)$-separable hereditary family of graphs with $f(n) \leqslant cn/(\log n \log \log n)^2$ for some constant c, then F is small.

Pach and Tóth [36] showed that the number of string graphs on n vertices is $2^{(\frac{3}{4}+o(1))\binom{n}{2}}$. The above result of Norine and Dvořák [8], together with Theorems 1.1 and 1.3, show that if H is bipartite, then the family of H-free string graphs is small. It is easy to check that if H is 2-connected, then the family of H-free string graphs is addable. We thus get the following corollary.

Corollary 8.2. *If H is a 2-connected bipartite graph, then there is a constant $c = c(H)$ such that the number of labelled H-free string graphs on n vertices is $n! c^{(1+o(1))n}$.*

Notice that every tree is a string graph. In the other direction, we have the following result, which says that there are not many more string graphs of large girth than trees on a given number of vertices. It can be proved by the same divide-and-conquer approach.

Corollary 8.3. *For each $\epsilon > 0$, there exists $g = g(\epsilon)$ such that the number of labelled string graphs on n vertices with girth at least g is at most $(1 + \epsilon)^n T(n)$, where $T(n) = n^{n-2}$ is the number of labelled trees on n vertices.*

Acknowledgement

References

[1] Alon, N., Seymour, P. and Thomas, R. (1990) A separator theorem for nonplanar graphs. *J. Amer. Math. Soc.* **3** 801–808.

[2] Alon, N. and Spencer, J. (2000) *The Probabilistic Method*, 2nd edn, Wiley.

[3] Bollobás, B. (1998) *Modern Graph Theory*, Springer.

[4] Böttcher, J., Pruessmann, K. P., Taraz, A. and Würfl, A. (2008) Bandwidth, treewidth, separators, expansion, and universality. In *Proc. Topological. and Geometric Graph Theory* (TGGT 08), *Electron. Notes. Discrete Math.* **31** 91–96.

[5] Capoyleas, V. and Pach, J. (1992) A Turán-type theorem on chords of a convex polygon. *J. Combin. Theory Ser. B* **56** 9–15.

[6] Chung, F. (1988) Labelings of graphs. In *Selected Topics in Graph Theory*, Academic Press, San Diego, pp. 151–168.

[7] Diestel, R. (2000) *Graph Theory*, 2nd edn, Springer.

[8] Dvořák, Z. and Norine, S. Small graph classes and bounded expansion. *J. Combin. Theory Ser. B*, to appear.

[9] Erdős, P. (1959) Graph theory and probability. *Canad. J. Math.* **11** 34–38.

[10] Fox, J. (2006) A bipartite analogue of Dilworth's theorem. *Order* **23** 197–209.

[11] Fox, J. and Pach, J. (2008) Separator theorems and Turán-type results for planar intersection graphs. *Adv. Math.* **219** 1070–1080.

[12] Fox, J. and Pach, J. (2008) Coloring K_k-free intersection graphs of geometric objects in the plane. In *Proc. 24th ACM Sympos. on Computational Geometry*, ACM Press, New York, pp. 346–354.

[13] Fox, J. and Pach, J. String graphs and incomparability graphs. Submitted.

[14] Fox, J., Pach, J. and Tóth, C. D. Intersection patterns of curves. *J. London Math. Soc.*, to appear.

[15] Fox, J., Pach, J. and Tóth, C. D. Turán-type results for partial orders and intersection graphs of convex sets. *Israel J. Math,* to appear.

[16] Gilbert, J. R., Hutchinson, J. P. and Tarjan, R. E. (1984) A separator theorem for graphs of bounded genus. *J. Algorithms* **5** 391–407.

[17] Koebe, P. (1936) Kontaktprobleme der konformen Abbildung. Berichte über die Verhandlungen der Sachsischen Akademie der Wissenschaften, Leipzig, *Mathematische-Physische Klasse* **88** 141–164.

[18] Kolman, P. and Matoušek, J. (2004) Crossing number, pair-crossing number, and expansion. *J. Combin. Theory Ser. B* **92** 99–113.

[19] Kostochka, A. V. and Nešetřil, J. (1998) Coloring relatives of intervals on the plane I: Chromatic number versus girth. *Europ. J. Combin.* **19** 103–110.

[20] Knopp, K. (1956) *Infinite Sequences and Series*, Dover, New York.

[21] Kratochvíl, J. and Matoušek, J. (1989) NP-hardness results for intersection graphs. *Comment. Math. Univ. Carolin.* **30** 761–773.

[22] Kratochvíl, J. and Matoušek, J. (1994) Intersection graphs of segments. *J. Combin. Theory Ser. B* **62** 289–315.

[23] Kühn, D. and Osthus, D. (2004) Induced subdivisions in $K_{s,s}$-free graphs of large average degree. *Combinatorica* **24** 287–304.

[24] Leighton, T. and Rao, S. (1999) Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. Assoc. Comput. Mach.* **46** 787–832.

[25] Lipton, R. J., Rose, D. J. and Tarjan, R. E. (1979) Generalized nested dissection. *SIAM J. Numer. Anal.* **16** 346–358.

[26] Lipton, R. J. and Tarjan, R. E. (1979) A separator theorem for planar graphs. *SIAM J. Appl. Math.* **36** 177–189.

[27] Lipton, R. J. and Tarjan, R. E. (1980) Applications of a planar separator theorem. *SIAM J. Comput.* **9** 615–627.

[28] McDiarmid, C., Steger, A. and Welsh, D. (2005) Random planar graphs. *J. Combin. Theory Ser. B* **93** 187–206.

[29] Miller, G. L., Teng, S.-H., Thurston, W. and Vavasis, S. A. (1997) Separators for sphere-packings and nearest neighbor graphs. *J. Assoc. Comput. Mach.* **44** 1–29.

[30] Norine, S., Seymour, P., Thomas, R. and Wollan, P. (2006) Proper minor-closed families are small. *J. Combin. Theory Ser. B* **96** 754–757.

[31] Pach, J. and Agarwal, P. (1995) *Combinatorial Geometry*, Wiley, New York.

[32] Pach, J., Pinchasi, R., Sharir, M. and Tóth, G. (2005) Topological graphs with no large grids. *Graphs Combin.* **21** 355–364.

[33] Pach, J., Shahrokhi, F. and Szegedy, M. (1996) Applications of the crossing number. *Algorithmica* **16** 111–117.

[34] Pach, J. and Sharir, M. (2009) On planar intersection graphs with forbidden subgraphs. *J. Graph Theory* **59** 205–214.

[35] Pach, J. and Tóth, G. (2000) Which crossing number is it anyway? *J. Combin. Theory Ser. B* **80** 225–246.

[36] Pach, J. and Tóth, G. (2006) Comment on Fox News. *Geombinatorics* **15** 150–154.

[37] Preparata, F. and Shamos, M. (1985) *Computational geometry: An Introduction*, Texts and Monographs in Computer Science, Springer, New York.

[38] Radoičić, R. and Tóth, G. (2008) The discharging method in combinatorial geometry and its application to Pach–Sharir conjecture on intersection graphs. In *Proc. Joint Summer Research Conference on Discrete and Computational Geometry* (J. E. Goodman, J. Pach and J. Pollack, eds), Vol. 453 of *Contemporary Mathematics*, AMS, pp. 319–342.

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