Putting the Pieces Together

From Graph Decompositions to Ramsey Theory



Lukas Michel

Worcester College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity Term 2026

Abstract

This thesis investigates the decomposition of graphs into a small number of well-structured subgraphs and the unavoidable substructures that arise in such decompositions.

A celebrated problem in this area is the Linear Arboricity Conjecture, which states that the edges of every graph with maximum degree Δ can be decomposed into at most $\lceil (\Delta+1)/2 \rceil$ linear forests, where a linear forest is a collection of vertex-disjoint paths. Although this conjecture remains wide open, we prove that every graph with n vertices admits a decomposition into at most $\Delta/2 + \mathcal{O}(\log n)$ linear forests. For $\Delta \geq \Omega(n^{\varepsilon})$, this is an exponential improvement over the previous best error term. We also show that every d-regular graph has a spanning linear forest with at most $2 \cdot n/(d+1)$ paths and a cycle-factor with at most $\mathcal{O}((n \log d)/d)$ cycles. This confirms a conjecture of Feige and Fuchs, resolves a conjecture of Magnant and Martin up to a factor of two, and yields an efficient algorithm for finding short tours in connected regular graphs.

We then turn to an extremal problem in graph colouring. Here, we disprove a conjecture of Scott and Seymour by constructing graphs of arbitrarily large chromatic number together with tournaments on the same vertex set such that every out-neighbourhood in the tournament induces a bipartite subgraph of the graph.

The thesis concludes with several contributions to Ramsey theory. First, we establish that the k-colour Ramsey number of an odd cycle of length $2\ell+1$ is at most $(4\ell)^k \cdot k^{k/\ell}$. This proves a conjecture of Fox and is the first improvement in the exponent beyond an absolute constant factor since the work of Bondy and Erdős. Next, we substantially improve bounds towards a conjecture of Lefmann, Rödl, and Thomas concerning monotone monochromatic and rainbow paths in edge-coloured complete graphs. Finally, we determine the density and structure of product-free sets in free semigroups and the density of product-free sets in free groups, confirming conjectures of Leader, Letzter, Narayanan, and Walters as well as Ortega, Rué, and Serra.

About this thesis

The research in this thesis was carried out at the Mathematical Institute of the University of Oxford between October 2022 and May 2026, supervised by Professor Alex Scott. This thesis is based on the following joint work.

- Chapter 2 is based on joint work with Micha Christoph, Nemanja Draganić, António Girão, Eoin Hurley, and Alp Müyesser [CDG+25b].
- Chapter 3 is based on joint work with Micha Christoph, Nemanja Draganić, António Girão, Eoin Hurley, and Alp Müyesser [CDG+25a].
- Chapter 4 is based on joint work with António Girão, Kevin Hendrey, Freddie Illingworth, Florian Lehner, Michael Savery, and Raphael Steiner [GHI+24].
- Chapter 5 is based on joint work with Maria Axenovich, Wouter Cames van Batenburg, Oliver Janzer, and Mathieu Rundström [ACJ+26].
- Chapter 6 is based on joint work with António Girão, Freddie Illingworth, Michael Savery, and Alex Scott [GIM+24].
- Chapter 7 is based on joint work with Freddie Illingworth and Alex Scott [IMS25].

The material in this thesis has not been submitted for any other degree or qualification at any university.

During the preparation of this thesis, ChatGPT 5.5 by OpenAI was used for proofreading purposes. This use complies with University, divisional, and departmental guidance on generative AI.

Acknowledgements

Thank you to Alex for all the advice and guidance throughout the last four years, for the unending supply of interesting research problems that kept me busy, and for cultivating a research environment where I could find and explore the mathematics I enjoy most. I discovered combinatorics through your lectures during my master's degree, and I stayed for the exciting research and the welcoming community you introduced me to. I will miss the free hot chocolate.

Thank you to the DPhil students in the combinatorics group and beyond who made these years so special. To Emma, for always being there, for all the hiking across peaks and through canyons, and for countless Puzzled Pint sessions. To Michael, for the road trips around the world and for ensuring that there was always cake. To Jane, for the dumplings and sushi, whether handmade or overcooked. To Youri, for all the fun board games. And to Emil, Dmitry, Kate, Mira, Veronica, and many others. Thank you all for the frequent two-hour lunch breaks and for persevering through many joint projects that would have been far less fun without you.

I am also grateful to the many friendly combinatorialists I was fortunate enough to meet along the way. To Freddie, for showing us how much we still have to learn at croquet. To António, for many memorable lunch conversations. To Eoin, for pushing our research projects even further. To Carla, Louigi, and Marthe, for several wonderful research visits. And to so many others, in Oxford and further afield, for great collaborations, for useful advice, and for making every workshop and conference a fun experience.

Thank you also to my housemates and friends for making Oxford feel like home. To Hannah, Maddy, Rohit, and Spencer, for the quests and campaigns. To Josh, for the woodland adventures. And to the Worcester College ultimate team, for the deep throws and hammers.

Finally, thank you to my family for your continued love, encouragement, and support throughout this journey.

Contents

1 Introduction **1**

1.1 Graph decompositions . . . 3

1.1.1 Linear arboricity . . . 5

1.1.2 Spanning linear forests with few paths . . . 7

1.1.3 Cycle-factors with few cycles . . . 8

1.2 Graph colouring . . . 9

1.2.1 Chromatic number is not tournament-local . . . 10

1.3 Ramsey theory . . . 12

1.3.1 Multicolour Ramsey numbers of odd cycles . . . 13

1.3.2 Flashes and rainbows in tournaments . . . 15

1.3.3 Product-free sets in free semigroups and free groups . . . 16

2 Linear arboricity **21**

2.1 Proof techniques . . . 25

2.2 Small linear forests in regular graphs . . . 26

2.3 Logarithmic error for linear arboricity . . . 36

2.4 Open problems . . . 40

3 Cycle-factors with few cycles **43**

3.1 Entropy . . . 46

3.2 Expected number of cycles in a random cycle-factor . . . 47

3.3 Open problems . . . 53

4 Chromatic number is not tournament-local **55**

4.1 Chromatic number . . . 57

4.2 Fractional chromatic number . . . 61

4.3 Degeneracy . . . 62

4.4 Open problems . . . 63

5 Multicolour Ramsey numbers of odd cycles **65**
5.1 Neighbourhoods with small chromatic number 68
5.2 Consequences for short monochromatic odd cycles 69

6 Flashes and rainbows in tournaments **71**
6.1 Upper bounds for arbitrary tournaments 74
6.2 Upper bounds for transitive tournaments 80
6.3 Colourings of arbitrary tournaments 84
6.4 Open problems 86

7 Product-free sets in free semigroups and free groups **89**
7.1 Density 93
7.1.1 Upper asymptotic and upper Banach density 93
7.1.2 Diagonalisation and relative density 95
7.1.3 Key technical density results 98
7.2 Structure of product-free sets in free semigroups 101
7.2.1 Structure of strongly k-product-free sets 101
7.2.2 Structure of k-product-free sets 105
7.3 Density of product-free sets in free semigroups 110
7.3.1 Steeplechases 111
7.3.2 Density of (strongly) k-product-free sets 116
7.4 Product-free sets in free groups 124
7.5 Open problems 129

Bibliography **133**

Chapter 1

Introduction

When can large and complex graphs be decomposed into simpler structures? This fundamental question from graph theory underpins a vast array of research in combinatorics. One of the earliest examples of such a graph decomposition problem was posed by Kirkman in 1850 in *The Lady's and Gentleman's Diary*.

> Fifteen young ladies in a school walk out three abreast for seven days in succession: it is required to arrange them daily so that no two shall walk twice abreast.

Put differently, Kirkman was asking for a decomposition of the edges of the complete graph on fifteen vertices into copies of a triangle and to group these triangles into seven groups of five vertex-disjoint triangles each.

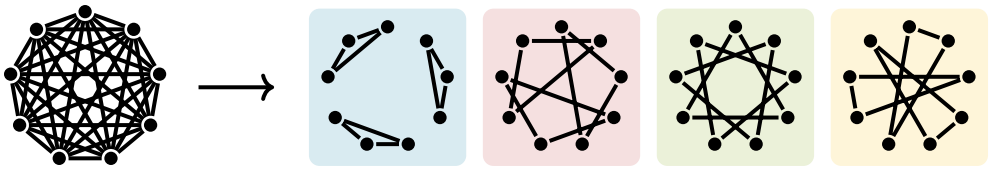


Figure 1.1: A decomposition of the complete graph on nine vertices into four groups of vertex-disjoint triangles.

What if instead of decomposing the complete graph into triangles, we want to decompose it into complete subgraphs or longer cycles? Besides decomposing complete graphs, can we also decompose other graphs into similar structures? And rather than decomposing the edges of a graph, what happens if we decompose its vertices? All of these variations lead to various interesting questions. As a result, the study of graph decompositions has developed into a vibrant area of active research in combinatorics with many long-standing open problems and exciting recent developments.

Graph decompositions can also be encoded using colours by assigning a different colour to each part of the decomposition. From this perspective, we want to colour the edges or vertices of a graph so that the monochromatic subgraphs satisfy certain structural properties. Graph colourings date back to 1852 when Guthrie noticed that the map of counties of England can be coloured with only four colours while ensuring that neighbouring counties receive different colours. This led to the following famous Four Colour Conjecture.

> The regions of every map can be coloured with at most four colours so that regions which share a border receive different colours.

In the language of graph theory, this conjecture posits that the chromatic number of every planar graph is at most four, where the chromatic number of a graph is the minimum number of colours required for a vertex-colouring of the graph in which adjacent vertices receive different colours.



Figure 1.2: A vertex-colouring of a planar graph with four colours.

The Four Colour Conjecture turned out to be surprisingly difficult. Despite sustained effort, it was only in 1976 that Appel and Haken proved this conjecture in one of the first computer-assisted proofs in mathematics. To this day, no elementary proof of this result is known.

If every planar graph has a small chromatic number, what causes the chromatic number of a graph to be large? While a large chromatic number tells us something about the global structure of a graph, does it also provide any information about the local structure? Understanding the connections between chromatic number and graph structure remains a major focus of graph theory and has inspired a large body of research in both mathematics and computer science.

Graph colouring can also be viewed through the lens of Ramsey theory, whose underlying principle states that order appears even amid disorder. This field was initiated in 1930 by Ramsey's theorem, which proves the following.

> If the edges of a very large complete graph are coloured with a fixed number of colours, then there exists a large complete subgraph whose edges all have the same colour.

This provides a complementary perspective on graph decompositions, stating that if we decompose the edges of a sufficiently large complete graph into a fixed number of parts, then one part must contain a large complete subgraph. As an example, if we colour the edges of a complete graph on six vertices with red and blue, there is always a monochromatic triangle.

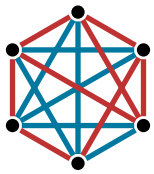


Figure 1.3: A colouring of the edges of the complete graph on six vertices with red and blue.

Once again, we may ask what happens if we colour other graphs, if we attempt to find a different unavoidable substructure, or if we vary the properties that the colouring of the obtained subgraph has to satisfy. The principle of finding order in every sufficiently large mathematical object also extends far beyond graphs and has found applications in number theory, geometry, and other areas.

This thesis studies combinatorial problems across all three topics introduced above. First, we study sparse graph decomposition problems, proving new upper bounds for linear arboricity as well as for spanning linear forests and cycle-factors in regular graphs. Second, we investigate locality phenomena for the chromatic number of graphs and tournaments on the same vertex set. Third, we prove new Ramsey-type results for multicolour Ramsey numbers of odd cycles, a canonical Ramsey problem for monotone paths, and product-free sets in free semigroups and free groups. Although these problems mostly revolve around the common theme of graph decompositions, we will also encounter connections to probability theory, information theory, group theory, and theoretical computer science.

1.1 GRAPH DECOMPOSITIONS

The area of graph decomposition focuses on partitioning the edges or vertices of a graph into a collection of smaller subgraphs that satisfy certain structural properties. We already described Kirkman's problem that asks for a decomposition of the edges of a complete graph into triangles, the smallest possible cycles in a graph. What about decompositions of the edges of a complete graph K_n into

Hamilton cycles, which are cycles that contain all n vertices? Since every Hamilton cycle would cover exactly two edges incident to every vertex, this would require the degree of each vertex to be even, and so n must be odd. It turns out that this condition is also sufficient. If n is odd, Walecki proved in 1882 that K_n can indeed be decomposed into Hamilton cycles.

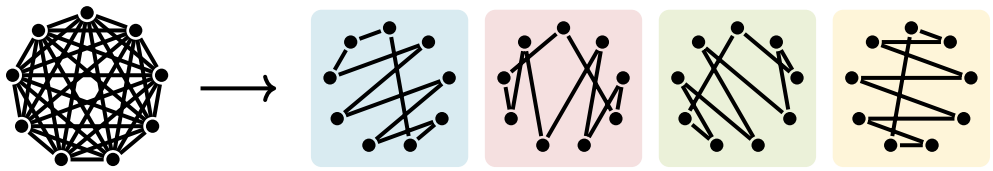


Figure 1.4: A decomposition of K_9 into Hamilton cycles.

What other graphs can the complete graph K_n be decomposed into? The answer to this question is perhaps counter-intuitive: if basic divisibility conditions are satisfied, then K_n can be decomposed into a wide variety of graphs. Examples include large trees in Ringel's conjecture [MPS21, KS25], complete graphs in the study of designs [Kee14, GKLO23], cycles of various lengths in the Oberwolfach problem [GJK+21, KS22], and many others. Notably, all of these central conjectures have been proved in the past few decades as the machinery associated with absorption and the probabilistic method has become more sophisticated and versatile [AS16, KKK+23, GKLO23, Mon24, KSSS24].

The complete graph, however, is a very special and highly symmetric object, and so it is unclear exactly which of its properties are necessary for these decomposition results. One way to investigate this is to consider dense graphs as a slight weakening of complete graphs. It turns out that here, too, a surprising variety of decompositions is possible, and significant progress has been made in the past few decades. Perhaps the largest body of work in this direction concerns the decomposition of graphs that satisfy minimum degree conditions. Prominent examples of this include progress on the Nash-Williams conjecture on triangle decompositions [BKLO16, DP21], generalising Kirkman's problem, and the Hamilton decomposition theorem for regular Dirac graphs [CKL+16], generalising Walecki's result. Another exciting recent direction concerns the decomposition of graphs that are dense and quasirandom. Here, results include generalisations of Ringel's conjecture [KS25] and of the Oberwolfach problem [KS22], as well as approximate decompositions into arbitrary graphs of bounded degree [KKOT19]. This work shares some machinery with the case of complete graphs and makes frequent use of Szemerédi's famous regularity lemma.

Overall, for the decomposition of dense graphs into graphs of fixed size or very sparse spanning structures, the following perspective has emerged.

> If there is no obvious barrier to the existence of a particular graph decomposition, then the decomposition likely exists.

But what if we go beyond complete and dense graphs? After all, many graphs arising in areas like statistical physics, computer science, or group theory are graphs of bounded degree. Is it still true that beyond obvious barriers, decompositions exist? Of course, in general we can no longer hope for decompositions into complete graphs or Hamilton cycles since sparse graphs may not even contain these structures. So, we have to ask for different decompositions.

Two classical and elegant results on sparse graph decompositions are Vizing's theorem [Viz64] and the Nash-Williams theorem [NW61, NW64]. Vizing's theorem states that any graph of maximum degree Δ can be decomposed into $\Delta + 1$ matchings. The Nash-Williams theorem states that a graph can be decomposed into k forests if and only if every set of ℓ vertices spans at most $k \cdot (\ell - 1)$ edges. Thus, in both cases, decompositions exist beyond the obvious barriers.

In recent decades, however, progress on sparse graph decomposition problems has lagged far behind that on their dense counterparts. This is largely because various tools—such as Szemerédi's regularity lemma—offer little leverage in sparse graphs. Among the most famous open problems in this area are the List Edge Colouring Conjecture [Viz76, ERT80, Kah00, MR00] on decomposing graphs into matchings where each edge is allowed only in certain matchings, the Erdős–Gallai Conjecture on decomposing graphs into cycles and edges [EGP66, BM24], and the Linear Arboricity Conjecture.

1.1.1 LINEAR ARBORICITY

A *linear forest* is a collection of vertex-disjoint paths, and the *linear arboricity* $\mathrm{la}(G)$ of a graph G is the minimum number of linear forests needed to decompose the edges of G. This graph invariant was introduced in 1970 by Harary [Har70].

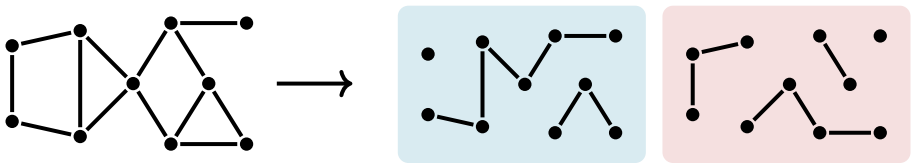


Figure 1.5: A decomposition of a graph into two linear forests.

Matchings are a special case of linear forests, and linear forests are a special case of forests, so the linear arboricity of a graph lies between the aforementioned results of Vizing and Nash-Williams. Since Hamilton paths are linear forests with the maximum possible number of edges, linear arboricity can also be viewed as a relaxation of a decomposition into Hamilton paths.

How large is the linear arboricity of a graph? Since a linear forest covers at most two edges incident to any vertex, every graph G with maximum degree Δ satisfies $\mathrm{la}(G) \geq \lceil \Delta/2 \rceil$. Moreover, a Δ-regular graph G on n vertices has $n \cdot \Delta/2$ edges while any linear forest of G has strictly less than n edges, and so $\mathrm{la}(G) \geq \lceil (\Delta+1)/2 \rceil$. The celebrated Linear Arboricity Conjecture of Akiyama, Exoo, and Harary [AEH80] asserts that, beyond these obvious barriers, there should always be a decomposition into linear forests.

Conjecture 2.1 (Linear Arboricity Conjecture)**.** *The linear arboricity of a graph with maximum degree Δ is at most $\lceil (\Delta+1)/2 \rceil$.*

This conjecture immediately received significant attention and has been proved in many special cases since then. In an influential paper from 1988, Alon [Alo88] also confirmed the Linear Arboricity Conjecture asymptotically, showing that

$$\mathrm{la}(G) \leq \frac{\Delta}{2} + \mathcal{O}\left(\frac{\Delta \log\log \Delta}{\log \Delta}\right) = (1+o(1)) \cdot \frac{\Delta}{2}.$$

This was one of the early applications of the Lovász Local Lemma, giving the Linear Arboricity Conjecture a prominent place in the history of the probabilistic method [AS16]. After several subsequent improvements, the error term of Alon's upper bound was recently reduced to $\mathcal{O}(\sqrt{\Delta}\log^4 \Delta)$ by Lang and Postle [LP23]. However, an error of $\Omega(\sqrt{\Delta})$ is a natural bottleneck for these asymptotic upper bounds due to the randomisation that was a central ingredient in all previous proofs and caused errors of order $\Theta(\sqrt{\Delta})$ due to its standard deviation.

In Chapter 2, we prove the following upper bound on the linear arboricity of a graph with an error that is logarithmic in the number of vertices of the graph.

Theorem 2.2. *Every graph G on n vertices with maximum degree Δ satisfies*

$$\mathrm{la}(G) \leq \frac{\Delta}{2} + \mathcal{O}(\log n).$$

This result breaks the square-root barrier for the Linear Arboricity Conjecture if $\Delta \geq \omega(\log^2 n)$ and represents an exponential improvement in the error term

for $\Delta \geq \Omega(n^{\varepsilon})$. Previously, there were no results that improved on the upper bound of Lang and Postle [LP23] even for very dense graphs with $\Delta \geq \Omega(n)$. In addition, we reduce the error term for the fractional relaxation of linear arboricity to $\mathcal{O}(\log \Delta)$, improving the previous best error term of $\mathcal{O}(\sqrt{\Delta})$ by Feige, Ravi, and Singh [FRS14] for all ranges of Δ. We achieve these results by generalising Pósa rotations from rotations of one endpoint of a path to simultaneous rotations of multiple endpoints of a linear forest.

These results are based on joint work with Micha Christoph, Nemanja Draganić, António Girão, Eoin Hurley, and Alp Müyesser [CDG+25b].

1.1.2 SPANNING LINEAR FORESTS WITH FEW PATHS

If the Linear Arboricity Conjecture holds, then, by averaging, it follows that every d-regular graph G on n vertices has a linear forest with at least

$$\frac{\frac{n \cdot d}{2}}{\mathrm{la}(G)} \geq \frac{n \cdot d}{d+2}$$

edges. If we add missing vertices as isolated vertices to this linear forest, this implies that G has a spanning linear forest with as few as $2 \cdot n/(d+2)$ paths. Equivalently, this means that the vertices of G can be decomposed into at most $2 \cdot n/(d+2)$ vertex-disjoint paths.

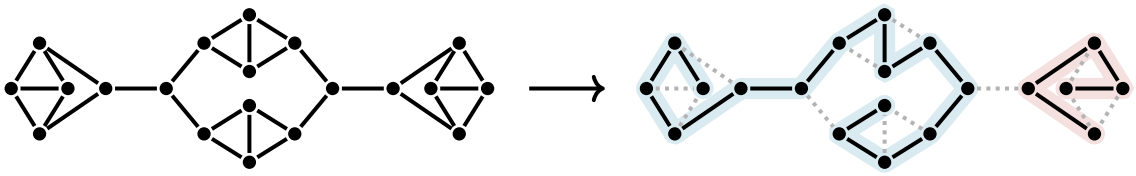


Figure 1.6: A spanning linear forest of a 3-regular graph with two paths.

This consequence of the Linear Arboricity Conjecture is a challenging and well-studied problem in its own right. In 2009, Magnant and Martin [MM09] made the following conjecture.

Conjecture 2.4 (Magnant and Martin)**.** *Every d-regular graph on n vertices has a spanning linear forest with at most $n/(d+1)$ paths.*

This conjecture would be tight, as witnessed by a disjoint union of complete graphs. However, despite recent interest, this conjecture is only known to hold in a few special cases. Due to the lack of progress, Feige and Fuchs [FF22] later conjectured the weaker statement that d-regular graphs have spanning linear forests

with $\mathcal{O}(n/d)$ paths, but even this conjecture remained open. In fact, the progress on this problem was stuck at the same square-root barrier as for the Linear Arboricity Conjecture. The best upper bound by Feige, Ravi, and Singh [FRS14] only provided a spanning linear forest with $\mathcal{O}(n/\sqrt{d})$ paths.

In Chapter 2, using the same methods that we develop for linear arboricity, we establish the conjecture of Magnant and Martin up to a factor of two. In particular, this proves the conjecture of Feige and Fuchs.

Theorem 2.5. *Every d-regular graph on n vertices has a spanning linear forest with at most $2 \cdot n/(d+1)$ paths.*

For regular bipartite graphs, we even confirm the conjecture of Magnant and Martin. Moreover, our results also imply that every connected d-regular graph has a tour of length at most $(1+\mathcal{O}(1/d)) \cdot n$, where a tour is a sequence of adjacent vertices that visits every vertex and that starts and ends at the same vertex. This is optimal and significantly improves the previous best bound of $(1+\mathcal{O}(1/\sqrt{d})) \cdot n$ by Feige, Ravi, and Singh [FRS14].

These results are again based on joint work with Micha Christoph, Nemanja Draganić, António Girão, Eoin Hurley, and Alp Müyesser [CDG+25b].

1.1.3 CYCLE-FACTORS WITH FEW CYCLES

We just saw that we can decompose the vertices of a regular graph into a small number of vertex-disjoint paths. Can we strengthen this result to a decomposition into a small number of vertex-disjoint cycles? A *cycle-factor* of a graph is a decomposition of its vertices into vertex-disjoint cycles, where we consider a single edge to be a cycle of length two. Using Hall's theorem, it can be shown that every regular graph has at least one cycle-factor.

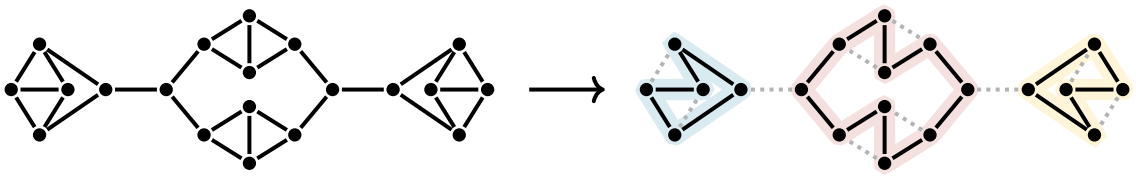


Figure 1.7: A cycle-factor of a 3-regular graph with three cycles.

If we delete one edge from each cycle in a cycle-factor, we obtain a spanning linear forest with as many paths as there were cycles in the cycle-factor. Since we know that every d-regular graph on n vertices has a spanning linear forest with $\mathcal{O}(n/d)$ paths, we may therefore ask whether there is also a cycle-factor with $\mathcal{O}(n/d)$

cycles. However, in contrast to spanning linear forests where an upper bound of $\mathcal{O}(n/\sqrt{d})$ was known [FRS14], for cycle-factors the previous best bound by Vishnoi [Vis12] only guaranteed at most $\mathcal{O}(n/\sqrt{\log d})$ cycles.

In Chapter 3, we significantly improve this upper bound and match our bound for spanning linear forests up to a logarithmic factor.

Theorem 3.2. *Every d-regular graph on n vertices has a cycle-factor with at most $\mathcal{O}((n \log d)/d)$ cycles.*

In fact, we prove that this bound holds for all directed regular graphs and that a uniformly random cycle-factor of such a graph has $\mathcal{O}((n \log d)/d)$ cycles on average. For random cycle-factors, this is tight up to the constant factor and generalises the classical result that a uniformly random permutation of n elements has $\mathcal{O}(\log n)$ cycles. In addition, our bound also implies that there exists an efficient algorithm for computing such a cycle-factor. We prove this bound using the language of entropy to exploit the fact that the upper and lower bounds on the number of perfect matchings in regular bipartite graphs are extremely close.

These results are once again based on joint work with Micha Christoph, Nemanja Draganić, António Girão, Eoin Hurley, and Alp Müyesser [CDG+25a].

1.2 GRAPH COLOURING

Perhaps the simplest possible graph decomposition we could ask for is to decompose the vertices of a graph into independent sets, where an independent set is a set of vertices that are pairwise non-adjacent. Usually, such decompositions are phrased as vertex-colourings of the graph where adjacent vertices need to receive different colours. A decomposition into few independent sets then corresponds to using few colours. This is directly linked to the *chromatic number* $\chi(G)$ of a graph G as the minimum number of colours required in such a vertex-colouring.

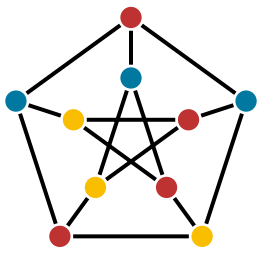


Figure 1.8: A vertex-colouring of the Petersen graph with three colours.

This very simple definition leads to surprising complexity. We already encountered the Four Colour Conjecture, which states that planar graphs have chromatic number at most four, but this was notoriously difficult to prove, and the attempts to do so sparked the entire field of graph colouring. The problem of determining whether a graph can be coloured with k colours was also one of the first problems shown to be NP-hard by Karp [Kar72] in 1972. This makes it very unlikely that there is an efficient algorithm that computes the chromatic number of a graph.

What structures cause the chromatic number of a graph to be large? One simple reason can be that the graph contains a large complete subgraph since this forces all vertices of the complete subgraph to receive a different colour. However, constructions from the 1940s and 1950s, for example by Tutte [Des54] and Zykov [Zyk49], show that there exist graphs of large chromatic number that do not even contain triangles. In fact, with one of the first applications of the probabilistic method in graph theory, Erdős [Erd59] showed that for all ℓ there are graphs of large chromatic number that contain no cycle of length at most ℓ. This shows that the chromatic number of a graph is predominantly a global property of the graph that cannot necessarily be explained by its local substructures.

1.2.1 CHROMATIC NUMBER IS NOT TOURNAMENT-LOCAL

The constructions of triangle-free graphs with large chromatic number show that there are graphs with large chromatic number where the neighbourhood of every vertex is extremely simple, namely just an independent set. Berger, Choromanski, Chudnovsky, Fox, Loebl, Scott, Seymour, and Thomassé [BCC+13] conjectured that the analogous phenomenon does not occur for tournaments. A *tournament* is an orientation of a complete graph, and the *chromatic number* $\chi(T)$ of a tournament T is the minimum number of parts into which the vertices of T can be partitioned so that each part induces an acyclic tournament. Equivalently, this is the minimum number of colours required to colour the vertices of T such that every monochromatic set of vertices contains no directed cycle.

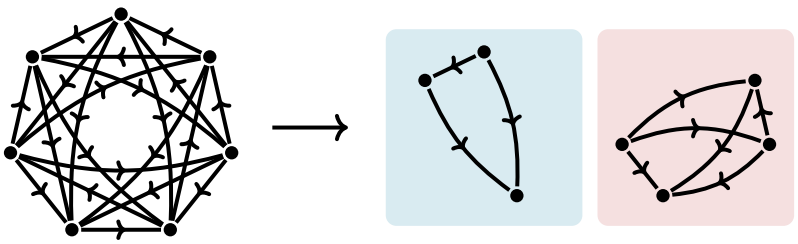


Figure 1.9: A vertex-partition of a tournament into two acyclic tournaments.

Harutyunyan, Le, Thomassé, and Wu [HLTW19] confirmed this conjecture by showing that for every k there exists a χ such that every tournament T with $\chi(T) \geq \chi$ contains a vertex v with $\chi(T[N_T^+(v)]) \geq k$, where $N_T^+(v)$ denotes the set of out-neighbours of v in T. That is, every tournament with large chromatic number contains a vertex whose out-neighbourhood has large chromatic number itself, and so this out-neighbourhood is not simple.

Note that in contrast to graphs, a typical out-neighbourhood of a vertex in a tournament is large. It seemed plausible that this might be the only reason why there is an out-neighbourhood with large chromatic number. Indeed, Scott and Seymour [SS16] conjectured that a similar result should also hold for a graph and a tournament on the same vertex set (see also [HLTW19, Conjecture 7]). This would imply that every tournament always provides a local reason for why a graph has large chromatic number.

Conjecture 4.1 (Scott and Seymour)**.** *For every positive integer k there exists a χ such that for every graph G with $\chi(G) \geq \chi$ and every tournament T on the same vertex set, there is a vertex v such that $\chi(G[N_T^+(v)]) \geq k$.*

This conjecture was supported by their observation [SS16] that the statement holds if both occurrences of chromatic number are replaced by their fractional relaxation. Harutyunyan, Le, Thomassé, and Wu [HLTW19] also asked whether a similar result holds for two tournaments on the same vertex set.

In Chapter 4, we disprove the conjecture of Scott and Seymour for $k \geq 3$. In fact, we show that G and T can be chosen so that for any small set of vertices U, the combined out-neighbourhood $N_T^+(U) := \bigcup_{u \in U} N_T^+(u)$ induces a bipartite graph.

Theorem 4.2. *For every positive integer χ there are arbitrarily large n for which there is a triangle-free graph G and a tournament T on the same vertex set of size n such that $\chi(G) = \chi$ and, for every set U of at most $(\log n)/(2\chi^2)$ vertices, $\chi(G[N_T^+(U)]) \leq 2$.*

We also use this result to answer the question of Harutyunyan, Le, Thomassé, and Wu [HLTW19] negatively and to disprove a related conjecture of Nguyen, Scott, and Seymour [NSS25]. The graph in our construction is the classical Schrijver graph [Sch78] for which we show that an appropriate tournament exists that proves this result. In addition, we show that the setting where chromatic number is replaced by degeneracy exhibits a quite different behaviour.

These results are based on joint work with António Girão, Kevin Hendrey, Freddie Illingworth, Florian Lehner, Michael Savery, and Raphael Steiner [GHI+24].

1.3 RAMSEY THEORY

The final perspective on graph decompositions that we will consider is that of Ramsey theory, which asks what substructures we cannot avoid if we decompose a graph into few pieces. A guiding principle of Ramsey theory is that complete disorder is impossible, and so one of those pieces should contain an ordered subgraph. However, this principle also applies to many other mathematical objects, which means that we do not have to restrict ourselves only to graphs.

One of the earliest results of this type, even before Ramsey's theorem, was proved by Schur [Sch17] in 1917. Schur showed that if you colour the integers from 1 to n with a fixed number of colours, then if n is sufficiently large, there will be three integers x, y, z with the same colour such that $x + y = z$. For example, for three colours, taking $n = 14$ suffices.

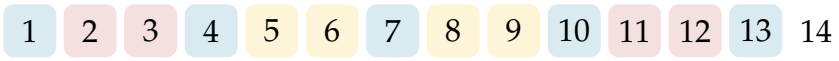


Figure 1.10: A colouring of the integers from 1 to 13 with three colours that avoids a monochromatic solution to $x + y = z$. This colouring cannot be extended to 14.

We can also translate this problem into a problem about graphs. Indeed, construct the complete graph on the integers from 1 to $n + 1$ and colour an edge between two integers $a < b$ with the colour of $b - a$. Then, if this complete graph contains a monochromatic triangle between $a < b < c$, the three integers $x := b - a$, $y := c - b$, and $z := c - a$ have the same colour and $x + y = z$, as required.

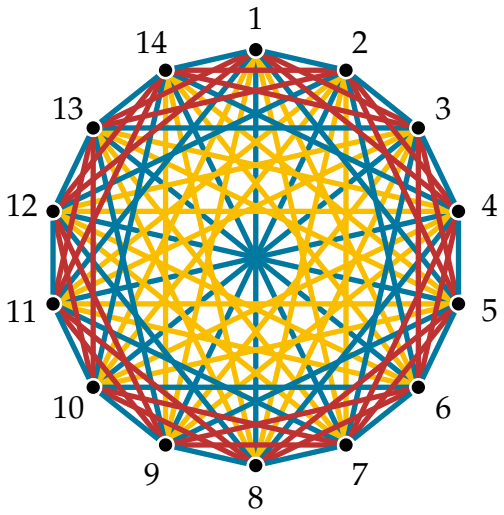


Figure 1.11: The colouring of the edges of the complete graph on the integers from 1 to 14 obtained from Figure 1.10.

How large must the complete graph be until a monochromatic triangle is guaranteed to exist? For any graph H, the *k-colour Ramsey number* $R_k(H)$ of H is the minimum integer n such that every colouring of the edges of K_n with k colours contains a monochromatic copy of H. In particular, in the above problem, as long as $n+1 \geq R_k(K_3)$, there exists a monochromatic triangle, and so we obtain a monochromatic solution to $x+y=z$.

Although Ramsey's theorem implies that Ramsey numbers exist, it is surprisingly difficult to determine good bounds for them. In the most fundamental case of two colours and complete graphs, for a long time the best bounds were

$$\Omega(2^{\ell/2}) \leq R_2(K_\ell) \leq \mathcal{O}(4^\ell).$$

The lower bound was established by Erdős [Erd47] in 1947 with a non-constructive proof that initiated the use of the probabilistic method in combinatorics. Since then, this lower bound has remained unchanged up to the constant factor. The upper bound was established by Erdős and Szekeres [ES35] in 1935, and it was only very recently that Campos, Griffiths, Morris, and Sahasrabudhe [CGMS26] gave the first exponential improvement for this upper bound.

Although the lower and upper bounds for $R_2(K_\ell)$ remain far apart, we know at least that $R_2(K_\ell)$ grows exponentially in ℓ. In contrast, if the graph is fixed and the number of colours grows, our understanding is often much worse. For example, in the case of the triangle, the notoriously difficult Schur–Erdős problem asks to determine the growth rate of $R_k(K_3)$. The arguments of Schur [Sch17] imply that

$$\Omega(3^k) \leq R_k(K_3) \leq \mathcal{O}(k!).$$

Erdős [CG98] conjectured that $R_k(K_3) = 2^{\Theta(k)}$ and offered monetary awards for proving this conjecture and solving some related problems. However, the upper bound on $R_k(K_3)$ has remained unchanged up to small improvements to the constant factor, and the lower bound has only been improved to $\Omega(3.28^k)$ [ACP+21].

1.3.1 MULTICOLOUR RAMSEY NUMBERS OF ODD CYCLES

As it seems very difficult to determine the multicolour Ramsey number of a triangle, what about longer cycles? Let C_ℓ denote the cycle of length ℓ. While cycles of even length behave quite differently compared to a triangle, for odd cycles,

Bondy and Erdős [BE73] and Erdős and Graham [EG75] obtained the bounds

$$\ell \cdot 2^k + 1 \leq R_k(C_{2\ell+1}) \leq 2\ell \cdot (k+2)!.$$

If k is fixed, the lower bound is sharp by a result of Jenssen and Skokan [JS21]. They proved that $R_k(C_{2\ell+1}) = \ell \cdot 2^k + 1$ for all sufficiently large ℓ. Unfortunately, if ℓ is fixed and k is large, our understanding is once again much worse. In that setting, the lower bound is no longer tight. This was shown by Day and Johnson [DJ17], who proved that for all ℓ there exists some $\delta := \delta(\ell) > 0$ such that $R_k(C_{2\ell+1}) \geq 2\ell \cdot (2+\delta)^{k-1}$ for all sufficiently large k. In terms of upper bounds when the number of colours is large, the existence of longer odd cycles should usually be easier to prove than the existence of shorter odd cycles. For this reason, Fox [Fox] made the following conjecture.

Conjecture 5.1 (Fox). *For every $\varepsilon > 0$ there exists some ℓ such that $R_k(C_{2\ell+1}) \leq k^{\varepsilon k}$ for all sufficiently large k.*

Despite the interest in upper bounds for $R_k(C_{2\ell+1})$, there was only very limited progress towards this conjecture. The best bound by Lin and Chen [LC19] showed that for $\ell \geq 2$ we have $R_k(C_{2\ell+1}) \leq c^k \cdot \sqrt{k!}$, where c is a constant that only depends on ℓ.

In Chapter 5, we significantly improve this upper bound and establish the conjecture of Fox. Our result also shows that a previous conditional upper bound of Li [Li09] holds unconditionally.

Theorem 5.2. *For all positive integers k and ℓ,*

$$R_k(C_{2\ell+1}) \leq (4\ell - 2)^k \cdot k^{k/\ell} + 1.$$

In addition to finding a monochromatic odd cycle of a specific length, we also make progress on finding any short monochromatic odd cycle. We prove these results by bounding the number of vertices in a complete graph whose edges are coloured in such a way that each monochromatic subgraph with small diameter has small chromatic number.

These results are based on joint work with Maria Axenovich, Wouter Cames van Batenburg, Oliver Janzer, and Mathieu Rundström [ACJ+26].

1.3.2 FLASHES AND RAINBOWS IN TOURNAMENTS

Perhaps one of the most important generalisations of Ramsey's theorem for multiple colours is the canonical Ramsey theorem of Erdős and Rado [ER50] from 1950. It states that every colouring of the edges of a very large complete graph on the integers from 1 to n with an arbitrary number of colours contains a large complete subgraph with one of the following four canonical types of colourings:

- rainbow — no two edges have the same colour,
- monochromatic — all edges have the same colour,
- upper lexical — two edges have the same colour if and only if their larger vertices coincide, or
- lower lexical — two edges have the same colour if and only if their smaller vertices coincide.

This theorem has an immediate corollary for monotone paths, which are paths where each vertex is larger than the vertex preceding it. It implies that there is an integer $f(\ell, k)$ such that if $n > f(\ell, k)$ then every colouring of the edges of a complete graph on the ordered integers from 1 to n either contains an *ℓ-flash*, which is a monotone path with ℓ edges that all have the same colour, or contains a *k-rainbow*, which is a monotone path with k edges that all have distinct colours.

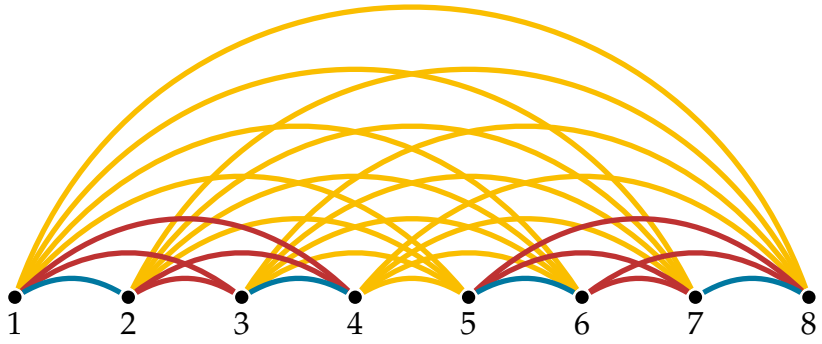


Figure 1.12: A colouring of the edges of the complete graph on the integers from 1 to 8 with no 2-flash and no 4-rainbow.

In 1992, Lefmann, Rödl, and Thomas [LRT92] provided a simple construction showing that $f(\ell, k) \geq \ell^{k-1}$. Their construction uses only $k-1$ colours and is optimal if only $k-1$ colours are used. However, they made the attractive conjecture that this lower bound is tight in general.

Conjecture 6.1 (Lefmann, Rödl, and Thomas)**.** *For all positive integers ℓ and k, we have $f(\ell, k) = \ell^{k-1}$.*

In support of their conjecture, they showed that it holds for $\ell \leq 2$, for $k \leq 4$, as well as for $\ell \geq (3k)^{2k}$. In 2000, Jiang and Mubayi [JM00] gave the first general upper bound on $f(\ell, k)$, showing that $f(\ell, k) \leq (1 + 1/\sqrt{\ell})^{k-4} \cdot \ell^{k-1}$ for $k \geq 4$. Although this confirms the conjecture asymptotically for $\ell \geq \omega(k^2)$, when ℓ is fixed and k is large, this upper bound differs from the lower bound by a factor that is exponential in k. Note that in the lower bound construction, the case where ℓ is fixed and k is large corresponds to having many colours, and so this is once again the case where our understanding is much worse.

In Chapter 6, we reduce the gap between the upper and lower bounds to a factor that is only polynomial in k.

Theorem 6.4. *For all $\varepsilon > 0$, there exists some D such that for all integers $\ell \geq 1$ and $k \geq 2$, we have*

$$f(\ell, k) \leq Dk(\log k)^{1+\varepsilon} \cdot \ell^{k-1}.$$

We also show that the conjecture of Lefmann, Rödl, and Thomas already holds when ℓ is only polynomially large in k.

Theorem 6.2. *For all $\varepsilon > 0$, there exists some C such that for all positive integers ℓ and k with $\ell \geq Ck^3(\log k)^{1+\varepsilon}$, we have $f(\ell, k) = \ell^{k-1}$.*

In addition, we generalise some of these results to tournaments. In that setting, we also investigate the question of how small a tournament can be while guaranteeing that every colouring of its edges contains an ℓ-flash or a k-rainbow. We prove many of these results by finding so-called robust vertices in the graph or tournament. These are vertices such that even if we forbid a colour, we can still find long rainbows without this colour that end at the robust vertex.

These results are based on joint work with António Girão, Freddie Illingworth, Michael Savery, and Alex Scott [GIM+24].

1.3.3 PRODUCT-FREE SETS IN FREE SEMIGROUPS AND FREE GROUPS

If we decompose a mathematical object into few pieces, then one of these pieces must be large. Sometimes, this property alone is sufficient to guarantee that the piece contains an ordered substructure. For example, in the problem investigated by Schur, any subset with more than $\lceil n/2 \rceil$ integers from 1 to n necessarily contains three integers x, y, z with $x + y = z$. This is tight since the subset of all odd integers contains no solution to $x + y = z$.

It turns out that this example is a classical problem from extremal group theory. We say that a subset S of a (semi)group G is *product-free* if $x \cdot y \notin S$ for all $x, y \in S$. The above problem then corresponds to the non-negative integers under addition, where the term 'sum-free' is used in place of 'product-free'. In this setting, every sum-free subset has density at most $1/2$, and Łuczak [Łuc95] showed that every sum-free subset of density $1/2$ is a subset of the odd integers.

What about other (semi)groups? How dense is a largest product-free subset? And what is the structure of a densest product-free subset? These problems have been extensively studied over the last fifty years [Ked09, TV17]. In the finite abelian case, this culminated in a solution to the density problem by Green and Ruzsa [GR05] and the structure problem by Balasubramanian, Prakash, and Ramana [BPR16]. The finite non-abelian case was first investigated by Babai and Sós [BS85]. This case behaves very differently since largest product-free subsets can have vanishing density as shown by the seminal work of Gowers [Gow08] on quasirandom groups. Recent breakthroughs include the alternating group where Eberhard [Ebe16] solved the density problem (up to logarithmic factors) and Keevash, Lifshitz, and Minzer [KLM24] solved the structure problem.

Since much less was known in the infinite case, Leader, Letzter, Narayanan, and Walters [LLNW20] recently considered these problems for the *free semigroup* $\mathcal{F}$ on a finite alphabet $\mathcal{A}$, which is the set of all finite words with letters from $\mathcal{A}$ equipped with the operation of concatenation. They solved the density problem for free semigroups by showing that any product-free subset $S \subseteq \mathcal{F}$ satisfies $d^*(S) \leq 1/2$, where $d^*(S)$ denotes the upper Banach density of S. This is best possible since for any non-empty subset $\Gamma \subseteq \mathcal{A}$, the *odd-occurrence set* $O_\Gamma \subseteq \mathcal{F}$, which consists of all words with an odd number of letters from Γ, is product-free with density $1/2$. Leader, Letzter, Narayanan, and Walters conjectured that these sets are the only examples of product-free sets with density $1/2$ in free semigroups.

Conjecture 7.2 (Leader, Letzter, Narayanan, and Walters)**.** *Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is product-free and $d^*(S) = 1/2$, then $S \subseteq O_\Gamma$ for some non-empty subset $\Gamma \subseteq \mathcal{A}$.*

In Chapter 7, we confirm this conjecture. This solves the structure problem for product-free sets in free semigroups. In fact, we prove our results for two natural extensions of the notion of being product-free. Define a subset S of a (semi)group to be *k-product-free* if $x_1 \cdot \ldots \cdot x_k \notin S$ for all $x_1, \ldots, x_k \in S$, and *strongly k-product-free* if it is ℓ-product-free for all $\ell = 2, \ldots, k$.

The density problem for strongly k-product-free sets in free semigroups was solved by Ortega, Rué, and Serra [ORS24]. They showed that any strongly k-product-free subset $S \subseteq \mathcal{F}$ satisfies $\bar{d}(S) \leq 1/k$, where $\bar{d}(S)$ denotes the upper asymptotic density of S, which is a slight weakening of upper Banach density that satisfies $\bar{d}(S) \leq d^*(S)$. We solve the structure problem for strongly k-product-free sets in free semigroups. The special case of this result with $k = 2$ establishes the conjecture of Leader, Letzter, Narayanan, and Walters.

Theorem 7.4. *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is strongly k-product-free and $d^*(S) = 1/k$, then the following holds. It is possible to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/k\mathbb{Z}$ such that S is a subset of the strongly k-product-free set*

$$T := \{w \in \mathcal{F} : \textit{the sum of the labels of letters in } w \textit{ is } 1 \bmod k\}.$$

In addition, we consider k-product-free sets in free semigroups. For the non-negative integers, Calkin and Erdős [CE96] conjectured that a k-sum-free subset has density at most $1/\rho(k)$, where

$$\rho(k) := \min\{\ell \in \mathbb{Z}^+ : \ell \nmid k-1\}.$$

Note that the integers which are $1 \bmod \rho(k)$ form a k-sum-free set, and so $1/\rho(k)$ is best possible. Łuczak and Schoen [ŁS97] confirmed this conjecture and also solved the structure problem for the non-negative integers.

We extend their results by solving the density problem (for all k) and the structure problem (for all but four values of k) for k-product-free sets in free semigroups. For the density problem, we prove the following.

Theorem 7.5. *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is k-product-free, then $d^*(S) \leq 1/\rho(k)$.*

For the structure problem, we show that the structure of the extremal k-product-free sets is exactly that of strongly $\rho(k)$-product-free sets.

Theorem 7.6. *Let $k \geq 2$ be an integer with $k \notin \{3, 5, 7, 13\}$ and $\rho := \rho(k)$. Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is k-product-free and $d^*(S) = 1/\rho$, then the following holds. It is possible to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/\rho\mathbb{Z}$ such that S is a subset of the k-product-free set*

$$T := \{w \in \mathcal{F} : \textit{the sum of the labels of letters in } w \textit{ is } 1 \bmod \rho\}.$$

Finally, we consider the setting of free groups. A conjecture of Leader, Letzter, Narayanan, and Walters [LLNW20] asserted that the density of product-free sets in free groups is at most $1/2$. This conjecture was confirmed by Ortega, Rué, and Serra [ORS24]. More generally, they solved the density problem for strongly k-product-free sets in this setting and formulated an analogue of the conjecture of Calkin and Erdős for k-product-free sets. We prove their conjecture, thereby solving the density problem for k-product-free sets in free groups.

Theorem 7.8. *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{G}$ be the free group on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{G}$ is k-product-free, then $d^*(S) \leq 1/\rho(k)$.*

These results are based on joint work with Freddie Illingworth and Alex Scott [IMS25].

Chapter 2

Linear arboricity

Two fundamental results on graph decompositions are Vizing's theorem [Viz64], which states that the edges of a graph with maximum degree Δ can be decomposed into at most $\Delta + 1$ matchings, and the Nash-Williams theorem [NW61, NW64], which determines the minimum number of forests needed to decompose the edges of a graph. Between these two extremes lies the notion of linear arboricity, introduced by Harary [Har70] in 1970. The *linear arboricity* $\mathrm{la}(G)$ of a graph G is the minimum number of linear forests that decompose the edges of G, where a *linear forest* is a forest in which every component is a path.

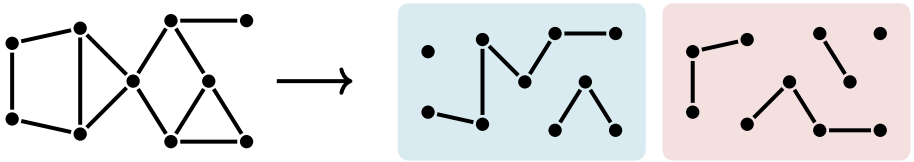


Figure 2.1: A decomposition of a graph into two linear forests.

Since every matching is a linear forest, Vizing's theorem [Viz64] implies that $\mathrm{la}(G) \leq \Delta + 1$. To obtain a lower bound, note that a linear forest has at most $n - 1$ edges, and so every Δ-regular graph G satisfies

$$\mathrm{la}(G) \geq \frac{\frac{n \cdot \Delta}{2}}{n-1} > \frac{\Delta}{2}.$$

Since $\mathrm{la}(G)$ is an integer, this implies that $\mathrm{la}(G) \geq \lceil (\Delta+1)/2 \rceil$. The celebrated Linear Arboricity Conjecture of Akiyama, Exoo, and Harary [AEH80] from 1980 asserts that this lower bound is tight.

Conjecture 2.1 (Linear Arboricity Conjecture)**.** *The linear arboricity of a graph with maximum degree Δ is at most $\lceil (\Delta+1)/2 \rceil$.*

Since every linear forest can be partitioned into two matchings, this conjecture would generalise Vizing's theorem up to rounding errors. Also note that every graph with maximum degree Δ is a subgraph of some Δ-regular graph, possibly on more vertices. Therefore, this conjecture is equivalent to the statement that every Δ-regular graph G satisfies $\mathrm{la}(G) = \lceil (\Delta+1)/2 \rceil$.

The Linear Arboricity Conjecture immediately received significant attention. By 1988, it was known to hold in the special cases $\Delta \in \{3,4,5,6,8,10\}$ [AEH80, AEH81, Eno81, Pér81, Tom82, EP84, Gul86a], and the best general upper bound was $\mathrm{la}(G) \leq \lceil (3\Delta+2)/5 \rceil$ [Gul86b]. In an influential result, Alon [Alo88] then confirmed the conjecture asymptotically, showing that

$$\mathrm{la}(G) \leq \frac{\Delta}{2} + \mathcal{O}\left(\frac{\Delta \log\log \Delta}{\log \Delta}\right) = (1+o(1)) \cdot \frac{\Delta}{2}.$$

Since then, some additional special cases of the conjecture have been solved, such as graphs with minimum degree at least $(1+o(1)) \cdot n/2$ [GKO16, GS24], planar graphs [Wu99, WW08], random or dense quasirandom graphs [GKO16, DGMS25], and random regular graphs of constant degree [MR90]. The error term of Alon's upper bound was subsequently improved to $\mathcal{O}(\Delta^{2/3} \log^{1/3} \Delta)$ by Alon and Spencer [AS16], to $\mathcal{O}(\Delta^{2/3-\varepsilon})$ by Ferber, Fox, and Jain [FFJ20], and to $\mathcal{O}(\sqrt{\Delta} \log^4 \Delta)$ by Lang and Postle [LP23]. A central ingredient in all of these upper bounds is a certain degree of randomisation, either through a random partition of the vertex set or through nibble-based strategies. For both approaches, $\Omega(\sqrt{\Delta})$ is a natural bottleneck for the error term, caused by concentration inequalities used throughout these proofs. This is a reflection of the fact that the standard deviation of a binomial random variable with Δ trials is $\Theta(\sqrt{\Delta})$.

In this chapter, we prove the following upper bound on the linear arboricity.

Theorem 2.2. *Every graph G on n vertices with maximum degree Δ satisfies*

$$\mathrm{la}(G) \leq \frac{\Delta}{2} + \mathcal{O}(\log n).$$

For $\Delta \geq \Omega(n^{\varepsilon})$, this result is an exponential improvement over the error term of Lang and Postle [LP23], and this breaks the square-root barrier for the Linear Arboricity Conjecture if $\Delta \geq \omega(\log^2 n)$.

For the fractional relaxation of linear arboricity, the previous best error term was also stuck at the same square-root barrier. Formally, the *fractional linear arboricity*

$\mathrm{fla}(G)$ of a graph G is the smallest real number k for which we can choose a random linear forest of G that contains every fixed edge of G with probability at least $1/k$. Clearly, $\Delta/2 \leq \mathrm{fla}(G) \leq \mathrm{la}(G)$. Feige, Ravi, and Singh [FRS14] showed that $\mathrm{fla}(G) \leq \Delta/2 + \mathcal{O}(\sqrt{\Delta})$. We improve the error term in this result to $\mathcal{O}(\log \Delta)$, providing an exponential improvement for all ranges of Δ.

Theorem 2.3. *Every graph G with maximum degree Δ satisfies*

$$\mathrm{fla}(G) \leq \frac{\Delta}{2} + \mathcal{O}(\log \Delta).$$

We prove these results using a novel variant of Pósa rotations for linear forests. Instead of just rotating one endpoint of a path, we simultaneously rotate multiple endpoints of a linear forest. We discuss these proof techniques in more detail in Section 2.1. In particular, this section outlines the key lemma that we use to prove these results, Lemma 2.9.

If the Linear Arboricity Conjecture holds, then every d-regular graph on n vertices has a spanning linear forest with at most $2 \cdot n/(d+2)$ paths.[1] Finding such a spanning linear forest with few paths is a difficult problem by itself. In 2009, Magnant and Martin [MM09] conjectured the following.

Conjecture 2.4 (Magnant and Martin)**.** *Every d-regular graph on n vertices has a spanning linear forest with at most $n/(d+1)$ paths.*

This conjecture would be tight for the disjoint union of $n/(d+1)$ complete graphs on $d+1$ vertices each. Later, Feige and Fuchs [FF22] conjectured that the weaker bound $\mathcal{O}(n/d)$ holds, which would follow from the Linear Arboricity Conjecture.

The conjecture of Magnant and Martin is only known to hold for $d \leq 6$ [MM09, FF22] and $d \geq \Omega(n)$ [Han18, GL21]. The upper bound on the fractional linear arboricity by Feige, Ravi, and Singh [FRS14] implies that there is a spanning linear forest with $\mathcal{O}(n/\sqrt{d})$ paths. In a different direction, Montgomery, Müyesser, Pokrovskiy, and Sudakov [MMPS25] showed that every d-regular graph has a linear forest with $n/(d+1)$ paths covering all but $o(n)$ vertices as $d \to \infty$. Under the stronger hypothesis that $d \geq \Omega((\log n)^{130})$, Letzter, Methuku, and Sudakov [LMS26] showed that the paths in this result can be replaced by cycles. In both papers, the conjecture of Feige and Fuchs was highlighted as being wide open.

These results employed a variety of different approaches. In addition to strategies

[1] See [FF22, Section 1.1] for a detailed discussion of these calculations.

also used for the Linear Arboricity Conjecture, such as random partitioning arguments [MMPS25], previous techniques included Szemerédi's regularity lemma [Han18] or robust expander decompositions [GL21, LMS26].

Using our new methods based on rotations of linear forests, we establish the conjecture of Magnant and Martin up to a factor of two and thereby prove the conjecture of Feige and Fuchs.

Theorem 2.5. *Every d-regular graph on n vertices has a spanning linear forest with at most $2 \cdot n/(d+1)$ paths.*

We also confirm the conjecture of Magnant and Martin for regular bipartite graphs (see Theorem 2.13), but note that the best lower bound for bipartite graphs is only $n/(2d)$. We prove these results as an immediate consequence of the key lemma that we use to make progress on the Linear Arboricity Conjecture.

Finally, our results imply that connected d-regular graphs have short tours, where a *tour* is a walk that ends at the starting vertex and visits every vertex at least once. The problem of finding such tours was introduced by Vishnoi [Vis12] as a restriction of the Travelling Salesman Problem to regular graphs. Feige, Ravi, and Singh [FRS14] showed that every connected d-regular graph on n vertices has a tour of length $(1 + \mathcal{O}(1/\sqrt{d})) \cdot n$ and observed that one cannot do better than $(1 + \mathcal{O}(1/d)) \cdot n$. As a direct consequence of Theorem 2.5 in combination with a result of Feige, Ravi, and Singh [FRS14, Corollary 1], we resolve this problem.

Corollary 2.6. *Every connected d-regular graph on n vertices has a tour of length at most $(1 + \mathcal{O}(1/d)) \cdot n$.*

A Hamilton cycle in a graph would form a tour of length n. Corollary 2.6 shows that regular graphs always contain tours that approximate Hamilton cycles with remarkable precision.

Chapter outline. We begin in Section 2.1 by discussing our proof techniques, which are based on rotations of linear forests and are crucial to proving all results in this chapter. In Section 2.2, we prove the key lemma that follows from these techniques, Lemma 2.9, and immediately deduce Theorem 2.5. We also use this lemma in Section 2.3 to prove our bounds on the (fractional) linear arboricity of a graph, Theorems 2.2 and 2.3. We finish in Section 2.4 with some open problems.

This chapter is based on joint work with Micha Christoph, Nemanja Draganić, António Girão, Eoin Hurley, and Alp Müyesser [CDG+25b].

2.1 PROOF TECHNIQUES

A rotation of a path or a linear forest is a simple operation that modifies the path or linear forest in such a way that exactly one of its endpoints moves to a different vertex of the graph (see also Figure 2.2). This idea can be traced back to Dirac's theorem [Dir52], which states that graphs on n vertices with minimum degree at least $n/2$ are Hamiltonian. It was also used by Pósa [Pós76] to determine the threshold for the Hamiltonicity of random graphs. Pósa's proof is notable for being the first to consider *the family of all possible sequences of rotations* that can be performed to a given path of maximal length. This allows one to show that many vertices can become an endpoint of such a path, which in turn is crucial for showing that sufficiently dense random graphs are Hamiltonian. Since this idea was highly influential, rotations are often referred to as *Pósa rotations* in the literature. Although the idea of rotations has been around for a long time, several breakthroughs in recent years were only made possible by introducing novel twists to this method [GMS24, DMS26, DMM+24].

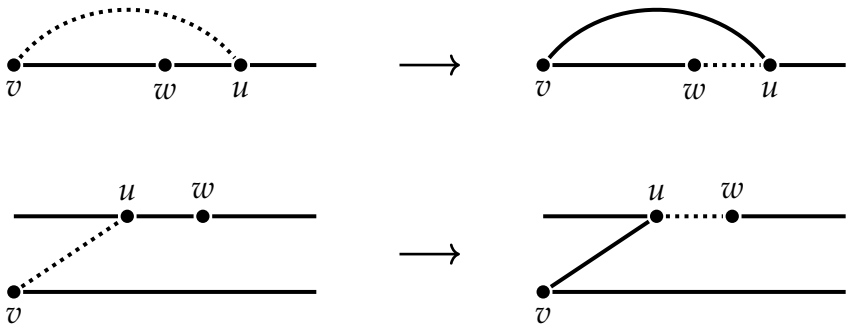


Figure 2.2: The two different types of rotations of a linear forest. Dotted edges represent edges of the graph that are not contained in the linear forest.

A common feature of these applications is some notion of expansion in the underlying graph. In contrast, we apply rotations to sparse regular graphs without any expansion conditions. Moreover, while Pósa rotations are only about rotations of *a single endpoint* of a path, our main novelty is to consider simultaneous rotations of *a set of endpoints* of a linear forest. Concretely, our key lemma, Lemma 2.9, proves the following.

> Suppose we are given a spanning linear forest of a d-regular graph with the minimum possible number of paths and a set of ℓ endpoints. If we allow any sequence of rotations that only involves this set of endpoints, then at least $\Omega(\ell \cdot d)$ vertices can become an endpoint.

As we make no assumptions on the structure of the underlying regular graph, this requires a highly delicate argument. We prove this lemma in Section 2.2, where we also concisely explain the key idea of its proof. For now, note that if there are at least $\Omega(\ell \cdot d)$ distinct potential endpoints, it immediately follows that $\ell \leq \mathcal{O}(n/d)$, which implies the conjecture of Feige and Fuchs and is how we prove Theorem 2.5. Moreover, in Section 2.3, we use this key lemma to create a distribution on the spanning linear forests of the graph in which every vertex has a probability of at most $\mathcal{O}(1/d)$ of being an endpoint. To prove our bounds on the (fractional) linear arboricity, Theorems 2.2 and 2.3, we then iteratively remove $\Delta/2$ such random linear forests from the graph and decompose the remaining edges into matchings using Vizing's theorem. Note that a vertex has a high degree in the remaining graph only if it was an endpoint in many of the removed linear forests. Since every vertex is unlikely to be an endpoint in a linear forest that we remove, we can apply standard concentration inequalities and a union bound to show that the maximum degree of the remaining graph is $\mathcal{O}(\log n)$. This gives the error term of $\mathcal{O}(\log n)$ in our result. We refer the reader to Section 2.4 for a discussion on the difficulties of improving this error term further.

Notation. If G is a graph and $X \subseteq V(G)$, then the *neighbourhood* of X in G is the set $N_G(X) := \{v \in V(G) : uv \in E(G) \text{ for some } u \in X\}$. Also, throughout the remainder of this chapter, every linear forest is a spanning linear forest. This can be achieved by adding isolated vertices to the linear forest as necessary.

2.2 SMALL LINEAR FORESTS IN REGULAR GRAPHS

In this section, we will show that any d-regular graph G on n vertices has a spanning linear forest with at most $2 \cdot n/(d+1)$ paths. Our main tool for achieving this is a rotation of a linear forest. A rotation modifies a linear forest by adding an edge incident to some endpoint of the linear forest and removing a different edge so that the resulting subgraph is still a linear forest. Intuitively, a rotation moves one endpoint of the linear forest to a different vertex of the graph. Figure 2.2 illustrates the two different types of possible rotations of a linear forest.

Definition 2.7. Let F be a linear forest of a graph G. Let v be an endpoint of F, and let u be a neighbour of v in G. If u and v are on the same path in F, let w be the neighbour of u that is closer to v on that path. Otherwise, let w be any neighbour of u on its path in F. Then, the linear forest F' obtained from F by removing the edge uw and adding the edge uv is called a *rotation* of F in G.

We say that such a rotation rotates the *old endpoint* v to the *new endpoint* w, using the *pivot* u. The restriction on w in the case where u and v are on the same path is necessary to ensure that the resulting subgraph is still a linear forest. Note that w is allowed to be the same vertex as v, which could happen if u is the unique neighbour of v on its path in F. Also, observe that if F' is a rotation of F, then F' has the same number of paths as F.

If we can use rotations to reach a linear forest where two endpoints of *distinct* paths are at adjacent vertices of the graph, then we can reduce the number of paths of the linear forest by adding the edge between these two endpoints. Suppose that F is a linear forest where we can no longer do this, for example because F has the minimum number of paths possible. To show that F has few paths, we will consider the set C of all vertices of the graph that could become an endpoint of F through some sequence of rotations. We claim that if F has ℓ endpoints, then the size of C is at least $\ell \cdot (d+1)/4$. Since the size of C is at most the total number of vertices of the graph, this claim implies that F has at most $4 \cdot n/(d+1)$ endpoints and therefore at most $2 \cdot n/(d+1)$ paths, as required.

We now sketch our proof of this claim. First, suppose for simplicity that G is bipartite. In this case, by only rotating endpoints in one part of the bipartition of G, we may assume that C is an independent set. Then, we will analyse the edges between C and its *boundary* $B := N_G(C) \setminus C$.

In particular, we exploit the tension between the following two facts. Firstly, each vertex in C has d neighbours in B while each vertex in B has at most d neighbours in C. Secondly, each vertex in C has at most two path-neighbours[2] in B while each vertex in B has at least one path-neighbour in C.[3] Using a double counting argument on these edges and combining this with the fact that C contains at least one endpoint, we can find a vertex $u \in B$ with $(d+1)/2$ neighbours in C, but which has only one path-neighbour in C.

Now, rotate one endpoint of F to an arbitrary neighbour v of u and *fix it there*. If we never allow v to be rotated again, we claim that no other endpoint can be rotated to a neighbour v' of u. Indeed, if that were possible, a short case analysis shows that either one of the two endpoints v or v' could be rotated to a vertex outside of C, contradicting the definition of C, or one of the two edges uv or uv' could be added to F, contradicting the minimality of F. This case analysis uses the fact that u has only one path-neighbour in C.

[2] A *path-neighbour* of a vertex v is a neighbour of v on its path in F.
[3] The existence of a path-neighbour in C follows from a standard argument for Pósa rotations.

Therefore, fixing an endpoint of F in the neighbourhood of u reduces the number of endpoints that we rotate by one while it reduces the number of vertices that could become an endpoint through a sequence of rotations by at least $(d+1)/2$. Iterating this argument shows that C has size at least $\ell \cdot (d+1)/2$, proving the result for bipartite graphs.

For non-bipartite graphs, we divide the argument into two cases based on the number of edges spanned by C. If C spans few edges, we can proceed exactly as for bipartite graphs. If instead C spans many edges, we iteratively fix endpoints at vertices that have a high degree within C and show that this reduces the size of C sufficiently. In both cases, we then iterate the argument to show that C has size at least $\ell \cdot (d+1)/4$, as required.

We now formalise this strategy. We begin by defining the set of all vertices of the graph that could become an endpoint of a linear forest through some sequence of rotations. In the following, we will use $\mathrm{End}(F)$ to denote the multiset of endpoints of a linear forest F, where an isolated vertex of F is contained twice in $\mathrm{End}(F)$. In particular, if F' is a rotation of F that rotates an old endpoint v to a new endpoint w, then $\mathrm{End}(F') = (\mathrm{End}(F) \setminus \{v\}) \cup \{w\}$.

Definition 2.8. Let F be a linear forest of a graph G, and let $X \subseteq \mathrm{End}(F)$ be a multiset of *fixed endpoints*. Then, $\mathcal{F}(F, X)$ denotes the family of linear forests obtained by a sequence of rotations starting from F that never rotates any endpoint of X. Define the set

$$C(F, X) := \bigcup_{F' \in \mathcal{F}(F,X)} (\mathrm{End}(F') \setminus X).$$

We call $C(F, X)$ a *rotation-component*. Intuitively, $C(F, X)$ consists of all vertices that could become an endpoint of F through a sequence of rotations that leaves all endpoints of X fixed. Note that if an endpoint v is contained once in X but twice in $\mathrm{End}(F)$, then we allow one of the endpoints at v to be rotated, but not both. In particular, $C(F, X)$ might not be disjoint from X. Since rotations are reversible, we have $\mathcal{F}(F, X) = \mathcal{F}(F', X)$ for all $F' \in \mathcal{F}(F, X)$, and therefore also $C(F, X) = C(F', X)$. Moreover, if $X \subseteq Y \subseteq \mathrm{End}(F)$, then $\mathcal{F}(F, X) \supseteq \mathcal{F}(F, Y)$, and so we also have $C(F, X) \supseteq C(F, Y)$.

A *minimum linear forest* F of G is a linear forest with the minimum number of paths among all linear forests of G. We now state the key lemma of this section, which asserts that for any minimum linear forest F, the size of the rotation-component $C(F, X)$ is at least $(d+1)/4$ times the number of endpoints that are not fixed.

Lemma 2.9. *Let F be a minimum linear forest of a d-regular graph, and let $X \subseteq \mathrm{End}(F)$. Then,*

$$|C(F,X)| \geq \frac{d+1}{4} \cdot |\mathrm{End}(F) \setminus X|.$$

This result suffices to prove that every d-regular graph on n vertices has a linear forest with at most $2 \cdot n/(d+1)$ paths.

Proof of Theorem 2.5. Let G be a d-regular graph on n vertices, and let F be a minimum linear forest of G. By Lemma 2.9, we have

$$|\mathrm{End}(F)| \leq \frac{4}{d+1} \cdot |C(F,\varnothing)| \leq \frac{4}{d+1} \cdot n.$$

Since the number of endpoints of a linear forest is twice the number of its paths, this implies that F has at most $2 \cdot n/(d+1)$ paths. □

To prove Lemma 2.9, we first make the following observations about rotations that are used extensively throughout our proofs. These observations follow directly from the definition of rotations together with a small amount of case analysis as indicated in Figure 2.3.

(R1) Suppose that v and u are endpoints of F and neighbours in G.
  (a) If w is a neighbour of u on its path in F, then we can rotate v to w using the pivot u.
  (b) Alternatively, if u and v are on different paths in F, we can reduce the number of paths of F by adding the edge uv to F. In particular, this applies if u is an isolated vertex of F.

(R2) Suppose that v is an endpoint of F which is a neighbour of u in G. If w and w' are neighbours of u on its path in F, then we can rotate v to one of the two vertices w or w' using the pivot u.

(R3) Suppose that v and v' are endpoints of F which are both neighbours of u in G. Note that we allow $v = v'$ if v is an isolated vertex of F.
  (a) If w is a neighbour of u on its path in F, then we can rotate one of the two endpoints v or v' to w using the pivot u.
  (b) Alternatively, if u is an endpoint of F, we can reduce the number of paths of F by adding one of the two edges uv or uv' to F.

Now, consider the *boundary* $B(F,X) := N_G(C(F,X)) \setminus C(F,X)$ of $C(F,X)$. Since $C(F,X) = C(F',X)$ for all $F' \in \mathcal{F}(F,X)$, we also have $B(F,X) = B(F',X)$ for all $F' \in \mathcal{F}(F,X)$. It is useful to classify each vertex of $C(F,X)$ and $B(F,X)$ based

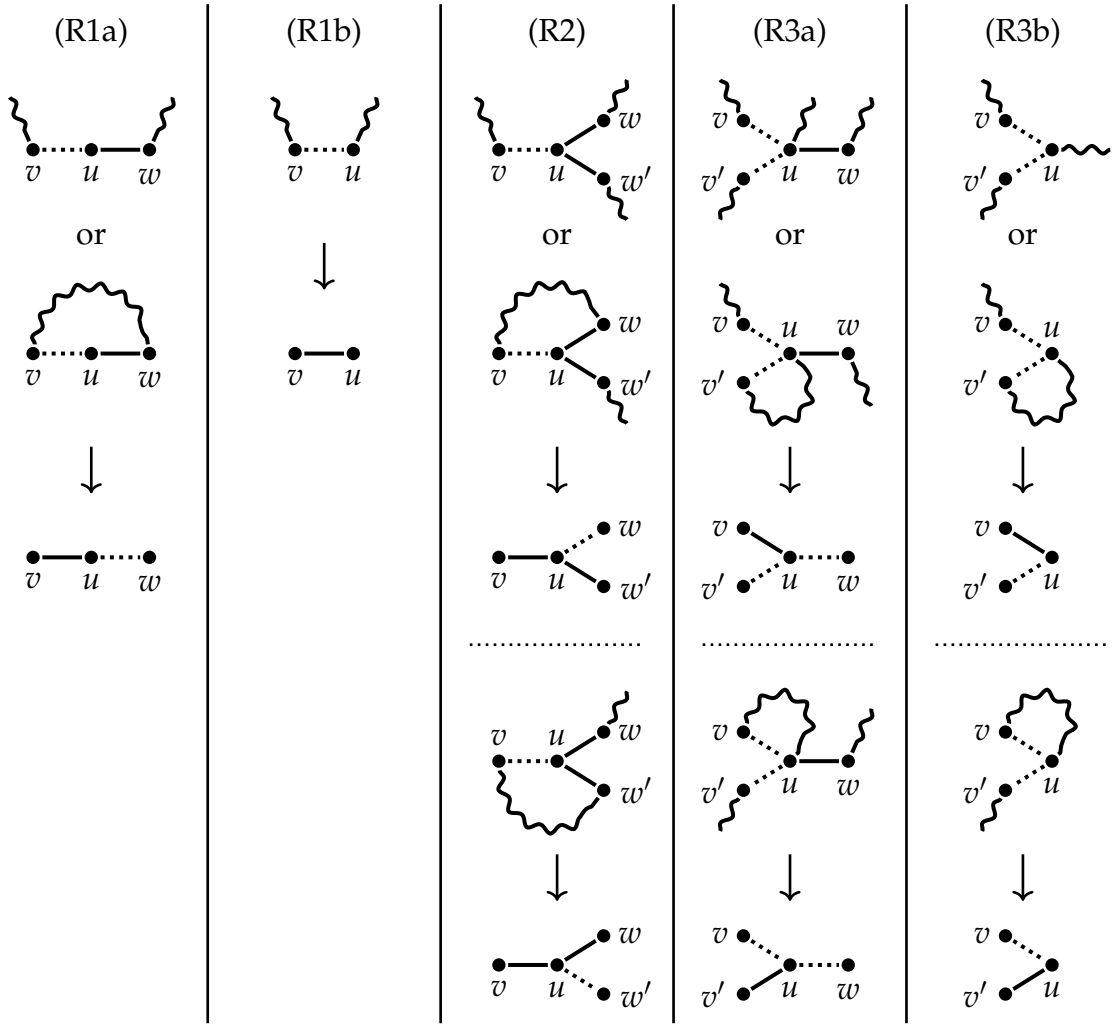


Figure 2.3: The case analysis for the observations about rotations of linear forests.

on how many of its neighbours on its path in F are contained in the opposite set. Therefore, define $C_i(F,X) := \{v \in C(F,X) : |N_F(v) \cap B(F,X)| = i\}$ and $B_i(F,X) := \{v \in B(F,X) : |N_F(v) \cap C(F,X)| = i\}$. Since every vertex is incident to at most two edges of F, we may assume that $i \in \{0,1,2\}$.

We start by proving some basic properties about these sets. The second of these properties, namely that $B_0(F,X) = \varnothing$, is a well-known fact for Pósa rotations.

Lemma 2.10. *Let F be a linear forest and let $X \subseteq \mathrm{End}(F)$. Then, for all $F' \in \mathcal{F}(F,X)$ and $i \in \{0,1,2\}$ we have $B_i(F,X) = B_i(F',X)$. Moreover, if F is a minimum linear forest, then $B_0(F,X) = \varnothing$.*

Proof. Note that F' is obtained by a sequence of rotations starting at F that never rotates any endpoint of X. It suffices to consider a single rotation of this sequence, so we may assume that F' is a rotation of F that rotates an old endpoint v to a new

endpoint w, using a pivot u. Note that $v, w \in C(F, X) = C(F', X)$. In particular, $|N_F(u) \cap C(F, X)| = |N_{F'}(u) \cap C(F', X)|$, and for each vertex in $B(F, X)$ apart from u, the edges of F and F' incident to that vertex are exactly the same. This shows that $B_i(F, X) = B_i(F', X)$.

Now, let F be a minimum linear forest. We want to show that $B_0(F, X) = \varnothing$, so suppose for a contradiction that instead there exists a vertex $u \in B_0(F, X)$. Let $v \in N_G(u) \cap C(F, X)$ be any neighbour of u in $C(F, X)$. Since $v \in C(F, X)$, there exists a linear forest $F' \in \mathcal{F}(F, X)$ with $X \cup \{v\} \subseteq \mathrm{End}(F')$. Because F' has the same number of paths as F, F' is a minimum linear forest, and from above we know that $u \in B_0(F, X) = B_0(F', X)$.

If u is an isolated vertex of F', then by (R1b) we could reduce the number of paths of F' by adding the edge uv, contradicting that F' is a minimum linear forest. If u is an endpoint of F' but has a neighbour $w \in N_{F'}(u)$ on its path in F', then by (R1a) we could rotate v to w using the pivot u. But this would imply that $w \in C(F', X)$, contradicting $u \in B_0(F', X)$. Finally, if u has two neighbours on its path in F', then by (R2) there exists at least one neighbour $w \in N_{F'}(u)$ such that we could rotate v to w using the pivot u, again contradicting $w \notin C(F', X)$. In all cases we get a contradiction, and so $B_0(F, X) = \varnothing$. □

We now show that we can significantly reduce the size of $C(F, X)$ by fixing an endpoint at a vertex in $C(F, X)$ that has a neighbour in $B_1(F, X)$ which itself has many neighbours in $C(F, X)$.

Lemma 2.11. *Let F be a minimum linear forest and let $X \subseteq \mathrm{End}(F)$. Suppose that $v \in \mathrm{End}(F) \setminus X$ and $u \in N_G(v) \cap B_1(F, X)$. Then,*

$$|C(F, X \cup \{v\})| \leq |C(F, X)| - |N_G(u) \cap C(F, X)|.$$

Proof. We claim that $N_G(u) \cap C(F, X \cup \{v\}) = \varnothing$. Suppose for a contradiction that there exists a vertex $v' \in N_G(u) \cap C(F, X \cup \{v\})$. Since $v' \in C(F, X \cup \{v\})$, there exists a linear forest $F' \in \mathcal{F}(F, X \cup \{v\})$ with $X \cup \{v, v'\} \subseteq \mathrm{End}(F')$. Note that F' is a minimum linear forest. Furthermore, by Lemma 2.10 we also know that $u \in B_1(F, X) = B_1(F', X)$.

If u is an endpoint of F', then by (R3b) we could reduce the number of paths of F' by adding one of the two edges uv or uv', contradicting that F' is a minimum linear forest. Otherwise, u has two neighbours on its path in F'. Since $u \in B_1(F', X)$, there must be a neighbour $w \in N_{F'}(u)$ that is not contained in $C(F', X)$. But

then by (R3a) we could rotate one of the two endpoints v or v' to w using the pivot u, contradicting $w \notin C(F', X)$. In both cases we get a contradiction, and so $N_G(u) \cap C(F, X \cup \{v\}) = \varnothing$, as claimed. Since $C(F, X \cup \{v\}) \subseteq C(F, X)$, this implies that $|C(F, X \cup \{v\})| \leq |C(F, X)| - |N_G(u) \cap C(F, X)|$. □

To use this result to prove a good lower bound on $|C(F, X)|$, we need to find a vertex in $B_1(F, X)$ with many neighbours in $C(F, X)$. This is easy if there are many edges between $C(F, X)$ and $B_1(F, X)$. We will use the following double counting argument to provide a lower bound on the number of such edges. Here, for $A, B \subseteq V(G)$, we denote the number of edges of G connecting a vertex in A to a vertex in B by $e(A, B)$, and we write $e(A) := e(A, A)$.

Lemma 2.12. *Let F be a minimum linear forest of a d-regular graph G. Then, for any $X \subseteq \mathrm{End}(F)$ we have*

$$e(C(F, X), B_1(F, X)) \geq \frac{d}{2} \cdot |B_1(F, X)| + \frac{d}{2} \cdot |C_1(F, X)| + d \cdot |C_0(F, X)| - 2 \cdot e(C(F, X)).$$

Proof. By double counting the edges of F between $C(F, X)$ and $B(F, X)$, we have

$$|C_1(F, X)| + 2 \cdot |C_2(F, X)| = |B_1(F, X)| + 2 \cdot |B_2(F, X)|.$$

The left side of this equation is equal to $2 \cdot |C(F, X)| - |C_1(F, X)| - 2 \cdot |C_0(F, X)|$. Moreover, note that the total number of edges of G between $C(F, X)$ and $B(F, X)$ is $d \cdot |C(F, X)| - 2 \cdot e(C(F, X))$. Because at most $d \cdot |B_2(F, X)|$ of these edges are incident with $B_2(F, X)$ and we know that $B_0(F, X) = \varnothing$ by Lemma 2.10, we get

$$\begin{aligned} e(C(F, X), B_1(F, X)) &\geq d \cdot |C(F, X)| - 2 \cdot e(C(F, X)) - d \cdot |B_2(F, X)| \\ &= \frac{d}{2} \cdot |B_1(F, X)| + \frac{d}{2} \cdot |C_1(F, X)| \\ &\quad + d \cdot |C_0(F, X)| - 2 \cdot e(C(F, X)). \end{aligned}$$

□

We now first prove that bipartite graphs have small linear forests. In that setting, observe that if we fix all endpoints that are contained in one part of the bipartition, then $C(F, X)$ will be contained in the other part of the bipartition. This implies that $e(C(F, X)) = 0$. As a result, Lemma 2.12 shows that there exists a vertex in $B_1(F, X)$ with at least $(d + 1)/2$ neighbours in $C(F, X)$. This allows us to apply Lemma 2.11 and use induction to complete the proof.

Theorem 2.13. *Every d-regular bipartite graph on n vertices has a spanning linear forest with at most $n/(d+1)$ paths.*

Proof. Let G be a d-regular bipartite graph on n vertices, and let U be one part of its bipartition. We claim that if F is a minimum linear forest of G and $X \subseteq \mathrm{End}(F)$ is such that $\mathrm{End}(F) \setminus X \subseteq U$, then $|C(F,X)| \geq |\mathrm{End}(F) \setminus X| \cdot (d+1)/2$ and $C(F,X) \subseteq U$. We prove the claim by induction on $\ell := |\mathrm{End}(F) \setminus X|$. The claim clearly holds for $\ell = 0$, so suppose that $\ell > 0$.

Observe that if a rotation of F rotates an old endpoint $v \in U$ to a new endpoint w, then there is a walk of length two from w to v. Since G is bipartite, it follows that $w \in U$. This shows that $C(F,X) \subseteq U$. In particular, $e(C(F,X)) = 0$. Since we also know that $\varnothing \neq \mathrm{End}(F) \setminus X \subseteq C_1(F,X) \cup C_0(F,X)$, this implies by Lemma 2.12 that $e(C(F,X), B_1(F,X)) > |B_1(F,X)| \cdot d/2$. So, there exists a vertex $u \in B_1(F,X)$ with $|N_G(u) \cap C(F,X)| \geq (d+1)/2$.

Let $v \in N_G(u) \cap C(F,X)$ be arbitrary, let $X' := X \cup \{v\}$, and let $F' \in \mathcal{F}(F,X)$ be any linear forest with $X' \subseteq \mathrm{End}(F')$. Using the fact that $C(F,X) = C(F',X)$, Lemma 2.11 implies that

$$|C(F,X)| = |C(F',X)| \geq |C(F',X')| + |N_G(u) \cap C(F',X)| \geq |C(F',X')| + \frac{d+1}{2}.$$

By the induction hypothesis we have $|C(F',X')| \geq |\mathrm{End}(F') \setminus X'| \cdot (d+1)/2$. So,

$$|C(F,X)| \geq \frac{d+1}{2} \cdot |\mathrm{End}(F') \setminus X'| + \frac{d+1}{2} = \frac{d+1}{2} \cdot |\mathrm{End}(F) \setminus X|.$$

This proves the claim. Finally, if $V(G) = V \cup W$ is the bipartition of G, the claim implies that

$$\begin{aligned} |\mathrm{End}(F)| &= |\mathrm{End}(F) \setminus (\mathrm{End}(F) \cap V)| + |\mathrm{End}(F) \setminus (\mathrm{End}(F) \cap W)| \\ &\leq \frac{2}{d+1} \cdot |C(F, \mathrm{End}(F) \cap V)| + \frac{2}{d+1} \cdot |C(F, \mathrm{End}(F) \cap W)| \leq \frac{2}{d+1} \cdot n. \end{aligned}$$

Since the number of endpoints of a linear forest is twice the number of its paths, this implies that F has at most $n/(d+1)$ paths. □

In non-bipartite graphs, we will define an iterative process that fixes endpoints of F one by one at vertices that have a high degree within $C(F,X)$. If this reduces the size of $C(F,X)$ sufficiently, we are done by induction. Otherwise, we show that $C(F,X)$ spans only few edges, and so we finish the proof as for bipartite graphs.

Proof of Lemma 2.9. Let G be a d-regular graph, let F be a minimum linear forest of G, and let $X \subseteq \mathrm{End}(F)$. We proceed by induction on $\ell := |\mathrm{End}(F) \setminus X|$. The result clearly holds for $\ell = 0$, so suppose that $\ell > 0$.

Consider the following iterative process. Let $X_1 := X$ and let $F_1 := F$. Then, given X_i and F_i, let $\Delta_i := \Delta(C(F_i, X_i))$ be the maximum degree of the subgraph of G induced by $C(F_i, X_i)$, and choose any vertex $v_i \in C(F_i, X_i)$ with degree Δ_i in $C(F_i, X_i)$. Define $X_{i+1} := X_i \cup \{v_i\}$ and let $F_{i+1} \in \mathcal{F}(F_i, X_i)$ be a linear forest with $X_{i+1} \subseteq \mathrm{End}(F_{i+1})$. This process gives a chain of rotation-components

$$C(F, X) = C(F_1, X_1) \supseteq C(F_2, X_2) \supseteq C(F_3, X_3) \supseteq \cdots \supseteq C(F_{\ell+1}, X_{\ell+1}) = \varnothing.$$

Since the maximum degree of the subgraph induced by $S_i := C(F_i, X_i)$ is Δ_i, this implies that

$$e(C(F, X)) \leq \sum_{i=1}^{\ell} |S_i \setminus S_{i+1}| \cdot \Delta_i. \tag{2.1}$$

We know that $|S_{i+1}| = |C(F_{i+1}, X_{i+1})| \geq |\mathrm{End}(F_{i+1}) \setminus X_{i+1}| \cdot (d+1)/4$ by the induction hypothesis. Therefore, if $|S_{i+1}| \leq |S_1| - i \cdot (d+1)/4$, then

$$\begin{aligned} |C(F, X)| = |S_1| &\geq \frac{d+1}{4} \cdot i + |S_{i+1}| \\ &\geq \frac{d+1}{4} \cdot i + \frac{d+1}{4} \cdot |\mathrm{End}(F_{i+1}) \setminus X_{i+1}| = \frac{d+1}{4} \cdot |\mathrm{End}(F) \setminus X|. \end{aligned}$$

So, we may assume that $|S_{i+1}| \geq |S_1| - i \cdot (d+1)/4$. Since $\Delta_1 \geq \cdots \geq \Delta_\ell$, this implies that

$$\begin{aligned} \sum_{i=1}^{\ell} |S_i \setminus S_{i+1}| \cdot \Delta_i &= |S_1| \cdot \Delta_1 + \sum_{i=1}^{\ell-1} |S_{i+1}| \cdot (\Delta_{i+1} - \Delta_i) - |S_{\ell+1}| \cdot \Delta_\ell \\ &\leq |S_1| \cdot \Delta_1 + \sum_{i=1}^{\ell-1} \left(|S_1| - i \cdot \frac{d+1}{4} \right) \cdot (\Delta_{i+1} - \Delta_i) \\ &\qquad - \left(|S_1| - \ell \cdot \frac{d+1}{4} \right) \cdot \Delta_\ell \\ &= \sum_{i=1}^{\ell} \frac{d+1}{4} \cdot \Delta_i = \frac{d+1}{4} \sum_{i=1}^{\ell} |N_G(v_i) \cap S_i|. \end{aligned} \tag{2.2}$$

We claim that

$$|\mathrm{End}(F_{\ell+1}) \setminus X| + \sum_{i=1}^{\ell} |N_G(v_i) \cap S_i| \leq |C_1(F_{\ell+1}, X)| + 2 \cdot |C_0(F_{\ell+1}, X)|. \tag{2.3}$$

To see this, observe that every set that appears on the left side of the inequality is a subset of $C(F, X) = C(F_{\ell+1}, X)$. We will show that the contribution of every vertex $u \in C(F_{\ell+1}, X)$ to the left side of the inequality is at most as large as its contribution to the right side. Indeed, consider the following cases:

(a) Suppose that $u \in C_2(F_{\ell+1}, X)$. Clearly $u \notin \mathrm{End}(F_{\ell+1}) \setminus X$ since u has two neighbours on its path in $F_{\ell+1}$. Moreover, we claim that $u \notin N_G(v_i) \cap S_i$ for all i. Indeed, if $u \in N_G(v_i)$ for some i, then by (R2) there exists at least one neighbour $w \in N_{F_{\ell+1}}(u)$ such that we could rotate v_i to w using the pivot u. This would imply $w \in C(F_{\ell+1}, X)$, contradicting $u \in C_2(F_{\ell+1}, X)$.

(b) Suppose that $u \in C_1(F_{\ell+1}, X)$. Then, we claim that either $u \in \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \notin N_G(v_i) \cap S_i$ for all i, or $u \notin \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i) \cap S_i$ for up to one value of i. Indeed, let $w \in N_{F_{\ell+1}}(u)$ be the unique neighbour of u in $B(F_{\ell+1}, X)$. If $u \in \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i)$ for some i, then by (R1a) we could rotate v_i to w using the pivot u, contradicting $w \notin C(F_{\ell+1}, X)$. Moreover, if $u \in N_G(v_i) \cap N_G(v_j)$ for some $i \neq j$, then by (R3a) we could rotate one of the two endpoints v_i or v_j to w using the pivot u, again contradicting $w \notin C(F_{\ell+1}, X)$. This proves the claim.

(c) Suppose that $u \in C_0(F_{\ell+1}, X)$. Then, we claim that either $u \in \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i) \cap S_i$ for up to one value of i, or we have $u \notin \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i) \cap S_i$ for up to two values of i. Indeed, if $u \in \mathrm{End}(F_{\ell+1}) \setminus X$ and $u \in N_G(v_i) \cap N_G(v_j)$ for some $i \neq j$, then by (R3b) we could reduce the number of paths of $F_{\ell+1}$ by adding one of the two edges uv_i or uv_j, contradicting the fact that $F_{\ell+1}$ is a minimum linear forest. Moreover, if $u \in N_G(v_i) \cap N_G(v_j)$ for some $i < j$, then $u \notin S_k = C(F_k, X_k)$ for any $k > j$. Otherwise, there would exist a linear forest $F' \in \mathcal{F}(F_k, X_k)$ with $\{u, v_i, v_j\} \subseteq \mathrm{End}(F')$, but then by (R3b) we could reduce the number of paths of F' by adding one of the two edges uv_i or uv_j, contradicting the fact that F' is a minimum linear forest. This proves the claim.

In all cases, the contribution of a vertex to the left side of inequality (2.3) is at most as large as its contribution to the right side, which proves this inequality.

Assume that $|\mathrm{End}(F_{\ell+1}) \setminus X| \cdot (d+1)/2 = |\mathrm{End}(F) \setminus X| \cdot (d+1)/2 > |C(F, X)|$, as we would be done otherwise. Since $C(F, X) = C(F_{\ell+1}, X)$, we therefore have

$$\frac{d+1}{2} \cdot |\mathrm{End}(F_{\ell+1}) \setminus X| > |C(F_{\ell+1}, X)| \geq \frac{1}{2} \cdot |C_1(F_{\ell+1}, X)| + |C_0(F_{\ell+1}, X)|. \quad (2.4)$$

By combining the inequalities (2.1) to (2.4) and using again the fact that $C(F, X) =$

$C(F_{\ell+1}, X)$, we get that

$$\begin{aligned}
2 \cdot e(C(F_{\ell+1}, X)) &\leq 2 \sum_{i=1}^{\ell} |S_i \setminus S_{i+1}| \cdot \Delta_i \\
&\leq \frac{d+1}{2} \sum_{i=1}^{\ell} |N_G(v_i) \cap S_i| \\
&\leq \frac{d+1}{2} \cdot |C_1(F_{\ell+1}, X)| + (d+1) \cdot |C_0(F_{\ell+1}, X)| \\
&\qquad - \frac{d+1}{2} \cdot |\mathrm{End}(F_{\ell+1}) \setminus X| \\
&< \frac{d}{2} \cdot |C_1(F_{\ell+1}, X)| + d \cdot |C_0(F_{\ell+1}, X)|.
\end{aligned}$$

By Lemma 2.12, it follows that $e(C(F_{\ell+1}, X), B_1(F_{\ell+1}, X)) > |B_1(F_{\ell+1}, X)| \cdot d/2$. Since we know that $F_{\ell+1} \in \mathcal{F}(F, X)$, by Lemma 2.10 this is equivalent to $e(C(F, X), B_1(F, X)) > |B_1(F, X)| \cdot d/2$. As in the proof of Theorem 2.13, we can then find a vertex $v \in C(F, X)$ such that if $X' := X \cup \{v\}$ and $F' \in \mathcal{F}(F, X)$ is a linear forest with $X' \subseteq \mathrm{End}(F')$, then $|C(F, X)| \geq |C(F', X')| + (d+1)/2$. By the induction hypothesis, it then follows that

$$|C(F, X)| \geq \frac{d+1}{4} \cdot |\mathrm{End}(F') \setminus X'| + \frac{d+1}{2} \geq \frac{d+1}{4} \cdot |\mathrm{End}(F) \setminus X|. \qquad \square$$

2.3 LOGARITHMIC ERROR FOR LINEAR ARBORICITY

We now use the results from the previous section to show that the linear arboricity of any graph G on n vertices with maximum degree Δ is at most $\Delta/2 + \mathcal{O}(\log n)$. We achieve this by repeatedly choosing linear forests with few paths and removing them from the graph. If the maximum degree of the remaining graph after $\Delta/2$ iterations is $\mathcal{O}(\log n)$, we can decompose the remaining edges into $\mathcal{O}(\log n)$ linear forests, say with Vizing's theorem, to obtain the desired upper bound on the linear arboricity.

Note that the degree of a vertex after the first $\Delta/2$ iterations can only be larger than $\mathcal{O}(\log n)$ if the vertex appeared as an endpoint in more than $\mathcal{O}(\log n)$ linear forests during those iterations. To avoid this, we will choose the linear forests in each iteration randomly in such a way that every vertex has a low probability of being an endpoint. Then, we can simply use standard concentration inequalities and a union bound to bound the probability that any vertex is an endpoint in more than $\mathcal{O}(\log n)$ linear forests, which completes the proof.

We start by showing that for every d-regular graph G, there exists a distribution on the minimum linear forests of G such that every vertex has a low probability of being an endpoint. Recall that $\mathrm{End}(F)$ denotes the multiset of endpoints of a linear forest F. In the proof, we will make extensive use of the notation introduced in Section 2.2.

Lemma 2.14. *Let G be a d-regular graph. Then, we can choose a random minimum linear forest F of G in such a way that for all vertices $v \in V(G)$ we have*

$$\mathbb{P}(v \in \mathrm{End}(F)) \leq \frac{16}{d+1}.$$

Proof. We claim that for all minimum linear forests F of G and all $X \subseteq \mathrm{End}(F)$, we can choose a random linear forest $F' \in \mathcal{F}(F, X)$ in such a way that for all vertices $v \in V(G)$ we have $\mathbb{P}(v \in \mathrm{End}(F') \setminus X) \leq 16/(d+1)$. For $X = \varnothing$, this implies the result. We prove the claim by induction on $\ell := |\mathrm{End}(F) \setminus X|$. The claim clearly holds for $\ell = 0$, so suppose that $\ell > 0$.

For all $u \in C(F, X)$, let $F_u \in \mathcal{F}(F, X)$ be a linear forest with $X \cup \{u\} \subseteq \mathrm{End}(F_u)$. By the induction hypothesis, we know that we can choose a random linear forest $F'_u \in \mathcal{F}(F_u, X \cup \{u\})$ in such a way that for all vertices $v \in V(G)$ we have $\mathbb{P}(v \in \mathrm{End}(F'_u) \setminus (X \cup \{u\})) \leq 16/(d+1)$. Now, choose a random linear forest $F'' \in \mathcal{F}(F, X)$ by first sampling $u \in C(F, X)$ uniformly at random and then sampling the random linear forest F'_u. By the linearity of expectation, we get that

$$\sum_{v \in C(F,X)} \mathbb{P}(v \in \mathrm{End}(F'') \setminus X) \leq \mathbb{E}(|\mathrm{End}(F'') \setminus X|) = |\mathrm{End}(F) \setminus X|.$$

By Lemma 2.9, we also have $|C(F, X)| \geq |\mathrm{End}(F) \setminus X| \cdot (d+1)/4$. This implies that the set $U := \{v \in C(F, X) : \mathbb{P}(v \in \mathrm{End}(F'') \setminus X) \leq 8/(d+1)\}$ must contain at least half of the vertices of $C(F, X)$.

Finally, choose the random linear forest $F' \in \mathcal{F}(F, X)$ by first sampling $u \in U$ uniformly at random and then sampling the random linear forest F'_u. For all vertices $v \in V(G) \setminus U$, we have

$$\begin{aligned}\mathbb{P}(v \in \mathrm{End}(F') \setminus X) &= \frac{1}{|U|} \sum_{u \in U} \mathbb{P}(v \in \mathrm{End}(F'_u) \setminus X) \\ &= \frac{1}{|U|} \sum_{u \in U} \mathbb{P}(v \in \mathrm{End}(F'_u) \setminus (X \cup \{u\})) \leq \frac{16}{d+1}.\end{aligned}$$

For all vertices $v \in U$, we have

$$\begin{aligned}\mathbb{P}(v \in \mathrm{End}(F') \setminus X) &= \frac{1}{|U|} \sum_{u \in U} \mathbb{P}(v \in \mathrm{End}(F'_u) \setminus X) \\ &\leq \frac{1}{|U|} \sum_{u \in C(F,X)} \mathbb{P}(v \in \mathrm{End}(F'_u) \setminus X) \\ &= \frac{|C(F,X)|}{|U|} \cdot \mathbb{P}(v \in \mathrm{End}(F'') \setminus X) \leq \frac{16}{d+1}. \qquad \square\end{aligned}$$

We can now prove Theorem 2.2. As described earlier in this section, our strategy is to repeatedly remove linear forests from the graph G using Lemma 2.14. Since that result only applies to regular graphs, we need to regularise G after every step, but the remainder of the proof works exactly as outlined before.

Proof of Theorem 2.2. Let G be a graph on n vertices with maximum degree Δ. Let $d := \lceil \Delta/2 \rceil$. By adding vertices and edges to G, we may extend G to a $2d$-regular graph G_d.

Consider the following iterative process. Start with the $2d$-regular graph G_d. Then, given a $2i$-regular graph G_i, let $C_i \subseteq G_i$ be a 2-regular spanning subgraph of G_i. This subgraph exists by Petersen's 2-factor theorem [Pet91] as G_i is regular with even degree. Choose a random minimum linear forest F_i of G_i according to Lemma 2.14 so that for all $v \in V(G_i)$ we have $\mathbb{P}(v \in \mathrm{End}(F_i)) \leq 16/(2i+1)$. Then, let $E_i \subseteq E(C_i)$ be the subset of all edges of C_i that are incident to some endpoint of F_i. Note that every vertex of G_i is incident to at least two edges of $F_i \cup E_i$. This implies that the subgraph $G_i \setminus (F_i \cup E_i)$ has maximum degree at most $2(i-1)$ and can therefore be extended to a $2(i-1)$-regular graph G_{i-1}.

We iterate this process until we reach the empty graph G_0. Note that the restriction of each linear forest F_i to the initial graph G is a linear forest. We also have that every edge of G is contained either in one of the linear forests F_i or in one of the sets E_i. So, the degree of a vertex $v \in V(G)$ in $G' := G \setminus (\bigcup_{i=1}^{d} F_i)$ is at most the number of edges of $\bigcup_{i=1}^{d} E_i$ that are incident to v.

Let D_v^i be the indicator random variable of the event that v is incident to an edge of E_i. Since v is incident to at most two edges of E_i, the degree of v in G' is at most $D_v := \sum_{i=1}^{d} 2D_v^i$. Moreover, v can only be incident to an edge of E_i if v or one of its two neighbours in C_i is an endpoint of F_i. By construction, the probability that a vertex is an endpoint of F_i is at most $16/(2i+1) \leq 8/i$, regardless of the choices of $F_{i+1}, \ldots, F_d$. It follows that $\mathbb{P}(D_v^i = 1) \leq 24/i$, even conditioned on

the values of $D_v^{i+1}, \dots, D_v^d$. So, the random variables $D_v^1, \dots, D_v^d$ can be coupled with a collection of independent Bernoulli random variables $X_1, \dots, X_d$ so that $D_v^i \leq X_i$ and $\mathbb{P}(X_i = 1) \leq 24/i$. In particular, $X := \sum_{i=1}^d X_i$ satisfies $D_v \leq 2X$. Since $\mu := \mathbb{E}(X) \leq \sum_{i=1}^d 24/i \leq 24(\log d + 1)$, a standard Chernoff bound with $\delta := 24(\log n + 1)/\mu \geq 1$ implies that

$$\mathbb{P}(D_v > 2(1+\delta)\mu) \leq \mathbb{P}(X > (1+\delta)\mu) \leq e^{-\frac{\delta^2\mu}{2+\delta}} \leq e^{-\frac{\delta\mu}{3}} = e^{-8(\log n+1)} < \frac{1}{n}.$$

In particular, a union bound shows that with positive probability, we have $D_v \leq 2(1+\delta)\mu = \mathcal{O}(\log n)$ for all $v \in V(G)$. In that case, the maximum degree of G' is at most $\mathcal{O}(\log n)$, which implies that G' can be decomposed into at most $\mathcal{O}(\log n)$ linear forests. Together with $F_1, \dots, F_d$, this provides a decomposition of G into at most $d + \mathcal{O}(\log n) = \Delta/2 + \mathcal{O}(\log n)$ linear forests. □

The proof for fractional linear arboricity follows almost exactly the same strategy, except that it ignores vertices with degree larger than $\mathcal{O}(\log \Delta)$ in the remaining graph. This produces a family of $\Delta/2 + \mathcal{O}(\log \Delta)$ linear forests such that each edge is contained in that family with high probability. If we then choose a uniformly random linear forest from that family, this proves the upper bound on the fractional linear arboricity.

Proof of Theorem 2.3. Let G be a graph on n vertices with maximum degree Δ. Define d, F_i, G', D_v, X, and μ as in the proof of Theorem 2.2. Then, a standard application of a Chernoff bound with $\delta := 24(\log d + 1)/\mu \geq 1$ implies that

$$\mathbb{P}(D_v > 2(1+\delta)\mu) \leq \mathbb{P}(X > (1+\delta)\mu) \leq e^{-\frac{\delta^2\mu}{2+\delta}} \leq e^{-\frac{\delta\mu}{3}} = e^{-8(\log d+1)} < \frac{1}{d^2}.$$

So, the probability that v has degree larger than $c := 2(1+\delta)\mu = \mathcal{O}(\log d)$ in G' is less than $1/d^2$. Let $E \subseteq E(G')$ be the subset of all edges of G' that are incident to a vertex of degree larger than c in G'. In particular, every edge of G is contained in E with probability at most $2/d^2$. Moreover, the graph $G'' := G' \setminus E$ has maximum degree c, which implies that G'' can be decomposed into at most $c+1$ linear forests $F_1', \dots, F_{c+1}'$. Choose a linear forest $F \in \{F_1, \dots, F_d, F_1', \dots, F_{c+1}'\}$ uniformly at random. Then, for every edge $e \in E(G)$, we have

$$\mathbb{P}(e \in F) = \mathbb{P}(e \notin E) \cdot \mathbb{P}(e \in F \mid e \notin E) \geq \left(1 - \frac{2}{d^2}\right) \frac{1}{d+c+1} \geq \frac{1}{d + \mathcal{O}(\log d)}.$$

This shows that $\mathrm{fla}(G) \leq d + \mathcal{O}(\log d) = \Delta/2 + \mathcal{O}(\log \Delta)$. □

2.4 OPEN PROBLEMS

By proving the conjecture of Feige and Fuchs, we removed a major obstacle towards resolving the Linear Arboricity Conjecture. However, the bound on the linear arboricity that we obtained in Theorem 2.2 has an error term of $\mathcal{O}(\log n)$, which depends on n. A natural next step is to replace this error term with a term polynomial in $\log \Delta$. To achieve this, one would like to replace the union bound that we use with a variant of the Lovász Local Lemma or a similar technique. Unfortunately, the dependency structure of the corresponding bad events is extremely intricate, and so it seems difficult to make this strategy work. We believe that such an improvement would already be interesting for bipartite graphs.

Problem 2.15. *Does every bipartite graph G with maximum degree Δ satisfy*

$$\mathrm{la}(G) \leq \Delta/2 + (\log \Delta)^{\mathcal{O}(1)}?$$

In bipartite graphs, we can sample a random linear forest more easily than for non-bipartite graphs. Indeed, as shown in the proof of Theorem 2.13, we can always find a vertex $u \in B_1(F, X)$ with at least $(d+1)/2$ neighbours in $C(F, X)$. If we then fix an endpoint of the linear forest at a uniformly random neighbour of u in $C(F, X)$, all neighbours of u will afterwards be removed from $C(F, X)$. Repeating this process results in a random linear forest where every vertex is an endpoint with probability at most $2/(d+1)$.

When we iteratively remove linear forests from a graph, in later iterations we would like to fix endpoints only at vertices that have not been endpoints in too many previous iterations. This would be possible if we never fix too many endpoints in the neighbourhood of a vertex since we could then simply avoid these vertices when generating a new linear forest. Note that the above random linear forest has a very low probability of fixing many endpoints in the neighbourhood of a given vertex. However, we cannot apply the Lovász Local Lemma because the vertex u as well as the subset of its neighbours that are contained in $C(F, X)$ can depend on previously fixed endpoints, creating large dependencies. Even just the fact that we only have a subset of the neighbours of u available when we fix an endpoint seems to make this problem difficult. More precisely, even if we knew in advance the sequence of vertices u obtained while fixing endpoints, but not the subset of available neighbours, we do not know how to ensure that we fix at most $(\log \Delta)^{\mathcal{O}(1)}$ endpoints in the neighbourhood of any vertex.

We believe that the following online matching game exhibits the same difficulty and could be of independent interest. Suppose that G is a bipartite graph with bipartition $A \cup B$ such that every vertex in A has degree $c \cdot d$ while every vertex in B has degree at most d. Let $A = \{a_1, \dots, a_n\}$. Now, for $i = 1, \dots, n$, we receive a subset $B_i \subseteq N_G(a_i)$ of size $c \cdot d/2$, and we immediately have to match a_i to an unmatched vertex in B_i. If the sets B_i were determined in advance and $c \geq 2$, Hall's theorem [Hal35] would imply that we could find a perfect matching.

Problem 2.16. *If c is a sufficiently large constant, is there a way to construct a perfect matching in this online matching game where an adversary chooses the subsets B_i based on the matching of $a_1, \dots, a_{i-1}$?*

An alternative approach to replacing the error term of $\mathcal{O}(\log n)$ with $(\log \Delta)^{\mathcal{O}(1)}$ could be to prove stronger properties about how $C(F, X)$ evolves when fixing endpoints. This might be enough to recover the limited dependencies necessary to apply the Lovász Local Lemma.

In terms of the existence of linear forests with few paths, the conjecture of Magnant and Martin remains open up to a factor of two, but we believe that our techniques could be sufficient to eventually resolve this conjecture completely. Note that even for regular bipartite graphs where we proved this conjecture, our result still differs by a factor of roughly two from the best lower bound of $n/(2d)$ that comes from a disjoint union of complete bipartite graphs.

There are also several potential generalisations of the conjecture of Feige and Fuchs that remain open. For instance, this conjecture might generalise to cycles and to directed regular graphs.

Conjecture 2.17. *Every directed d-regular graph on n vertices has a collection of $\mathcal{O}(n/d)$ vertex-disjoint cycles that cover all vertices of the graph.*

In Chapter 3, we will prove that there exists a collection of $\mathcal{O}((n \log d)/d)$ cycles that satisfy this, which is the current best upper bound. In addition, this will provide a randomised polynomial-time algorithm that computes such a collection. As a corollary, it follows that there is an efficient algorithm for computing a tour of length $(1 + \mathcal{O}((\log d)/d)) \cdot n$ in a connected d-regular graph on n vertices.

In contrast, all of the results in this chapter are non-constructive, and so we do not obtain efficient algorithms. This raises the question whether there are efficient algorithms for finding spanning linear forests with $\mathcal{O}(n/d)$ paths in d-regular graphs. This would also imply the existence of efficient algorithms for comput-

ing a tour of length $(1 + \mathcal{O}(1/d)) \cdot n$ in connected regular graphs. An efficient algorithm would likely improve our understanding of small linear forests significantly, which may lead to further progress on the Linear Arboricity Conjecture.

Problem 2.18. *Is there an efficient algorithm for computing a spanning linear forest with at most $\mathcal{O}(n/d)$ paths in a d-regular graph on n vertices?*

In particular, what is the computational complexity of finding a rotation-component $C(F, X)$? This could plausibly be NP-hard. In extremal combinatorics it is unusual for a proof to rely on hard-to-compute objects, but this does occur with algebraic methods such as the method of interlacing polynomials [MSS15a, MSS15b] and topological methods such as applications of the Borsuk-Ulam Theorem [Mat03]. Note that for Hamiltonian cubic graphs where minimum linear forests are Hamilton paths, it is known that the number of rotations required to rotate one endpoint to some fixed vertex of the graph may be exponential in n. This follows, for example, from a construction in [BS22].

Chapter 3

Cycle-factors with few cycles

From the preceding chapter we know that the vertices of a regular graph can be decomposed into few vertex-disjoint paths. Does this also hold for cycles instead of paths? A *cycle-factor* of a graph is a decomposition of its vertices into vertex-disjoint cycles, where we consider a single edge to be a cycle of length two.

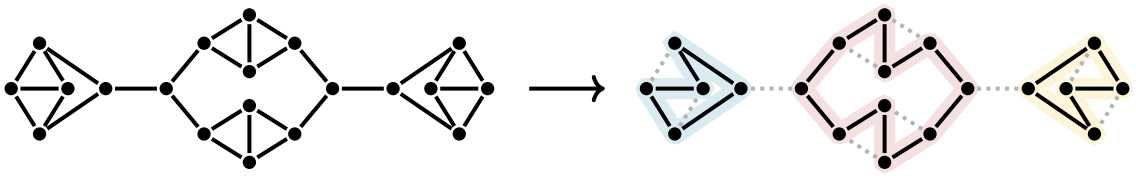


Figure 3.1: A cycle-factor of a 3-regular graph with three cycles.

Every d-regular graph G has a cycle-factor. To see this, construct an auxiliary d-regular bipartite graph H with bipartition $V(H) := A \cup B$, where A and B are both copies of $V(G)$. For each edge uv of G, add an edge between $u \in A$ and $v \in B$ and an edge between $v \in A$ and $u \in B$ to H. A cycle-factor of G then corresponds to a perfect matching of H. Since H is d-regular, Hall's theorem [Hal35] implies that H has a perfect matching, and so G has a cycle-factor, as required.

By proving the conjecture of Feige and Fuchs [FF22], we showed that every d-regular graph on n vertices has a spanning linear forest with at most $\mathcal{O}(n/d)$ paths. We conjecture that the same result should hold for cycle-factors.

Conjecture 3.1. *Every d-regular graph on n vertices has a cycle-factor with at most $\mathcal{O}(n/d)$ cycles.*

Some results towards the conjecture of Magnant and Martin [MM09] provided decompositions into cycles. Gruslys and Letzter [GL21] showed that if $d \geq \Omega(n)$, then there is a cycle-factor with at most $n/(d+1)$ cycles. If $d \geq \Omega((\log n)^{130})$,

Letzter, Methuku, and Sudakov [LMS26] also showed that there is a collection of $n/(d+1)$ vertex-disjoint cycles that cover all but $o(n)$ vertices. However, while Feige, Ravi, and Singh [FRS14] showed that there is a spanning linear forest with $\mathcal{O}(n/\sqrt{d})$ paths, the best upper bound for cycle-factors by Vishnoi [Vis12] only guaranteed a cycle-factor with $\mathcal{O}(n/\sqrt{\log d})$ cycles.

In this chapter, we significantly improve on this result by establishing the following upper bound.

Theorem 3.2. *Every d-regular graph on n vertices has a cycle-factor with at most $\mathcal{O}((n \log d)/d)$ cycles.*

In fact, we prove this result for directed regular graphs[1] and show that a random cycle-factor achieves this bound. A *cycle-factor* of a directed graph is a decomposition of its vertices into vertex-disjoint directed cycles. Equivalently, this corresponds to a permutation of the vertices such that every vertex is mapped to one of its out-neighbours. In particular, a random cycle-factor of the complete directed graph on n vertices with a loop at every vertex is the same as a random permutation of n elements.

The study of random permutations is one of the most classical topics in discrete probability. The probability that a random permutation has a fixed point was considered as early as 1704 in Pierre-Rémond de Montmort's *Essai d'Analyse sur les Jeux de Hasard*. Cauchy [Cau40] was the first to study permutations for their own sake. In the 1840s, he derived a formula for the number of permutations of n elements with a given cycle structure. This formula implies that, on average, a random permutation contains $\sum_{k=1}^{n} 1/k = \log n + \mathcal{O}(1)$ cycles. The latter result is now a standard exercise in discrete probability.

We generalise this result to all directed d-regular graphs on n vertices. By the discussion above, a random cycle-factor of the disjoint union of n/d complete directed graphs on d vertices has, on average, about $(n \log d)/d$ cycles. We prove that, up to constant factors, this is an upper bound on the expected number of cycles in a random cycle-factor of any directed d-regular graph.

Theorem 3.3. *The expected number of cycles in a uniformly random cycle-factor of a directed d-regular graph on n vertices is at most $\mathcal{O}((n \log d)/d)$.*

This recovers the aforementioned fact that, on average, uniformly random permutations of n elements have $\mathcal{O}(\log n)$ cycles. The constant factor in Theorem 3.3

[1] A directed graph can have loops and directed cycles of length 2, but no parallel edges. A directed graph is *d-regular* if the in-degree and out-degree of every vertex are both exactly d.

can be taken to be 4. Note that Theorem 3.2 follows from Theorem 3.3 if we direct each edge of the undirected regular graph in both directions.

In addition to proving that cycle-factors with few cycles exist, we can also find such cycle-factors efficiently. This follows from the fact that we can efficiently sample almost-uniform random perfect matchings of bipartite graphs [JSV04].

Corollary 3.4. *There exists a randomised polynomial-time algorithm that computes, with high probability, a cycle-factor of a directed d-regular graph on n vertices with at most $\mathcal{O}((n \log d)/d)$ cycles.*

If we remove one edge from each cycle in a cycle-factor, we obtain a spanning linear forest with the same number of paths as there were cycles in the cycle-factor. This implies that we can also efficiently compute spanning linear forests with few paths. Note that the results in the preceding chapter did not yield efficient algorithms, and the previous best algorithm by Feige, Ravi, and Singh [FRS14] only yielded spanning linear forests with $\mathcal{O}(n/\sqrt{d})$ paths.

Corollary 3.5. *There exists a randomised polynomial-time algorithm that computes, with high probability, a spanning linear forest of a d-regular graph on n vertices with at most $\mathcal{O}((n \log d)/d)$ paths.*

Finally, we can use our results to find short tours of connected regular graphs. Observe that a cycle-factor with c cycles in a connected graph can be efficiently transformed into a tour of length $n + 2(c-1)$. Therefore, we also obtain an efficient algorithm that computes a short tour of a connected regular graph.

Corollary 3.6. *There exists a randomised polynomial-time algorithm that computes, with high probability, a tour of a connected d-regular graph on n vertices of length at most $(1 + \mathcal{O}((\log d)/d)) \cdot n$.*

From the preceding chapter we know that a tour of length $(1 + \mathcal{O}(1/d)) \cdot n$ exists, but this was a non-algorithmic result. The previous best algorithm by Feige, Ravi, and Singh [FRS14] only yielded tours of length $(1 + \mathcal{O}(1/\sqrt{d})) \cdot n$.

Chapter outline. In Section 3.2, we prove our main results, Theorem 3.3 and Corollary 3.4. This section also includes an overview of our proof strategy. Since the proof of Theorem 3.3 uses the language of entropy, we first provide a short introduction to entropy in Section 3.1. We also use this section to introduce a small amount of notation. We finish in Section 3.3 with some open problems.

This chapter is based on joint work with Micha Christoph, Nemanja Draganić, António Girão, Eoin Hurley, and Alp Müyesser [CDG+25a].

3.1 ENTROPY

In this section, we recall some basic definitions and properties of Shannon entropy that will be essential for our proofs. Suppose that X is a random variable that takes values in a finite set $\mathcal{X}$. Then, the *entropy* of X is defined as

$$\mathcal{H}(X) := -\sum_{x\in\mathcal{X}} \mathbb{P}(X = x)\log_2(\mathbb{P}(X = x)).$$

Here, we use the convention $p\log_2 p = 0$ for $p = 0$. One basic property of entropy is that $\mathcal{H}(X) \leq \log_2|\mathcal{X}|$ with equality if and only if X is distributed uniformly at random in $\mathcal{X}$. For $0 \leq p \leq 1$, we write $\mathcal{H}(p) := -p\log_2 p - (1-p)\log_2(1-p)$ for the entropy of a Bernoulli random variable with probability p.

The definition of entropy easily extends to conditional probability. If E is some event, then the entropy of X conditioned on E occurring is

$$\mathcal{H}(X \mid E) := -\sum_{x\in\mathcal{X}} \mathbb{P}(X = x \mid E)\log_2(\mathbb{P}(X = x \mid E)).$$

If Y is another random variable that takes values in a finite set $\mathcal{Y}$, then the *conditional entropy* of X given Y is defined as

$$\mathcal{H}(X \mid Y) := \sum_{y\in\mathcal{Y}} \mathbb{P}(Y = y)\cdot\mathcal{H}(X \mid Y = y).$$

We remark that $\mathcal{H}(X \mid Y) = 0$ if and only if Y completely determines X.

The only fact about conditional entropy that we will need is the *chain rule*, which says that if $X_1,\dots,X_n$ are random variables, then

$$\mathcal{H}(X_1,\dots,X_n) = \mathcal{H}(X_1) + \mathcal{H}(X_2 \mid X_1) + \cdots + \mathcal{H}(X_n \mid X_1,\dots,X_{n-1})$$

where the tuple $X_1,\dots,X_i$ is viewed as a single random variable whose distribution is the joint distribution of $X_1,\dots,X_i$.

Notation. Let $[n] := \{1,\dots,n\}$. A permutation of $[n]$ is a bijection $\sigma : [n] \to [n]$. We denote the family of all permutations of $[n]$ by S_n. If $\sigma \in S_n$ and $\mathcal{I} \subseteq [n]$, we write $\sigma(\mathcal{I}) := \{\sigma(i) : i \in \mathcal{I}\}$. A cycle of $\sigma \in S_n$ is a minimal non-empty subset $\mathcal{I} \subseteq [n]$ such that $\sigma(\mathcal{I}) = \mathcal{I}$. We denote the number of cycles of σ by $|\sigma|$.

Finally, if G is a directed graph and $v \in V(G)$, then the *out-neighbourhood* of v in G is the set $N^+_G(v) := \{u \in V(G) : (v,u) \in E(G)\}$.

3.2 EXPECTED NUMBER OF CYCLES IN A RANDOM CYCLE-FACTOR

Suppose that G is a directed d-regular graph on n vertices. We wish to show that the expected number of cycles of a uniformly random cycle-factor of G is $\mathcal{O}((n \log d)/d)$. To achieve this, we will reveal the edges of the random cycle-factor one by one in a random order and bound the probability that each new revealed edge closes a cycle. We do this by going through the vertices in a random order and revealing their out-neighbours in the cycle-factor. If the revealed out-neighbour of a vertex were always uniformly distributed among all remaining available out-neighbours, then we could show the desired bound. Indeed, since the edges are revealed in a random order, it is possible to show that the number t of remaining available out-neighbours of a vertex is uniformly distributed in $[d]$. Because at most one of these choices closes a cycle, this implies that the probability that each new revealed edge closes a cycle would be at most

$$\frac{1}{d}\sum_{t=1}^{d}\frac{1}{t} = \mathcal{O}\left(\frac{\log d}{d}\right).$$

To handle the general case where out-neighbours are not uniformly distributed, we connect our process to Radhakrishnan's entropy proof [Rad97] of the Bregman–Minc inequality [Bre73]. Towards that end, construct a d-regular bipartite graph H with bipartition $V(H) = A \cup B$, where A and B are both copies of $V(G)$, and for each directed edge (u, v) of G we add an edge between $u \in A$ and $v \in B$ to H. It is easy to see that cycle-factors of G correspond to perfect matchings of H.[2] The Bregman–Minc inequality gives the following upper bound on the number of perfect matchings in regular bipartite graphs.

Theorem 3.7 (Bregman–Minc inequality)**.** *Every d-regular bipartite graph on $2n$ vertices has at most $(d!)^{n/d}$ perfect matchings.*

On a high level, Radhakrishnan's proof of Theorem 3.7 picks a uniformly random perfect matching and shows that its entropy is at most $\log_2((d!)^{n/d})$. The idea is to reveal the edges of the random matching one by one in a random order[3] and to bound the amount of entropy added by each new revealed edge. Note that the

[2] This observation was also used by Vishnoi in [Vis12] to show that a typical cycle-factor in a d-regular graph has $\mathcal{O}(n/\sqrt{\log d})$ cycles.

[3] To the best of our knowledge, the random order was first used by Alon and Spencer [AS16] to prove the Bregman–Minc inequality.

amount of entropy added is maximised if each new revealed edge is uniformly distributed among its remaining options. This can be used to prove Theorem 3.7.

We pursue a very similar strategy to analyse random cycle-factors of G. Conveniently, revealing an edge of a random perfect matching of H corresponds precisely to revealing an edge of the random cycle-factor of G. As mentioned before, we wish to understand the case where a revealed edge of the cycle-factor is not uniformly distributed among its remaining options. If this happens, the amount of entropy added for the corresponding edge in H is lower than the entropy of the uniform distribution. This, in turn, translates into a better upper bound on the number of perfect matchings of H. However, this upper bound cannot drop too low. This is because the Van der Waerden inequality, proved independently by Egorychev [Ego81] and Falikman [Fal81], gives the following lower bound on the number of perfect matchings in regular bipartite graphs.

Theorem 3.8 (Van der Waerden inequality)**.** *Every d-regular bipartite graph on $2n$ vertices has at least $n! \cdot d^n / n^n$ perfect matchings.*

This lower bound allows us to quantify how much entropy can be lost while revealing the edges of the random cycle-factor of G. We will see that the total loss in entropy is at most $\mathcal{O}((n \log d)/d)$ compared to what the entropy would have been had the edges been uniformly distributed throughout the process. The key fact that ties everything together is that any loss in entropy increases the probability of closing a cycle at most by an additive linear term. Together, this implies Theorem 3.3.

We start with the following simple lemma that formalises the relation between a loss in entropy and an increase in the probability of a random variable attaining a specific value. In our case, this will be the probability of closing a cycle.

Lemma 3.9. *Let X be a random variable taking values in a finite set $\mathcal{X}$, and define $\ell := \log_2|\mathcal{X}| - \mathcal{H}(X)$. Then, for all $x \in \mathcal{X}$ we have*

$$\mathbb{P}(X = x) \leq \frac{2}{|\mathcal{X}|} + \ell.$$

Proof. Let $x \in \mathcal{X}$ and let $p := \mathbb{P}(X = x)$. Consider the indicator random variable $\mathbb{1}_{X=x}$ of the event that $X = x$. By the chain rule for entropy,

$$\begin{aligned}\mathcal{H}(X) &= \mathcal{H}(\mathbb{1}_{X=x}) + \mathcal{H}(X \mid \mathbb{1}_{X=x}) \\ &= \mathcal{H}(p) + p \cdot \mathcal{H}(X \mid X = x) + (1-p) \cdot \mathcal{H}(X \mid X \neq x).\end{aligned}$$

Clearly, $\mathcal{H}(X \mid X = x) = 0$. Also, if $X \neq x$, then X takes at most $|\mathcal{X}| - 1$ distinct values, and so $\mathcal{H}(X \mid X \neq x) \leq \log_2(|\mathcal{X}| - 1)$. This implies that

$$\mathcal{H}(X) \leq \mathcal{H}(p) + (1-p)\log_2(|\mathcal{X}| - 1).$$

Let $h(q) := \mathcal{H}(q) + (1-q)\log_2(|\mathcal{X}| - 1)$. Note that $h(q) = \mathcal{H}(Y)$ for the random variable Y that satisfies $\mathbb{P}(Y = x) = q$ and $\mathbb{P}(Y = y) = (1-q)/(|\mathcal{X}| - 1)$ for all $y \in \mathcal{X} \setminus \{x\}$, and so $h(q) \leq \log_2|\mathcal{X}|$. Moreover,

$$h'(q) = -\log_2 q + \log_2(1-q) - \log_2(|\mathcal{X}| - 1) = \log_2\left(\frac{1-q}{q(|\mathcal{X}| - 1)}\right).$$

Since this is decreasing in q, we get for $q \geq 2/|\mathcal{X}|$ that

$$h'(q) \leq h'\left(\frac{2}{|\mathcal{X}|}\right) = \log_2\left(\frac{|\mathcal{X}| - 2}{2(|\mathcal{X}| - 1)}\right) \leq \log_2\left(\frac{1}{2}\right) = -1.$$

So, for $q \geq 2/|\mathcal{X}|$, we have

$$h(q) \leq h\left(\frac{2}{|\mathcal{X}|}\right) - \left(q - \frac{2}{|\mathcal{X}|}\right) \leq \log_2|\mathcal{X}| - \left(q - \frac{2}{|\mathcal{X}|}\right).$$

Since $h(p) \geq \mathcal{H}(X) = \log_2|\mathcal{X}| - \ell$, it follows that $p \leq 2/|\mathcal{X}| + \ell$. □

We can now prove our main result using the strategy outlined earlier in this section. The first part of the proof closely follows Radhakrishnan's proof [Rad97] of the Bregman–Minc inequality (see also [AZ18, Chapter 37] for a nice exposition of this proof), with the twist being that we keep track of the total loss in entropy. This is used in the second part to prove an upper bound on the expected number of cycles, by passing through Lemma 3.9.

Proof of Theorem 3.3. We identify the vertices of G with $[n]$. Let $\mathcal{C}$ be the set of all permutations σ of $[n]$ such that $(i, \sigma(i)) \in E(G)$ for all $i \in [n]$. Note that $\mathcal{C}$ corresponds to the collection of all cycle-factors of G. Since cycle-factors of G correspond to perfect matchings in the auxiliary bipartite graph H from before, Theorem 3.8 implies that $|\mathcal{C}| \geq n! \cdot d^n/n^n$. As $n! \geq (n/e)^n$, we get $|\mathcal{C}| \geq (d/e)^n$.

Pick $\sigma \in \mathcal{C}$ uniformly at random, so $\mathcal{H}(\sigma) = \mathcal{H}(\sigma(1), \ldots, \sigma(n)) = \log_2|\mathcal{C}|$. For every permutation $\tau \in S_n$, the chain rule for entropy implies that

$$\mathcal{H}(\sigma) = \sum_{j=1}^{n} \mathcal{H}(\sigma(\tau(j)) \mid \sigma(\tau(1)), \ldots, \sigma(\tau(j-1))).$$

For $i \in [n]$, let $\tau_{<i} := \tau(1), \ldots, \tau(j-1)$, where $j := \tau^{-1}(i)$. We will use the notation $\sigma(\tau_{<i}) := \sigma(\tau(1)), \ldots, \sigma(\tau(j-1))$. If we reorder the above sum, we get

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \mathcal{H}(\sigma(i) \mid \sigma(\tau_{<i})).$$

Expanding the definition of conditional entropy, this yields

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \sum_{\pi} \mathbb{P}(\sigma(\tau_{<i}) = \pi) \cdot \mathcal{H}(\sigma(i) \mid \sigma(\tau_{<i}) = \pi),$$

where π ranges over all possible values of $\sigma(\tau_{<i})$. Since σ is chosen uniformly at random, observe that $\mathbb{P}(\sigma(\tau_{<i}) = \pi) = |\{\sigma' \in \mathcal{C} : \sigma'(\tau_{<i}) = \pi\}|/|\mathcal{C}|$. Therefore,

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathcal{H}(\sigma(i) \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i})).$$

For all $i \in [n]$, $\sigma' \in \mathcal{C}$, and $\tau \in S_n$, let $a(i, \sigma', \tau) := |N_G^+(i) \setminus \sigma'(\tau_{<i})|$ be the number of *available out-neighbours* of i, that is, the out-neighbours of i in G that are not contained in $\sigma'(\tau_{<i})$. If $\sigma(\tau_{<i}) = \sigma'(\tau_{<i})$, then $\sigma(i)$ must be one of these available out-neighbours, and so $\mathcal{H}(\sigma(i) \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i})) \leq \log_2 a(i, \sigma', \tau)$. Define $\ell(i, \sigma', \tau) := \log_2 a(i, \sigma', \tau) - \mathcal{H}(\sigma(i) \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i}))$. Then,

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} (\log_2 a(i, \sigma', \tau) - \ell(i, \sigma', \tau)).$$

Since this holds for all $\tau \in S_n$, we may pick $\tau \in S_n$ uniformly at random. This yields

$$\begin{aligned} \mathcal{H}(\sigma) &= \mathbb{E}_\tau \left(\sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} (\log_2 a(i, \sigma', \tau) - \ell(i, \sigma', \tau)) \right) \\ &= \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathbb{E}_\tau (\log_2 a(i, \sigma', \tau)) - \mathcal{L} \end{aligned}$$

where

$$\mathcal{L} := \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathbb{E}_\tau (\ell(i, \sigma', \tau))$$

denotes the loss in entropy compared to a uniform distribution.

For a fixed i and σ', note that $a(i, \sigma', \tau)$ is the number of out-neighbours $j \in N_G^+(i)$ of i in G such that $(\sigma')^{-1}(j)$ does not appear before i in τ. Since τ is a uniformly

random permutation, this number is distributed uniformly at random in $[d]$, and so it follows that

$$\mathcal{H}(\sigma) = \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \frac{1}{d} \sum_{t=1}^{d} \log_2 t - \mathcal{L} = \frac{n}{d} \log_2(d!) - \mathcal{L}.$$

Since $|\mathcal{C}| \geq (d/e)^n$ and $d! \leq ed(d/e)^d$, this implies that

$$n \log_2 \left(\frac{d}{e}\right) \leq \log_2 |\mathcal{C}| = \mathcal{H}(\sigma) = \frac{n}{d} \log_2(d!) - \mathcal{L} \leq n \log_2 \left(\frac{d}{e}\right) + \frac{n}{d} \log_2(ed) - \mathcal{L},$$

and therefore $\mathcal{L} \leq (n \log_2(ed))/d \leq 2n(\log d + 1)/d$.

We now bound the number of cycles $|\sigma|$ of σ using a very similar calculation. For every permutation $\tau \in S_n$, we say that $\sigma(i)$ closes a cycle if $\tau_{<i}$ contains all other vertices that are on the same cycle as i in σ. That is, adding $\sigma(i)$ to $\sigma(\tau_{<i})$ creates one additional cycle. Note that $|\sigma|$ is equal to the number of vertices $i \in [n]$ for which $\sigma(i)$ closes a cycle, and so

$$\begin{aligned}
\mathbb{E}_\sigma(|\sigma|) &= \sum_{i=1}^{n} \mathbb{P}(\sigma(i) \text{ closes a cycle}) \\
&= \sum_{i=1}^{n} \sum_{\pi} \mathbb{P}(\sigma(\tau_{<i}) = \pi) \cdot \mathbb{P}(\sigma(i) \text{ closes a cycle} \mid \sigma(\tau_{<i}) = \pi) \\
&= \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathbb{P}(\sigma(i) \text{ closes a cycle} \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i})).
\end{aligned}$$

If $\sigma(\tau_{<i}) = \sigma'(\tau_{<i})$, then either $\sigma(i)$ cannot close a cycle, or $\sigma(i)$ needs to be one specific available out-neighbour $j \in N_G^+(i) \setminus \sigma'(\tau_{<i})$ in order to close a cycle. So, by Lemma 3.9 we know that

$$\begin{aligned}
\mathbb{P}(\sigma(i) \text{ closes a cycle} \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i})) &\leq \max_{j \in N_G^+(i)} \mathbb{P}(\sigma(i) = j \mid \sigma(\tau_{<i}) = \sigma'(\tau_{<i})) \\
&\leq \frac{2}{a(i, \sigma', \tau)} + \ell(i, \sigma', \tau).
\end{aligned}$$

This implies that

$$\mathbb{E}_\sigma(|\sigma|) \leq \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \left(\frac{2}{a(i, \sigma', \tau)} + \ell(i, \sigma', \tau) \right).$$

Since this holds for all $\tau \in S_n$, we may again pick $\tau \in S_n$ uniformly at random.

As noted before, $a(i, \sigma', \tau)$ is then distributed uniformly at random in $[d]$, and so

$$\begin{aligned}
\mathbb{E}_\sigma(|\sigma|) &\leq \mathbb{E}_\tau\left(\sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \left(\frac{2}{a(i,\sigma',\tau)} + \ell(i,\sigma',\tau)\right)\right) \\
&= \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \mathbb{E}_\tau\left(\frac{2}{a(i,\sigma',\tau)}\right) + \mathcal{L} \\
&= \sum_{i=1}^{n} \frac{1}{|\mathcal{C}|} \sum_{\sigma' \in \mathcal{C}} \frac{1}{d} \sum_{t=1}^{d} \frac{2}{t} + \mathcal{L} \\
&\leq 2 \cdot \frac{n}{d}(\log d + 1) + \mathcal{L} \leq 4 \cdot \frac{n}{d}(\log d + 1). \qquad \square
\end{aligned}$$

We finish by showing that there is a randomised polynomial-time algorithm that finds a cycle-factor of G with at most $\mathcal{O}((n \log d)/d)$ cycles with high probability. For this, we will use the fact that there exists an algorithm that samples an almost-uniform[4] random perfect matching in bipartite graphs, which was proved by Jerrum, Sinclair, and Vigoda [JSV04].

Theorem 3.10. *For any bipartite graph H, there exists a randomised polynomial-time algorithm that samples an almost-uniform random perfect matching of H.*

Using the fact that cycle-factors correspond to perfect matchings in the bipartite graph H constructed before, Corollary 3.4 is now an easy consequence.

Proof of Corollary 3.4. Recall that cycle-factors of G correspond to perfect matchings in the auxiliary bipartite graph H constructed before. By Theorem 3.10 we can sample an almost-uniform random perfect matching of H in polynomial time, and so we can sample an almost-uniform random cycle-factor σ of G in polynomial time.

By Theorem 3.3 and Markov's inequality, we know that a uniformly random cycle-factor of G has at most $\mathcal{O}((n \log d)/d)$ cycles with probability at least $1/2$. Since σ is almost-uniform with some error δ, this implies that σ has at most $\mathcal{O}((n \log d)/d)$ cycles with probability at least $1/2 - \delta$. If we now sample σ independently at least $\Omega(\log n)$ times and return the cycle-factor with the least number of cycles, then the returned cycle-factor has at most $\mathcal{O}((n \log d)/d)$ cycles with high probability. $\square$

[4] If X is a random variable that takes values in a finite set $\mathcal{X}$, then X is *almost-uniform* with error δ if the total variation distance between X and the uniform distribution is at most δ, that is, $\sum_{x \in \mathcal{X}} |1/|\mathcal{X}| - \mathbb{P}(X = x)|/2 \leq \delta$.

3.3 OPEN PROBLEMS

We believe that Theorem 3.3 should hold with a constant factor of $1 + o(1)$. Note, however, that Gajjala [Gaj26] recently showed that the expected number of cycles in a uniformly random cycle-factor is not maximised by the disjoint union of n/d complete directed graphs on d vertices with a loop at every vertex.

Conjecture 3.11. *The expected number of cycles in a uniformly random cycle-factor of a directed d-regular graph on n vertices is at most $(1 + o(1))(n \log d)/d$ as $d \to \infty$.*

It is natural to ask what other properties of random permutations might generalise to random cycle-factors of directed regular graphs. As an example, de Montmort's problem about the probability of a fixed point does not generalise. Indeed, in the complete directed bipartite graph between a set A of d vertices and a set B of $d - 1$ vertices with a loop at every vertex of A, every cycle-factor has exactly one fixed point, and so the probability that a random cycle-factor has a fixed point is 1. In contrast, the probability that a random permutation has a fixed point is approximately $1 - 1/e$. However, it remains possible that the expected number of fixed points generalises to random cycle-factors.

Problem 3.12. *Is the expected number of fixed points in a uniformly random cycle-factor of a directed d-regular graph on n vertices at most $\mathcal{O}(n/d)$?*

Another property of random permutations is that every element is equally likely to be mapped to every other element. The corresponding generalisation to directed regular graphs would have answered Problem 3.12 affirmatively. Unfortunately, this generalisation is completely false since there are directed d-regular graphs with an edge that is contained in a uniformly random cycle-factor with probability $1 - o(1)$. This follows from a construction in [ACFK16, Theorem 1.7]. We note that this generalisation turns out to be correct in the special case of dense and robustly expanding regular bipartite graphs [GJ24, Theorem 3.3].

In general, it would be interesting to identify further statistics of the cycle structure of random permutations that admit analogues for random cycle-factors of directed regular graphs.

Chapter 4

Chromatic number is not tournament-local

Since the 1940s and 1950s, it has been well-known that there are graphs of large chromatic number in which each neighbourhood is independent [Zyk49, Des54]. Harutyunyan, Le, Thomassé, and Wu [HLTW19] recently showed that the situation is different for tournaments. A *tournament* T is an orientation of a complete graph, and its *chromatic number* $\chi(T)$ is the minimum k for which there is a partition of the vertices of T into k parts each of which induces an acyclic subtournament of T. Harutyunyan, Le, Thomassé, and Wu showed that for every k there exists a χ such that every tournament T with $\chi(T) \geq \chi$ contains a vertex v with $\chi(T[N_T^+(v)]) \geq k$.[1] This confirmed a conjecture of Berger, Choromanski, Chudnovsky, Fox, Loebl, Scott, Seymour, and Thomassé [BCC+13].

Separately, Scott and Seymour [SS16] made a similar conjecture for a graph and a tournament on the same vertex set (see also [HLTW19, Conjecture 7]).

Conjecture 4.1 (Scott and Seymour)**.** *For every positive integer k there exists a χ such that for every graph G with $\chi(G) \geq \chi$ and every tournament T on the same vertex set, there is a vertex v such that $\chi(G[N_T^+(v)]) \geq k$.*

They also observed that the statement holds if both occurrences of chromatic number are replaced by fractional chromatic number (see Section 4.2).

In this chapter, we disprove the conjecture of Scott and Seymour for $k \geq 3$. In fact, we show that G and T can be chosen in such a way that the out-neighbourhood of any set of vertices U of size at most $(\log|V(T)|)/(2\chi^2)$ is bipartite.[2]

[1] The *out-neighbourhood* of a vertex v in a tournament T is $N_T^+(v) \coloneqq \{u \in V(T) : (v,u) \in E(T)\}$.
[2] The *out-neighbourhood* of U in T is $N_T^+(U) \coloneqq \bigcup_{v \in U} N_T^+(v)$. This need not be disjoint from U.

Theorem 4.2. *For every positive integer χ there are arbitrarily large n for which there is a triangle-free graph G and a tournament T on the same vertex set of size n such that $\chi(G) = \chi$ and, for every set U of at most $(\log n)/(2\chi^2)$ vertices, $\chi(G[N_T^+(U)]) \leq 2$.*

We prove this by taking G to be a Schrijver graph [Sch78], which is known to have both large chromatic number and many induced bipartite subgraphs that each cover far more than half of the vertices of G. For each vertex v, we carefully select one of these subgraphs as the out-neighbourhood of v in the tournament T. The fact that G is triangle-free will be useful for the proof of Corollary 4.3.

We make two remarks concerning the optimality of Theorem 4.2.

- It is not possible to replace 2 by 1 in the bound on the chromatic number of the out-neighbourhood, even when U consists of a single vertex. Indeed, suppose that $G[N_T^+(v)]$ is independent for every vertex v. Let xy be an edge of G. No out-neighbourhood of a vertex of T contains both x and y, so $\{x, y\}$ dominates T. But then G is 3-colourable: one colour for each of $N_T^+(x)$ and $N_T^+(y)$, and a final colour for whichever of x and y has not been coloured.
- The bound on the size of U is very close to being best possible. Let S be a dominating set of T of size at most $\lceil \log_2 n \rceil$ (such a set can be constructed greedily). Then $N_T^+(S)$ contains all vertices of G except perhaps one and so, for any $1 \leq \ell \leq \chi - 2$, there is some subset $U \subseteq S$ of size at most $\lceil (\log_2 n)/\lfloor (\chi-2)/\ell \rfloor \rceil$ with $\chi(G[N_T^+(U)]) > \ell$.

Theorem 4.2 has the following corollary, which resolves in a strong sense a question of Harutyunyan, Le, Thomassé, and Wu [HLTW19] concerning the analogous problem for two tournaments on the same vertex set.

Corollary 4.3. *For every positive integer χ there are arbitrarily large n for which there are tournaments T_1 and T_2 on the same vertex set of size n such that $\chi(T_1) \geq \chi$ and, for every set U of at most $(\log n)/(8\chi^2)$ vertices, $\chi(T_1[N_{T_2}^+(U)]) \leq 2$.*

In turn, Corollary 4.3 has the following immediate consequence, which disproves a conjecture of Nguyen, Scott, and Seymour [NSS25].

Corollary 4.4. *For every positive integer χ there are arbitrarily large n for which there is a tournament T with n vertices and disjoint subsets $A, B \subseteq V(T)$ such that $\chi(T[A]), \chi(T[B]) \geq \chi$ and the following holds. For all $A' \subseteq A$ and $B' \subseteq B$ of size at most $(\log n)/(32\chi^2)$, both $\chi(T[A \cap N_T^+(B')]) \leq 2$ and $\chi(T[B \cap N_T^+(A')]) \leq 2$.*

Finally, we consider the setting where chromatic number is replaced by degeneracy (or equivalently maximum average degree). Since every graph of high chro-

matic number has high degeneracy, Theorem 4.2 shows that for every positive integer d there is a graph G and a tournament T on the same vertex set such that the degeneracy of G is at least d, but the subgraph of G induced on each out-neighbourhood of T is bipartite. The following result strengthens this statement by ensuring that the graph induced on each out-neighbourhood is 1-degenerate.

Proposition 4.5. *For every positive integer d, there is a d-regular graph G and a tournament T on the same vertex set such that $G[N_T^+(v)]$ is a forest for every vertex v.*

Despite this result, and in contrast to Theorem 4.2, if G has high degeneracy and T is a tournament on the same vertex set, then there is a set of two vertices whose out-neighbourhood has high degeneracy.

Theorem 4.6. *For every positive integer d, every graph G with degeneracy at least $12d$, and every tournament T on the same vertex set as G, there exist two vertices u and v such that $G[N_T^+(\{u,v\})]$ has degeneracy at least $d-1$.*

Chapter outline. In Section 4.1, we prove our main result, Theorem 4.2, and deduce Corollaries 4.3 and 4.4. In Section 4.2, we provide a short argument for why Conjecture 4.1 holds for fractional chromatic number instead of chromatic number. We then transition from chromatic number to degeneracy in Section 4.3 and prove Proposition 4.5 and Theorem 4.6. We finish in Section 4.4 with some open problems.

Notation. Let $[n] := \{1,\dots,n\}$ and $\binom{[n]}{k} := \{S \subseteq [n] : |S| = k\}$.

This chapter is based on joint work with António Girão, Kevin Hendrey, Freddie Illingworth, Florian Lehner, Michael Savery, and Raphael Steiner [GHI+24].

4.1 CHROMATIC NUMBER

Our construction that disproves the conjecture of Scott and Seymour is based on the classical Schrijver graphs [Sch78].

Definition 4.7. Let $k \geq 1$ and $\ell \geq 2k$ be integers. The *Kneser graph* $\mathrm{KG}(\ell,k)$ is the graph whose vertex set is $\binom{[\ell]}{k}$ and in which two sets $S_1, S_2 \in \binom{[\ell]}{k}$ are adjacent if and only if $S_1 \cap S_2 = \varnothing$. The *Schrijver graph* $\mathrm{SG}(\ell,k)$ is the induced subgraph of $\mathrm{KG}(\ell,k)$ whose vertex set consists of all sets that are *stable* in the cycle C_ℓ, that is, all sets $S \in \binom{[\ell]}{k}$ that do not include two cyclically consecutive[3] elements of $[\ell]$.

[3] By this we mean a pair $(i, i+1)$ where $1 \leq i < \ell$ or the pair $(\ell, 1)$.

Kneser [Kne55] conjectured that the chromatic number of $\mathrm{KG}(\ell,k)$ is $\ell-2k+2$. Observe that we can colour $\mathrm{KG}(\ell,k)$ with $\ell-2k+2$ colours by assigning colour $c\in[\ell-2k+1]$ to each set whose smallest element is c, and assigning colour $\ell-2k+2$ to all remaining sets. Kneser's conjecture remained open for two decades and was first proved by Lovász [Lov78] using homotopy theory (see also Bárány [Bár78] and Greene [Gre02] for very short proofs). Shortly afterwards, Schrijver [Sch78] introduced $\mathrm{SG}(\ell,k)$ and proved that $\mathrm{SG}(\ell,k)$ is vertex-critical with chromatic number $\chi(\mathrm{SG}(\ell,k))=\chi(\mathrm{KG}(\ell,k))=\ell-2k+2$.

We will show that for every $\chi\geq 3$ and every sufficiently large k, there is a tournament T on the vertex set of $\mathrm{SG}(2k+\chi-2,k)$ such that the out-neighbourhood of every sufficiently small set of vertices in T induces a bipartite subgraph of $\mathrm{SG}(2k+\chi-2,k)$. As $\chi(\mathrm{SG}(2k+\chi-2,k))=\chi$, this will prove Theorem 4.2.

Our construction relies on the existence of large induced bipartite subgraphs in $\mathrm{SG}(\ell,k)$. Specifically, for any $r\in[\ell]$, observe that the family of stable sets $S\in\binom{[\ell]}{k}$ with $S\cap\{r,r+1\}\neq\varnothing$ induces a bipartite subgraph of $\mathrm{SG}(\ell,k)$ that can contain an arbitrarily large fraction of the vertices. The corresponding family for $\mathrm{KG}(\ell,k)$ contains only about $3/4$ of the vertices, which is why our construction requires Schrijver graphs. For any set $S\in\binom{[\ell]}{k}$, let $\mathrm{gap}(S)$ be the set of all $r\in[\ell]$ such that $S\cap\{r,r+1\}=\varnothing$, where addition is modulo ℓ (that is, $\ell+1$ is identified with 1). Then, the family above is precisely the family of sets S with $r\notin\mathrm{gap}(S)$.

To motivate our construction, suppose for now that we only want to ensure that the out-neighbourhood of each individual vertex is bipartite. To this end, we can assign an element $f(\mathrm{gap}(S))\in[\ell]$ to each stable set $S\in\binom{[\ell]}{k}$ and define the out-neighbourhood of S in T to be the family of all stable sets $S'\in\binom{[\ell]}{k}$ with $f(\mathrm{gap}(S))\notin\mathrm{gap}(S')$. As observed above, this out-neighbourhood induces a bipartite subgraph of $\mathrm{SG}(\ell,k)$. However, in order for T to be a tournament, there must be a directed edge between each pair of stable sets $S_1,S_2\in\binom{[\ell]}{k}$. This requires that either $f(\mathrm{gap}(S_1))\notin\mathrm{gap}(S_2)$ or $f(\mathrm{gap}(S_2))\notin\mathrm{gap}(S_1)$.

The following result shows that such an assignment exists. In fact, it gives an assignment satisfying a stronger condition which will allow us to ensure that the combined out-neighbourhood of every small set of vertices is bipartite. Crucially, instead of assigning a single element to each stable set S, we will assign a set of elements $f(\mathrm{gap}(S))\subseteq[\ell]$ to S. Later, the out-neighbourhood of S in T will then consist only of stable sets S' with $f(\mathrm{gap}(S))\cap\mathrm{gap}(S')=\varnothing$. We deduce this result from the existence of tournaments with high domination number.

Lemma 4.8. *For every positive integer t there is some ℓ_0 such that for all integers $\ell \geq \ell_0$ there exists a function $f\colon \binom{[\ell]}{t} \to 2^{[\ell]}$ with the following two properties:*

- *for every $A, B \in \binom{[\ell]}{t}$, at least one of $A \cap f(B)$ and $B \cap f(A)$ is empty, and*
- *for every collection $(A_i)_{i\in I}$ of at most $(\log \ell)/(2t)$ sets from $\binom{[\ell]}{t}$,*

$$\bigcap_{i\in I} f(A_i) \neq \varnothing.$$

Proof. By a classical probabilistic result of Erdős [Erd63] (see [GS71] for an explicit construction), for every sufficiently large ℓ there is a tournament on ℓ vertices in which every set of at most $(\log \ell)/2$ vertices is dominated by a vertex outside the set. Let ℓ be large enough so that this holds and $(\log \ell)/2 \geq t$, and let T be the corresponding tournament. Identify $V(T)$ with $[\ell]$. For $A \in \binom{[\ell]}{t}$, define

$$f(A) := \{v \in [\ell] \setminus A : v \text{ dominates } A\}.$$

We claim that f satisfies the two properties of the lemma statement. First, let $A, B \in \binom{[\ell]}{t}$ and suppose for a contradiction that $A \cap f(B)$ and $B \cap f(A)$ are both non-empty. Then there is some $a \in A \setminus B$ that dominates B and some $b \in B \setminus A$ that dominates A. This implies that a and b are distinct and the edge between them is oriented in both directions, which is a contradiction. Next, let $(A_i)_{i\in I}$ be a collection of at most $(\log \ell)/(2t)$ sets from $\binom{[\ell]}{t}$. Define $A := \bigcup_{i\in I} A_i$, which is a set of size at most $(\log \ell)/2$. By the definition of T, some vertex $x \notin A$ dominates A, but then $x \in \bigcap_{i\in I} f(A_i)$, as required. □

Next, we will deduce Theorem 4.2 using the construction outlined before.

Proof of Theorem 4.2. The result is trivial for $\chi \leq 2$, so suppose $\chi \geq 3$, define $t := \chi - 2$, and let ℓ_0 be given by Lemma 4.8. Pick some positive integer $k > t$ such that $2k + t \geq \ell_0$, let $\ell := 2k + t$, and let $G := \mathrm{SG}(\ell, k)$. Note that G is triangle-free and has chromatic number χ. Also, every $S \in V(\mathrm{SG}(\ell, k))$ satisfies $\mathrm{gap}(S) \in \binom{[\ell]}{t}$ and is uniquely determined by $\mathrm{gap}(S)$, and so $n := |V(G)| \leq \binom{\ell}{t} \leq \ell^t$.

Let $f\colon \binom{[\ell]}{t} \to 2^{[\ell]}$ be the function given by Lemma 4.8. Define a directed graph D on the same vertex set as G that has a directed edge from a vertex S_1 to a vertex S_2 if and only if $f(\mathrm{gap}(S_1)) \cap \mathrm{gap}(S_2) = \varnothing$. Note that by the first property of f guaranteed by Lemma 4.8, any two distinct vertices of D are connected by a directed edge in at least one of the two possible directions. Hence, there exists a spanning subdigraph T of D which is a tournament.

Let U be any set of at most $(\log n)/(2\chi^2) \leq (\log n)/(2t^2) \leq (\log \ell)/(2t)$ vertices. To finish the proof we will show that the out-neighbourhood $N_D^+(U)$ induces a bipartite subgraph of G, and so the same is true for the out-neighbourhood $N_T^+(U) \subseteq N_D^+(U)$. Write $U = \{S_1, \ldots, S_{|U|}\}$. By the second property of f guaranteed by Lemma 4.8, there is some $r \in [\ell]$ common to all the $f(\text{gap}(S_i))$. By the definition of D, any $S \in N_D^+(U)$ satisfies $r \notin \text{gap}(S)$ and so $S \cap \{r, r+1\} \neq \varnothing$. Colouring all the vertices in the out-neighbourhood that include the element r with one colour and all the remaining vertices (which necessarily contain $r+1$) with another colour provides a proper 2-colouring of $G[N_D^+(U)]$. This concludes the proof of the theorem. □

We can convert the graph G from Theorem 4.2 to a tournament: pick any linear order on the vertices of G and construct a tournament T_1 whose back-edge graph is G. We will show that $\chi(G)$ and $\chi(T_1)$ are closely related, and this suffices to prove Corollary 4.3.

Proof of Corollary 4.3. By Theorem 4.2, there are arbitrarily large n for which there is a triangle-free graph G with chromatic number 2χ and a tournament T_2 on the same vertex set of size n such that, for every set U of at most $(\log n)/(8\chi^2)$ vertices, $\chi(G[N_{T_2}^+(U)]) \leq 2$. Let $\prec$ be a linear order on $V(G)$ and define a tournament T_1 on the vertex set $V(G)$ by adding a directed edge from a vertex u to a vertex v if either $v \prec u$ and $uv \in E(G)$ or $u \prec v$ and $uv \notin E(G)$.

Let $W \subseteq V(G)$ be any set of vertices such that $T_1[W]$ is acyclic. Note that if $v_1v_2v_3$ is a path in G with $v_1 \prec v_2 \prec v_3$, then $v_1v_3 \notin E(G)$ since G is triangle-free, so $v_1v_2v_3$ is a cyclic triangle in T_1, and this implies that v_1, v_2, and v_3 are not all in W. In particular, the partition $W = W_0 \cup W_1$ where

$$W_0 \coloneqq \{w \in W : \text{there is no } w' \in W \text{ such that } w' \prec w \text{ and } w'w \in E(G)\} \text{ and}$$
$$W_1 \coloneqq \{w \in W : \text{there is some } w' \in W \text{ such that } w' \prec w \text{ and } w'w \in E(G)\}$$

gives a proper 2-colouring of the vertices of $G[W]$. Because this holds for any $W \subseteq V(G)$ where $T_1[W]$ is acyclic, it follows that $\chi(T_1) \geq \chi(G)/2 = \chi$.

To finish the proof, consider any set U of at most $(\log n)/(8\chi^2)$ vertices. Since $G[N_{T_2}^+(U)]$ is bipartite, there are two disjoint independent sets I_1 and I_2 in G such that $N_{T_2}^+(U) = I_1 \cup I_2$. Consider two vertices $u, v \in I_i$ for some $i \in \{1,2\}$ and note that since $uv \notin E(G)$, there is a directed edge from u to v in T_1 if and only if $u \prec v$. Hence, $T_1[I_1]$ and $T_1[I_2]$ are acyclic tournaments and so $\chi(T_1[N_{T_2}^+(U)]) \leq 2$. □

To prove Corollary 4.4, we can now take the two tournaments T_1 and T_2 from Corollary 4.3 and combine them appropriately: we simply orient the edges within A and B according to T_1, and the edges between A and B according to T_2.

Proof of Corollary 4.4. By Corollary 4.3, there are arbitrarily large n for which there exist tournaments T_1 and T_2 on the same vertex set of size n with $\chi(T_1) \geq 2\chi$ and $\chi(T_1[N^+_{T_2}(U)]) \leq 2$ for every set U of at most $(\log n)/(32\chi^2)$ vertices. Partition $V(T_1)$ into two sets A and B such that $\chi(T_1[A]), \chi(T_1[B]) \geq \chi$ and construct a new tournament T on $V(T_1)$ by orienting an edge between two vertices u and v as in T_1 if $u, v \in A$ or $u, v \in B$, and orienting the edge as in T_2 otherwise. Then, for all $A' \subseteq A$ of size at most $(\log n)/(32\chi^2)$, we have $T[B \cap N^+_T(A')] = T_1[B \cap N^+_{T_2}(A')]$ and therefore $\chi(T[B \cap N^+_T(A')]) \leq \chi(T_1[N^+_{T_2}(A')]) \leq 2$. Similarly, for all $B' \subseteq B$ of size at most $(\log n)/(32\chi^2)$, we get $\chi(T[A \cap N^+_T(B')]) \leq 2$. □

4.2 FRACTIONAL CHROMATIC NUMBER

In this section, we demonstrate that the modified version of Conjecture 4.1 with chromatic number replaced by fractional chromatic number is true, as observed by Scott and Seymour [SS16] without proof. The *fractional chromatic number* $\chi_f(G)$ of a graph G is the smallest real number k for which we can choose a random independent set I of G in such a way that $\mathbb{P}(v \in I) \geq 1/k$ for all $v \in V(G)$.

Theorem 4.9. *For every $k \geq 1$, every graph G with $\chi_f(G) > 2(k+1)$, and every tournament T on the same vertex set, there is a vertex v such that $\chi_f(G[N^+_T(v)]) > k$.*

To prove this result, we need the following fact about the existence of a probability distribution on the vertices of a tournament that assigns large probability to each closed in-neighbourhood.[4] This lemma follows from the linear programming duality (see [FR95, Section 1.2]).

Lemma 4.10. *Let T be a tournament. Then, there exists a probability distribution p on $V(T)$ such that for every vertex $v \in V(T)$, we have $p(N^-_T[v]) \geq 1/2$.*

Proof of Theorem 4.9. Suppose for each vertex v we have $\chi_f(G[N^+_T(v)]) \leq k$, so there is a random independent set I_v of $G[N^+_T(v)]$ such that $\mathbb{P}(u \in I_v) \geq 1/k$ for every $u \in N^+_T(v)$. By Lemma 4.10, there exists a probability distribution p on $V(T)$ such that $p(N^-_T[u]) \geq 1/2$ for all $u \in V(G)$.

[4] The *closed out-neighbourhood* of a vertex v in a tournament T is $N^+_T[v] := \{v\} \cup N^+_T(v)$. The *closed in-neighbourhood* is $N^-_T[v] := \{v\} \cup N^-_T(v)$ where $N^-_T(v) := \{u \in V(T) : (u, v) \in E(T)\}$.

We sample a random independent set I of G as follows. First, choose a random vertex $v \in V(G)$ according to p. Then, with probability $1/(k+1)$ let $I := \{v\}$, and otherwise with the remaining probability $k/(k+1)$ let $I := I_v$. Note that, for any vertex $u \in V(G)$, if $v \in N_T^-[u]$, then $u \in I$ with probability at least $1/(k+1)$. Hence, by the choice of p, $\mathbb{P}(u \in I) \geq 1/(2k+2)$ and so $\chi_f(G) \leq 2(k+1)$. □

4.3 DEGENERACY

Finally, we consider the setting in which degeneracy replaces chromatic number. We first show that there is a tournament on the vertex set of the d-dimensional hypercube such that each out-neighbourhood induces a forest in the hypercube, proving Proposition 4.5. This shows that having high degeneracy does not imply that some out-neighbourhood has high degeneracy.

Proof of Proposition 4.5. Let Q_d be the hypercube on $\{0,1\}^d$ where two vertices are adjacent if they differ in only one coordinate. We claim that there is a tournament T_d on the vertex set of Q_d such that the closed in- and out-neighbourhoods $Q_d[N_{T_d}^-[v]]$ and $Q_d[N_{T_d}^+[v]]$ are both forests for every vertex $v \in V(Q_d)$. Note that this implies the desired result. We prove the claim by induction on d. It clearly holds for $d = 1$. Suppose now that we have proved the claim for $d \geq 1$. We will view Q_{d+1} as the union of two copies of Q_d, namely $Q_d \times \{0\}$ and $Q_d \times \{1\}$, connected via the matching consisting of all edges of the form $(v,0)(v,1)$, where $(v,0) \in V(Q_d \times \{0\})$ and $(v,1) \in V(Q_d \times \{1\})$ denote the copies of $v \in V(Q_d)$.

Define a tournament T_{d+1} on $V(Q_{d+1})$ as follows. First, orient the edges within each of $V(Q_d \times \{0\})$ and $V(Q_d \times \{1\})$ according to T_d, in the canonical way. Then, for each $v \in V(Q_d)$, orient every edge between $(v,0)$ and $N_{T_d}^-[v] \times \{1\}$ away from $(v,0)$ and every edge between $(v,0)$ and $N_{T_d}^+(v) \times \{1\}$ towards $(v,0)$. Observe that this orients the edges between $(v,1)$ and $N_{T_d}^-(v) \times \{0\}$ away from $(v,1)$ and the edges between $(v,1)$ and $N_{T_d}^+[v] \times \{0\}$ towards $(v,1)$.

Let $v \in V(Q_d)$ and note that $N_{T_{d+1}}^+[(v,0)] = (N_{T_d}^+[v] \times \{0\}) \cup (N_{T_d}^-[v] \times \{1\})$. By the induction hypothesis, $N_{T_d}^+[v]$ and $N_{T_d}^-[v]$ both induce forests in Q_d, and so $N_{T_d}^+[v] \times \{0\}$ and $N_{T_d}^-[v] \times \{1\}$ do the same in Q_{d+1}. Since there is exactly one edge in Q_{d+1} between these two sets, namely the edge $(v,0)(v,1)$, it follows that $Q_{d+1}[N_{T_{d+1}}^+[(v,0)]]$ is acyclic. Similar arguments show that $Q_{d+1}[N_{T_{d+1}}^-[(v,0)]]$, $Q_{d+1}[N_{T_{d+1}}^+[(v,1)]]$, and $Q_{d+1}[N_{T_{d+1}}^-[(v,1)]]$ are all acyclic as well. This completes the induction step, and so this proves the claim. □

However, in contrast to the setting with chromatic number, we will show that high degeneracy implies that there exist two vertices u and v such that the out-neighbourhood of $\{u, v\}$ has high degeneracy.

Proof of Theorem 4.6. Since G has degeneracy at least $12d$, it contains a subgraph G' with minimum degree $\delta(G') \geq 12d$. Consider a partition $V(G') = A \cup B$ that maximises the number of edges of G' that cross between A and B. The bipartite subgraph H of G' that consists of exactly those edges satisfies $\delta(H) \geq 6d$. Without loss of generality, we may assume that $|A| \geq |B|$.

Define $T_1 := T[A]$ and $T_2 := T[B]$. Let $u \in A$ be such that $|N^+_{T_1}[u]| \geq |A|/2$, and define $A' := N^+_{T_1}[u]$ and $H_1 := H[A' \cup B]$. By Lemma 4.10, there exists a probability distribution p on $V(T_2)$ such that $p(N^-_{T_2}[w]) \geq 1/2$ for all $w \in B$. Let $v \in B$ be a random vertex chosen according to p, and so $\mathbb{P}(w \in N^+_{T_2}[v]) \geq 1/2$ for every $w \in B$. If $H_2 := H_1[A' \cup N^+_{T_2}[v]]$, this implies that $\mathbb{P}(e \in E(H_2)) \geq 1/2$ for every $e \in E(H_1)$ and so $\mathbb{E}(e(H_2)) \geq e(H_1)/2 \geq 3d|A'| \geq d(|A'| + |B|)$. Since $|N^+_{T_2}[v]| \leq |B|$, it follows that there exists some $v \in B$ such that $e(H_2) \geq d|V(H_2)|$.

Finally, if $H_3 := H_2 \setminus \{u, v\}$, then $e(H_3) \geq e(H_2) - 2|V(H_2)| \geq (d-2)|V(H_3)|$, and so H_3 has degeneracy larger than $d-2$. Note that $H_3 \subseteq G[N^+_T(\{u, v\})]$. This implies that $G[N^+_T(\{u, v\})]$ has degeneracy larger than $d-2$. □

4.4 OPEN PROBLEMS

From Theorem 4.2 and Proposition 4.5 we know that high chromatic number does not force an out-neighbourhood with high chromatic number, and high degeneracy does not force an out-neighbourhood with high degeneracy. It would be interesting to know whether high chromatic number forces an out-neighbourhood with high degeneracy.

Problem 4.11. *Does there exist, for each integer d, an integer χ such that for every graph G with $\chi(G) \geq \chi$ and every tournament T on the same vertex set, there is a vertex v such that $G[N^+_T(v)]$ has degeneracy at least d?*

We suspect that this is true at least for $d = 2$. That is, it should be possible to force some out-neighbourhood to contain a cycle.

Conjecture 4.12. *There exists an integer χ such that for every graph G with $\chi(G) \geq \chi$ and every tournament T on the same vertex set, there is a vertex v such that $G[N^+_T(v)]$ contains a cycle.*

Also note that Theorem 4.2 only shows that for certain very structured tournaments T there are graphs on the same vertex set with large chromatic number in which every out-neighbourhood of T induces a bipartite subgraph. We conjecture that (with high probability) we cannot replace T with a random tournament.

Conjecture 4.13. *For every positive integer k, there exists a χ such that if T is the uniformly random tournament on n vertices, then with high probability (as $n \to \infty$), for every graph G on $V(T)$ with $\chi(G) \geq \chi$ there is a vertex v such that $\chi(G[N_T^+(v)]) \geq k$.*

Finally, by the remark after the statement of Theorem 4.2, if $\chi(G) \geq \chi$, then there is a collection of at most $\lceil (\log_2 n)/\lfloor \chi/2 - 1 \rfloor \rceil$ out-neighbourhoods whose union induces a subgraph of chromatic number at least 3. It would be interesting to know if $o((\log n)/\chi)$ (as $\chi \to \infty$) out-neighbourhoods suffice here. In particular, we conjecture the following.

Conjecture 4.14. *There exists a function $f(n)$ satisfying $f(n) \leq o(\log n)$ such that for every graph G with n vertices and $\chi(G) \geq f(n)$ and every tournament T on the same vertex set, there is a vertex v such that $\chi(G[N_T^+(v)]) \geq 3$.*

Chapter 5

Multicolour Ramsey numbers of odd cycles

The *k-colour Ramsey number* $R_k(H)$ of a graph H is the smallest integer n such that every k-edge-colouring of the complete graph K_n contains a monochromatic copy of H. Multicolour Ramsey numbers have been studied extensively. If $H = K_2$, we clearly have $R_k(K_2) = 2$. The next case is when H has three vertices, so H is either the path P_3 or the triangle C_3. For P_3, it is relatively easy to see that

$$R_k(P_3) = \begin{cases} k+1 & \text{if } k \text{ is even, and} \\ k+2 & \text{if } k \text{ is odd.} \end{cases}$$

In contrast, for C_3, it is a wide-open problem to determine even just the asymptotic growth rate of $R_k(C_3)$. While Erdős [CG98] conjectured that $R_k(C_3) = 2^{\Theta(k)}$, the current best bounds are only

$$\Omega(3.28^k) \leq R_k(C_3) \leq \mathcal{O}(k!).$$

The lower bound was proved by Ageron, Casteras, Pellerin, Portella, Rimmel, and Tomasik [ACP+21], while the upper bound dates back to the work of Schur [Sch17] from 1917. Since then, this upper bound has remained unchanged up to the constant factor [XXC02, Eli20].

What about longer paths and cycles? Let P_ℓ and C_ℓ be the path and cycle on ℓ vertices, respectively. For paths of any length as well as cycles of large even length, it is known that their multicolour Ramsey numbers are linear in k and ℓ.

More precisely, for a fixed k, the current best bounds are

$$(k-1-o(1))2\ell \leq R_k(P_{2\ell}) \leq R_k(C_{2\ell}) \leq \left(k-\frac{1}{2}+o(1)\right)2\ell.$$

The lower bound follows from a construction of Sun, Yang, Xu, and Li [SYXL06] while the upper bound was shown by Knierim and Su [KS19]. When ℓ is fixed and k is large, we still have $R_k(P_{2\ell}) \leq 2k\ell$, but for even cycles the upper bound increases to $R_k(C_{2\ell}) \leq c \cdot k^{\ell/(\ell-1)}$ for some constant c that only depends on ℓ. This upper bound is known to be tight when $\ell \in \{2,3,5\}$ [LL09].

For cycles of odd length, the behaviour is much closer to that of a triangle. Bondy and Erdős [BE73] and Erdős and Graham [EG75] proved that

$$\ell \cdot 2^k + 1 \leq R_k(C_{2\ell+1}) \leq 2\ell \cdot (k+2)!.$$

For a fixed k, Jenssen and Skokan [JS21] showed that the lower bound is sharp, namely that $R_k(C_{2\ell+1}) = \ell \cdot 2^k + 1$ for all sufficiently large ℓ. However, this is not true if ℓ is fixed. Indeed, Day and Johnson [DJ17] showed that for all ℓ there exists some $\delta := \delta(\ell) > 0$ so that $R_k(C_{2\ell+1}) \geq 2\ell \cdot (2+\delta)^{k-1}$ for all sufficiently large k.

Usually, for large k, longer odd cycles should be easier to find than shorter odd cycles. For instance, Erdős and Graham [EG75] conjectured that

$$\lim_{k\to\infty} \frac{R_k(C_{2\ell+1})}{R_k(C_3)} = 0$$

for all $\ell \geq 2$. In a similar direction, Fox [Fox] conjectured that the upper bound on $R_k(C_{2\ell+1})$ should be significantly lower for longer odd cycles.

Conjecture 5.1 (Fox). *For every $\varepsilon > 0$ there exists some ℓ such that $R_k(C_{2\ell+1}) \leq k^{\varepsilon k}$ for all sufficiently large k.*

Li [Li09] also made the stronger conjecture that $R_k(C_{2\ell+1}) \leq o(k!^{1/\ell})$ as $k \to \infty$. Nevertheless, the gap between these conjectures and the best upper bound known remained large. There was only limited progress, first by Li [Li09] who proved that $R_k(C_5) \leq c^k \cdot \sqrt{k!}$ for some constant c, and then by Lin and Chen [LC19] who extended this result to longer odd cycles by showing that for $\ell \geq 2$ we have $R_k(C_{2\ell+1}) \leq c^k \cdot \sqrt{k!}$ for some constant c that only depends on ℓ.[1] Only conditional on the unproved and restrictive assumption that each Ramsey graph

[1] These results were proved independently by Fox [Fox], but were never published.

for $R_k(C_{2\ell+1})$ is nearly regular,[2] Li [Li09] showed that this upper bound could be improved to $R_k(C_{2\ell+1}) \leq c^k \cdot k!^{1/\ell}$.

In this chapter, we prove the conjecture of Fox by establishing the following explicit upper bound on $R_k(C_{2\ell+1})$.

Theorem 5.2. *For all positive integers k and ℓ,*

$$R_k(C_{2\ell+1}) \leq (4\ell-2)^k \cdot k^{k/\ell} + 1.$$

As $(k/e)^k < k!$, it also follows that $R_k(C_{2\ell+1}) \leq c^k \cdot k!^{1/\ell} + 1$ for $c := (4\ell-2)e^{1/\ell}$. This proves the upper bound of Li without the near-regularity assumption.

In addition to finding a monochromatic odd cycle of a specific length, there has also been considerable interest in finding any short monochromatic odd cycle. It is easy to see that there is a k-edge-colouring of K_{2^k} without any monochromatic odd cycles, but that every k-edge-colouring of K_{2^k+1} contains such a cycle. Motivated by this, in 1973, Erdős and Graham [EG75, Question (iii) in Section 6] asked for the smallest integer $L(k)$ such that every k-edge-colouring of K_{2^k+1} contains a monochromatic odd cycle of length at most $L(k)$.

Day and Johnson [DJ17] proved that $L(k) \geq 2^{\Omega(\sqrt{\log k})}$. Recently, Girão and Hunter [GH24] obtained the first non-trivial upper bound for this problem, showing that $L(k) \leq (2^k+1)/k^{1-o(1)}$. Using an algebraic approach, Janzer and Yip [JY26] improved this to $L(k) \leq \mathcal{O}(k^{3/2} \cdot 2^{k/2})$.

Both Girão and Hunter [GH24] and Janzer and Yip [JY26] also discussed the more general problem of finding short monochromatic odd cycles in colourings of complete graphs with more vertices. In the regime where the number of vertices is $(2+\delta)^k$ for some small $\delta > 0$, the methods of both papers guarantee a monochromatic odd cycle of length at most $\mathcal{O}_\delta(k)$. The dependence on δ is better in [JY26], where the length of the cycle is $\mathcal{O}(\delta^{-1/2} \cdot k)$ if $\delta < 1$. With our method, we can find significantly shorter monochromatic odd cycles, unless δ tends to 0 very quickly.

Theorem 5.3. *For every integer $k \geq 2$ and real number $b > 2$, every k-edge-colouring of K_n with $n > b^k$ contains a monochromatic odd cycle of length at most $2\lceil \log_{b/2} k \rceil + 1$.*

In particular, if $b = 2+\delta$ for $0 < \delta < 1$, this result guarantees a cycle of length $\mathcal{O}(\delta^{-1} \cdot \log k)$. In addition, the result also applies to larger values of b. If $b = k^\varepsilon$,

[2] That is, there is an absolute constant ε such that for all sufficiently large k and every k-edge-colouring of K_n with $n := R_k(C_{2\ell+1}) - 1$ that contains no monochromatic $C_{2\ell+1}$, every colour induces a subgraph of K_n whose minimum degree is at least an ε-fraction of its average degree.

this yields a result similar to Theorem 5.2, but without controlling the exact cycle length. We remark that Theorem 5.3 does not give any non-trivial bound for the Erdős–Graham problem, where $\delta \approx 1/(k \cdot 2^{k-1})$.

Chapter outline. In Section 5.1, we prove a key lemma, Lemma 5.4, which bounds the number of vertices in an edge-coloured complete graph where each monochromatic subgraph with small diameter has small chromatic number. This suffices to deduce our main results, Theorems 5.2 and 5.3, in Section 5.2.

Notation. Let $[n] := \{1, \dots, n\}$. For an edge-coloured graph, let $N_c^i(v)$ be the set of all vertices u such that the shortest path of colour c from v to u has length i, and write $N_c(v) := N_c^1(v)$ and $N_c^{\leq \ell}(v) := \bigcup_{i=0}^{\ell} N_c^i(v)$. Note that $N_c^0(v) = \{v\}$. For an uncoloured graph, $N^i(v)$ denotes the set of vertices at distance exactly i from v.

This chapter is based on joint work with Maria Axenovich, Wouter Cames van Batenburg, Oliver Janzer, and Mathieu Rundström [ACJ+26].

5.1 NEIGHBOURHOODS WITH SMALL CHROMATIC NUMBER

We prove Theorems 5.2 and 5.3 using the following key lemma. This lemma bounds the number of vertices of a k-edge-coloured complete graph where for every vertex v and every colour c, the subgraph of colour c induced by the neighbourhood $N_c^{\leq \ell}(v)$ has small chromatic number. This is the case if the colouring does not contain a monochromatic odd cycle of length $2\ell + 1$. In fact, we prove this lemma for *k-local-edge-colourings*, which are edge-colourings where at most k colours are incident to each vertex.

Lemma 5.4. *Let k, ℓ, and χ be positive integers. Consider a k-local-edge-colouring of a complete graph K_n such that for every vertex v and every colour c, the subgraph of colour c induced by $N_c^{\leq \ell}(v)$ in K_n has chromatic number at most χ. Then, $n \leq \chi^k \cdot k^{k/\ell}$.*

Proof. Let $G := K_n$. Define the weight of a vertex v as $w(v) := (\chi \cdot k^{1/\ell})^{-d_{\mathrm{col}}(v)}$, where $d_{\mathrm{col}}(v)$ denotes the number of colours incident to v, and define the weight of a subset of vertices $U \subseteq V(G)$ as $w(U) := \sum_{v \in U} w(v)$. Since $d_{\mathrm{col}}(v) \leq k$ for every vertex v, we get that $w(v) \geq \chi^{-k} \cdot k^{-k/\ell}$ and so $w(V(G)) \geq n \cdot \chi^{-k} \cdot k^{-k/\ell}$. To show that $n \leq \chi^k \cdot k^{k/\ell}$, it therefore suffices to prove that $w(V(G)) \leq 1$. We will prove this by induction on the number of vertices of G. This is trivial if $|V(G)| = 1$, so suppose that $|V(G)| \geq 2$.

Let $v \in V(G)$ be an arbitrary vertex. Since at most k colours are incident to v, there exists a colour c such that $w(N_c(v)) \geq w(V(G) \setminus \{v\})/k$. This implies that $w(N_c^{\leq 1}(v)) = w(\{v\} \cup N_c(v)) \geq w(V(G))/k$. We claim that there exists some $i \in [\ell]$ such that $w(N_c^{\leq i+1}(v)) \leq k^{1/\ell} \cdot w(N_c^{\leq i}(v))$. Indeed, otherwise we have $w(N_c^{\leq i+1}(v)) > k^{1/\ell} \cdot w(N_c^{\leq i}(v))$ for all $i \in [\ell]$ and so it follows that

$$w(N_c^{\leq \ell+1}(v)) > (k^{1/\ell})^\ell \cdot w(N_c^{\leq 1}(v)) = k \cdot w(N_c^{\leq 1}(v)) \geq w(V(G)),$$

which is a contradiction since $N_c^{\leq \ell+1}(v) \subseteq V(G)$.

Let $i \in [\ell]$ be given by this claim, so if $S := N_c^{\leq i}(v)$ and $T := N_c^{\leq i+1}(v)$, then $w(T) \leq k^{1/\ell} \cdot w(S)$. By assumption, the subgraph of colour c induced by S in G has chromatic number at most χ. Fix such a χ-vertex-colouring of S and let $S' \subseteq S$ be one of its colour classes with maximum weight. Then, S' spans no edge of colour c and satisfies $w(S') \geq w(S)/\chi$.

Delete all vertices of $T \setminus S'$ from G and update the weights of the remaining vertices. Note that the weight of every vertex in S' increases by a multiplicative factor of at least $\chi \cdot k^{1/\ell}$ since colour c is no longer incident to any of these vertices. On the other hand, the weight of every vertex in $V(G) \setminus T$ is either unchanged or increases. Therefore, the weight of the entire graph increases by at least

$$(\chi \cdot k^{1/\ell}) \cdot w(S') - w(T) \geq k^{1/\ell} \cdot w(S) - k^{1/\ell} \cdot w(S) = 0.$$

Since we know by the induction hypothesis that the weight of the new graph is at most 1, it follows that the weight of the original graph was at most 1. □

We remark that essentially the same proof also works if we replace chromatic number by fractional chromatic number (see Section 4.2 for a definition of fractional chromatic number).

5.2 CONSEQUENCES FOR SHORT MONOCHROMATIC ODD CYCLES

To deduce Theorems 5.2 and 5.3 for cycles of length at most $2\ell + 1$, it remains to bound the chromatic number of the subgraphs of colour c induced by the neighbourhoods $N_c^{\leq \ell}(v)$. This is very easy for Theorem 5.3 since none of the neighbourhoods $N_c^i(v)$ for $i \leq \ell$ can span an edge of colour c.

Proof of Theorem 5.3. Let $\ell := \lceil \log_{b/2} k \rceil$. Note that if a k-edge-colouring of K_n contains no monochromatic odd cycle of length at most $2\ell + 1$, then for every vertex v, every colour c, and every $i \in [\ell]$ we have that $N_c^i(v)$ spans no edge of colour c. This implies that the subgraph of colour c induced by $N_c^{\leq \ell}(v)$ in K_n is bipartite. By Lemma 5.4, it follows that $n \leq 2^k \cdot k^{k/\ell} \leq 2^k \cdot (b/2)^k = b^k$. □

To prove Theorem 5.2, we can employ a known bound on the chromatic number of the subgraph induced by $N^i(v)$ for $i \leq \ell$ in graphs that contain no cycle of length exactly $2\ell + 1$.

Proof of Theorem 5.2. An argument of Erdős, Faudree, Rousseau, and Schelp [EFRS78, Proof of Theorem 1] shows that if a graph G contains no cycle of length $2\ell + 1$, then for every vertex v and every $i \in [\ell]$ we have that $G[N^i(v)]$ has chromatic number at most $2\ell - 1$. So, if a k-edge-colouring of K_n contains no monochromatic cycle of length $2\ell + 1$, then for every vertex v and every colour c, the subgraph of colour c induced by $N_c^{\leq \ell}(v)$ in K_n has chromatic number at most $4\ell - 2$. By Lemma 5.4, it follows that $n \leq (4\ell - 2)^k \cdot k^{k/\ell}$. □

Chapter 6

Flashes and rainbows in tournaments

A classical theorem of Erdős and Szekeres [ES35] shows that any sequence of numbers $a_1, \dots, a_n$ with $n > \ell^2$ contains a monotone subsequence of length $\ell + 1$. This result can also be expressed in terms of edge-colourings of ordered complete graphs. Indeed, given such a sequence, construct the complete graph on the ordered integers from 1 to n and colour an edge uv where $u < v$ with red if $a_u \leq a_v$, and blue otherwise. A monotone subsequence of length $\ell + 1$ then corresponds to a monochromatic path with ℓ edges whose vertices appear in increasing order. We call such a path an *ℓ-flash*. Formally, this is a sequence of vertices $x_0 < \cdots < x_\ell$ such that the edges $x_0x_1, x_1x_2, \dots, x_{\ell-1}x_\ell$ all have the same colour.

In this language, the Erdős–Szekeres theorem follows from the more general fact that any colouring of the edges of a complete graph on $\{1, \dots, n\}$ with fewer than k colours contains an ℓ-flash whenever $n > \ell^{k-1}$. This bound is sharp: if $n = \ell^{k-1}$, we may label the vertices with strings from $\{1, 2, \dots, \ell\}^{k-1}$ in lexicographic order and colour an edge uv where $u < v$ by any coordinate at which the label of u is smaller than the label of v. This construction contains no ℓ-flash.

In 1992, Lefmann, Rödl, and Thomas [LRT92] asked whether this result still holds even if the number of colours is unbounded, provided we forbid the existence of a *k-rainbow*, which is a sequence of vertices $y_0 < \cdots < y_k$ such that the edges $y_0y_1, y_1y_2, \dots, y_{k-1}y_k$ all have distinct colours. Let $f(\ell, k)$ be the smallest integer such that if $n > f(\ell, k)$ then every colouring of the edges of the ordered complete graph on $\{1, \dots, n\}$ with an arbitrary number of colours contains an ℓ-flash or a k-rainbow. The existence of $f(\ell, k)$ follows, for example, from the canonical Ramsey

theorem of Erdős and Rado [ER50], whose current best bounds by Lefmann and Rödl [LR95] imply that $f(k,k) \leq k^{\mathcal{O}(k^2)}$.

The discussion above implies that $f(\ell,k) \geq \ell^{k-1}$ and that this is best possible if only $k-1$ colours are used. Lefmann, Rödl, and Thomas [LRT92] made the attractive conjecture that this lower bound is tight in general.

Conjecture 6.1 (Lefmann, Rödl, and Thomas)**.** *For all positive integers ℓ and k, we have $f(\ell,k) = \ell^{k-1}$.*

As evidence for their conjecture, they proved that it is true for $\ell \leq 2$ or $k \leq 4$ as well as when $\ell \geq (3k)^{2k}$.

The first main result in this chapter proves the conjecture of Lefmann, Rödl, and Thomas provided $\ell \geq k^3(\log k)^{1+o(1)}$.

Theorem 6.2. *For all $\varepsilon > 0$, there exists some C such that for all positive integers ℓ and k with $\ell \geq Ck^3(\log k)^{1+\varepsilon}$, we have $f(\ell,k) = \ell^{k-1}$.*

The best general upper bound on $f(\ell,k)$ that was previously known was given by Jiang and Mubayi [JM00] in 2000.

Theorem 6.3 (Jiang and Mubayi)**.** *For all integers $\ell \geq 1$ and $k \geq 4$,*

$$f(\ell,k) \leq \left(1 + \frac{1}{\sqrt{\ell}}\right)^{k-4} \cdot \ell^{k-1}.$$

This confirms Conjecture 6.1 asymptotically when $\ell \geq \omega(k^2)$. However, for ℓ fixed and k large, this differs from the lower bound by a factor exponential in k. We provide an upper bound in which this factor is only polynomial in k.

Theorem 6.4. *For all $\varepsilon > 0$, there exists some D such that for all integers $\ell \geq 1$ and $k \geq 2$, we have*

$$f(\ell,k) \leq Dk(\log k)^{1+\varepsilon} \cdot \ell^{k-1}.$$

Theorems 6.2 to 6.4 are all best in different ranges of ℓ and k. Namely, Theorem 6.2 is optimal for $\ell \geq k^3(\log k)^{1+o(1)}$, Theorem 6.4 gives the smallest upper bound known for $\ell \leq \mathcal{O}((k/\log k)^2)$, and Theorem 6.3 is best in the intermediate regime.

We further generalise flashes and rainbows to tournaments, where a *tournament* is an orientation of a complete graph. A *directed walk* in a tournament is a sequence of vertices $x_0, \ldots, x_\ell$ such that $x_{i-1}x_i$ is a directed edge for all $1 \leq i \leq \ell$. Note that walks are allowed to repeat vertices and edges.

Definition 6.5. Let ℓ and k be non-negative integers. An *ℓ-flash* in a tournament is a directed walk $x_0, \dots, x_\ell$ such that the edges $x_0x_1, \dots, x_{\ell-1}x_\ell$ all have the same colour. A *k-rainbow* is a directed walk $y_0, \dots, y_k$ such that the edges $y_0y_1, \dots, y_{k-1}y_k$ all have distinct colours.

Note that a directed monochromatic cycle yields an ℓ-flash for every ℓ. Let $t(\ell, k)$ denote the smallest integer such that every colouring of the edges of any tournament with more than $t(\ell, k)$ vertices contains an ℓ-flash or a k-rainbow. Conjecture 6.1 can now be viewed as the restriction of this problem to transitive tournaments, and so we have

$$t(\ell, k) \geq f(\ell, k) \geq \ell^{k-1}.$$

We conjecture that the tournament strengthening of Conjecture 6.1 holds.

Conjecture 6.6. *For all positive integers ℓ and k, we have $t(\ell, k) = \ell^{k-1}$.*

We prove two upper bounds on $t(\ell, k)$. The first, Theorem 6.7, is better when ℓ is large compared to k and confirms Conjecture 6.6 asymptotically for $\ell \geq \omega(k^2)$. Although the bound looks very similar to Theorem 6.3, the proof is quite different and uses a probabilistic argument.

Theorem 6.7. *For all positive integers ℓ and k, we have $t(\ell, k) \leq (1 + 2/\sqrt{\ell})^k \cdot \ell^{k-1}$.*

Our second upper bound, Theorem 6.8, is better when k is large compared to ℓ and, in particular, improves upon Theorem 6.7 for $\ell \leq \mathcal{O}((k/\log k)^2)$. Note that this bound directly implies Theorem 6.4.

Theorem 6.8. *For every $\varepsilon > 0$, there exists some D such that for all integers $\ell \geq 1$ and $k \geq 2$, we have*

$$t(\ell, k) \leq Dk(\log k)^{1+\varepsilon} \cdot \ell^{k-1}.$$

Conjecture 6.6 posits that if one wants to colour the edges of a large tournament while avoiding ℓ-flashes and k-rainbows, then transitive tournaments are best in the sense of having the most vertices. We observe that there are tournaments that are strictly worse than transitive tournaments.

Proposition 6.9. *For all integers $\ell \geq 2$ and $k \geq 3$, there is a tournament on $f(\ell, k) - 1$ vertices such that every colouring of its edges contains an ℓ-flash or a k-rainbow.*

Finally, it is interesting to ask how bad a tournament can be: how few vertices can a tournament have if every colouring of its edges contains an ℓ-flash or a k-rainbow? We provide a simple argument showing that the number of vertices in such a tournament must be within a factor $\ell\sqrt{k}$ of ℓ^{k-1}.

Proposition 6.10. *There exists some $C > 0$ such that every tournament with at most $C\ell^{k-2}/\sqrt{k}$ vertices has an edge colouring with neither ℓ-flashes nor k-rainbows.*

Chapter outline. In Section 6.1, we consider arbitrary tournaments and prove Theorems 6.7 and 6.8. We then specialise to transitive tournaments in Section 6.2 and provide a proof of Theorem 6.2. In Section 6.3 we explore the differences between these two settings, proving Propositions 6.9 and 6.10. We finish in Section 6.4 with some open problems.

Notation. Let $[n] := \{1, \dots, n\}$. Also, throughout this chapter, ℓ and k are always positive integers.

This chapter is based on joint work with António Girão, Freddie Illingworth, Michael Savery, and Alex Scott [GIM+24].

6.1 UPPER BOUNDS FOR ARBITRARY TOURNAMENTS

In this section, we prove Theorems 6.7 and 6.8. For both, we rely on the following observation: if a tournament contains no k-rainbow and a $(k-1)$-rainbow ends at some vertex v, then v has at most $k-1$ outgoing colours. Indeed, otherwise any $(k-1)$-rainbow ending at v could be extended to a k-rainbow. This will allow us to restrict ourselves to tournaments where each vertex only has few outgoing (and incoming) colours. In this setting, we show that the following lemma holds.

Lemma 6.11. *Let T be an edge-coloured tournament containing no ℓ-flash, and assume that every vertex has at most c incoming colours and at most c outgoing colours. Then,*

$$|V(T)| \leq \left(1 + \frac{2}{\sqrt{\ell}}\right)^c \cdot \ell^c.$$

We prove this result with a probabilistic argument. We will use the following definitions throughout this chapter. For any edge $e \in E(T)$, let $c(e)$ be the colour of e. For any vertex v, we write $C^-(v) := \{c(uv) : uv \in E(T)\}$ for the set of incoming colours of v, $C^+(v) := \{c(vu) : vu \in E(T)\}$ for the set of outgoing colours of v, and $C(v) := C^-(v) \cup C^+(v)$ for the set of colours incident to v. Moreover, for every vertex v and every colour a, let

$$\ell_a(v) := \begin{cases} \ell - 1 & \text{if } a \in C(v) \setminus C^+(v), \\ \text{length of a longest flash of colour } a \text{ ending at } v & \text{otherwise.} \end{cases}$$

This is defined for every colour a in our colouring; however, $\ell_a(v)$ is only non-zero for incoming colours of v. Note also that if T has no ℓ-flash and an edge uv has colour a, then $\ell_a(u) < \ell_a(v)$. The first condition in this definition is only needed for convenience in the following proof.

Proof of Lemma 6.11. Let p be the positive root of $(\ell-2)p^2 + 2p = 1$, and so $p = 1/(1+\sqrt{\ell-1})$. For every colour a, choose $\ell_a \in \{0,\dots,\ell-1\}$ independently at random such that $\mathbb{P}(\ell_a = 0) = \mathbb{P}(\ell_a = \ell-1) = p$ and $\mathbb{P}(\ell_a = i) = p^2$ for every $i \in [\ell-2]$. Define

$$U := \{v \in V(T) : \ell_a(v) = \ell_a \text{ for all } a \in C(v)\}.$$

If $T[U]$ contains an edge uv, say of colour a, then $a \in C(u) \cap C(v)$ and we have $\ell_a(u) < \ell_a(v)$, contradicting the definition of U. Thus, $|U| \leq 1$. Moreover, for every vertex v and every colour $a \in C(v)$ we have $\ell_a(v) \in \{0, \ell-1\}$ if and only if $a \in C^-(v) \triangle C^+(v)$, and so

$$\mathbb{P}(v \in U) = p^{|C^-(v)\triangle C^+(v)|} p^{2|C^-(v)\cap C^+(v)|} = p^{|C^-(v)|+|C^+(v)|} \geq p^{2c}.$$

It follows that the expected size of U is at least $p^{2c}|V(T)|$. Since $|U| \leq 1$, we get

$$|V(T)| \leq p^{-2c} = \left(1+\sqrt{\ell-1}\right)^{2c} = \left(\ell + 2\sqrt{\ell-1}\right)^c \leq \left(1+\frac{2}{\sqrt{\ell}}\right)^c \cdot \ell^c. \qquad \square$$

To obtain Theorem 6.7, we will use induction on k to bound the number of vertices in T where no $(k-1)$-rainbow starts and no $(k-1)$-rainbow ends. By the observation from the beginning of this section, the number of remaining vertices can then be bounded via the preceding lemma.

Proof of Theorem 6.7. We claim that $t(\ell,k) \leq 2 \cdot t(\ell,k-1) + (\ell + 2\sqrt{\ell})^{k-1}$ for all $k \geq 2$. Indeed, let T be a tournament containing no ℓ-flash and no k-rainbow. Define

$$R := \{v \in V(T) : \text{no } (k-1)\text{-rainbow ends at } v\} \text{ and}$$
$$S := \{v \in V(T) : \text{no } (k-1)\text{-rainbow starts at } v\}.$$

Clearly $|R|, |S| \leq t(\ell,k-1)$. Every vertex $v \in V(T) \setminus R$ has at most $k-1$ outgoing colours (or else any $(k-1)$-rainbow ending at v could be extended to a k-rainbow), and similarly every vertex in $V(T) \setminus S$ has at most $k-1$ incoming

colours. By Lemma 6.11, it follows that $|V(T) \setminus (R \cup S)| \leq (1 + 2/\sqrt{\ell})^{k-1} \cdot \ell^{k-1}$ and therefore

$$|V(T)| \leq |R| + |S| + |V(T) \setminus (R \cup S)| \leq 2 \cdot t(\ell, k-1) + (\ell + 2\sqrt{\ell})^{k-1}.$$

This proves the claim. A straightforward induction on k with base case $t(\ell, 1) = 1$ then yields

$$\begin{aligned} t(\ell,k) &\leq \sum_{i=0}^{k-1} \left(\frac{2}{\ell + 2\sqrt{\ell}}\right)^i (\ell + 2\sqrt{\ell})^{k-1} \leq \frac{1}{1 - \frac{2}{\ell+2\sqrt{\ell}}} (\ell + 2\sqrt{\ell})^{k-1} \\ &= \frac{\ell + 2\sqrt{\ell}}{\ell + 2\sqrt{\ell} - 2} \left(1 + \frac{2}{\sqrt{\ell}}\right)^{k-1} \cdot \ell^{k-1} \leq \left(1 + \frac{2}{\sqrt{\ell}}\right)^k \cdot \ell^{k-1}. \end{aligned}$$ □

Next, we prove Theorem 6.8. To this end, we will consider so-called "robust" vertices, which were introduced by Lefmann, Rödl, and Thomas [LRT92] under a different name. Loosely, we call a vertex v robust if we can find long rainbows ending or starting at v that avoid any given colour.

Definition 6.12. Let r be a non-negative integer. A vertex v is *r-in-robust* if, for every colour a, some r-rainbow without colour a ends at v.

In-robust vertices are very useful for constructing long rainbows. Indeed, assume that u is an in-robust vertex with an edge of colour a to some other vertex v. Then, we can pick a long rainbow without colour a ending at u and try to connect this rainbow to some rainbow starting at v. Since the edge uv cannot interfere with the rainbow ending at u, choosing an appropriate rainbow starting at v might result in a long rainbow.

Of course, in-robust vertices are only useful if they are forced to occur in large enough tournaments. As above, it will suffice to consider tournaments with few incoming and outgoing colours at each vertex. Therefore, let $g(\ell, k, r)$ denote the smallest integer such that if T is a tournament with more than $g(\ell, k, r)$ vertices, then every edge colouring of T with at most $k-1$ outgoing colours at each vertex contains an ℓ-flash, a k-rainbow, or an r-in-robust vertex.

It turns out that $g(\ell, k, r)$ is not much larger than $t(\ell, r)$. Therefore, in large tournaments we will be able to find r-in-robust vertices for large r, which we can then use to find k-rainbows as outlined above. This strategy yields the following bound on $t(\ell, k)$.

Lemma 6.13. *For all $k \geq 2$ and $1 \leq r < k$,*

$$t(\ell, k) \leq t(\ell, k-1) + g(\ell, k, r-1) + 1 + 2(k-1)\ell^r \cdot t(\ell, k-r).$$

Proof. Let T be an edge-coloured tournament with no ℓ-flash or k-rainbow. Let

$$R := \{v \in V(T) : \text{no } (k-1)\text{-rainbow ends at } v\}.$$

Then $|R| \leq t(\ell, k-1)$, and every vertex in $V(T) \setminus R$ has at most $k-1$ outgoing colours.

Order the vertices in $V(T) \setminus R$ in a way that maximises the number of forward edges of $T[V(T) \setminus R]$. In particular, every vertex in $V(T) \setminus R$ has outgoing edges to at least half of the vertices following it. Let $P \subseteq V(T) \setminus R$ be a minimal initial segment of this ordering such that $T[P]$ contains an $(r-1)$-in-robust vertex v (if no such vertex exists, then we already have $|V(T)| \leq |R| + |V(T) \setminus R| \leq t(\ell, k-1) + g(\ell, k, r-1)$). Note that $|P| \leq g(\ell, k, r-1) + 1$. Since v has outgoing edges to at least half of the vertices following it, we get that its out-neighbourhood $U := \{u \in V(T) : vu \in E(T)\}$ satisfies $|V(T) \setminus (R \cup P)| \leq 2|U|$.

For every colour $a \in C^+(v)$, let $U_a := \{u \in U : c(vu) = a\}$. Since v is $(r-1)$-in-robust, we can choose an $(r-1)$-rainbow W_a without colour a that ends at v, and extend it to each of the vertices in U_a by the edges of colour a. Let C_a be the set containing a and the $r-1$ colours appearing in W_a.

Next, for every $m \in \{0, \dots, \ell-1\}^{C_a}$ define

$$U_a^m := \{u \in U_a : \ell_b(u) = m_b \text{ for all } b \in C_a\}.$$

If the colour of an edge uw in $T[U_a^m]$ is b, then $\ell_b(u) < \ell_b(w)$, and so $b \notin C_a$. Thus, if $T[U_a^m]$ contained a $(k-r)$-rainbow starting at some vertex u, then this rainbow would use no colour from C_a and could therefore be extended by the r-rainbow ending at u to a k-rainbow. It follows that $T[U_a^m]$ contains no $(k-r)$-rainbow, so $|U_a^m| \leq t(\ell, k-r)$ and

$$|U_a| = \sum_{m \in \{0,\dots,\ell-1\}^{C_a}} |U_a^m| \leq \sum_{m \in \{0,\dots,\ell-1\}^{C_a}} t(\ell, k-r) \leq \ell^r \cdot t(\ell, k-r).$$

Since v has at most $k-1$ outgoing colours, this implies that

$$|U| = \sum_{a \in C^+(v)} |U_a| \leq \sum_{a \in C^+(v)} \ell^r \cdot t(\ell, k-r) \leq (k-1)\ell^r \cdot t(\ell, k-r).$$

Putting everything together, we get

$$\begin{aligned}|V(T)| &= |R| + |P| + |V(T) \setminus (R \cup P)| \\ &\leq |R| + |P| + 2|U| \\ &\leq t(\ell, k-1) + g(\ell, k, r-1) + 1 + 2(k-1)\ell^r \cdot t(\ell, k-r). \qquad \square\end{aligned}$$

To apply this lemma, we need a good bound on $g(\ell, k, r)$. Fortunately, we can use a strategy similar to that in the preceding proof to obtain the following result.

Lemma 6.14. *For all $r \geq 1$,*

$$g(\ell, k, r) \leq g(\ell, k, r-1) + 1 + 2(k-1)r\ell \cdot t(\ell, r).$$

Proof. Let T be an edge-coloured tournament with no ℓ-flash, no k-rainbow, no r-in-robust vertex, and at most $k-1$ outgoing colours at each vertex. Order the vertices of T in a way that maximises the number of forward edges, and let $P \subseteq V(T)$ be a minimal initial segment of this ordering such that $T[P]$ contains an $(r-1)$-in-robust vertex v. As before, $|P| \leq g(\ell, k, r-1) + 1$ and the out-neighbourhood $U := \{u \in V(T) : vu \in E(T)\}$ of v satisfies $|V(T) \setminus P| \leq 2|U|$. For each $a \in C^+(v)$, define U_a, W_a, and C_a as in the previous proof.

Next, for every colour $b \in C_a$, define

$$U_{a,b} := \{u \in U_a : \text{all } r\text{-rainbows ending at } u \text{ contain colour } b\}.$$

By assumption, no vertex is r-in-robust, but for each $u \in U_a$ there is an r-rainbow using only the colours of C_a ending at u. It follows that $U_a = \bigcup_{b \in C_a} U_{a,b}$. Finally, for every $m \in \{0, \dots, \ell-1\}$, let

$$U_{a,b}^m := \{u \in U_{a,b} : \ell_b(u) = m\}.$$

Now, if an edge uw in $T[U_{a,b}^m]$ has colour b, then $\ell_b(u) < \ell_b(w)$, which gives a contradiction. Thus, if $T[U_{a,b}^m]$ contained an r-rainbow ending at some vertex u, then this rainbow would not use the colour b, contradicting $u \in U_{a,b}$. It follows that $T[U_{a,b}^m]$ contains no r-rainbow, so $|U_{a,b}^m| \leq t(\ell, r)$ and

$$|U_a| \leq \sum_{b \in C_a} |U_{a,b}| = \sum_{b \in C_a} \sum_{m \in \{0,\dots,\ell-1\}} |U_{a,b}^m| \leq \sum_{b \in C_a} \sum_{m \in \{0,\dots,\ell-1\}} t(\ell, r) = r\ell \cdot t(\ell, r).$$

Since v has at most $k-1$ outgoing colours, this implies $|U| \leq (k-1)r\ell \cdot t(\ell, r)$,

and so

$$\begin{aligned}|V(T)| &= |P| + |V(T) \setminus P| \\ &\leq |P| + 2|U| \\ &\leq g(\ell, k, r-1) + 1 + 2(k-1)r\ell \cdot t(\ell, r).\end{aligned}$$ □

Finally, to obtain good bounds on $t(\ell, k)$ and $g(\ell, k, r)$, we apply these recursive bounds repeatedly. This is handled by the following corollary, which proves Theorem 6.8 and gives an upper bound on $g(\ell, k, r)$ that will be useful in Section 6.2.

Corollary 6.15. *For every $\varepsilon > 0$, there exists some D such that for $k \geq 2$ and $0 \leq r < k$,*

$$t(\ell, k) \leq Dk(\log k)^{1+\varepsilon} \cdot \ell^{k-1} \quad \text{and} \quad g(\ell, k, r) \leq 4Dk^3(\log k)^{1+\varepsilon} \cdot \ell^r.$$

Proof. From Lemma 6.13 and Lemma 6.14, we can obtain the inequalities

$$\begin{aligned}t(\ell, k) &\leq t(\ell, k-1) + g(\ell, k, r-1) + 2k\ell^r \cdot t(\ell, k-r) \text{ and} \\ g(\ell, k, r) &\leq g(\ell, k, r-1) + 2k^2\ell \cdot t(\ell, r)\end{aligned}$$

for $k \geq 2$ and $1 \leq r < k$.

Let $k_0 \geq 2$ be such that $8(3\log k + 5)(\log(3\log k + 5))^{1+\varepsilon} \leq (\log k)^{1+\varepsilon}$ for all $k \geq k_0$. By Theorem 6.7, there is some D such that $t(\ell, k) \leq Dk(\log k)^{1+\varepsilon} \cdot \ell^{k-1}$ for all $2 \leq k < k_0$.

We prove the corollary by induction on k. Assume that we have already proved it for all $2 \leq k' < k$. We may also assume that $\ell \geq 2$ since $t(1, k) = g(1, k, r) = 1$. Then, for $0 \leq r < k$, a simple induction on r with base case $g(\ell, k, 0) = 0$ yields

$$\begin{aligned}g(\ell, k, r) &\leq g(\ell, k, r-1) + 2k^2\ell \cdot t(\ell, r) \\ &\leq 4Dk^3(\log k)^{1+\varepsilon} \cdot \ell^{r-1} + 2Dk^3(\log k)^{1+\varepsilon} \cdot \ell^r \overset{\ell \geq 2}{\leq} 4Dk^3(\log k)^{1+\varepsilon} \cdot \ell^r.\end{aligned}$$

To bound $t(\ell, k)$, we may assume that $k \geq k_0$ by our choice of D. Note that

$$t(\ell, k-1) \leq Dk(\log k)^{1+\varepsilon} \cdot \ell^{k-2} \overset{\ell \geq 2}{\leq} \frac{Dk(\log k)^{1+\varepsilon} \cdot \ell^{k-1}}{2}.$$

Choose

$$r := \left\lfloor k - \frac{2\log k + \log 16}{\log \ell} \right\rfloor,$$

then

$$g(\ell,k,r-1) \leq 4Dk^3(\log k)^{1+\varepsilon}\cdot \ell^{k-1-\frac{2\log k+\log 16}{\log \ell}} = \frac{Dk(\log k)^{1+\varepsilon}\cdot \ell^{k-1}}{4}.$$

Also,

$$r \geq k - \frac{2\log k+\log 16}{\log \ell} - 1 \geq k - \frac{2\log k+\log 16}{\log 2} - 1 \geq k - (3\log k + 5)$$

and so, using the fact that $k \geq k_0$, we have

$$8(k-r)(\log(k-r))^{1+\varepsilon} \leq 8(3\log k+5)(\log(3\log k+5))^{1+\varepsilon} \leq (\log k)^{1+\varepsilon}.$$

It follows by induction that

$$2k\ell^r \cdot t(\ell,k-r) \leq 2kD(k-r)(\log(k-r))^{1+\varepsilon}\cdot \ell^{k-1} \leq \frac{Dk(\log k)^{1+\varepsilon}\cdot \ell^{k-1}}{4}.$$

Combining, we obtain

$$\begin{aligned} t(\ell,k) &\leq t(\ell,k-1) + g(\ell,k,r-1) + 2k\ell^r\cdot t(\ell,k-r) \\ &\leq \left(\frac{1}{2}+\frac{1}{4}+\frac{1}{4}\right) Dk(\log k)^{1+\varepsilon}\cdot \ell^{k-1} \\ &= Dk(\log k)^{1+\varepsilon}\cdot \ell^{k-1}. \end{aligned}$$

□

6.2 UPPER BOUNDS FOR TRANSITIVE TOURNAMENTS

In this section we prove Theorem 6.2, that is, we show that $f(\ell,k) = \ell^{k-1}$ if ℓ is sufficiently large in terms of k. We will follow a similar strategy to that used by Lefmann, Rödl, and Thomas [LRT92] when they showed this for $\ell \geq (3k)^{2k}$, but our bounds and arguments from the previous section will allow us to improve the dependence on k.

The idea of the proof is to find a *strongly $(k-2)$-robust vertex*, that is, a vertex v such that there exists a set C of $k-1$ colours which contains all colours incident to v and which has the property that for all $a \in C$, there is a $(k-2)$-rainbow with colour set $C \setminus \{a\}$ that ends at v, and also one that starts at v. The existence of such a vertex will allow us to show that every 2-flash in the entire tournament takes its colour from C, which will in turn allow us to reduce the problem to the

$\ell = 2$ case, for which Lefmann, Rödl, and Thomas [LRT92] have already shown that the result holds.

Theorem 6.16 (Lefmann, Rödl, and Thomas)**.** *For all positive integers k, we have $f(2,k) = 2^{k-1}$.*

This final ingredient, Theorem 6.16, is the only part of the proof where we need our tournament to be transitive: our arguments will show that if $t(2,k) = 2^{k-1}$ for all k, then Theorem 6.2 holds with $t(\ell,k)$ in place of $f(\ell,k)$. For this reason, we will state and prove most of the results in this section for arbitrary tournaments.

We begin by explaining how the fact that few colours appear in 2-flashes allows us to reduce the problem to the $\ell = 2$ case. Write $C^{m\text{-flash}}(v)$ for the set of colours a such that there exists an m-flash of colour a containing v.

Lemma 6.17. *Let $1 \leq m \leq \ell$, let T be an edge-coloured tournament containing no ℓ-flash or k-rainbow, and suppose that $|C^{m\text{-flash}}(v)| \leq c$ for every vertex v. Then,*

$$|V(T)| \leq t(m,k)\left(\frac{\ell}{m}\right)^c.$$

We will prove this result with a probabilistic argument similar to that used to prove Lemma 6.11. Here, it suffices if $\ell_a(v)$ simply denotes the length of the longest flash of colour a ending at v.

Proof. For every colour a, choose a subset $L_a \subseteq \{0,\dots,\ell-1\}$ of size m uniformly and independently at random. Define

$$U := \Big\{v \in V(T) : \ell_a(v) \in L_a \text{ for all } a \in C^{m\text{-flash}}(v)\Big\},$$

and note that $T[U]$ contains no m-flash $v_0,\dots,v_m$ in any colour a. Indeed, otherwise $\ell_a(v_0) < \cdots < \ell_a(v_m)$ with $a \in C^{m\text{-flash}}(v_i)$ for every i, so L_a has size at least $m+1$, which is a contradiction. Hence, $|U| \leq t(m,k)$. On the other hand, for every vertex v we have $\mathbb{P}(v \in U) = (m/\ell)^{|C^{m\text{-flash}}(v)|} \geq (m/\ell)^c$, and so

$$|V(T)|\left(\frac{m}{\ell}\right)^c \leq \sum_{v \in V(T)} \mathbb{P}(v \in U) = \mathbb{E}(|U|) \leq t(m,k),$$

which implies that $|V(T)| \leq t(m,k)(\ell/m)^c$. □

Next, we show how the existence of a strongly $(k-2)$-robust vertex implies that few colours occur in 2-flashes.

Lemma 6.18. *Let $k \geq 2$ and let T be an edge-coloured tournament containing no k-rainbow. Suppose that T contains a strongly $(k-2)$-robust vertex. Then at most $k-1$ colours occur in 2-flashes.*

Proof. Let u be a strongly $(k-2)$-robust vertex in T, and let C be the corresponding set of $k-1$ colours. Suppose for a contradiction that there exists a 2-flash v_0, v_1, v_2 of some colour $b \notin C$ in T. Consider the case where the edge between u and v_1 is directed towards v_1, and let $a \in C$ be the colour of that edge. Since u is strongly $(k-2)$-robust, some $(k-2)$-rainbow with colour set $C \setminus \{a\}$ ends at u. Extending this $(k-2)$-rainbow to v_1 and v_2 yields a k-rainbow in T, which is a contradiction. The case where the edge between u and v_1 is directed towards u is similar. Hence, no 2-flash of a colour not in C can exist in T. Since $|C| = k-1$, this implies the result. □

The last step is to show that every sufficiently large tournament with no ℓ-flash or k-rainbow contains a strongly $(k-2)$-robust vertex. The proof of this result is similar to the proofs of Lemmas 6.13 and 6.14.

Lemma 6.19. *Let $k \geq 2$ and let T be an edge-coloured tournament containing no ℓ-flash and no k-rainbow. If the number of vertices of T is at least*

$$t(\ell, k-1) + g(\ell, k, k-2) + 2(k-1)\big(t(\ell, k-1) + 2(k-1)\ell \cdot t(\ell, k-2)\big) + 2,$$

then T contains a strongly $(k-2)$-robust vertex.

Proof. Define

$$R := \{v \in V(T) : \text{no } (k-1)\text{-rainbow ends at } v\}.$$

Then $|R| \leq t(\ell, k-1)$, and every vertex in $V(T) \setminus R$ has at most $k-1$ outgoing colours.

Order the vertices of $T[V(T) \setminus R]$ in a way that maximises the number of forward edges, and let $P \subseteq V(T) \setminus R$ be a minimal initial segment of this ordering such that $T[P]$ contains a $(k-2)$-in-robust vertex v. Then $|P| \leq g(\ell, k, k-2) + 1$ and we get that the out-neighbourhood $U := \{u \in V(T) : vu \in E(T)\}$ of v satisfies $2|U| \geq |V(T) \setminus (R \cup P)|$. So,

$$2|U| \geq |V(T)| - |R| - |P| \geq 2(k-1)\big(t(\ell, k-1) + 2(k-1)\ell \cdot t(\ell, k-2)\big) + 1$$

which implies that

$$|U| \geq (k-1)\big(t(\ell,k-1) + 2(k-1)\ell \cdot t(\ell,k-2)\big) + 1.$$

For every colour $a \in C^+(v)$, let $U_a := \{u \in U : c(vu) = a\}$. Since v has at most $k-1$ outgoing colours, there must exist a colour a such that $|U_a| \geq |U|/(k-1)$ and so

$$|U_a| \geq t(\ell,k-1) + 2(k-1)\ell \cdot t(\ell,k-2) + 1.$$

Since v is $(k-2)$-in-robust, we can choose a $(k-2)$-rainbow W without colour a that ends at v, and extend it to each of the vertices in U_a by the edges of colour a. Let C be the set containing a and the $k-2$ colours appearing in W. Then, every outgoing colour of every vertex $u \in U_a$ must be a colour from C. In particular, all edges in $T[U_a]$ have one of the $k-1$ colours from C.

Define

$$S := \{u \in U_a : \text{no } (k-1)\text{-rainbow in } T[U_a] \text{ starts at } u\}.$$

Then $|S| \leq t(\ell,k-1)$. For every vertex $u \in U_a \setminus S$, some $(k-1)$-rainbow in $T[U_a]$ starts at u. Since every edge in $T[U_a]$ has a colour from C, this $(k-1)$-rainbow uses only colours from C, and so every incoming colour of u in T must be a colour from C.

Next, for every colour $b \in C$, define

$U_{a,b} := \{u \in U_a : \text{all } (k-2)\text{-rainbows in } T[U_a] \text{ ending at } u \text{ contain colour } b\}$ and
$V_{a,b} := \{u \in U_a : \text{all } (k-2)\text{-rainbows in } T[U_a] \text{ starting at } u \text{ contain colour } b\}$.

For every $m \in \{0,\dots,\ell-1\}$, let

$$U^m_{a,b} := \{u \in U_{a,b} : \ell_b(u) = m\}.$$

If an edge uw in $T[U^m_{a,b}]$ has colour b, then $\ell_b(u) < \ell_b(w)$, which gives a contradiction. Thus, $T[U^m_{a,b}]$ contains no $(k-2)$-rainbow, so $|U^m_{a,b}| \leq t(\ell,k-2)$ and $|U_{a,b}| \leq \ell \cdot t(\ell,k-2)$. Similarly, $|V_{a,b}| \leq \ell \cdot t(\ell,k-2)$.

Finally, consider $U'_a := U_a \setminus (S \cup \bigcup_{b\in C}(U_{a,b} \cup V_{a,b}))$, so

$$|U'_a| \geq |U_a| - |S| - \sum_{b\in C}(|U_{a,b}| + |V_{a,b}|) \geq 1.$$

Let $u \in U'_a$ be arbitrary. For every colour $b \in C$, there is a $(k-2)$-rainbow in $T[U_a]$

without colour b that ends at u, and also one that starts at u. Since every edge in $T[U_a]$ has a colour from C, the colour set of these $(k-2)$-rainbows is $C \setminus \{b\}$. This shows that u is strongly $(k-2)$-robust in T. □

We can now combine the preceding three lemmas and insert the bounds from the previous section to prove the following.

Corollary 6.20. *For every $\varepsilon > 0$, there exists some C such that for all $k \geq 2$,*

$$t(\ell,k) \leq \max\left\{ Ck^3(\log k)^{1+\varepsilon} \cdot \ell^{k-2}, t(2,k)\left(\frac{\ell}{2}\right)^{k-1} \right\}.$$

Proof. Let D be given by Corollary 6.15, and let $C := 12D$. Suppose that T is an edge-coloured tournament containing no ℓ-flash and no k-rainbow that satisfies $n := |V(T)| \geq Ck^3(\log k)^{1+\varepsilon} \cdot \ell^{k-2}$. Then, we have

$$\begin{aligned} n &\geq 12Dk^3(\log k)^{1+\varepsilon} \cdot \ell^{k-2} \\ &\geq 4Dk^3(\log k)^{1+\varepsilon} \cdot \ell^{k-2} + 2Dk^2(\log k)^{1+\varepsilon} \cdot \ell^{k-2} + 4\ell Dk^3(\log k)^{1+\varepsilon} \cdot \ell^{k-3} + 2 \\ &\geq g(\ell,k,k-2) + (2(k-1)+1)t(\ell,k-1) + 4(k-1)^2\ell \cdot t(\ell,k-2) + 2 \\ &= t(\ell,k-1) + g(\ell,k,k-2) + 2(k-1)\big(t(\ell,k-1) + 2(k-1)\ell \cdot t(\ell,k-2)\big) + 2. \end{aligned}$$

Hence, by Lemma 6.19, T contains a strongly $(k-2)$-robust vertex. Applying Lemma 6.18 yields that at most $k-1$ colours appear in 2-flashes in T, and the result now follows from Lemma 6.17. □

Specialising these proofs to transitive tournaments, the preceding corollary tells us that $f(\ell,k) \leq \max\{Ck^3(\log k)^{1+\varepsilon} \cdot \ell^{k-2}, f(2,k)(\ell/2)^{k-1}\}$. Using this upper bound, we can now prove Theorem 6.2.

Proof of Theorem 6.2. Let C be given by Corollary 6.20. The result is trivial for $k = 1$, so assume $k \geq 2$. If $\ell \geq Ck^3(\log k)^{1+\varepsilon}$, then $Ck^3(\log k)^{1+\varepsilon} \cdot \ell^{k-2} \leq \ell^{k-1}$, and by Theorem 6.16 we know that $f(2,k)(\ell/2)^{k-1} = \ell^{k-1}$. Therefore, the specialisation of Corollary 6.20 to transitive tournaments implies that $f(\ell,k) \leq \ell^{k-1}$. □

6.3 COLOURINGS OF ARBITRARY TOURNAMENTS

In this section, we ask how small a tournament can be while still having the property that every colouring of its edges contains an ℓ-flash or a k-rainbow. For $k = 2$, any non-transitive tournament is an example of this: if the tournament

has a directed cycle, this cycle either contains two consecutive edges with different colours, creating a 2-rainbow, or the cycle is monochromatic, creating flashes of arbitrary lengths. For $\ell \geq 2$ and $k \geq 3$, we show that there are tournaments on fewer than $f(\ell, k)$ vertices with the desired property.

Proof of Proposition 6.9. Set $n := f(\ell, k)$ and let T be the "increasing" transitive tournament on $[n-1]$ with the edge from 1 to $n-1$ reversed, that is, $V(T) := [n-1]$ and

$$E(T) := \{uv : 1 \leq u < v \leq n-1,\ uv \neq (1, n-1)\} \cup \{(n-1, 1)\}.$$

Suppose that T has an edge colouring c with neither ℓ-flashes nor k-rainbows. Define an edge colouring c' of the "increasing" transitive tournament T' on the vertex set $\{0\} \cup [n]$ as follows. For any edge uv of T with $uv \neq (n-1, 1)$, let $c'(uv) = c(uv)$. Next, for any $v \in [n-1]$, let $c'(0, v)$ be the colour of any incoming edge of v in T, and let $c'(v, n)$ be the colour of any outgoing edge of v in T (every vertex in T has an incoming and an outgoing edge). Finally, let $c'(1, n-1)$ and $c'(0, n)$ be distinct and entirely new colours.

Note first that the edge $(0, n)$ is not contained in an ℓ-flash or k-rainbow since it is in no walk of length greater than 1. Next, the edge $(1, n-1)$ is not contained in an ℓ-flash in T' since there are no other edges with the same colour. It is also not contained in a k-rainbow since the only 3-walk containing that edge is $0, 1, n-1, n$, but $c'(0, 1) = c(n-1, 1) = c'(n-1, n)$. On the other hand, for every walk w in T' containing neither $(0, n)$ nor $(1, n-1)$, there is a walk in T with the same colours. Indeed, if w uses the edge $(0, v)$ for $v \in [n-1]$, then by the construction of c' this can be replaced by an edge of T incident to v of the same colour, and the same holds for edges (v, n). Since T contains no ℓ-flash and no k-rainbow, it follows that T' also contains no ℓ-flash and no k-rainbow. However, T' is a transitive tournament with $n + 1 = f(\ell, k) + 1$ vertices, contradicting the definition of $f(\ell, k)$. □

In the positive direction, we prove that all tournaments with $\mathcal{O}(\ell^{k-2}/\sqrt{k})$ vertices can be coloured in a way that avoids ℓ-flashes and k-rainbows. Our construction is very similar to the construction from the introduction which showed that $f(\ell, k) \geq \ell^{k-1}$. However, it only uses those strings whose entries sum to a fixed value. This ensures that no matter how an edge is directed, we can always pick an index where the first string has a smaller value than the second string.

Proof of Proposition 6.10. Let

$$X := \left\{ x \in [\ell]^{k-1} : \sum_{i=1}^{k-1} x_i = \left\lfloor \frac{(\ell+1)(k-1)}{2} \right\rfloor \right\}.$$

A result of Anderson [And69] says that there is an absolute constant $C > 0$ such that $|X| \geq C\ell^{k-2}/\sqrt{k}$. Let T be a tournament with at most $C\ell^{k-2}/\sqrt{k}$ vertices.

As $|V(T)| \leq |X|$, we can assign to every vertex v a unique $x(v) \in X$. Define an edge colouring of T by picking, for an edge $uv \in E(T)$, a colour $c(uv) \in [k-1]$ such that $x(u)_{c(uv)} < x(v)_{c(uv)}$. By the construction of X, this is always possible.

Since this edge colouring uses at most $k-1$ colours, there is no k-rainbow. If there were an ℓ-flash $v_0, \ldots, v_\ell$ of colour a, then $x(v_0)_a < x(v_1)_a < \cdots < x(v_\ell)_a$, implying $x(v_\ell)_a > \ell$, which is a contradiction. Thus, this gives an edge colouring of T without ℓ-flashes and k-rainbows. □

Note that X is the largest antichain in the grid poset [BEK51] and so $C\ell^{k-2}/\sqrt{k}$ is the best bound that can be attained by this simple strategy. With additional work, it is possible to slightly strengthen Proposition 6.10. We sketch the argument, which combines our proof with an argument from Bucić, Letzter, and Sudakov [BLS19] and improves the bound to $C\ell^{k-2}(\log \ell)^{1/(k-1)}/\sqrt{k}$. The idea is to partition a tournament of size n into $\mathcal{O}(n/\log \ell)$ transitive tournaments of size $\mathcal{O}(\log \ell)$ each. From the introduction, we know that each of the transitive tournaments can be coloured with $k-1$ colours and no a-flash if $a^{k-1} = \Omega(\log \ell)$. Moreover, the previous proof provides a way to colour the edges between the transitive tournaments with $k-1$ colours and no b-flash provided $b^{k-2}/\sqrt{k} = \Omega(n/\log \ell)$. The tournament will then contain no (ab)-flash, so we need $ab \leq \ell$, and this is satisfied if $n = \mathcal{O}(\ell^{k-2}(\log \ell)^{1/(k-1)}/\sqrt{k})$.

6.4 OPEN PROBLEMS

We have proved that Lefmann, Rödl, and Thomas's conjecture, Conjecture 6.1, holds for $\ell \geq k^3(\log k)^{1+o(1)}$, but the conjecture remains open when ℓ is small compared to k. The case $\ell = 3$ (that is, proving $f(3,k) = 3^{k-1}$) is already of significant interest. In this setting, there are edge-coloured transitive tournaments with more than ℓ^{k-1} vertices that contain no ℓ-flash and no k-rainbow starting at the first vertex. The existence of k-rainbows starting at the first vertex was a crucial ingredient in Lefmann, Rödl, and Thomas's proof that $f(2,k) = 2^{k-1}$.

Conjecture 6.6 remains wide open for any $\ell \geq 2$ and $k \geq 3$. Establishing this conjecture for $\ell = 2$ would be very useful as our results would then imply that the conjecture also holds for $\ell \geq k^3(\log k)^{1+o(1)}$.

In terms of general upper bounds on $f(\ell, k)$ and $t(\ell, k)$, an interesting next step would be to reduce the gap between the upper and the lower bounds to a multiplicative factor that is sublinear in k. All of our arguments relied on certain vertices having at most $k-1$ outgoing colours, which always added a factor of at least k to our upper bounds.

Problem 6.21. *Is it true that $f(\ell, k) \leq o(k) \cdot \ell^{k-1}$ as $k \to \infty$?*

We have also constructed a tournament with $f(\ell, k) - 1$ vertices, every edge colouring of which contains an ℓ-flash or a k-rainbow. Are there tournaments with $o(f(\ell, k))$ vertices that satisfy this? We believe random tournaments are good candidates.

Problem 6.22. *What is the slowest growing function $h(\ell, k)$ such that for a uniformly random tournament T on $h(\ell, k)$ vertices, with high probability (as $h(\ell, k) \to \infty$) every edge colouring of T contains an ℓ-flash or a k-rainbow?*

Finally, we defined flashes and rainbows in tournaments to be walks, but it is possible to consider the same problem with paths instead. This adds some technical difficulties. For example, when we split a set of vertices according to the longest flash of colour a ending at those vertices, it is no longer guaranteed that each resulting set contains no edge of colour a: the vertices in a directed cycle of colour a could all end up in the same set. However, Conjecture 6.6 could still hold in this more restrictive setting.

Conjecture 6.23. *Let ℓ and k be positive integers and let T be a tournament with more than ℓ^{k-1} vertices. Then every edge colouring of T contains a directed monochromatic path of length ℓ or a directed rainbow path of length k.*

Chapter 7

Product-free sets in free semigroups and free groups

A subset $S \subseteq \mathbb{N}$ of the non-negative integers is *sum-free* if $x + y \notin S$ for all $x, y \in S$. Two very natural questions present themselves.

> **Density:** How dense is the largest sum-free subset of $\mathbb{N}$?
> **Structure:** What is the structure of the densest sum-free subsets of $\mathbb{N}$?

To measure the density of a subset $S \subseteq \mathbb{N}$, define its *upper asymptotic density* as $\bar{d}(S) := \limsup_{n\to\infty} |S \cap \{1, \dots, n\}|/n$. Clearly, the odd integers form a sum-free subset with density $1/2$, and it is relatively easy to show that any sum-free subset $S \subseteq \mathbb{N}$ satisfies $\bar{d}(S) \leq 1/2$. Additionally, Łuczak [Łuc95] showed that every sum-free subset $S \subseteq \mathbb{N}$ with $\bar{d}(S) = 1/2$ is a subset of the odd integers.

In 1996, Calkin and Erdős [CE96] asked what happens for sums of more than two integers. For $k \geq 2$, a subset $S \subseteq \mathbb{N}$ is *k-sum-free* if $x_1 + \cdots + x_k \notin S$ for all $x_1, \dots, x_k \in S$, and *strongly k-sum-free* if it is ℓ-sum-free for every $\ell = 2, \dots, k$. For strongly k-sum-free subsets, one can show that the maximum density is $1/k$, which is achieved by the integers which are 1 mod k. For k-sum-free subsets, Calkin and Erdős [CE96] conjectured that the maximum density is $1/\rho(k)$ where

$$\rho(k) := \min\{\ell \in \mathbb{Z}^+ : \ell \nmid k - 1\}.$$

This would be tight because the integers which are 1 mod $\rho(k)$ form a k-sum-free set: the sum of k such integers is congruent to k mod $\rho(k)$ and $k \not\equiv 1 \bmod \rho(k)$. Łuczak and Schoen [ŁS97] confirmed this conjecture and solved the structure problem for k-sum-free and strongly k-sum-free sets in the non-negative integers.

What about other (semi)groups G? Here, we call a subset $S \subseteq G$ *product-free* if $x \cdot y \notin S$ for all $x, y \in S$. More generally, S is *k-product-free* if $x_1 \cdot \ldots \cdot x_k \notin S$ for all $x_1, \ldots, x_k \in S$, and *strongly k-product-free* if it is ℓ-product-free for every $\ell = 2, \ldots, k$. Product-free sets in finite groups have been extensively studied over the last few decades [Ked09, TV17]. This resulted in a complete solution of the density and structure problems for finite abelian groups [GR05, BPR16] and a large range of results for finite non-abelian groups [BS85, Gow08, Ebe16, KLM24].

In contrast, much less is known in the infinite setting. In this chapter, we will focus on the *free semigroup* $\mathcal{F}$ and the *free group* $\mathcal{G}$ on a finite alphabet $\mathcal{A}$.[1] To define density on $\mathcal{F}$ and $\mathcal{G}$, it is natural to use the measure that, for each n, assigns total weight 1 to the set of all (reduced) words of length n, and gives equal weight to all such words. For a set of words S, we write $\bar{d}(S)$ for its upper asymptotic density and $d^*(S)$ for its upper Banach density under this measure (see Section 7.1 for details). Since $\bar{d}(S) \leq d^*(S)$ for every set S, any result proved for upper Banach density also applies to upper asymptotic density. Note that in the special case $|\mathcal{A}| = 1$, the free semigroup $\mathcal{F}$ on alphabet $\mathcal{A}$ is isomorphic to the non-negative integers under addition, and so this generalises the setting studied by Calkin and Erdős [CE96] as well as Łuczak and Schoen [ŁS97].

The density problem for product-free sets in free semigroups was solved by Leader, Letzter, Narayanan, and Walters [LLNW20].

Theorem 7.1 (Leader, Letzter, Narayanan, and Walters)**.** *Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is product-free, then $d^*(S) \leq 1/2$.*

There is a simple class of examples that show that the constant $1/2$ is best possible here. For a non-empty subset $\Gamma \subseteq \mathcal{A}$, the *odd-occurrence set* $O_\Gamma \subseteq \mathcal{F}$ generated by Γ is the set of all words in which the total number of occurrences of letters from Γ is odd (note that if $\Gamma = \mathcal{A}$, then O_Γ consists of all words of odd length). It is easy to see that these sets are product-free with density $1/2$. Leader, Letzter, Narayanan, and Walters conjectured that these are the only examples of product-free sets with density $1/2$ in free semigroups.

Conjecture 7.2 (Leader, Letzter, Narayanan, and Walters)**.** *Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is product-free and $d^*(S) = 1/2$, then $S \subseteq O_\Gamma$ for some non-empty subset $\Gamma \subseteq \mathcal{A}$.*

[1] The free semigroup on alphabet $\mathcal{A}$ is the set of all finite words with letters from $\mathcal{A}$ (including the empty word) equipped with the operation of concatenation. The free group on alphabet $\mathcal{A}$ is the group generated by $\mathcal{A}$ whose only relations are that aa^{-1} and $a^{-1}a$ are the empty word for every $a \in \mathcal{A}$.

In this chapter, we confirm this conjecture. In fact, we study the structure problem in the more general setting of k-product-free and strongly k-product-free sets. Clearly, solving the structure problem for these sets first requires solving the density problem. For strongly k-product-free sets, this was previously achieved by Ortega, Rué, and Serra [ORS24], both in free semigroups and in free groups.

Theorem 7.3 (Ortega, Rué, and Serra)**.** *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{H}$ be the free semigroup or the free group on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{H}$ is strongly k-product-free, then $\bar{d}(S) \leq 1/k$.*

Our first main result solves the structure problem for strongly k-product-free sets in free semigroups. Before stating this result, note first that the odd-occurrence set O_Γ can also be described as follows: label each letter in Γ with a 1 and every other letter with a 0 and let the sum of a word be the sum of the labels of its letters; then O_Γ is the set of words with odd sum. The natural generalisation of this to $k \geq 3$ provides strongly k-product-free subsets of $\mathcal{F}$ with density $1/k$. We prove that these are the only examples.

Theorem 7.4. *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is strongly k-product-free and $d^*(S) = 1/k$, then the following holds. It is possible to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/k\mathbb{Z}$ such that S is a subset of the strongly k-product-free set*

$$T := \{w \in \mathcal{F} : \textit{the sum of the labels of letters in } w \textit{ is } 1 \bmod k\}.$$

The set T defined in Theorem 7.4 is strongly k-product-free and has density $1/k$ provided the labelling is non-degenerate (otherwise T is empty; see Remark 7.19 for details). The special case of Theorem 7.4 with $k = 2$ confirms the conjecture of Leader, Letzter, Narayanan, and Walters.

Next, we turn to k-product-free sets. Recall that for the non-negative integers, Łuczak and Schoen [ŁS97] showed that any k-sum-free set has density at most $1/\rho(k)$, which is optimal. We generalise this result to free semigroups.

Theorem 7.5. *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is k-product-free, then $d^*(S) \leq 1/\rho(k)$.*

In addition, we solve the structure problem in this setting for all but four values of k. More precisely, we show that the structure of the extremal k-product-free sets is very similar to that of strongly k-product-free sets except everything is modulo $\rho(k)$ (see Section 7.5 for further discussion of the cases when k is 3, 5, 7, or 13).

Theorem 7.6. *Let $k \geq 2$ be an integer with $k \notin \{3, 5, 7, 13\}$ and $\rho := \rho(k)$. Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is k-product-free and $d^*(S) = 1/\rho$, then the following holds. It is possible to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/\rho\mathbb{Z}$ such that S is a subset of the k-product-free set*

$$T := \{w \in \mathcal{F} : \textit{the sum of the labels of letters in } w \textit{ is } 1 \bmod \rho\}.$$

Finally, we consider k-product-free sets in the more challenging setting of free groups. Ortega, Rué, and Serra [ORS24] solved the density problem for strongly k-product-free sets in this setting (see Theorem 7.3) and made a conjecture corresponding to Calkin and Erdős's for k-product-free sets.

Conjecture 7.7 (Ortega, Rué, and Serra)**.** *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{G}$ be the free group on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{G}$ is k-product-free, then $\bar{d}(S) \leq 1/\rho(k)$.*

We prove a strengthening of this conjecture, with upper Banach density in place of upper asymptotic density. See Section 7.5 for a discussion and conjecture on what the structure of the extremal sets might be.

Theorem 7.8. *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{G}$ be the free group on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{G}$ is k-product-free, then $d^*(S) \leq 1/\rho(k)$.*

Chapter outline. In Section 7.1 we provide the formal definitions of density. Sections 7.1.1 and 7.1.2 introduce the notions of upper asymptotic and upper Banach density as well as some important technical lemmas. In Section 7.1.3, we then state our main density result, Theorem 7.14. This result immediately implies Theorem 7.5 and we also use it to prove two key technical results, Lemmas 7.17 and 7.18, that we need for our structural proofs. In addition, we use this density result to deduce Corollary 7.15 which shows that the density bounds of Theorems 7.3 and 7.5 hold locally down all subtrees.

Before proving our main density result, we obtain our structural results in Section 7.2 whose proofs are simpler and already contain some of the key ideas. Sections 7.2.1 and 7.2.2 prove Theorems 7.4 and 7.6, respectively. After that, Section 7.3 is devoted to our main density result. In Section 7.3.1 we build the machinery that we use to prove Theorem 7.14 in Section 7.3.2. In Section 7.4 we adapt our machinery to free groups. Using an observation of Ortega, Rué, and Serra, this implies Theorem 7.8. We finish in Section 7.5 with some open problems.

This chapter is based on joint work with Freddie Illingworth and Alex Scott [IMS25].

7.1 DENSITY

Throughout this chapter $\mathcal{F}$ will be a free semigroup on a finite alphabet $\mathcal{A}$. That is, $\mathcal{F}$ is the set of all finite words whose letters are in $\mathcal{A}$ equipped with the associative operation of concatenation and whose identity is the empty word.

In this section, we will formally define notions of density for subsets of $\mathcal{F}$. We start in Section 7.1.1 by introducing the upper asymptotic and upper Banach density. Although we want to prove upper bounds for these densities, they are difficult to use directly in our proofs for various technical reasons, including that they are not additive.

Instead, in Section 7.1.2, we will define an adapted notion of density that is closely related but much easier to handle. We use this directly to prove some basic properties about the density and to define a notion of density for subtrees of $\mathcal{F}$.

With these notions of density in hand, we state our main density result, Theorem 7.14, in Section 7.1.3. While we defer the proof of this result to Section 7.3 due to its complexity, we will immediately deduce some properties about the density of a product-free set in subtrees of $\mathcal{F}$ which will be crucial for proving our structural results in Section 7.2.

7.1.1 UPPER ASYMPTOTIC AND UPPER BANACH DENSITY

To motivate and provide intuition for the notation, we view things from the perspective of a randomly generated word. Let $\mathbf{W} := \alpha_1\alpha_2\cdots$ be a random infinite word where each letter α_i is chosen independently and uniformly at random from $\mathcal{A}$. With $\mathbf{W}_n := \alpha_1\alpha_2\cdots\alpha_n$, we may view $(\mathbf{W}_n)$ as a random walk on an infinite $|\mathcal{A}|$-ary tree. We say $\mathbf{W}$ *hits* a set $B \subseteq \mathcal{F}$ if the random walk hits B (or equivalently if $\mathbf{W}$ has a prefix in B) and $\mathbf{W}$ *avoids* B otherwise. We equip $\mathcal{F}$ with a measure μ satisfying, for every word $w \in \mathcal{F}$,

$$\mu(w) := \mathbb{P}(\mathbf{W} \text{ hits } w) = |\mathcal{A}|^{-|w|}.$$

Note that, for $B \subseteq \mathcal{F}$, $\mu(B) := \sum_{w\in B} \mu(w)$ is the expected number of times that $\mathbf{W}$ hits B. This has a useful corollary. A set $C \subseteq \mathcal{F}$ is *prefix-free* if there are no distinct words $a, b \in C$ where a is a prefix of b. $\mathbf{W}$ can hit a prefix-free set at most once, and so we get the following observation, also known as Kraft's inequality.

Observation 7.9. *If $C \subseteq \mathcal{F}$ is prefix-free, then $\mu(C) \leq 1$.*

For a non-negative integer n and a set $B \subseteq \mathcal{F}$, the *layer* n of B is

$$B(n) := \{w \in B : |w| = n\},$$

while, for an interval $I \subseteq \mathbb{N}$,

$$B(I) := \{w \in B : |w| \in I\}.$$

Note that the measure μ is defined so that $\mu(\mathcal{F}(n)) = 1$. The density of B on layer n is $|B(n)|/|\mathcal{F}(n)| = \mu(B(n))$, which is the probability that $\mathbf{W}_n$ is in B. The *density* of B on interval I is

$$d^I(B) := \frac{\mu(B(I))}{\mu(\mathcal{F}(I))} = \frac{1}{|I|}\sum_{n\in I}\mu(B(n)).$$

We can now give standard notions of density. The *upper asymptotic density* of B is

$$\bar{d}(B) := \limsup_{m\to\infty} d^{\{1,2,\ldots,m\}}(B) = \limsup_{m\to\infty}\frac{1}{m}\sum_{n=1}^{m}\mu(B(n)).$$

The *upper Banach density* of B is

$$d^*(B) := \limsup_{I\to\infty} d^I(B) = \limsup_{I\to\infty}\frac{1}{|I|}\sum_{n\in I}\mu(B(n)),$$

where I is an interval and the notation $I \to \infty$ denotes that both $|I|$ and $\min I$ tend to infinity.[2] In the following we will exclusively use upper Banach density, but since $\bar{d}(B) \leq d^*(B)$ for any set B, our results apply equally well to upper asymptotic density. We remark that these densities always lie in the interval $[0, 1]$.

It should be noted that limit superiors are only subadditive (and not additive). In particular, for disjoint sets $A, B \subseteq \mathcal{F}$ we have $d^*(A \cup B) \leq d^*(A) + d^*(B)$ and equality may not hold. For example, the sets

$$A := \bigcup_{n\in\mathbb{Z}^+}\mathcal{F}(\{(2n-1)!+1, (2n-1)!+2, \ldots, (2n)!\}) \text{ and}$$
$$B := \bigcup_{n\in\mathbb{Z}^+}\mathcal{F}(\{(2n)!+1, (2n)!+2, \ldots, (2n+1)!\})$$

are disjoint and both have density 1.

[2]The condition $\min I \to \infty$ is often omitted from the definition. However, some simple analysis shows that, whether or not this condition is included, the resulting density is the same.

Despite this, in the semigroup of non-negative integers $\mathcal{F} = \mathbb{N}$, the upper Banach density $d^*(B)$ of a set $B \subseteq \mathbb{N}$ satisfies some useful properties. For example, we have $|d^I(x + B) - d^I(B)| \leq x/|I|$. This implies that $d^*(x + B) = d^*(B)$. Even more importantly, if $x_1, \ldots, x_n \in \mathbb{N}$ are such that $x_1 + B, \ldots, x_n + B$ are disjoint, then $d^I(x_1 + B) + \cdots + d^I(x_n + B) \leq 1$, implying that

$$n \cdot d^I(B) \leq \sum_{i=1}^{n} \left(d^I(x_i + B) + \frac{x_i}{|I|} \right) \leq 1 + \sum_{i=1}^{n} \frac{x_i}{|I|}$$

and so $n \cdot d^*(B) \leq 1$. Not only can this provide upper bounds on the density of B, but if $d^*(B) > 1/n$ we could conclude that $x_1 + B, \ldots, x_n + B$ cannot all be disjoint and thereby deduce some structural information about B. Such arguments were used by Łuczak and Schoen [Łuc95, ŁS97] for their results about sum-free subsets of the non-negative integers and also of right cancellative semigroups [ŁS01].

In a free semigroup $\mathcal{F}$ where $|\mathcal{A}| > 1$, these arguments no longer work. For example, if $w \in \mathcal{F}$, then $d^*(wB) = |\mathcal{A}|^{-|w|} \cdot d^*(B)$, where $wB := \{wb : b \in B\}$. Also, the fact that $w_1 B, \ldots, w_n B$ are disjoint gives no general upper bound on the density of B. Even for nested sets $B, wB, \ldots, w^n B$, taking $B := \mathcal{F} \setminus (w\mathcal{F})$ provides an example where these sets are pairwise disjoint, but $d^*(B) = 1 - |\mathcal{A}|^{-|w|}$ which can be arbitrarily close to 1.

We address these issues in Sections 7.1.2 and 7.1.3. By modifying the density that we consider, we can ensure that the density is additive. Importantly, the density of the set $S \subseteq \mathcal{F}$ whose upper Banach density we want to bound will not change. Moreover, in certain situations, we prove that n disjoint nested copies of B imply that the density of B is at most $1/n$. This will be crucial for proving all of our structural results.

7.1.2 DIAGONALISATION AND RELATIVE DENSITY

Let $S \subseteq \mathcal{F}$ be a fixed set whose upper Banach density we wish to bound (for example, S might be k-product-free). By definition of $d^*(S)$, there is a sequence of intervals (I_j) such that $I_j \to \infty$ and

$$d^{I_j}(S) \to d^*(S), \quad \text{as } j \to \infty.$$

Let $B \subseteq \mathcal{F}$ be another set. The sequence $(d^{I_j}(B))$ is bounded as all terms are in $[0, 1]$ and so, by the Bolzano–Weierstrass theorem, has a convergent subsequence. In particular, by passing to a subsequence of (I_j), we may assume that

$d^{I_j}(S) \to d^*(S)$ and $(d^{I_j}(B))$ converges to some limit that we will call $d^{I_\infty}(B)$. Given a countable collection of subsets of $\mathcal{F}$, we may, by a diagonalisation argument, assume there is a subsequence (I_j) such that $d^{I_j}(B) \to d^{I_\infty}(B)$ for every B in the collection and $d^{I_\infty}(S) = d^*(S)$. Throughout this chapter, we will only ever consider countably many sets and so we may assume that this convergence occurs for all sets we consider. These limits, unlike the corresponding upper Banach densities, are additive. Indeed, if sets A and B are disjoint, then $d^{I_j}(A \cup B) = d^{I_j}(A) + d^{I_j}(B)$ and so $d^{I_\infty}(A \cup B) = d^{I_\infty}(A) + d^{I_\infty}(B)$. It should be noted that while $d^{I_\infty}(S) = d^*(S)$, we only have $d^{I_\infty}(B) \leq d^*(B)$ for the other sets.

For our proofs we will need not only to bound the density of a product-free set S but also to bound the density of S on subtrees. We now begin to define this. The *product* AB of two sets $A, B \subseteq \mathcal{F}$ is

$$AB \coloneqq \{ab : a \in A, b \in B\},$$

and the set B^k is the product of k copies of B. Note that B is k-product-free exactly if $B \cap B^k = \varnothing$. A particularly important example of a product is $w\mathcal{F}$ for a word $w \in \mathcal{F}$: this is the subtree of $\mathcal{F}$ consisting of all words starting with w. Similarly, $B\mathcal{F}$ is the set of words that have a prefix in B.

For a finite set $B \subseteq \mathcal{F}$ we write $\min B$ and $\max B$ for the length of the shortest and longest words in B, respectively. Note that if B is finite, then for $n \geq \max B$ the random infinite word $\mathbf{W}$ hits $(B\mathcal{F})(n)$ if and only if it hits B.

Observation 7.10. *If $n \geq |w|$, then $\mu((w\mathcal{F})(n)) = \mu(w)$. If $C \subseteq \mathcal{F}$ is prefix-free and finite, then $\mu((C\mathcal{F})(n)) = \mu(C)$ for $n \geq \max C$.*

Definition 7.11 (Relative density)**.** Let $w \in \mathcal{F}$ and $B \subseteq \mathcal{F}$. For $n \geq |w|$, the relative density of B in $w\mathcal{F}$ on layer n is

$$\frac{|B(n) \cap w\mathcal{F}|}{|\mathcal{F}(n) \cap w\mathcal{F}|} = \frac{\mu(B(n) \cap w\mathcal{F})}{\mu(\mathcal{F}(n) \cap w\mathcal{F})} = \frac{\mu(B(n) \cap w\mathcal{F})}{\mu(w)},$$

which is the probability that $\mathbf{W}_n$ is in B conditioned on the event that $\mathbf{W}$ hits w. If $n < |w|$, we will take the relative density to be 0 by convention. Furthermore, if I is an interval with $\min I \geq |w|$, then the *relative density* of B in $w\mathcal{F}$ on interval I is

$$d^I_{w\mathcal{F}}(B) \coloneqq \frac{\mu(B(I) \cap w\mathcal{F})}{\mu(\mathcal{F}(I) \cap w\mathcal{F})} = \frac{1}{|I| \cdot \mu(w)} \sum_{n \in I} \mu(B(n) \cap w\mathcal{F}) = \frac{d^I(B \cap w\mathcal{F})}{\mu(w)}.$$

If $\min I < |w|$, then we will take the relative density to be 0 by convention.

Note that if w is the empty word, then this relative density is just $d^I(B)$.

Consider the sequence of intervals (I_j) given above, where $d^{I_j}(B) \to d^{I_\infty}(B)$ for every set B in a countable collection. For each word $w \in \mathcal{F}$ and each set in the collection, the sequence $(d^{I_j}_{w\mathcal{F}}(B))$ is bounded as all terms are in $[0,1]$ and so, by the Bolzano–Weierstrass theorem, has a convergent subsequence. Since $\mathcal{F}$ is countable, we may, via a diagonalisation argument, pass to a subsequence (I_j) such that, for every $w \in \mathcal{F}$ and every B in the countable collection, $(d^{I_j}_{w\mathcal{F}}(B))$ converges to some limit $d^{I_\infty}_{w\mathcal{F}}(B)$. In conclusion, we may assume throughout the chapter that for any set B we encounter and all $w \in \mathcal{F}$ we have

$$d^{I_j}_{w\mathcal{F}}(B) \to d^{I_\infty}_{w\mathcal{F}}(B),$$

where $d^{I_\infty}(B) \leq d^*(B)$ and $d^{I_\infty}(S) = d^*(S)$ for one fixed set S. Note that when w is the empty word, $d^{I_\infty}_{w\mathcal{F}}$ is the same as d^{I_∞}. As before, these limits are additive. They satisfy the useful property that we may strip away prefixes.

Lemma 7.12. *If $w, v \in \mathcal{F}$, then $d^{I_\infty}_{wv\mathcal{F}}(wB) = d^{I_\infty}_{v\mathcal{F}}(B)$.*

Proof. Let I be any interval with $\min I \geq |wv|$, so

$$d^I_{wv\mathcal{F}}(wB) = \frac{1}{|I| \cdot \mu(wv)} \sum_{n \in I} \mu((wB)(n) \cap wv\mathcal{F}).$$

Removing the leading w from each word in $(wB)(n) \cap wv\mathcal{F}$ shows that $\mu((wB)(n) \cap wv\mathcal{F}) = \mu(w) \cdot \mu(B(n - |w|) \cap v\mathcal{F})$. Also $\mu(wv) = \mu(w)\mu(v)$ and so

$$d^I_{wv\mathcal{F}}(wB) = \frac{1}{|I| \cdot \mu(v)} \sum_{n \in I - |w|} \mu(B(n) \cap v\mathcal{F}),$$

where $I - |w|$ is the interval obtained by subtracting $|w|$ from each element of I. Thus,

$$\left| d^I_{wv\mathcal{F}}(wB) - d^I_{v\mathcal{F}}(B) \right| = \frac{1}{|I| \cdot \mu(v)} \left| \sum_{n \in I - |w|} \mu(B(n) \cap v\mathcal{F}) - \sum_{n \in I} \mu(B(n) \cap v\mathcal{F}) \right|.$$

But, for each non-negative integer n, we have $\mu(B(n) \cap v\mathcal{F}) \in [0,1]$ and so

$$\left| d^I_{wv\mathcal{F}}(wB) - d^I_{v\mathcal{F}}(B) \right| \leq \frac{|w|}{|I| \cdot \mu(v)}.$$

Setting $I = I_j$ and taking j to infinity gives the required result. □

We are now ready to make an important definition that captures how dense a set B can be within a subtree.

Definition 7.13 (Sup density)**.** For a set $B \subseteq \mathcal{F}$ in the countable collection, the *sup density* of B is defined as

$$d^{I_\infty}_{\sup}(B) := \sup_{w \in \mathcal{F}} d^{I_\infty}_{w\mathcal{F}}(B).$$

Of course, the sup density satisfies $d^{I_\infty}_{\sup}(B) \geq d^{I_\infty}(B)$ (note that the empty word is in $\mathcal{F}$) and so $d^{I_\infty}_{\sup}(S) \geq d^*(S)$.

7.1.3 KEY TECHNICAL DENSITY RESULTS

We will prove the following strengthening of Theorems 7.3 and 7.5 in Section 7.3. Note that while it would not be strictly necessary to state this result for an arbitrary sequence of intervals (I_j) as we could just use the sequence of intervals described in Section 7.1.2 for the proofs of all our main theorems, the extra flexibility allows us to prove Corollary 7.15: the bounds $1/k$ and $1/\rho(k)$ for the density of (strongly) k-product-free sets in a free semigroup $\mathcal{F}$ hold not only globally but also locally down each subtree $w\mathcal{F}$.

Theorem 7.14. *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. Let (I_j) be a sequence of intervals, with $I_j \to \infty$, on which all relevant densities converge.*

(a) If $S \subseteq \mathcal{F}$ is strongly k-product-free, then $d^{I_\infty}_{\sup}(S) \leq 1/k$.

(b) If $S \subseteq \mathcal{F}$ is k-product-free, then $d^{I_\infty}_{\sup}(S) \leq 1/\rho(k)$.

Note that Theorems 7.3 and 7.5 follow immediately from this: taking (I_j) as the sequence described in Section 7.1.2 we have $d^*(S) = d^{I_\infty}(S) \leq d^{I_\infty}_{\sup}(S)$ and so Theorem 7.14 gives the required upper bounds.

We can also directly prove that the bounds for the upper Banach density hold locally down all subtrees. We will not use this result elsewhere.

Corollary 7.15. *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. Let w be any word of $\mathcal{F}$.*

(a) If $S \subseteq \mathcal{F}$ is strongly k-product-free, then $\limsup_{I\to\infty} d^I_{w\mathcal{F}}(S) \leq 1/k$.

(b) If $S \subseteq \mathcal{F}$ is k-product-free, then $\limsup_{I\to\infty} d^I_{w\mathcal{F}}(S) \leq 1/\rho(k)$.

Proof. Let $S \subseteq \mathcal{F}$ be (strongly) k-product-free. By definition of the limit superior, there is a sequence of intervals (I_j) such that $I_j \to \infty$ and

$$d^{I_j}_{w\mathcal{F}}(S) \to \limsup_{I\to\infty} d^I_{w\mathcal{F}}(S), \quad \text{as } j \to \infty.$$

Start with this sequence of intervals and apply the diagonalisation argument outlined in Section 7.1.2, repeatedly passing to subsequences on which all densities we need to consider converge. By the initial choice of the sequence (I_j), even after passing to subsequences we have that $d^{I_\infty}_{w\mathcal{F}}(S) = \limsup_{I\to\infty} d^I_{w\mathcal{F}}(S)$. Therefore,

$$\limsup_{I\to\infty} d^I_{w\mathcal{F}}(S) = d^{I_\infty}_{w\mathcal{F}}(S) \leq d^{I_\infty}_{\sup}(S).$$

Theorem 7.14 then gives the required upper bounds. □

We will also need Theorem 7.14 for our structural results, Theorems 7.4 and 7.6. For example, if $S \subseteq \mathcal{F}$ is strongly k-product-free with $d^*(S) = 1/k$, then applying part (a) gives $1/k = d^*(S) = d^{I_\infty}(S) \leq d^{I_\infty}_{\sup}(S) \leq 1/k$ and so $d^{I_\infty}(S) = d^{I_\infty}_{\sup}(S)$. This suggests that S is uniformly distributed down all subtrees. We make this precise with the following lemma.

Lemma 7.16. *If $d^{I_\infty}(B) = d^{I_\infty}_{\sup}(B)$, then $d^{I_\infty}_{w\mathcal{F}}(B) = d^{I_\infty}(B)$ for every $w \in \mathcal{F}$.*

Proof. Let n be a non-negative integer and let I be an interval with $\min I > n$. Every word of length greater than n is in exactly one $w\mathcal{F}$ with $w \in \mathcal{F}(n)$. Hence,

$$d^I(B) = \sum_{w\in\mathcal{F}(n)} d^I(B \cap w\mathcal{F}) = \sum_{w\in\mathcal{F}(n)} \mu(w) \cdot d^I_{w\mathcal{F}}(B).$$

Setting $I = I_j$ and taking j to infinity gives

$$d^{I_\infty}(B) = \sum_{w\in\mathcal{F}(n)} \mu(w) \cdot d^{I_\infty}_{w\mathcal{F}}(B).$$

Now $\sum_{w\in\mathcal{F}(n)} \mu(w) = \mu(\mathcal{F}(n)) = 1$ and $d^{I_\infty}_{w\mathcal{F}}(B) \leq d^{I_\infty}_{\sup}(B) = d^{I_\infty}(B)$ for each $w \in \mathcal{F}(n)$. So, we must have $d^{I_\infty}_{w\mathcal{F}}(B) = d^{I_\infty}(B)$ for every $w \in \mathcal{F}(n)$. Since n was arbitrary, this proves the lemma. □

The next two lemmas are the key technical results for our structural proofs. We remark that for the non-negative integers (that is, when $|\mathcal{A}| = 1$) they are much more obvious.

Lemma 7.17. *Let $B \subseteq \mathcal{F}$ be such that $d^{I_\infty}(B) = d^{I_\infty}_{\sup}(B) > 1/n$. Then, for any $w_1, \ldots, w_n \in \mathcal{F}$, the sets*

$$w_1 B, \quad w_1 w_2 B, \quad \ldots, \quad w_1 w_2 \cdots w_{n-1} B, \quad w_1 w_2 \cdots w_n B$$

cannot be pairwise disjoint.

Proof. Assume that these sets are pairwise disjoint. Then, for any $w \in \mathcal{F}$,

$$d^{I_\infty}_{w\mathcal{F}}(w_1 B) + \cdots + d^{I_\infty}_{w\mathcal{F}}(w_1 \cdots w_n B) = d^{I_\infty}_{w\mathcal{F}}((w_1 B) \cup \cdots \cup (w_1 \cdots w_n B)) \leq 1.$$

Choose $w = w_1 \cdots w_n$. Applying Lemma 7.12 to each term gives

$$d^{I_\infty}_{w_2 \cdots w_n \mathcal{F}}(B) + \cdots + d^{I_\infty}_{w_n \mathcal{F}}(B) + d^{I_\infty}_{\mathcal{F}}(B) \leq 1.$$

By Lemma 7.16, each term is $d^{I_\infty}(B)$ which contradicts $d^{I_\infty}(B) > 1/n$. □

Lemma 7.18. *Let $B \subseteq \mathcal{F}$ be such that $d^{I_\infty}(B) = d^{I_\infty}_{\sup}(B) > 2/(2n-1)$. Then, for any $w_1, \ldots, w_n, v_1, \ldots, v_n \in \mathcal{F}$ and $C \subseteq B$, either the sets*

$$w_1 B, \quad w_1 w_2 B, \quad \ldots, \quad w_1 \cdots w_{n-1} B, \quad w_1 \cdots w_n C$$

or the sets

$$v_1 B, \quad v_1 v_2 B, \quad \ldots, \quad v_1 \cdots v_{n-1} B, \quad v_1 \cdots v_n (B \setminus C)$$

are not pairwise disjoint.

Proof. Assume that both collections of sets are pairwise disjoint. Then, as in the proof of Lemma 7.17, we have that

$$d^{I_\infty}_{w_2 \cdots w_n \mathcal{F}}(B) + \cdots + d^{I_\infty}_{w_n \mathcal{F}}(B) + d^{I_\infty}_{\mathcal{F}}(C) \leq 1$$

and

$$d^{I_\infty}_{v_2 \cdots v_n \mathcal{F}}(B) + \cdots + d^{I_\infty}_{v_n \mathcal{F}}(B) + d^{I_\infty}_{\mathcal{F}}(B \setminus C) \leq 1.$$

Note that $d^{I_\infty}_{\mathcal{F}}(B \setminus C) = d^{I_\infty}_{\mathcal{F}}(B) - d^{I_\infty}_{\mathcal{F}}(C)$. By applying this and adding both inequalities together, we get that

$$d^{I_\infty}_{w_2 \cdots w_n \mathcal{F}}(B) + \cdots + d^{I_\infty}_{w_n \mathcal{F}}(B) + d^{I_\infty}_{v_2 \cdots v_n \mathcal{F}}(B) + \cdots + d^{I_\infty}_{\mathcal{F}}(B) \leq 2.$$

By Lemma 7.16, each term is $d^{I_\infty}(B)$ which contradicts $d^{I_\infty}(B) > 2/(2n-1)$. □

7.2 STRUCTURE OF PRODUCT-FREE SETS IN FREE SEMIGROUPS

After formally defining density, we now turn to the structure of product-free sets. In this section, we determine the structure of (strongly) k-product-free sets in free semigroups and prove Theorems 7.4 and 7.6 assuming Theorem 7.14.

We start in Section 7.2.1 by considering strongly k-product-free sets $S \subseteq \mathcal{F}$ with density $d^*(S) = 1/k$. Recall that we want to show that we can label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/k\mathbb{Z}$ such that S is a subset of

$$T := \{a \in \mathcal{F} : \text{the sum of the labels of letters in } a \text{ is } 1 \bmod k\}.$$

We say that the label of a word $a \in \mathcal{F}$ is the sum of the labels of letters in a. Our strategy will be to label a with $\ell \in \mathbb{Z}/k\mathbb{Z}$ if and only if $S \cap aS^{k-\ell+1} \neq \varnothing$. To prove that this is a valid labelling, we will repeatedly construct nested copies of S and apply Lemmas 7.17 and 7.18. For example, given any $x \in S$, consider the sets

$$S, \quad aS, \quad axS, \quad ax^2S, \quad \ldots, \quad ax^{k-1}S.$$

By Lemma 7.17, these cannot all be disjoint, but since S is strongly k-product-free we know that the sets $aS, axS, \ldots, ax^{k-1}S$ are disjoint. Therefore, $S \cap ax^rS \neq \varnothing$ for some $r \in \mathbb{Z}/k\mathbb{Z}$ and so $S \cap aS^{r+1} \neq \varnothing$. Thus, we can assign a label to a. With more complicated versions of such arguments, we eventually show that $S \subseteq T$.

To prove our structural result for k-product-free sets $S \subseteq \mathcal{F}$ with $k \notin \{3, 5, 7, 13\}$, we will show in Section 7.2.2 that if $d^*(S) = 1/\rho(k)$, S is strongly $\rho(k)$-product-free. This uses arguments similar to the strongly k-product-free case. Theorem 7.6 then follows from our structural result for strongly $\rho(k)$-product-free sets.

7.2.1 STRUCTURE OF STRONGLY k-PRODUCT-FREE SETS

Let $S \subseteq \mathcal{F}$ be a strongly k-product-free set with $d^*(S) = 1/k$. By Theorem 7.14 we know that $d^{I_\infty}(S) = 1/k = d^{I_\infty}_{\sup}(S)$ and so we may and will frequently apply Lemmas 7.17 and 7.18 with $n = k + 1$.

To prove Theorem 7.4, we need to show that there exists a labelling of the letters of $\mathcal{A}$ such that S is a subset of the strongly k-product-free set

$$T := \{a \in \mathcal{F} : \text{the label of } a \text{ is } 1 \bmod k\}.$$

Remark 7.19. *If some prime divides k and every label given to letters in $\mathcal{A}$, then T will be empty. If there is no such prime, then T will be non-empty by Bezout's lemma. If T is non-empty, then $d^*(T) = 1/k$. Indeed, consider the infinite random word $\mathbf{W} = \alpha_1\alpha_2\cdots$ from Section 7.1.1 and let X_n be the sum of the labels of $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then (X_n) is a Markov chain on $\mathbb{Z}/k\mathbb{Z}$ which is irreducible since $T \neq \varnothing$. The uniform distribution π on $\mathbb{Z}/k\mathbb{Z}$ is stationary for this chain. If p_n denotes the distribution of X_n, the convergence of Cesàro averages in finite irreducible Markov chains (see, for example, [GS06, Theorem 11.11]) implies that the averages $|I|^{-1}\sum_{n\in I} p_n$ over long intervals converge to π. In particular, $d^*(T) = \pi(1) = 1/k$.*

To construct this labelling, we need to identify the labels for all $a \in \mathcal{F}$. Clearly, everything in S should be labelled 1 mod k. For any other $a \in \mathcal{F}$, appending a word from S should increase the label by 1. So, if a has label ℓ and we append i words from S to a where $i \equiv -\ell \bmod k$, we should get the label 0, and appending one more word from S should give the label 1, which might itself be a word from S. On the other hand, for any $j \not\equiv i \bmod k$, appending $j+1$ words from S to a should give a label different from 1 and therefore never yield a word from S.

Based on this intuition, for $i \in \{0, 1, \ldots, k-1\}$ define

$$T_i := \{a \in \mathcal{F} : S \cap aS^{i+1} \neq \varnothing\}.$$

Then, everything in T_i should have the label $-i \bmod k$. In particular, we expect that $S \subseteq T_{k-1}$ and that $T_iT_j \subseteq T_{i+j}$. This is exactly what we will show, and this will allow us to deduce the structure of T_{k-1}, which will be the set T from above.

Remark 7.20. *Throughout we will view the indices of the T_i as elements of $\mathbb{Z}/k\mathbb{Z}$ and, in particular, all addition of indices is modulo k.*

Note that our definition of T_i is slightly arbitrary. Whether we append or prepend words from S to some $a \in \mathcal{F}$, the change in the label of a should always be the same. So, we could also have defined T_i as the set $\{a \in \mathcal{F} : S \cap S^{i+1}a \neq \varnothing\}$. Fortunately, the following result tells us that these definitions are equivalent.

Proposition 7.21. *For every positive integer r and every $a \in \mathcal{F}$,*

$$S \cap S^r a \neq \varnothing \Leftrightarrow S \cap S^{r-1}aS \neq \varnothing \Leftrightarrow \cdots \Leftrightarrow S \cap SaS^{r-1} \neq \varnothing \Leftrightarrow S \cap aS^r \neq \varnothing.$$

Proof. We first prove the case $r = 1$. Suppose that $S \cap Sa \neq \varnothing$, so there is some $x \in S$ with $xa \in S$. Consider the sets $S, xS, x^2S, \ldots, x^{k-1}S, x^{k-1}aS = x^{k-2}(xa)S$. By Lemma 7.17, these sets cannot all be pairwise disjoint. Since S is strongly

k-product-free and $x \in S$, the sets $S, xS, \ldots, x^{k-1}S$ are pairwise disjoint. Since S is strongly k-product-free and $xa \in S$, the sets $S, xS, \ldots, x^{k-2}S, x^{k-2}(xa)S$ are pairwise disjoint. Thus $x^{k-1}S$ and $x^{k-1}aS$ are not disjoint and so $S \cap aS \neq \varnothing$.

Let $f\colon \mathcal{F} \to \mathcal{F}$ be the reverse map that reverses each word of $\mathcal{F}$ (that is, reads them from right to left). The function f is a measure-preserving involution. Let $\overline{S} \coloneqq f(S)$ and note that $\overline{S}$ is a strongly k-product-free subset of $\mathcal{F}$ that satisfies $d^*(\overline{S}) = d^*(S) = 1/k$. In particular, from the previous paragraph it follows that $\overline{S} \cap \overline{S}a \neq \varnothing \Rightarrow \overline{S} \cap a\overline{S} \neq \varnothing$. Now, $\overline{S} \cap \overline{S}a = f(S \cap aS)$ and $\overline{S} \cap a\overline{S} = f(S \cap Sa)$ and so $S \cap aS \neq \varnothing \Rightarrow S \cap Sa \neq \varnothing$, concluding the case $r = 1$.

For the general case it suffices to prove that for all non-negative integers i, j we have that $S \cap S^{i+1}aS^j \neq \varnothing \Leftrightarrow S \cap S^iaS^{j+1} \neq \varnothing$. Suppose that $S \cap S^{i+1}aS^j \neq \varnothing$. Then there is $x_i \in S^i$ and $x_j \in S^j$ such that $S \cap Sx_iax_j \neq \varnothing$. Applying the $r = 1$ case to the word x_iax_j shows that $S \cap x_iax_jS \neq \varnothing$ and so $S \cap S^iaS^{j+1} \neq \varnothing$. The other direction is analogous. □

If the sets $T_0, T_1, \ldots, T_{k-1}$ are supposed to correctly identify the labels of all words $a \in \mathcal{F}$, then every a should be in exactly one of these sets, and S should satisfy $S \subseteq T_{k-1}$. This is proved by the following proposition.

Proposition 7.22. *The sets $T_0, T_1, \ldots, T_{k-1}$ partition $\mathcal{F}$ and $S \subseteq T_{k-1}$.*

Proof. Let $a \in \mathcal{F}$ and $x \in S$. Consider the sets S, aS, axS, ax^2S, …, $ax^{k-1}S$. By Lemma 7.17, these sets cannot all be pairwise disjoint. Since S is strongly k-product-free and $x \in S$, the sets $aS, axS, \ldots, ax^{k-1}S$ are pairwise disjoint. Hence, there is some $i \in \{0, 1, \ldots, k-1\}$ such that $S \cap ax^iS \neq \varnothing$ and so $S \cap aS^{i+1} \neq \varnothing$. That is, we have shown that $\bigcup_{i=0}^{k-1} T_i = \mathcal{F}$.

We next show that the sets T_i are pairwise disjoint (and so partition $\mathcal{F}$). Suppose that $a \in T_i \cap T_j$ where $0 \leq i < j \leq k-1$. Since $a \in T_i$, Proposition 7.21 implies $S \cap SaS^i \neq \varnothing$ and so there exist $x \in S$ and $y \in S^i$ such that $xay \in S$. Let

$$C \coloneqq \{s \in S : ays \in S\} \subseteq S.$$

Consider the $k+1$ sets

$$S, \quad xS, \quad x^2S, \quad \ldots, \quad x^{k-1}S, \quad x^{k-1}ay(S \setminus C).$$

As S is strongly k-product-free and $x \in S$, the first k of these sets are pairwise disjoint. Similarly, noting that $x^{k-1}ay = x^{k-2}(xay)$ and $xay \in S$, we have that the

last set is disjoint from each of the first $k-1$. Finally, the last two sets are disjoint by the definition of C. Hence, all $k+1$ sets are pairwise disjoint.

Since $a \in T_j$, Proposition 7.21 implies that $S \cap S^{j+1}a \neq \varnothing$ and so there exist $z_1, \ldots, z_{j+1} \in S$ such that $z_1 \cdots z_{j+1} a \in S$. Consider the $k+1$ sets

$$S, \quad z_1 S, \quad \ldots, \quad z_1^{k-j} S, \quad z_1^{k-j} z_2 S, \quad \ldots, \quad z_1^{k-j} z_2 \cdots z_j S, \quad z_1^{k-j} z_2 \cdots z_{j+1} a y C.$$

The first k of these sets are pairwise disjoint as S is strongly k-product-free. Similarly, noting that $z_1 z_2 \cdots z_{j+1} a \in S$, $y \in S^i$, and $i < j$, the last set is disjoint from each of S, $z_1 S$, $\ldots$, $z_1^{k-j-1} S$. Now, by the definition of C, $ayC \subseteq S$. Using this and the fact that S is strongly k-product-free shows that the last set is disjoint from each of $z_1^{k-j} S$, $z_1^{k-j} z_2 S$, $\ldots$, $z_1^{k-j} z_2 \cdots z_j S$. Hence, all $k+1$ sets are pairwise disjoint. This contradicts Lemma 7.18, and so the T_i do partition $\mathcal{F}$.

It remains to show that $S \subseteq T_{k-1}$. Since S is strongly k-product-free, for any $x \in S$ the set S is disjoint from each of xS, xS^2, $\ldots$, xS^{k-1} and so $x \notin T_0 \cup \cdots \cup T_{k-2}$. Since the T_i partition $\mathcal{F}$, we must have $x \in T_{k-1}$, as required. □

Given these two results, we know that $a \in T_i$ should be labelled by $-i \in \mathbb{Z}/k\mathbb{Z}$. Next, we want to show that the label of a product ab is the sum of the labels of a and b. We begin by proving that this is true whenever we append a word from S.

Proposition 7.23. *The following hold for all $j \in \mathbb{Z}/k\mathbb{Z}$.*

(a) If $ax \in T_j$ and $x \in S$, then $a \in T_{j+1}$.
(b) $T_{j+1} S \subseteq T_j$.

Proof. We first prove (a). Suppose that $0 \leq j \leq k-2$. We have $S \cap axS^{j+1} \neq \varnothing$ and $x \in S$, so $S \cap aS^{j+2} \neq \varnothing$ and so $a \in T_{j+1}$.

Now suppose that $j = k-1$. Consider the sets S, aS, axS, ax^2S, $\ldots$, $ax^{k-1}S$. By Lemma 7.17, these sets cannot all be pairwise disjoint. Since S is strongly k-product-free and $x \in S$, the sets $aS, axS, \ldots, ax^{k-1}S$ are pairwise disjoint. Also, as $ax \in T_{k-1}$ and so ax is not in $T_0 \cup T_1 \cup \cdots \cup T_{k-2}$ by Proposition 7.22, S is disjoint from each of $axS, ax^2S, \ldots, ax^{k-1}S$. Thus S and aS are not disjoint and so $a \in T_0$.

We now prove (b). Let $a \in T_{j+1}$ and $x \in S$. Suppose that $ax \in T_i$, where we know that such an i exists by Proposition 7.22. By (a), $i+1 \equiv j+1 \bmod k$ and so $i \equiv j \bmod k$, as required. □

It is now an easy consequence that the labels of all T_i are very well-behaved with respect to products.

Proposition 7.24. *For all $i, j \in \mathbb{Z}/k\mathbb{Z}$ we have $T_i T_j \subseteq T_{i+j}$.*

Proof. Let $a \in T_i$ and $b \in T_j$. As $b \in T_j$ there are $x_1, x_2, \ldots, x_{j+1} \in S$ such that $bx_1x_2 \cdots x_{j+1} \in S$. By Proposition 7.23 (b),

$$abx_1 \cdots x_{j+1} = a(bx_1 \cdots x_{j+1}) \in T_{i-1}.$$

Applying Proposition 7.23 (a) $j+1$ times, once to remove each x_ℓ, implies that $ab \in T_{i-1+(j+1)} = T_{i+j}$, as required. □

Finally, this allows us to complete the proof of Theorem 7.4.

Proof of Theorem 7.4. For each letter $\alpha \in \mathcal{A}$, we know by Proposition 7.22 that there exists a unique $i \in \mathbb{Z}/k\mathbb{Z}$ such that $\alpha \in T_i$. Label α with i. Then, by Proposition 7.24, for each $i \in \mathbb{Z}/k\mathbb{Z}$ we have

$$T_i = \{w \in \mathcal{F} : \text{the sum of the labels of letters in } w \text{ is } i \bmod k\}.$$

In particular, by Proposition 7.22,

$$S \subseteq T_{k-1} = \{w \in \mathcal{F} : \text{the sum of the labels of letters in } w \text{ is } -1 \bmod k\}.$$

Note that T_{k-1} is strongly k-product-free: if w is the concatenation of $\ell \geq 2$ words from T_{k-1}, then the sum of the labels of letters in w is $-\ell \bmod k$.

To obtain the result given in the statement of Theorem 7.4, i.e. with $1 \bmod k$ instead of $-1 \bmod k$, simply multiply the label of each letter by -1. □

7.2.2 STRUCTURE OF k-PRODUCT-FREE SETS

Let $k \geq 2$ be an integer with $k \notin \{3, 5, 7, 13\}$, let $\rho := \rho(k)$, and let $S \subseteq \mathcal{F}$ be k-product-free satisfying $d^*(S) = 1/\rho$. By Theorem 7.14 we have $d^{I_\infty}(S) = 1/\rho = d^{I_\infty}_{\sup}(S)$ and so we may and will frequently apply Lemma 7.17 with $n = \rho + 1$.

To prove Theorem 7.6, we will show that, in fact, S is strongly ρ-product-free and so the result follows from Theorem 7.4. To this end we make the following definition.

Definition 7.25. For a set $A \subseteq \mathbb{Z}^+$, the set $S_A \subseteq \mathcal{F}$ is defined as

$$S_A := \bigcap_{i \in A} S^i.$$

We will omit set parentheses, so, for example, $S_1 = S$ and $S_{1,3} = S \cap S^3$.

Since S is k-product-free, $S_{1,k} = \varnothing$. To show that S is strongly ρ-product-free, we need to show that $S_{1,2} = S_{1,3} = \cdots = S_{1,\rho} = \varnothing$. Note that the case $k = 2$ is immediate and so we assume that $k \geq 3$ from now on.

We need a quick technical lemma about the size of ρ.

Lemma 7.26. *Let $k \geq 3$ be an integer with $k \notin \{3, 5, 7, 13\}$ and let $\rho := \rho(k)$. Then*

$$k - 1 \geq \max\{(\rho - t)t(t+1) : t \in \{1, 2, \ldots, \rho - 1\}\}. \tag{7.1}$$

Proof. By the arithmetic mean–geometric mean inequality, for any $t \in [0, \rho]$,

$$(\rho - t)t(t+1) = 4(\rho - t)\frac{t}{2}\frac{t+1}{2} \leq 4\left(\frac{\rho + 1/2}{3}\right)^3 = \frac{4}{27} \cdot \left(\rho + \frac{1}{2}\right)^3.$$

On the other hand, Lev [Lev03, Lemma 18] proved that, for all integers $k \geq 2$,

$$\rho(k) \leq 2\log_2 k + 2.$$

Now, for all $k \geq 2400$,

$$k - 1 \geq \frac{4}{27} \cdot \left(2\log_2 k + \frac{5}{2}\right)^3,$$

and so (7.1) holds. If $\rho \geq 10$, then the definition of ρ gives $k - 1 \geq 5 \times 7 \times 8 \times 9 = 2520$, and so (7.1) holds. We are left to check the remaining cases.

- If $\rho = 2$, the right side of (7.1) is 2. The smallest $k \geq 3$ with $\rho = 2$ is 4.
- If $\rho = 3$, the right side of (7.1) is 6. The only $k \leq 6$ with $\rho = 3$ are 3 and 5.
- If $\rho = 4$, the right side of (7.1) is 12. The only $k \leq 12$ with $\rho = 4$ is 7.
- If $\rho = 5$, the right side of (7.1) is 24. The only $k \leq 24$ with $\rho = 5$ is 13.
- ρ is always a power of a prime, and so there are no k with $\rho = 6$.
- If $\rho = 7$, the right side of (7.1) is 60. The smallest k with $\rho = 7$ is 61.
- If $\rho = 8$, the right side of (7.1) is 90. The smallest k with $\rho = 8$ is 421.
- If $\rho = 9$, the right side of (7.1) is 126. The smallest k with $\rho = 9$ is 841. □

We first show that $S_{1,\rho}$ is empty.

Proposition 7.27. $S_{1,\rho} = \varnothing$.

Proof. Suppose that $S_{1,\rho} \neq \varnothing$ and let t be the maximum positive integer with $S_{1,\rho,2\rho-1,\ldots,t(\rho-1)+1} \neq \varnothing$, where the indices form an arithmetic progression with common difference $\rho - 1$. Such a t must exist as $k \equiv 1 \bmod \rho - 1$ and $S_{1,k} = \varnothing$. Let $w \in S_{1,\rho,2\rho-1,\ldots,t(\rho-1)+1}$. By Lemma 7.26, we have $k - 1 \geq \rho(\rho - 1)$. We split into two cases based on the size of k.

First suppose that $k - 1 > 2\rho(\rho - 1)$. Let α be the minimum positive integer such that $(\alpha - 1)\rho(\rho - 1) \geq k - 1$. Note that $\alpha \geq 4$. Write $\rho = 2a + b$, where $a := \lfloor \rho/2 \rfloor$ and $b := \rho - 2a \in \{0, 1\}$. Consider the following sets

$$S, \quad w^{\rho-1}S, \quad w^{\alpha(\rho-1)}S, \quad w^{(\alpha+1)(\rho-1)}S, \quad w^{2\alpha(\rho-1)}S, \quad w^{(2\alpha+1)(\rho-1)}S, \quad \ldots,$$
$$w^{(a-1)\alpha(\rho-1)}S, \quad w^{((a-1)\alpha+1)(\rho-1)}S, \quad w^{a\alpha(\rho-1)}S, \quad w^{(a\alpha+b)(\rho-1)}S.$$

We remark that these sets are formed by starting with S and then alternating between prepending $w^{\rho-1}$ and $w^{(\alpha-1)(\rho-1)}$. Two sets that differ only by prepending $w^{\rho-1}$ are called a pair: the pairs are the first and second sets, the third and fourth sets, and so on. The number of sets listed is $2a + b + 1 = \rho + 1$, and so these sets cannot all be pairwise disjoint by Lemma 7.17.

We first show that sets in different pairs are disjoint. If two such sets meet, then $S \cap w^{\ell(\rho-1)}S \neq \varnothing$ for some integer ℓ satisfying $\alpha - 1 \leq \ell \leq a\alpha + b$. We will show that $w^{\ell(\rho-1)} \in S^{k-1}$ which contradicts $S_{1,k} = \varnothing$. As $w \in S_{1,\rho,2\rho-1,\ldots,t(\rho-1)+1}$, we have $w^{\ell(\rho-1)} \in S_{\ell(\rho-1),(\ell+1)(\rho-1),\ldots,\ell(\rho-1)(t(\rho-1)+1)}$, where the indices form an arithmetic progression with common difference $\rho - 1$. It suffices to show that $k - 1$ is in this arithmetic progression. Since $k - 1$ is a multiple of $\rho - 1$, it is enough to show that $\ell(\rho - 1) \leq k - 1 \leq \ell(\rho - 1)(t(\rho - 1) + 1)$ for all integers ℓ satisfying $\alpha - 1 \leq \ell \leq a\alpha + b$. Now,

$$\ell(\rho - 1)(t(\rho - 1) + 1) \geq \ell(\rho - 1)\rho \geq (\alpha - 1)\rho(\rho - 1) \geq k - 1$$

and

$$\begin{aligned}\ell(\rho - 1) &\leq (a\alpha + b)(\rho - 1) = (a\alpha + \rho - 2a)(\rho - 1) \\ &= a(\alpha - 2)(\rho - 1) + \rho(\rho - 1) \leq \rho/2 \cdot (\alpha - 2)(\rho - 1) + \rho(\rho - 1) \\ &= \alpha/2 \cdot \rho(\rho - 1) \leq (\alpha - 2)\rho(\rho - 1) < k - 1,\end{aligned}$$

where we used the minimality of α and the fact that $\alpha \geq 4$ in the final and penultimate inequality, respectively.

Therefore, two sets in the same pair are not disjoint and thus $S \cap w^{\rho-1}S \neq \varnothing$. But $w^{\rho-1} \in S_{\rho-1,2(\rho-1),\ldots,(t(\rho-1)+1)(\rho-1)}$, and so given that $S \cap w^{\rho-1}S \neq \varnothing$, it follows that $S_{1,\rho,2\rho-1,\ldots,(t(\rho-1)+1)(\rho-1)+1} \neq \varnothing$. This contradicts the maximality of t, which concludes the case $k-1 > 2\rho(\rho-1)$.

In the remaining case where $2\rho(\rho-1) \geq k-1 \geq \rho(\rho-1)$, consider the following $\rho+1$ sets

$$S, \quad w^{\rho-1}S, \quad w^{2(\rho-1)}S, \quad \ldots, \quad w^{\rho(\rho-1)}S.$$

Since t is maximal, consecutive sets are disjoint as before. If non-consecutive sets are not disjoint, then $S \cap w^{\ell(\rho-1)}S \neq \varnothing$ for some integer ℓ satisfying $2 \leq \ell \leq \rho$. As before, $w^{\ell(\rho-1)} \in S_{\ell(\rho-1),(\ell+1)(\rho-1),\ldots,\ell(\rho-1)(t(\rho-1)+1)}$ and so it suffices to show that $\ell(\rho-1) \leq k-1 \leq \ell(\rho-1)(t(\rho-1)+1)$ for such ℓ. This is the case as

$$\ell(\rho-1)(t(\rho-1)+1) \geq \ell(\rho-1)\rho \geq 2\rho(\rho-1) \geq k-1$$

and $\ell(\rho-1) \leq \rho(\rho-1) \leq k-1$. Thus, all $\rho+1$ sets are disjoint. This contradicts Lemma 7.17, and so it follows that $S_{1,\rho} = \varnothing$. □

We now show that $S_{1,2}, \ldots, S_{1,\rho-1}$ are all empty.

Proposition 7.28. *For all $1 \leq d \leq \rho-1$, we have $S_{1,d+1} = \varnothing$.*

Proof. We argue via downwards induction on d with the base case $d = \rho-1$ given by Proposition 7.27. Suppose for a contradiction that the result is false, let $1 \leq d \leq \rho-2$ be the maximum integer with $S_{1,d+1} \neq \varnothing$, and let t be the maximum positive integer with $S_{1,d+1,2d+1,\ldots,td+1} \neq \varnothing$. Such a t exists as $k \equiv 1 \bmod d$ and $S_{1,k} = \varnothing$. Let $w \in S_{1,d+1,2d+1,\ldots,td+1}$.

By the definition of ρ, both d and $d+1$ divide $k-1$. Since d and $d+1$ are coprime, we may write $k-1 = \alpha d(d+1)$ for some positive integer α. Let s be the maximum positive integer with $ds \leq \rho-1$. Write $\rho = a(s+1)+b$, where $a := \lfloor \rho/(s+1) \rfloor$ and $b := \rho - a(s+1) \in \{0,1,\ldots,s\}$. Consider the following $\rho+1$ sets

$$\begin{aligned} &S, \quad w^{d}S, \quad \ldots, \quad w^{sd}S, \qquad\qquad w^{(\alpha+s)d}S, \quad w^{(\alpha+s)d+d}S, \quad \ldots, \quad w^{(\alpha+s)d+sd}S, \\ &w^{2(\alpha+s)d}S, \quad \ldots, \quad w^{2(\alpha+s)d+sd}S, \qquad \ldots, \qquad w^{a(\alpha+s)d}S, \quad \ldots, \quad w^{a(\alpha+s)d+bd}S. \end{aligned}$$

We remark that these sets are formed by starting with S, then prepending w^d a

total of s times, prepending $w^{\alpha d}$, then prepending w^d a total of s times, prepending $w^{\alpha d}$, and so on. We group up the sets: the first $s+1$ sets are in the first group, the next $s+1$ sets are in the second group, and so on.

We first show that sets in different groups are disjoint. If two such sets meet, then $S \cap w^{\ell d} S \neq \varnothing$ for some integer ℓ satisfying $\alpha \leq \ell \leq a(\alpha+s)+b$. Now $w^{\ell d} \in S_{\ell d,(\ell+1)d,\ldots,\ell d(td+1)}$ and so, since $k-1$ is a multiple of d, it suffices to show that $\ell d \leq k-1 \leq \ell d(td+1)$ for all such ℓ. Firstly,

$$\ell d(td+1) \geq \alpha d(d+1) = k-1.$$

Also,

$$\ell d \leq (a(\alpha+s)+b)d = (a(\alpha+s)+\rho-a(s+1))d = (a(\alpha-1)+\rho)d.$$

We wish to show this is at most $k-1 = \alpha d(d+1)$, and so it is enough to show that $a(\alpha-1)+\rho \leq \alpha(d+1)$. By the maximality of s, we know that $d(s+1) \geq \rho$ and so $d \geq \rho/(s+1) \geq a$. Hence, it suffices to show that $d(\alpha-1)+\rho \leq \alpha(d+1)$, or equivalently $\rho \leq \alpha+d$. But, by Lemma 7.26,

$$\begin{aligned}\alpha d(d+1) = k-1 &\geq \max\{(\rho-t)t(t+1) : t \in \{1,2,\ldots,\rho-1\}\} \\ &\geq (\rho-d)d(d+1),\end{aligned}$$

and so we indeed have $\rho \leq \alpha+d$.

Next, we show that sets in the same group are disjoint. If two consecutive sets in the same group meet, then $S \cap w^d S \neq \varnothing$. But $w^d \in S_{d,2d,\ldots,(td+1)d}$ and so if $S \cap w^d S \neq \varnothing$, then $S_{1,d+1,2d+1,\ldots,(td+1)d+1} \neq \varnothing$, which contradicts the maximality of t. If two non-consecutive sets in the same group meet, then $S \cap w^{\ell d} S \neq \varnothing$ for some integer ℓ with $2 \leq \ell \leq s$. But $w^{\ell d} \in S_{\ell d}$ and so $S_{1,\ell d+1} \neq \varnothing$. However, $d < 2d \leq \ell d \leq ds \leq \rho-1$, and so this contradicts the maximality of d.

Hence, all $\rho+1$ sets are pairwise disjoint. This contradicts Lemma 7.17, and so it follows that $S_{1,d+1} = \varnothing$ for all $1 \leq d \leq \rho-1$. □

Propositions 7.27 and 7.28 together show that S is strongly ρ-product-free. Because we also have $d^*(S) = 1/\rho$, our structural result for strongly product-free sets, Theorem 7.4, applies to S. This immediately gives Theorem 7.6.

7.3 DENSITY OF PRODUCT-FREE SETS IN FREE SEMIGROUPS

In this section, we will bound the density of a (strongly) k-product-free set $S \subseteq \mathcal{F}$ and prove Theorem 7.14. To motivate our approach, assume that S is strongly 3-product-free and that we want to bound $d^{I_\infty}(S)$. We might hope that $d^{I_\infty}(S) = d^{I_\infty}(S^2) = d^{I_\infty}(S^3)$. Because all of these sets are disjoint, this would imply that $d^{I_\infty}(S \cup S^2 \cup S^3) = 3 \cdot d^{I_\infty}(S)$ and so $d^{I_\infty}(S) \leq 1/3$, as required.

If $S\mathcal{F}$ covers all of $\mathcal{F}$, such an argument works. Indeed, note that for all $w \in S$ we have $d^{I_\infty}_{w\mathcal{F}}(S^2) \geq d^{I_\infty}_{w\mathcal{F}}(wS) = d^{I_\infty}(S)$. If $S\mathcal{F}$ covers all of $\mathcal{F}$, an averaging argument would then show that $d^{I_\infty}(S^2) \geq d^{I_\infty}(S)$, and so $d^{I_\infty}(S \cup S^2) \geq 2 \cdot d^{I_\infty}(S)$. To include S^3 in the union, we can just repeat the argument. Namely, for $w \in S$ we would have $d^{I_\infty}_{w\mathcal{F}}(S^2 \cup S^3) \geq d^{I_\infty}_{w\mathcal{F}}(w(S \cup S^2)) = d^{I_\infty}(S \cup S^2) \geq 2 \cdot d^{I_\infty}(S)$, giving $d^{I_\infty}(S^2 \cup S^3) \geq 2 \cdot d^{I_\infty}(S)$ and thus $d^{I_\infty}(S \cup S^2 \cup S^3) \geq 3 \cdot d^{I_\infty}(S)$, as required.

If $S\mathcal{F}$ fails to cover some part of $\mathcal{F}$, we would like to simply ignore the uncovered part. Within $S\mathcal{F}$, the relative density[3] of S is at least $d^{I_\infty}(S)$, and so it would suffice to show that the relative densities of $S \cup S^2$ and $S \cup S^2 \cup S^3$ are at least $2 \cdot d^{I_\infty}(S)$ and $3 \cdot d^{I_\infty}(S)$ respectively. While the relative density of $S \cup S^2$ can be computed as before, this fails for $S \cup S^2 \cup S^3$. We only know that $S \cup S^2$ has a high density within $S\mathcal{F}$, for example $d^{I_\infty}_{v\mathcal{F}}(S \cup S^2) \geq 2 \cdot d^{I_\infty}(S)$ for some $v \in S$. This does not suffice to get a lower bound on $d^{I_\infty}_{w\mathcal{F}}(S^2 \cup S^3)$ in the calculation above.

Instead, note that $d^{I_\infty}_{wv\mathcal{F}}(S^2 \cup S^3) \geq d^{I_\infty}_{wv\mathcal{F}}(w(S \cup S^2)) = d^{I_\infty}_{v\mathcal{F}}(S \cup S^2) \geq 2 \cdot d^{I_\infty}(S)$. This tells us that the relative density of $S^2 \cup S^3$ in $Sv\mathcal{F}$ is at least $2 \cdot d^{I_\infty}(S)$. If we could show that $Sv\mathcal{F}$ covers essentially all of $S\mathcal{F}$, this would imply that the relative density of $S^2 \cup S^3$ in $S\mathcal{F}$ is at least $2 \cdot d^{I_\infty}(S)$ which in turn would suffice to show that $d^{I_\infty}(S) \leq 1/3$.

The technical arguments in Section 7.3.1 are mostly devoted to showing that this is true up to some small error. In Section 7.3.2 we will then use this machinery to prove Theorem 7.14. The proof of the upper bound on the density of a strongly k-product-free set works exactly as outlined above, just with more sets. For k-product-free sets, the proof is more involved and requires some ideas similar to those from Section 7.2.2.

[3] We define relative density properly after Lemma 7.31.

7.3.1 STEEPLECHASES

We begin by showing that for all $v \in S$, the set $Sv\mathcal{F}$ covers almost all of $S\mathcal{F}$ up to some small error. The main idea for proving this is to approximate S by finite prefix-free subsets (C_t) of S such that, for all t, $C_{t+1} \subseteq C_t\mathcal{F}$ and $C_t\mathcal{F}$ covers almost all of S. Here, almost all means up to some set of density ε. In particular, for all large t and all $\ell > t$, we will have that $C_\ell\mathcal{F}$ covers almost all of $C_t\mathcal{F}$.

Suppose that we have constructed these sets. Then, observe that $C_tv\mathcal{F}$ covers an $|\mathcal{A}|^{-|v|}$-fraction of $C_t\mathcal{F}$. Since $C_{t+1}\mathcal{F}$ covers almost all of $C_t\mathcal{F}$, it covers almost all of $C_t\mathcal{F} \setminus C_tv\mathcal{F}$. So $C_{t+1}v\mathcal{F}$ will cover almost an $|\mathcal{A}|^{-|v|}$-fraction of $C_t\mathcal{F} \setminus C_tv\mathcal{F}$. Repeating this argument with $C_{t+2}\mathcal{F}$, we get that $C_{t+2}v\mathcal{F}$ will cover almost an $|\mathcal{A}|^{-|v|}$-fraction of $C_t\mathcal{F} \setminus (C_tv\mathcal{F} \cup C_{t+1}v\mathcal{F})$. Iterating this further shows that $\bigcup_{\ell \geq t} C_\ell v\mathcal{F}$ covers almost all of $C_t\mathcal{F}$ (now up to some set of density 2ε). Since $C_\ell \subseteq S$, this implies that $Sv\mathcal{F}$ covers almost all of $S\mathcal{F}$, which is what we needed.

We will establish the existence of the prefix-free subsets (C_t) that approximate S in Lemma 7.30. Afterwards, Lemma 7.31 shows that $\bigcup_{\ell \geq t} C_\ell v\mathcal{F}$ covers almost all of $C_t\mathcal{F}$. Finally, Lemma 7.34, our key technical lemma, uses these to give a result that will allow us to carry out the density arguments sketched above.

We start by defining the properties of the sets (C_t) from the preceding discussion.

Definition 7.29 (Steeplechase)**.** An infinite sequence (C_t) of subsets of $\mathcal{F}$ is a *steeplechase* if, for each positive integer t,

- C_t is prefix-free and finite, and
- every word in C_{t+1} has a proper prefix in C_t (in particular, $C_{t+1}\mathcal{F} \subseteq C_t\mathcal{F}$).

We say that a steeplechase (C_t) is *spread* if $\max C_t < \min C_{t+1}$ for all t and *ε-tight* if for all m, n we have that $|\mu(C_m) - \mu(C_n)| \leq \varepsilon$.

Observe that every steeplechase contains a spread steeplechase. Indeed, note that $\min C_t \geq t$ as every word in C_ℓ has a proper prefix in $C_{\ell-1}$. Let $\ell_1 := \max C_1$. Then $\min C_{\ell_1+1} > \ell_1 = \max C_1$. Let $\ell_2 := \max C_{\ell_1+1}$, so $\min C_{\ell_2+1} > \max C_{\ell_1+1}$. Continuing this iteratively, the sets $C_1, C_{\ell_1+1}, C_{\ell_2+1}, \ldots$ form a spread steeplechase.

We also claim that every steeplechase contains an ε-tight steeplechase. Indeed, since C_t is prefix-free, $\mu(C_t) \in [0,1]$. Also, for each t, $C_{t+1}\mathcal{F} \subseteq C_t\mathcal{F}$ and so, by Observation 7.10, the sequence $(\mu(C_t))$ is non-increasing. In particular, this sequence converges and so there exists an ℓ such that for $m, n \geq \ell$ we have $|\mu(C_m) - \mu(C_n)| \leq \varepsilon$. Thus, ignoring the first few C_t gives an ε-tight steeplechase.

Therefore, given any steeplechase (C_t) we may, by passing to a subsequence, assume that (C_t) is both spread and ε-tight. The following lemma shows that, for any set $B \subseteq \mathcal{F}$, there is a steeplechase that approximates B very well.

Lemma 7.30. *Let $\varepsilon > 0$ and $B \subseteq \mathcal{F}$. Then, there exists an ε-tight spread steeplechase (C_t) such that*

- $C_1 \cup C_2 \cup \cdots \subseteq B$,
- *for all t and all n sufficiently large in terms of t, $\mu((B \setminus C_t\mathcal{F})(n)) \leq \varepsilon$, and*
- *for all t, $\mu(C_t) \geq d^{I_\infty}(B) - \varepsilon$.*

Proof. For $w \in B$, define the *headcount* of w as

$$h(w) := |\{b \in B : b \text{ is a prefix of } w\}|.$$

For each positive integer t, let $D_t := \{w \in B : h(w) = t\}$. Note that each D_t is prefix-free and so $\mu(D_t) \leq 1$. Iteratively, we perform the following procedure for each positive integer t.

1. Let ℓ_t be such that $\mu(D_t(\{\ell_t + 1, \ell_t + 2, \ldots\})) \leq \varepsilon/2^t$.
2. Let $C_t := D_t(\{1, 2, \ldots, \ell_t\})$.
3. Remove $(D_t \setminus C_t)\mathcal{F}$ from B and from all D_ℓ with $\ell > t$.

Let B' be the set remaining at the end of this procedure. Note that in each application of step 3, for every $w \in B$, either $h(w)$ does not change or w is removed from B. In particular, every word in C_t has a proper prefix in C_{t-1}. Also, by construction, C_t is a finite subset of B. Thus (C_t) is a steeplechase and $C_1 \cup C_2 \cup \cdots \subseteq B$.

Note that step 3 removes all words with headcount t and length greater than ℓ_t from B. So, if we fix t and let $n > \max\{\ell_1, \ldots, \ell_t\}$, then any word of length n in B' has headcount greater than t and so is in $C_t\mathcal{F}$. Thus, $B'(n) \subseteq (C_t\mathcal{F})(n)$. Next, note that B' is obtained from B by deleting all the $(D_r \setminus C_r)\mathcal{F}$ and so,

$$\mu((B \setminus C_t\mathcal{F})(n)) \leq \mu((B \setminus B')(n)) \leq \sum_r \mu(((D_r \setminus C_r)\mathcal{F})(n)) \leq \sum_r \mu(D_r \setminus C_r) \leq \varepsilon.$$

Finally, this implies that $\mu(B(n)) \leq \mu((C_t\mathcal{F})(n)) + \varepsilon \leq \mu(C_t) + \varepsilon$. Averaging this over $n \in I_j$ and taking $j \to \infty$ gives $d^{I_\infty}(B) \leq \mu(C_t) + \varepsilon$.

Hence, (C_t) is a steeplechase satisfying all three conditions. As noted above, we may, by passing to a subsequence, assume that (C_t) is spread and ε-tight. Passing to a subsequence does not affect the three conditions. □

We call the steeplechase (C_t) given by Lemma 7.30 an *ε-capturing steeplechase* for B. Next, we show for any word w that $\bigcup_{\ell \geq 1} C_\ell w \mathcal{F}$ covers almost all of $C_1\mathcal{F}$.

Lemma 7.31. *Let (C_t) be an ε-tight spread steeplechase. Then, for every $w \in \mathcal{F}$ there is an ℓ such that if $C := C_1 \cup C_2 \cup \cdots \cup C_\ell$ and $n \geq \max C_\ell + |w|$, we have that*

$$\mu((C_1\mathcal{F} \setminus Cw\mathcal{F})(n)) \leq 2\varepsilon.$$

Proof. Let ℓ be sufficiently large in terms of ε, $|w|$, and $|\mathcal{A}|$. Let $n \geq \max C_\ell + |w|$. Note that

$$\mu((C_1\mathcal{F} \setminus C_\ell\mathcal{F})(n)) = \mu((C_1\mathcal{F})(n)) - \mu((C_\ell\mathcal{F})(n)) = \mu(C_1) - \mu(C_\ell) \leq \varepsilon,$$

since $C_\ell\mathcal{F} \subseteq C_1\mathcal{F}$ and (C_t) is ε-tight. Hence, it suffices to prove that

$$\mu((C_\ell\mathcal{F} \setminus Cw\mathcal{F})(n)) \leq \varepsilon.$$

Let D be the following finite prefix-free set

$$D := \{v \in C_\ell : v \text{ has no prefix in } Cw\}.$$

Note that $(C_\ell\mathcal{F} \setminus Cw\mathcal{F})(n) \subseteq (D\mathcal{F})(n)$ and so

$$\mu((C_\ell\mathcal{F} \setminus Cw\mathcal{F})(n)) \leq \mu((D\mathcal{F})(n)) = \mu(D).$$

Recall the random infinite word $\mathbf{W} = \alpha_1\alpha_2\cdots$ and its associated random walk from Section 7.1. As D is prefix-free, $\mu(D) = \mathbb{P}(\mathbf{W} \text{ hits } D)$. It suffices to show this probability is at most ε. Let m be the largest integer with $1 + m|w| \leq \ell - |w|$. If $\mathbf{W}$ hits D, then $\mathbf{W}$ hits C_ℓ and thus, since (C_t) is a steeplechase, $\mathbf{W}$ hits each of $C_1, C_{1+|w|}, \ldots, C_{1+m|w|}$. Also, $\mathbf{W}$ must avoid each of $C_1w, C_{1+|w|}w, \ldots, C_{1+m|w|}w$ in order to hit D.

We reveal the letters of $\mathbf{W}$ one by one. We wait until $\mathbf{W}$ hits or avoids C_1, which will certainly be known by the time the length of $\mathbf{W}$ is $\max C_1$. If $\mathbf{W}$ avoids C_1, then $\mathbf{W}$ avoids D. If $\mathbf{W}$ hits C_1, then we reveal the next $|w|$ letters of $\mathbf{W}$ and check if they spell w, which has probability $|\mathcal{A}|^{-|w|}$. If they do, then $\mathbf{W}$ avoids D. If they do not, then we wait until $\mathbf{W}$ hits or avoids $C_{1+|w|}$: note that this has not already happened since (C_t) is spread and so $\min C_{1+|w|} \geq \max C_1 + |w|$. If $\mathbf{W}$ avoids $C_{1+|w|}$, then $\mathbf{W}$ avoids D. If $\mathbf{W}$ hits $C_{1+|w|}$, then we reveal the next $|w|$ letters of $\mathbf{W}$ and check if they spell w, which has probability $|\mathcal{A}|^{-|w|}$. We continue this

procedure with the final check being whether the next $|w|$ letters of $\mathbf{W}$ after it hits $C_{1+m|w|}$ spell w. Note that each check has probability $|\mathcal{A}|^{-|w|}$ and is independent of the previous checks as each check involves only new letters. If $\mathbf{W}$ hits D, then $\mathbf{W}$ must fail each of these checks and so the probability that $\mathbf{W}$ hits D is at most

$$(1 - |\mathcal{A}|^{-|w|})^{m+1}.$$

By taking ℓ (and so m) sufficiently large in terms of ε, $|w|$, and $|\mathcal{A}|$ we may ensure this is at most ε, as required. □

Before proving our key technical result for our density proofs, Lemma 7.34, we will need to define the relative density of B on $C\mathcal{F}$. If $C \subseteq \mathcal{F}$ is finite and I is an interval with $\min I \geq \max C$, then the *relative density* of B in $C\mathcal{F}$ on interval I is

$$d^I_{C\mathcal{F}}(B) := \frac{\mu(B(I) \cap C\mathcal{F})}{\mu(\mathcal{F}(I) \cap C\mathcal{F})}.$$

Suppose that C is also prefix-free. Then, by Observation 7.10, we have that $\mu(\mathcal{F}(I) \cap C\mathcal{F}) = |I| \cdot \mu(C)$. Also, the sets $c\mathcal{F}$ for $c \in C$ partition $C\mathcal{F}$. In particular,

$$\begin{aligned}
d^I_{C\mathcal{F}}(B) &= \frac{1}{|I| \cdot \mu(C)} \sum_{n \in I} \mu(B(n) \cap C\mathcal{F}) \\
&= \sum_{c \in C} \frac{1}{|I| \cdot \mu(C)} \sum_{n \in I} \mu(B(n) \cap c\mathcal{F}) \\
&= \sum_{c \in C} \frac{\mu(c)}{\mu(C)} \cdot d^I_{c\mathcal{F}}(B) \\
&= \frac{1}{\mu(C)} \sum_{c \in C} d^I(B \cap c\mathcal{F}) \\
&= \frac{d^I(B \cap C\mathcal{F})}{\mu(C)}.
\end{aligned}$$

Recall that for every set B that we consider in this chapter and every word $w \in \mathcal{F}$, the sequence $(d^{I_j}_{w\mathcal{F}}(B))$ converges to $d^{I_\infty}_{w\mathcal{F}}(B)$ as $j \to \infty$. Therefore, the third line of the displayed equation implies that the sequence $(d^{I_j}_{C\mathcal{F}}(B))$ converges to some limit $d^{I_\infty}_{C\mathcal{F}}(B)$. Again, these limits are additive.

Observation 7.32. *Let $C \subseteq \mathcal{F}$ be finite and prefix-free. Then*

$$d^{I_\infty}_{C\mathcal{F}}(B) = \sum_{c \in C} \frac{\mu(c)}{\mu(C)} \cdot d^{I_\infty}_{c\mathcal{F}}(B) = \frac{d^{I_\infty}(B \cap C\mathcal{F})}{\mu(C)}.$$

In particular, the first equality of Observation 7.32 implies the following.

Observation 7.33. *Let $C \subseteq \mathcal{F}$ be finite and prefix-free. Then $d^{I_\infty}_{C\mathcal{F}}(B) \leq d^{I_\infty}_{\sup}(B)$.*

We now prove the key technical lemma for our density results.

Lemma 7.34. *Let $\varepsilon > 0$ and let $A, B \subseteq \mathcal{F}$. If (C_t) is an ε-capturing steeplechase for A with $\mu(C_1) \geq 2\varepsilon + \varepsilon^{1/3}$, then*

$$d^{I_\infty}_{C_1\mathcal{F}}(AB) \geq d^{I_\infty}_{\sup}(B) - 3\varepsilon^{1/3} \geq d^{I_\infty}_{C_1\mathcal{F}}(B) - 3\varepsilon^{1/3}.$$

Proof. Let $w \in \mathcal{F}$ be such that

$$d^{I_\infty}_{w\mathcal{F}}(B) \geq d^{I_\infty}_{\sup}(B) - \varepsilon^{1/3}.$$

Apply Lemma 7.31 to (C_t) and w to obtain an ℓ such that if $C := C_1 \cup \cdots \cup C_\ell$ and $n \geq \max C_\ell + |w|$, then $\mu((C_1\mathcal{F} \setminus Cw\mathcal{F})(n)) \leq 2\varepsilon$.

We may greedily choose $\widetilde{C} \subseteq C$ (starting with shorter words first) such that $\widetilde{C}w$ is prefix-free and $\widetilde{C}w\mathcal{F} = Cw\mathcal{F}$. Note that

$$\begin{aligned} 2\varepsilon &\geq \mu((C_1\mathcal{F} \setminus \widetilde{C}w\mathcal{F})(n)) \geq \mu((C_1\mathcal{F})(n)) - \mu((\widetilde{C}w\mathcal{F})(n)) \\ &= \mu(C_1) - \mu(\widetilde{C}w) \geq 2\varepsilon + \varepsilon^{1/3} - \mu(\widetilde{C}w) \end{aligned}$$

and so $\mu(\widetilde{C}w) \geq \varepsilon^{1/3}$.

Let I be an interval with $\min I \geq \max C_\ell + |w|$. Note that every word in $\widetilde{C}w\mathcal{F} = Cw\mathcal{F}$ has a prefix in $C_1 \cup \cdots \cup C_\ell$. Since (C_t) is a steeplechase, every such word has a prefix in C_1. Thus $\widetilde{C}w\mathcal{F} \subseteq C_1\mathcal{F}$ and so

$$\begin{aligned} d^{I}_{C_1\mathcal{F}}(AB) &= \frac{1}{|I| \cdot \mu(C_1)} \sum_{n \in I} \mu((AB)(n) \cap C_1\mathcal{F}) \\ &\geq \frac{1}{|I| \cdot \mu(C_1)} \sum_{n \in I} \mu((AB)(n) \cap \widetilde{C}w\mathcal{F}) \\ &\geq \frac{1}{|I|} \sum_{n \in I} \frac{\mu((AB)(n) \cap \widetilde{C}w\mathcal{F})}{\mu(\widetilde{C}w) + 2\varepsilon}. \end{aligned}$$

Using the fact that $x/(y + 2\varepsilon) \geq x/y - 2\varepsilon x/y^2 \geq x/y - 2\varepsilon^{1/3}$ for all $\varepsilon > 0$, $x \in [0, 1]$, and $y \geq \varepsilon^{1/3}$, we get that

$$d^{I}_{C_1\mathcal{F}}(AB) \geq d^{I}_{\widetilde{C}w\mathcal{F}}(AB) - 2\varepsilon^{1/3}.$$

Setting $I := I_j$ and taking j to infinity gives

$$d^{I_\infty}_{C_1\mathcal{F}}(AB) \geq d^{I_\infty}_{\widetilde{C}w\mathcal{F}}(AB) - 2\varepsilon^{1/3}. \tag{7.2}$$

Finally,

$$\begin{aligned} d^{I_\infty}_{\widetilde{C}w\mathcal{F}}(AB) &= \sum_{c\in\widetilde{C}} \frac{\mu(c)}{\mu(\widetilde{C})} \cdot d^{I_\infty}_{cw\mathcal{F}}(AB) \\ &\geq \sum_{c\in\widetilde{C}} \frac{\mu(c)}{\mu(\widetilde{C})} \cdot d^{I_\infty}_{cw\mathcal{F}}(cB) \\ &= \sum_{c\in\widetilde{C}} \frac{\mu(c)}{\mu(\widetilde{C})} \cdot d^{I_\infty}_{w\mathcal{F}}(B) \\ &= d^{I_\infty}_{w\mathcal{F}}(B) \geq d^{I_\infty}_{\sup}(B) - \varepsilon^{1/3}, \end{aligned}$$

where the first equality used Observation 7.32, the first inequality used the fact that $c \in \widetilde{C} \subseteq C \subseteq A$, the second equality used Lemma 7.12, and the second inequality is due to the choice of w. Combining this with (7.2) gives the first inequality claimed in the lemma, while Observation 7.33 gives the second. □

7.3.2 DENSITY OF (STRONGLY) k-PRODUCT-FREE SETS

We now prove Theorem 7.14, making use of the machinery developed in the preceding section. Part (a) has a simple iterative proof which uses the fact that a strongly k-product-free S is disjoint from each of $S^2, S^3, \ldots, S^k$.

Proof of Theorem 7.14 (a). Let $S \subseteq \mathcal{F}$ be strongly k-product-free and let w be any word of $\mathcal{F}$. It suffices to show that $d^{I_\infty}_{w\mathcal{F}}(S) \leq 1/k$. Since $d^{I_\infty}_{w\mathcal{F}}(S) = d^{I_\infty}_{w\mathcal{F}}(S \cap w\mathcal{F})$, we may restrict S to $S \cap w\mathcal{F}$ which remains strongly k-product-free. Thus, we may and will assume that $S \subseteq w\mathcal{F}$. If $d^{I_\infty}_{w\mathcal{F}}(S) = 0$, then we are done. Otherwise, by the definition of $d^{I}_{w\mathcal{F}}$ and the fact that $S \subseteq w\mathcal{F}$, $d^{I_\infty}(S) = \mu(w) \cdot d^{I_\infty}_{w\mathcal{F}}(S) > 0$.

Let $\varepsilon > 0$ be sufficiently small compared to $d^{I_\infty}(S)$ and let (C_t) be an ε-capturing steeplechase for S, as given by Lemma 7.30. Then $\mu(C_1) \geq d^{I_\infty}(S) - \varepsilon \geq 2\varepsilon + \varepsilon^{1/3}$ and so we may use Lemma 7.34 applied to S and any other subset $B \subseteq \mathcal{F}$. We will show by induction that, for $1 \leq i \leq k$,

$$d^{I_\infty}_{C_1\mathcal{F}}(S \cup S^2 \cup \cdots \cup S^i) \geq i \cdot d^{I_\infty}_{C_1\mathcal{F}}(S) - 3(i-1) \cdot \varepsilon^{1/3}.$$

For $i = 1$ this holds with equality. Suppose now that it holds for some $1 \leq i < k$.

Lemma 7.34 implies that

$$\begin{aligned} d^{I_\infty}_{C_1\mathcal{F}}(S^2 \cup S^3 \cup \cdots \cup S^{i+1}) &\geq d^{I_\infty}_{C_1\mathcal{F}}(S \cup S^2 \cup \cdots \cup S^i) - 3\varepsilon^{1/3} \\ &\geq i \cdot d^{I_\infty}_{C_1\mathcal{F}}(S) - 3i \cdot \varepsilon^{1/3}. \end{aligned}$$

Since $i < k$ and S is strongly k-product-free, S is disjoint from $S^2 \cup S^3 \cup \cdots \cup S^{i+1}$ and so

$$d^{I_\infty}_{C_1\mathcal{F}}(S \cup S^2 \cup \cdots \cup S^{i+1}) \geq (i+1) \cdot d^{I_\infty}_{C_1\mathcal{F}}(S) - 3i \cdot \varepsilon^{1/3},$$

which completes the induction. Taking $i = k$ we obtain

$$1 \geq d^{I_\infty}_{C_1\mathcal{F}}(S \cup S^2 \cup \cdots \cup S^k) \geq k \cdot d^{I_\infty}_{C_1\mathcal{F}}(S) - 3(k-1) \cdot \varepsilon^{1/3}.$$

Note that $C_1 \subseteq S \subseteq w\mathcal{F}$ is prefix-free, so $\mu(C_1) \leq \mu(w)$. Since (C_t) is ε-capturing for S, Observation 7.32 and the second bullet point of Lemma 7.30 imply that

$$d^{I_\infty}_{C_1\mathcal{F}}(S) = \frac{d^{I_\infty}(S \cap C_1\mathcal{F})}{\mu(C_1)} \geq \frac{d^{I_\infty}(S \cap C_1\mathcal{F})}{\mu(w)} \geq \frac{d^{I_\infty}(S) - \varepsilon}{\mu(w)} = d^{I_\infty}_{w\mathcal{F}}(S) - \frac{\varepsilon}{\mu(w)},$$

where the final equality used the fact that $d^{I_\infty}(S) = \mu(w) \cdot d^{I_\infty}_{w\mathcal{F}}(S)$ as noted at the beginning of this proof. Putting everything together, we have that

$$1 \geq k \cdot d^{I_\infty}_{w\mathcal{F}}(S) - k \cdot \frac{\varepsilon}{\mu(w)} - 3(k-1) \cdot \varepsilon^{1/3}.$$

Since ε may be arbitrarily small, it follows that $d^{I_\infty}_{w\mathcal{F}}(S) \leq 1/k$, as required. □

The argument for part (b), that is, for k-product-free sets, is more involved. Note that it is not necessary to keep track of the error term depending on ε as we eventually take ε to zero. We therefore introduce some notation to simplify the argument. Write $x \lesssim y$ to mean that $x \leq y + f(\varepsilon)$ where the error term $f(\varepsilon)$ goes to zero as ε goes to zero (in all cases $f(\varepsilon)$ will be a polynomial in $\varepsilon^{1/3}$).

To improve the clarity of the argument and to motivate the proof for arbitrary k-product-free sets, we first sketch a proof which shows that if S is 3-product-free, then $d^{I_\infty}(S) \leq 1/3$.

Proof that $d^{I_\infty}(S) \leq 1/3$ if S is 3-product-free. Let $S \subseteq \mathcal{F}$ be 3-product-free and let $\varepsilon > 0$ be sufficiently small. Let (C_t) be an ε-capturing steeplechase for $S_{1,2} = S \cap S^2$ (recall Definition 7.25), as given by Lemma 7.30. Since S is 3-product-free, $S_{1,2}$ is strongly 3-product-free.

We claim that $d^{I_\infty}(S \cap C_1\mathcal{F}) \lesssim 1/3 \cdot \mu(C_1)$. If $\mu(C_1) < 2\varepsilon + \varepsilon^{1/3}$, then this is immediate. Otherwise $\mu(C_1) \geq 2\varepsilon + \varepsilon^{1/3}$ and so, by Lemma 7.34, we have that $d^{I_\infty}_{C_1\mathcal{F}}(S_{1,2}S) \gtrsim d^{I_\infty}_{C_1\mathcal{F}}(S)$. Since S is 3-product-free, S and $S_{1,2}S$ are disjoint and so

$$d^{I_\infty}_{C_1\mathcal{F}}(S \cup S_{1,2}S) \gtrsim 2 \cdot d^{I_\infty}_{C_1\mathcal{F}}(S).$$

Then, applying Lemma 7.34 again, we get that

$$d^{I_\infty}_{C_1\mathcal{F}}(S_{1,2}S \cup S^2_{1,2}S) = d^{I_\infty}_{C_1\mathcal{F}}(S_{1,2}(S \cup S_{1,2}S)) \gtrsim d^{I_\infty}_{C_1\mathcal{F}}(S \cup S_{1,2}S) \gtrsim 2 \cdot d^{I_\infty}_{C_1\mathcal{F}}(S).$$

Since S is 3-product-free, S is disjoint from $S_{1,2}S \cup S^2_{1,2}S$ and so

$$1 \geq d^{I_\infty}_{C_1\mathcal{F}}(S \cup S_{1,2}S \cup S^2_{1,2}S) \gtrsim 3 \cdot d^{I_\infty}_{C_1\mathcal{F}}(S).$$

Observation 7.32 then gives $d^{I_\infty}(S \cap C_1\mathcal{F}) \lesssim 1/3 \cdot \mu(C_1)$, as claimed.

So far we have bounded the density of S on the part of $\mathcal{F}$ where $S_{1,2}$ is dense. We now bound the density of S on the rest. Let $S' := S \setminus (S_{1,2} \cup C_1\mathcal{F})$ and (D_t) be an ε-capturing steeplechase for S', as given by Lemma 7.30. By passing to a subsequence we may and will assume that $\min D_1 > \max C_1$. Since S is 3-product-free and $S' \cap S^2 = \varnothing$, we know that S' is strongly 3-product-free.

We claim that $d^{I_\infty}(S' \cap D_1\mathcal{F}) \lesssim 1/3 \cdot \mu(D_1)$. If $\mu(D_1) < 2\varepsilon + \varepsilon^{1/3}$, this is immediate. Otherwise, by Lemma 7.34, we have that $d^{I_\infty}_{D_1\mathcal{F}}(S'S) \gtrsim d^{I_\infty}_{D_1\mathcal{F}}(S) \geq d^{I_\infty}_{D_1\mathcal{F}}(S')$. Now S' and $S'S$ are disjoint since $S' \cap S^2 = \varnothing$. Thus

$$d^{I_\infty}_{D_1\mathcal{F}}(S' \cup S'S) \gtrsim 2 \cdot d^{I_\infty}_{D_1\mathcal{F}}(S').$$

Then, applying Lemma 7.34 again, we get that

$$d^{I_\infty}_{D_1\mathcal{F}}((S')^2 \cup (S')^2S) = d^{I_\infty}_{D_1\mathcal{F}}(S'(S' \cup S'S)) \gtrsim 2 \cdot d^{I_\infty}_{D_1\mathcal{F}}(S').$$

Since S is 3-product-free and $S' \cap S^2 = \varnothing$, S' is disjoint from $(S')^2 \cup (S')^2S$ and so

$$1 \geq d^{I_\infty}_{D_1\mathcal{F}}(S' \cup (S')^2 \cup (S')^2S) \gtrsim 3 \cdot d^{I_\infty}_{D_1\mathcal{F}}(S').$$

Observation 7.32 then gives $d^{I_\infty}(S' \cap D_1\mathcal{F}) \lesssim 1/3 \cdot \mu(D_1)$, as claimed.

By the definition of S' and since $\min D_1 > \max C_1$, it follows that $C_1\mathcal{F}$ and $D_1\mathcal{F}$ are disjoint (see proof of Proposition 7.35 for more details). In particular, C_1 and

D_1 are disjoint and their union is prefix-free. Hence $\mu(C_1) + \mu(D_1) \leq 1$. Thus,

$$d^{I_\infty}(S \cap C_1\mathcal{F}) + d^{I_\infty}(S' \cap D_1\mathcal{F}) \lesssim 1/3 \cdot (\mu(C_1) + \mu(D_1)) \leq 1/3.$$

Since (C_t) and (D_t) are ε-capturing, it follows (see the proof of Claim 7.37 below) that very little of S lies outside $(S \cap C_1\mathcal{F}) \cup (S' \cap D_1\mathcal{F})$. In particular, we get that $d^{I_\infty}(S) \lesssim 1/3$. Since ε may be arbitrarily small, it follows that $d^{I_\infty}(S) \leq 1/3$. □

For general k, the argument is a more involved version of the argument above. We will first consider some S_{A_1}, take an ε-capturing steeplechase (C_t) for S_{A_1}, and show that the relative density of S in $C_1\mathcal{F}$ is at most $1/\rho$. We then repeat this step for some S_{A_2}, S_{A_3}, and so on. In future steps we may use the fact that we have dealt with previous S_{A_i}. Proposition 7.35 says that if we have found a suitable sequence $A_1, A_2, \ldots$, then we obtain the required upper bound on $d^{I_\infty}_{\sup}(S)$. Proposition 7.38 then shows that for each k there is a suitable sequence. The combination of these two results proves Theorem 7.14 (b).

Note that in the statement below, the set dA_ℓ is the sumset

$$dA_\ell := \{a_1 + \cdots + a_d : a_1, \ldots, a_d \in A_\ell\}.$$

Proposition 7.35. *Let $k \geq 2$ be an integer and $\rho := \rho(k)$. Let $A_1, \ldots, A_m \subseteq \mathbb{N}$ be a sequence of sets with $A_m = \{1\}$. Suppose that for all $\ell \in [m]$ there exist positive integers $d_1, \ldots, d_{\rho-1}$ (which may depend on ℓ) such that, for all $1 \leq i \leq j \leq \rho - 1$, either*

- *$k \in \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$ or*
- *$A_{\ell'} \subseteq \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$ for some $1 \leq \ell' < \ell$.*

Then, if $S \subseteq \mathcal{F}$ is k-product-free, we have $d^{I_\infty}_{\sup}(S) \leq 1/\rho(k)$.

Proof. Let w be any word of $\mathcal{F}$. It suffices to show that $d^{I_\infty}_{w\mathcal{F}}(S) \leq 1/\rho$. Since $d^{I_\infty}_{w\mathcal{F}}(S) = d^{I_\infty}_{w\mathcal{F}}(S \cap w\mathcal{F})$, we may restrict S to $S \cap w\mathcal{F}$ which remains k-product-free. Thus, we may and will assume that $S \subseteq w\mathcal{F}$.

Let $\varepsilon > 0$ be sufficiently small. We define the following sets and steeplechases. Let $S^{(1)} := S$, $R^{(1)} := S_{A_1} \cap S^{(1)}$, and let $(C_t^{(1)})$ be an ε-capturing steeplechase for $R^{(1)}$, as given by Lemma 7.30. For $\ell = 2, 3, \ldots, m$, do the following:

- Let $S^{(\ell)} := S \setminus (S_{A_1} \cup \cdots \cup S_{A_{\ell-1}} \cup (C_1^{(1)} \cup \cdots \cup C_1^{(\ell-1)})\mathcal{F})$.
- Let $R^{(\ell)} := S_{A_\ell} \cap S^{(\ell)}$.

- Let $(C_t^{(\ell)})$ be an ε-capturing steeplechase for $R^{(\ell)}$. By passing to a subsequence we may and will assume that $\min C_1^{(\ell)} > \max C_1^{(\ell-1)}$.

Claim 7.36. *For each $\ell \in [m]$, we have $\mu(C_1^{(\ell)}) \gtrsim \rho \cdot d^{I_\infty}(S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})$.*

Proof. For clarity, let $R := R^{(\ell)}$ in this proof. First, if $\mu(C_1^{(\ell)}) < 2\varepsilon + \varepsilon^{1/3}$, then

$$\rho \cdot d^{I_\infty}(S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F}) \leq \rho \cdot d^{I_\infty}(C_1^{(\ell)}\mathcal{F}) = \rho \cdot \mu(C_1^{(\ell)}) \lesssim \mu(C_1^{(\ell)}).$$

Hence, we may and will assume that $\mu(C_1^{(\ell)}) \geq 2\varepsilon + \varepsilon^{1/3}$. We will prove by downwards induction on $1 \leq i \leq \rho$ that

$$d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S \cup R^{d_i}S \cup R^{d_i+d_{i+1}}S \cup \cdots \cup R^{d_i+\cdots+d_{\rho-1}}S) \gtrsim (\rho - i + 1) \cdot d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)}).$$

This holds for $i = \rho$ because $S^{(\ell)} \subseteq S$. Suppose now that it holds for some $1 < i \leq \rho$. Since $\mu(C_1^{(\ell)}) \geq 2\varepsilon + \varepsilon^{1/3}$, repeatedly applying Lemma 7.34 shows that $d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(R^{d_{i-1}}B) \gtrsim d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(B)$ for any set $B \subseteq \mathcal{F}$. In particular,

$$\begin{aligned} d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}&(R^{d_{i-1}}S \cup R^{d_{i-1}+d_i}S \cup \cdots \cup R^{d_{i-1}+\cdots+d_{\rho-1}}S) \\ &\gtrsim d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S \cup R^{d_i}S \cup R^{d_i+d_{i+1}}S \cup \cdots \cup R^{d_i+\cdots+d_{\rho-1}}S) \\ &\gtrsim (\rho - i + 1) \cdot d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)}). \end{aligned}$$

Consider the sets $S^{(\ell)}$ and $R^{d_{i-1}}S \cup R^{d_{i-1}+d_i}S \cup \cdots \cup R^{d_{i-1}+\cdots+d_{\rho-1}}S$. First note that $R^{d_{i-1}+\cdots+d_j} \subseteq S_{(d_{i-1}+\cdots+d_j)A_\ell}$. Hence, if the sets $S^{(\ell)}$ and $R^{d_{i-1}+\cdots+d_j}S$ meet, then $S^{(\ell)} \cap S_{1+(d_{i-1}+\cdots+d_j)A_\ell} \neq \varnothing$. By the assumption from the proposition statement, this implies that either $S^{(\ell)} \cap S_k \neq \varnothing$ or $S^{(\ell)} \cap S_{A_{\ell'}} \neq \varnothing$ for some $\ell' < \ell$. The k-product-freeness of S rules out the former and the definition of $S^{(\ell)}$ the latter. Therefore, $S^{(\ell)}$ and $R^{d_{i-1}}S \cup R^{d_{i-1}+d_i}S \cup \cdots \cup R^{d_{i-1}+\cdots+d_{\rho-1}}S$ are disjoint and so,

$$d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)} \cup R^{d_{i-1}}S \cup R^{d_{i-1}+d_i}S \cup \cdots \cup R^{d_{i-1}+\cdots+d_{\rho-1}}S) \gtrsim (\rho - i + 2) \cdot d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)}).$$

Finally, noting that $S^{(\ell)} \subseteq S$ completes the induction. Now, observe that the case $i = 1$ shows that

$$1 \geq d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S \cup R^{d_1}S \cup R^{d_1+d_2}S \cup \cdots \cup R^{d_1+\cdots+d_{\rho-1}}S) \gtrsim \rho \cdot d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)}).$$

The claim then follows from Observation 7.32. □

We next show that very little of S has not been captured by the previous claim.

Claim 7.37. *For all sufficiently large n, we have $\mu(S(n) \setminus \bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})) \leq m \cdot \varepsilon$.*

Proof. For each ℓ we know that $(C_t^{(\ell)})$ is an ε-capturing steeplechase for $R^{(\ell)}$ and so, for all sufficiently large n,

$$\mu((R^{(\ell)} \setminus C_1^{(\ell)}\mathcal{F})(n)) \leq \varepsilon.$$

Hence, it is enough to show that $S \setminus \bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F}) \subseteq \bigcup_\ell (R^{(\ell)} \setminus C_1^{(\ell)}\mathcal{F})$. Fix $w \in S \setminus \bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})$ and let ℓ' be maximal such that $w \in S^{(\ell')}$. Such an ℓ' exists because $S^{(1)} = S$. Since $w \in S \setminus \bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})$, we have $w \notin C_1^{(\ell')}\mathcal{F}$. We claim that $w \in S_{A_{\ell'}}$. If $\ell' = m$, then this is immediate since $w \in S = S_1 = S_{A_m}$. If $\ell' < m$, then, by the maximality of ℓ', we must have $w \in S_{A_{\ell'}} \cup C_1^{(\ell')}\mathcal{F}$ and so $w \in S_{A_{\ell'}}$. Thus, $w \in (S_{A_{\ell'}} \cap S^{(\ell')}) \setminus C_1^{(\ell')}\mathcal{F} = R^{(\ell')} \setminus C_1^{(\ell')}\mathcal{F}$, as required. □

We claim that $C_1^{(1)}\mathcal{F}, \ldots, C_1^{(m)}\mathcal{F}$ are pairwise disjoint. Indeed, if not, then there exists some $w_i \in C_1^{(i)}$ that is a prefix of some $w_j \in C_1^{(j)}$ for $i \neq j$. By construction, $\min C_1^{(\ell)} > \max C_1^{(\ell-1)}$ for all ℓ and so $i < j$. However, $w_j \in C_1^{(j)} \subseteq R^{(j)} \subseteq S^{(j)}$ and so $w_j \notin C_1^{(i)}\mathcal{F}$ by definition of $S^{(j)}$, which gives a contradiction.

In particular, $C_1^{(1)}, \ldots, C_1^{(m)}$ are pairwise disjoint and their union is prefix-free. Since these sets are all subsets of $S \subseteq w\mathcal{F}$, we get that $\sum_\ell \mu(C_1^{(\ell)}) \leq \mu(w)$.

We can now show that $d^{I_\infty}_{w\mathcal{F}}(S) \leq 1/\rho$. Summing Claim 7.36 over ℓ and then applying Claim 7.37 gives

$$\mu(w) \gtrsim \rho \cdot d^{I_\infty}\left(\bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})\right) \gtrsim \rho \cdot d^{I_\infty}(S).$$

By the definition of $d^I_{w\mathcal{F}}$ and the fact that $S \subseteq w\mathcal{F}$, $d^{I_\infty}_{w\mathcal{F}}(S) = d^{I_\infty}(S)/\mu(w) \lesssim 1/\rho$. Since ε may be arbitrarily small, it follows that $d^{I_\infty}_{w\mathcal{F}}(S) \leq 1/\rho$, as required. □

We now show that there is always a sequence of sets satisfying the requirements of Proposition 7.35. The sequence chosen here is motivated by the proofs of Propositions 7.27 and 7.28, together with a small amount of case work for small values of k.

Proposition 7.38. *For every integer $k \geq 2$ there is a sequence $A_1, \ldots, A_m \subseteq \mathbb{N}$ satisfying the hypothesis of Proposition 7.35.*

Proof. We deal with the cases $k = 2, 3, 5, 7, 13$ first.

- $k = 2$: take $A_1 = \{1\}$ (with $d_1 = 1$),
- $k = 3$: take $A_1 = \{1, 2\}$ (with $d_1 = d_2 = 1$) and $A_2 = \{1\}$ (with $d_1 = d_2 = 1$),
- $k = 5$: take $A_1 = \{1, 3\}$ (with $d_1 = d_2 = 2$) and $A_2 = \{1\}$ (with $d_1 = d_2 = 2$),
- $k = 7$: take

$$\begin{aligned} A_1 &= \{1, 3\} && \text{(with } d_1 = d_2 = d_3 = 2\text{)}, \\ A_2 &= \{1, 2\} && \text{(with } d_1 = d_2 = d_3 = 1\text{)}, \\ A_3 &= \{1, 4\} && \text{(with } d_1 = d_2 = d_3 = 1\text{), and} \\ A_4 &= \{1\} && \text{(with } d_1 = d_2 = d_3 = 1\text{),} \end{aligned}$$

- $k = 13$: take

$$\begin{aligned} A_1 &= \{1, 4\} && \text{(with } d_1 = d_2 = d_3 = d_4 = 3\text{)}, \\ A_2 &= \{1, 2\} && \text{(with } d_1 = d_2 = d_3 = d_4 = 3\text{)}, \\ A_3 &= \{1, 3, 5, 7\} && \text{(with } d_1 = d_2 = d_3 = d_4 = 1\text{)}, \\ A_4 &= \{1, 3\} && \text{(with } d_1 = d_2 = d_3 = d_4 = 1\text{)}, \\ A_5 &= \{1, 5\} && \text{(with } d_1 = d_2 = d_3 = d_4 = 1\text{), and} \\ A_6 &= \{1\} && \text{(with } d_1 = d_2 = d_3 = d_4 = 1\text{).} \end{aligned}$$

We now turn to $k \notin \{2, 3, 5, 7, 13\}$. For positive integers d and t, let

$$B_{d,t} := \{1, d + 1, 2d + 1, \ldots, td + 1\}.$$

We construct $A_1, A_2, \ldots$ by taking all the sets $B_{d,t}$ for which $1 \leq d \leq \rho - 1$ and $1 \leq t < (k - 1)/d$ in the order of decreasing d and then decreasing t, and adding the set $\{1\}$ to the end of the sequence.

Consider a set $A_\ell = B_{d,t}$. We need to show that A_ℓ satisfies Proposition 7.35. Let s be the maximum positive integer with $ds \leq \rho - 1$, and let α be the minimum positive integer with $\alpha d(d + 1) \geq k - 1$. For $1 \leq i \leq \rho - 1$, define

$$d_i := \begin{cases} \alpha d & \text{if } i \equiv 0 \bmod s + 1 \text{ and } \alpha \neq 2, \text{ and} \\ d & \text{otherwise.} \end{cases}$$

For any $1 \leq i \leq j \leq \rho - 1$, we have that $d_i + d_{i+1} + \cdots + d_j = \beta d$ for some integer β satisfying $1 \leq \beta \leq \rho - 1 + (\alpha - 1)(\rho - 1)/(s + 1)$. Moreover, by definition of

the d_i, either $\beta \geq \alpha$ or $\beta \leq s$. Note that

$$1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell = \{1 + \beta d, 1 + (\beta + 1)d, \ldots, 1 + \beta d(td + 1)\}.$$

If $\beta = 1$, then $B_{d,t+1} \subseteq \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$. Now, either $k \in B_{d,t+1}$ or $B_{d,t+1} = A_{\ell'}$ for some $\ell' < \ell$ and so A_ℓ satisfies Proposition 7.35.

If $1 < \beta \leq s$, then $B_{\beta d,1} \subseteq \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$. Note that because $d < \beta d \leq ds \leq \rho - 1$, we have $B_{\beta d,1} = A_{\ell'}$ for some $\ell' < \ell$ and so A_ℓ satisfies Proposition 7.35.

If $\beta \geq \alpha$, we claim that $k \in 1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell$. Since $k - 1$ is a multiple of d, it suffices to show that $\beta d \leq k - 1 \leq \beta d(td + 1)$. Firstly,

$$\beta d(td + 1) \geq \alpha d(d + 1) \geq k - 1.$$

For the second inequality, if $d \leq \rho - 2$, we have $\alpha d(d + 1) = k - 1$ as observed in the proof of Proposition 7.28. Furthermore, we have

$$\beta d \leq (\rho + (\alpha - 1)\rho/(s + 1))d \leq (\rho + (\alpha - 1)d)d$$

where the second inequality follows from $d(s + 1) \geq \rho$. This is less than $k - 1 = \alpha d(d + 1)$ if $d(\alpha - 1) + \rho \leq \alpha(d + 1)$, which is true for $k \notin \{2, 3, 5, 7, 13\}$ as shown in the proof of Proposition 7.28. On the other hand, if $d = \rho - 1$, we have

$$\beta d \leq (d + (\alpha - 1)d/(s + 1))d \leq ((\alpha + 1)/2)d^2 \leq ((\alpha + 1)/2)d(d + 1).$$

If $\alpha \geq 3$, this is at most $(\alpha - 1)d(d + 1) < k - 1$. If $\alpha \leq 2$, we can observe that $\beta \leq \rho - 1$ to obtain $\beta d \leq \rho(\rho - 1) \leq k - 1$ where the last inequality was shown in the proof of Proposition 7.27. In all cases, we have $\beta d \leq k - 1$ as required. Hence, $k \in 1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell$, and so A_ℓ satisfies Proposition 7.35.

Finally, for $A_\ell = \{1\}$, we can simply pick $d_1 = \cdots = d_{\rho-1} = 1$. We then get that $B_{j-i+1,1} \subseteq \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$. Since $1 \leq j - i + 1 \leq \rho - 1$, we have $B_{j-i+1,1} = A_{\ell'}$ for some $\ell' < \ell$ and so A_ℓ satisfies Proposition 7.35. □

Propositions 7.35 and 7.38 together imply Theorem 7.14 (b), and so this completes the proof of Theorem 7.14, as promised.

7.4 PRODUCT-FREE SETS IN FREE GROUPS

We now adapt our methods to free groups and prove Theorem 7.8. Throughout, $\mathcal{G}$ denotes a free group on a finite alphabet $\mathcal{A}$ and $S \subseteq \mathcal{G}$ is a k-product-free set whose density we want to bound. We always assume that all words are in reduced form. Moreover, AB denotes the *product* of two sets $A, B \subseteq \mathcal{G}$ without cancellation, that is,

$$AB := \{w \in \mathcal{G} : w = ab \text{ as a substring decomposition with } a \in A \text{ and } b \in B\}.$$

For example, if $A := \{ba^{-1}, c\}$ and $B := \{a, d\}$, then $AB = \{ba^{-1}d, ca, cd\}$ does not contain b even though $b = ba^{-1} \cdot a$, since b cannot be decomposed into two substrings where the first is in A and the second is in B. In particular, $C\mathcal{G}$ consists of all words with a prefix in C. We equip $\mathcal{G}$ with the measure μ defined for all $w \in \mathcal{G}$ as $\mu(w) := 1/|\mathcal{G}(|w|)|$. If $\mathbf{W} := \alpha_1\alpha_2\cdots$ is a random infinite word where each letter α_{i+1} is chosen uniformly at random from $(\mathcal{A} \cup \mathcal{A}^{-1}) \setminus \{\alpha_i^{-1}\}$, then

$$\mu(w) = \mathbb{P}(\mathbf{W} \text{ hits } w) = \begin{cases} (2|\mathcal{A}|)^{-1}(2|\mathcal{A}|-1)^{-(|w|-1)} & \text{if } |w| > 0, \text{ and} \\ 1 & \text{if } |w| = 0. \end{cases}$$

As before, for $B \subseteq \mathcal{G}$, $\mu(B) := \sum_{w\in B} \mu(w)$ is the expected number of times that $\mathbf{W}$ hits B. In particular, we can make the following observations corresponding to Observations 7.9 and 7.10.

Observation 7.39. *If $C \subseteq \mathcal{G}$ is prefix-free, then $\mu(C) \leq 1$, $\mu((C\mathcal{G})(n)) \leq \mu(C)$ for all n, and $\mu((C\mathcal{G})(n)) = \mu(C)$ if C is finite and $n \geq \max C$.*

We now define the relative density of subsets of $\mathcal{G}$ as follows.

Definition 7.40 (Relative density)**.** Let $B \subseteq \mathcal{G}$ and $\mathcal{H}$ be a subsemigroup of $\mathcal{G}$ with $\mathcal{H}(n) \neq \varnothing$ for all sufficiently large n. If I is an interval, then the *relative density* of B in $\mathcal{H}$ on interval I is

$$d^I_{\mathcal{H}}(B) := \frac{\mu(B(I) \cap \mathcal{H})}{\mu(\mathcal{H}(I))} = \frac{1}{\mu(\mathcal{H}(I))} \sum_{n\in I} \mu(B(n) \cap \mathcal{H}).$$

If $\mathcal{H}(I) = \varnothing$, we will take the relative density to be 0 by convention. We also define $d^I(B) := d^I_{\mathcal{G}}(B)$.

The *upper Banach density* of B is then $d^*(B) := \limsup_{I\to\infty} d^I(B)$. At this point, we can again diagonalise to obtain a sequence (I_j) so that, for every $\mathcal{H}$ and B that we

consider in our proofs, $(d^{I_j}_{\mathcal{H}}(B))$ converges to a limit $d^{I_\infty}_{\mathcal{H}}(B)$, and $d^{I_\infty}(S) = d^*(S)$. These limits are again additive. Also note that $d^I_{\mathcal{H}}(B) = d^I(B \cap \mathcal{H})/d^I(\mathcal{H})$ and so $d^{I_\infty}_{\mathcal{H}}(B) = d^{I_\infty}(B \cap \mathcal{H})/d^{I_\infty}(\mathcal{H})$ if $d^{I_\infty}(\mathcal{H}) > 0$. We define sup density as follows.

Definition 7.41 (Sup density)**.** For a set $B \subseteq \mathcal{G}$, the *sup density* of B in a subsemigroup $\mathcal{H}$ of $\mathcal{G}$ is defined as

$$d^{I_\infty}_{\sup \mathcal{H}}(B) := \sup_{w \in \mathcal{H}} d^{I_\infty}_{w\mathcal{G} \cap \mathcal{H}}(B).$$

From now on, let $\mathcal{H} := \mathcal{G}^{\alpha\beta} \subseteq \mathcal{G}$ be the subsemigroup of $\mathcal{G}$ consisting of all words starting with α and ending with β, where $\alpha, \beta \in \mathcal{A} \cup \mathcal{A}^{-1}$ are such that $\alpha \neq \beta^{-1}$. A random sequence argument shows the following.

Observation 7.42. *Let $C \subseteq \mathcal{H}$ be finite and prefix-free. Then, for all $n \geq \max C + 2$,*

$$\mu((C\mathcal{G} \cap \mathcal{H})(n)) \geq \frac{\mu(C)}{(2|\mathcal{A}| - 1)^2}.$$

In particular, $d^{I_\infty}(w\mathcal{G} \cap \mathcal{H}) > 0$ for all $w \in \mathcal{H}$. As in Lemma 7.12, subtree densities of $\mathcal{G}$ satisfy the property that we may strip away prefixes.

Lemma 7.43. *If $w, v \in \mathcal{H}$, then $d^{I_\infty}_{wv\mathcal{G} \cap \mathcal{H}}(wB) = d^{I_\infty}_{v\mathcal{G} \cap \mathcal{H}}(B)$.*

Proof. For $u \in \mathcal{H}$, $\mu(wu) = a \cdot \mu(u)$, where $a := (2|\mathcal{A}| - 1)^{-|w|}$. So, if $C \subseteq \mathcal{H}$ is finite, then $\mu(wC) = a \cdot \mu(C)$. Let I be any interval with $\min I \geq |wv|$. Then

$$d^I_{wv\mathcal{G} \cap \mathcal{H}}(wB) = \frac{\mu((wB)(I) \cap wv\mathcal{G} \cap \mathcal{H})}{\mu((wv\mathcal{G} \cap \mathcal{H})(I))} = \frac{\mu(B(I - |w|) \cap v\mathcal{G} \cap \mathcal{H})}{\mu((v\mathcal{G} \cap \mathcal{H})(I - |w|))},$$

where we used that $wv\mathcal{G} \cap \mathcal{H} = w(v\mathcal{G} \cap \mathcal{H})$. For any $C \subseteq \mathcal{H}$, note that

$$|\mu(C(I)) - \mu(C(I - |w|))| = \left| \sum_{n \in I} \mu(C(n)) - \sum_{n \in I - |w|} \mu(C(n)) \right| \leq |w|,$$

where we used the fact that $\mu(C(n)) \in [0, 1]$. Therefore,

$$\frac{\mu(B(I) \cap v\mathcal{G} \cap \mathcal{H}) - |w|}{\mu((v\mathcal{G} \cap \mathcal{H})(I)) + |w|} \leq d^I_{wv\mathcal{G} \cap \mathcal{H}}(wB) \leq \frac{\mu(B(I) \cap v\mathcal{G} \cap \mathcal{H}) + |w|}{\mu((v\mathcal{G} \cap \mathcal{H})(I)) - |w|}.$$

Set $I = I_j$ and take j to infinity. As $d^{I_\infty}(v\mathcal{G} \cap \mathcal{H}) > 0$, we get $\mu((v\mathcal{G} \cap \mathcal{H})(I_j)) \to \infty$. Hence, both bounds tend to $d^{I_\infty}_{v\mathcal{G} \cap \mathcal{H}}(B)$ and so $d^{I_\infty}_{wv\mathcal{G} \cap \mathcal{H}}(wB) = d^{I_\infty}_{v\mathcal{G} \cap \mathcal{H}}(B)$. □

We can also obtain the following analogues to Observations 7.32 and 7.33.

Observation 7.44. *Let $C \subseteq \mathcal{H}$ be finite and prefix-free. Then*

$$d^{I_\infty}_{C\mathcal{G}\cap\mathcal{H}}(B) = \sum_{c\in C} \frac{d^{I_\infty}(c\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(C\mathcal{G}\cap\mathcal{H})} \cdot d^{I_\infty}_{c\mathcal{G}\cap\mathcal{H}}(B) = \frac{d^{I_\infty}(B\cap C\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(C\mathcal{G}\cap\mathcal{H})}.$$

Proof. Since $d^{I_\infty}(c\mathcal{G}\cap\mathcal{H}) > 0$ for all $c \in C$ and therefore also $d^{I_\infty}(C\mathcal{G}\cap\mathcal{H}) > 0$, we have

$$d^{I_\infty}_{c\mathcal{G}\cap\mathcal{H}}(B) = \frac{d^{I_\infty}(B\cap c\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(c\mathcal{G}\cap\mathcal{H})} \quad \text{and} \quad d^{I_\infty}_{C\mathcal{G}\cap\mathcal{H}}(B) = \frac{d^{I_\infty}(B\cap C\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(C\mathcal{G}\cap\mathcal{H})}.$$

Since d^{I_∞} is additive, this implies that

$$\begin{aligned} d^{I_\infty}_{C\mathcal{G}\cap\mathcal{H}}(B) &= \frac{d^{I_\infty}(B\cap C\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(C\mathcal{G}\cap\mathcal{H})} \\ &= \sum_{c\in C} \frac{d^{I_\infty}(B\cap c\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(C\mathcal{G}\cap\mathcal{H})} = \sum_{c\in C} \frac{d^{I_\infty}(c\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(C\mathcal{G}\cap\mathcal{H})} \cdot d^{I_\infty}_{c\mathcal{G}\cap\mathcal{H}}(B). \end{aligned}$$

□

The first equality of Observation 7.44 implies the following.

Observation 7.45. *Let $C \subseteq \mathcal{H}$ be finite and prefix-free. Then $d^{I_\infty}_{C\mathcal{G}\cap\mathcal{H}}(B) \leq d^{I_\infty}_{\sup\mathcal{H}}(B)$.*

Steeplechases in free groups can be defined exactly as for free semigroups.

Definition 7.46 (Steeplechase)**.** An infinite sequence (C_t) of subsets of $\mathcal{H}$ is a *steeplechase* if, for each positive integer t,

- each C_t is prefix-free and finite, and
- every word in C_{t+1} has a proper prefix in C_t (in particular, $C_{t+1}\mathcal{G} \subseteq C_t\mathcal{G}$).

We say that a steeplechase (C_t) is *spread* if $\max C_t < \min C_{t+1}$ for all t and *ε-tight* if for all m, n we have that $|\mu(C_m) - \mu(C_n)| \leq \varepsilon$.

The following lemma is an analogue to Lemma 7.30, showing that any set $B \subseteq \mathcal{H}$ can be approximated by an ε-tight and spread steeplechase.

Lemma 7.47. *Let $\varepsilon > 0$ and $B \subseteq \mathcal{H}$. Then, there exists an ε-tight spread steeplechase (C_t) such that*

- *$C_1 \cup C_2 \cup \cdots \subseteq B$,*
- *for all t and all n sufficiently large in terms of t, $\mu((B \setminus C_t\mathcal{G})(n)) \leq \varepsilon$, and*
- *for all t, $\mu(C_t)/d^{I_\infty}(\mathcal{H}) \geq d^{I_\infty}_{\mathcal{H}}(B) - \varepsilon$.*

Proof. This is very similar to the proof of Lemma 7.30. Recall that the headcount $h(w)$ of $w \in B$ is the number of $b \in B$ that are prefixes of w. For each positive integer t, let $D_t := \{w \in B : h(w) = t\}$. Iteratively, we perform the following for each positive integer t.

1. Let ℓ_t be such that $\mu(D_t(\{\ell_t + 1, \ell_t + 2, \dots\})) \leq \varepsilon \cdot d^{I_\infty}(\mathcal{H})/2^t$.
2. Let $C_t := D_t(\{1, 2, \dots, \ell_t\})$.
3. Remove $(D_t \setminus C_t)\mathcal{G}$ from B and from all D_ℓ with $\ell > t$.

This ensures that (C_t) is a steeplechase and $C_1 \cup C_2 \cup \cdots \subseteq B$. Fix t and let $n > \max\{\ell_1, \dots, \ell_t\}$. Then,

$$\mu((B \setminus C_t\mathcal{G})(n)) \leq \sum_r \mu(((D_r \setminus C_r)\mathcal{G})(n)) \leq \sum_r \mu(D_r \setminus C_r) \leq \varepsilon \cdot d^{I_\infty}(\mathcal{H}) \leq \varepsilon.$$

Finally, we get that $\mu(B(n)) \leq \mu((C_t\mathcal{G})(n)) + \varepsilon \cdot d^{I_\infty}(\mathcal{H}) = \mu(C_t) + \varepsilon \cdot d^{I_\infty}(\mathcal{H})$. Averaging this over $n \in I_j$ and taking $j \to \infty$ gives $d^{I_\infty}(B) \leq \mu(C_t) + \varepsilon \cdot d^{I_\infty}(\mathcal{H})$ and therefore $d^{I_\infty}_{\mathcal{H}}(B) = d^{I_\infty}(B)/d^{I_\infty}(\mathcal{H}) \leq \mu(C_t)/d^{I_\infty}(\mathcal{H}) + \varepsilon$. By passing to a subsequence, we may assume that (C_t) is spread and ε-tight. □

We call a steeplechase (C_t) given by Lemma 7.47 an *ε-capturing steeplechase* for B. There is also an analogue to Lemma 7.31, showing for any word $w \in \mathcal{H}$ that $\bigcup_{\ell \geq 1}(C_\ell w\mathcal{G} \cap \mathcal{H})$ covers almost all of $C_1\mathcal{G} \cap \mathcal{H}$.

Lemma 7.48. *Let (C_t) be an ε-tight spread steeplechase. Then, for every $w \in \mathcal{H}$ there is an ℓ such that if $C := C_1 \cup C_2 \cup \cdots \cup C_\ell$ and $n \geq \max C_\ell + |w|$, we have that*

$$\mu(((C_1\mathcal{G} \cap \mathcal{H}) \setminus (Cw\mathcal{G} \cap \mathcal{H}))(n)) \leq 2\varepsilon.$$

Proof. Note that $(C_1\mathcal{G} \cap \mathcal{H}) \setminus (Cw\mathcal{G} \cap \mathcal{H}) = (C_1\mathcal{G} \setminus Cw\mathcal{G}) \cap \mathcal{H} \subseteq C_1\mathcal{G} \setminus Cw\mathcal{G}$, and so it suffices to show that $\mu((C_1\mathcal{G} \setminus Cw\mathcal{G})(n)) \leq 2\varepsilon$.

We can show this exactly as in the proof of Lemma 7.31, with the only change being that $\mathbf{W}$ needs to be the random walk from the beginning of this section. As a consequence, if $\mathbf{W}$ hits C_t, the probability that the next $|w|$ letters of $\mathbf{W}$ spell w is $(2|\mathcal{A}| - 1)^{-|w|}$. Importantly, this uses the fact that the last letter of a word in C_t is β and the first letter of w is α, and $\alpha \neq \beta^{-1}$. □

We now prove the key technical lemma for our density results in free groups. This corresponds to Lemma 7.34.

Lemma 7.49. *Let $\varepsilon > 0$ and let $A, B \subseteq \mathcal{H}$. If (C_t) is an ε-capturing steeplechase for A with $\mu(C_1) \geq (2|\mathcal{A}| - 1)^2(2\varepsilon + \varepsilon^{1/3})$, then*

$$d^{I_\infty}_{C_1\mathcal{G}\cap\mathcal{H}}(AB) \geq d^{I_\infty}_{\sup \mathcal{H}}(B) - 3\varepsilon^{1/3} \geq d^{I_\infty}_{C_1\mathcal{G}\cap\mathcal{H}}(B) - 3\varepsilon^{1/3}.$$

Proof. By Observation 7.42, we know that for $n \geq \max C_1 + 2$ it holds that $\mu((C_1\mathcal{G} \cap \mathcal{H})(n)) \geq \mu(C_1)/(2|\mathcal{A}| - 1)^2 \geq 2\varepsilon + \varepsilon^{1/3}$. We proceed as in the proof of Lemma 7.34. Let $w \in \mathcal{H}$ be such that

$$d^{I_\infty}_{w\mathcal{G}\cap\mathcal{H}}(B) \geq d^{I_\infty}_{\sup \mathcal{H}}(B) - \varepsilon^{1/3}.$$

Apply Lemma 7.48 to (C_t) and w to obtain an ℓ such that if $C \coloneqq C_1 \cup \cdots \cup C_\ell$ and $n \geq \max C_\ell + |w|$, then $\mu(((C_1\mathcal{G} \cap \mathcal{H}) \setminus (Cw\mathcal{G} \cap \mathcal{H}))(n)) \leq 2\varepsilon$. We may greedily choose $\widetilde{C} \subseteq C$ such that $\widetilde{C}w$ is prefix-free and $\widetilde{C}w\mathcal{G} = Cw\mathcal{G}$. Note that

$$\begin{aligned} 2\varepsilon &\geq \mu(((C_1\mathcal{G} \cap \mathcal{H}) \setminus (\widetilde{C}w\mathcal{G} \cap \mathcal{H}))(n)) \\ &\geq \mu((C_1\mathcal{G} \cap \mathcal{H})(n)) - \mu((\widetilde{C}w\mathcal{G} \cap \mathcal{H})(n)) \\ &\geq 2\varepsilon + \varepsilon^{1/3} - \mu((\widetilde{C}w\mathcal{G} \cap \mathcal{H})(n)) \end{aligned}$$

and so $\mu((\widetilde{C}w\mathcal{G} \cap \mathcal{H})(n)) \geq \varepsilon^{1/3}$ and $\mu((C_1\mathcal{G} \cap \mathcal{H})(n)) \leq \mu((\widetilde{C}w\mathcal{G} \cap \mathcal{H})(n)) + 2\varepsilon$.

Let I be an interval with $\min I \geq \max C_\ell + |w|$, so $\mu((\widetilde{C}w\mathcal{G} \cap \mathcal{H})(I)) \geq |I|\varepsilon^{1/3}$ and $\mu((C_1\mathcal{G} \cap \mathcal{H})(I)) \leq \mu((\widetilde{C}w\mathcal{G} \cap \mathcal{H})(I)) + |I|2\varepsilon$. Note that $\widetilde{C}w\mathcal{G} \subseteq C_1\mathcal{G}$ and so

$$\begin{aligned} d^{I}_{C_1\mathcal{G}\cap\mathcal{H}}(AB) &= \frac{\mu((AB)(I) \cap C_1\mathcal{G} \cap \mathcal{H})}{\mu((C_1\mathcal{G} \cap \mathcal{H})(I))} \\ &\geq \frac{\mu((AB)(I) \cap \widetilde{C}w\mathcal{G} \cap \mathcal{H})}{\mu((C_1\mathcal{G} \cap \mathcal{H})(I))} \\ &\geq \frac{\mu((AB)(I) \cap \widetilde{C}w\mathcal{G} \cap \mathcal{H})}{\mu((\widetilde{C}w\mathcal{G} \cap \mathcal{H})(I)) + |I|2\varepsilon}. \end{aligned}$$

Using the fact that $x/(y + |I|2\varepsilon) \geq x/y - |I|2\varepsilon x/y^2 \geq x/y - 2\varepsilon^{1/3}$ for all $\varepsilon > 0$, $0 \leq x \leq |I|$, and $y \geq |I|\varepsilon^{1/3}$, we get that

$$d^{I}_{C_1\mathcal{G}\cap\mathcal{H}}(AB) \geq d^{I}_{\widetilde{C}w\mathcal{G}\cap\mathcal{H}}(AB) - 2\varepsilon^{1/3}.$$

Setting $I \coloneqq I_j$ and taking j to infinity gives

$$d^{I_\infty}_{C_1\mathcal{G}\cap\mathcal{H}}(AB) \geq d^{I_\infty}_{\widetilde{C}w\mathcal{G}\cap\mathcal{H}}(AB) - 2\varepsilon^{1/3}. \tag{7.3}$$

Finally,

$$\begin{aligned} d^{I_\infty}_{\widetilde{C}w\mathcal{G}\cap\mathcal{H}}(AB) &= \sum_{c\in\widetilde{C}} \frac{d^{I_\infty}(cw\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(\widetilde{C}w\mathcal{G}\cap\mathcal{H})}\cdot d^{I_\infty}_{cw\mathcal{G}\cap\mathcal{H}}(AB) \\ &\geq \sum_{c\in\widetilde{C}} \frac{d^{I_\infty}(cw\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(\widetilde{C}w\mathcal{G}\cap\mathcal{H})}\cdot d^{I_\infty}_{cw\mathcal{G}\cap\mathcal{H}}(cB) \\ &= \sum_{c\in\widetilde{C}} \frac{d^{I_\infty}(cw\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(\widetilde{C}w\mathcal{G}\cap\mathcal{H})}\cdot d^{I_\infty}_{w\mathcal{G}\cap\mathcal{H}}(B) \\ &= d^{I_\infty}_{w\mathcal{G}\cap\mathcal{H}}(B) \geq d^{I_\infty}_{\sup\mathcal{H}}(B) - \varepsilon^{1/3}, \end{aligned}$$

where the first equality used Observation 7.44, the first inequality used the fact that $c \in \widetilde{C} \subseteq C \subseteq A$, the second equality used Lemma 7.43, and the second inequality is due to the choice of w. Combining this with (7.3) gives the first inequality claimed in the lemma, while Observation 7.45 gives the second. □

At this point, we have recovered all important technical results that we needed to bound the density of (strongly) k-product-free sets in free semigroups. We can now simply use exactly the same arguments as in Section 7.3.2. We only need to replace $w\mathcal{F}$ by $w\mathcal{G}\cap\mathcal{H}$, $\mu(w)$ by $d^{I_\infty}(w\mathcal{G}\cap\mathcal{H})$, $C_i\mathcal{F}$ by $C_i\mathcal{G}\cap\mathcal{H}$, $\mu(C_i)$ by $d^{I_\infty}(C_i\mathcal{G}\cap\mathcal{H})$, $d^{I_\infty}_{\sup}$ by $d^{I_\infty}_{\sup\mathcal{H}}$, and all references by references to the corresponding results in this section to prove the following analogue of Theorem 7.14.

Theorem 7.50. *Let $k \geq 2$ be an integer, $\mathcal{A}$ be a finite set, $\mathcal{G}$ be the free group on alphabet $\mathcal{A}$, and $\mathcal{H} := \mathcal{G}^{\alpha\beta}$ be the subsemigroup of $\mathcal{G}$ consisting of all words starting with α and ending with β, where $\alpha, \beta \in \mathcal{A}\cup\mathcal{A}^{-1}$ are such that $\alpha \neq \beta^{-1}$. Let (I_j) be a sequence of intervals, with $I_j \to \infty$, on which all relevant densities converge.*

(a) If $S \subseteq \mathcal{H}$ is strongly k-product-free, then $d^{I_\infty}_{\sup\mathcal{H}}(S) \leq 1/k$.

(b) If $S \subseteq \mathcal{H}$ is k-product-free, then $d^{I_\infty}_{\sup\mathcal{H}}(S) \leq 1/\rho(k)$.

The arguments from Ortega, Rué, and Serra [ORS24] show that a density bound on (strongly) k-product-free sets in $\mathcal{G}^{\alpha\beta}$ immediately translates to a density bound in $\mathcal{G}$. Therefore, Theorems 7.3 and 7.8 are immediate corollaries of Theorem 7.50.

7.5 OPEN PROBLEMS

A first natural problem left open is to determine the structure of the extremal k-product-free sets in free semigroups for $k \in \{3,5,7,13\}$. For $k \in \{5,7,13\}$, we conjecture that the extremal sets are exactly as in Theorem 7.6. The extremal sets

for $k = 3$ will be slightly more complicated. Indeed, while $1 + 3\mathbb{N}$ and $2 + 3\mathbb{N}$ are maximal 3-sum-free subsets of the non-negative integers of density $1/3$, so are $\{1, 2\} + 6\mathbb{N}$ and $\{5, 6\} + 6\mathbb{N}$. Łuczak and Schoen [ŁS97] showed that there are no others. We conjecture that the corresponding result holds for free semigroups.

Conjecture 7.51. *Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{F}$ is 3-product-free and $d^*(S) = 1/3$, then one of the following holds. Either it is possible to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/3\mathbb{Z}$ such that S is a subset of*

$$\{w \in \mathcal{F} : \textit{the sum of the labels of letters in } w \textit{ is } 1 \bmod 3\},$$

or it is possible to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/6\mathbb{Z}$ such that S is a subset of

$$\{w \in \mathcal{F} : \textit{the sum of the labels of letters in } w \textit{ is } 1 \textit{ or } 2 \bmod 6\}.$$

Łuczak [Luc95] also proved that every sum-free subset of the non-negative integers with density greater than $2/5$ is a subset of the odd integers. Later, Łuczak and Schoen [ŁS97] proved similar results for (strongly) k-sum-free sets. However, such strengthenings for strongly k-product-free subsets of free semigroups are false as the constant $1/k$ in Theorem 7.4 cannot be replaced by anything smaller. For example, let T be the set of words of length 1 mod k, let v be any word, and let $p > k$ be a prime. Define

$$\begin{aligned} T' := \{&w \in T : \text{neither } v \text{ nor } w \text{ is a prefix or suffix of the other}\} \\ &\cup \{vwv : w \in \mathcal{F} \text{ and } vwv \text{ has length 0 mod } k \text{ and 1 mod } p\}. \end{aligned}$$

Then T' is strongly k-product-free, has density at least $1/k - 2 \cdot |\mathcal{A}|^{-|v|}$, and there is no labelling of the letters of $\mathcal{A}$ by $\mathbb{Z}/k\mathbb{Z}$ in which every word of T' has sum 1 mod k. In fact, a set of positive density would need to be removed before this happens. Nonetheless, T' is a small perturbation from T. Hence, it is natural to ask whether there is some form of stability. We conjecture the following for $k = 2$.

Conjecture 7.52. *Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. For each $\delta > 0$, there is some $\varepsilon > 0$ such that if $S \subseteq \mathcal{F}$ is product-free and $d^*(S) > 1/2 - \varepsilon$, then there exists an odd-occurrence set O_Γ such that $d^*(S \setminus O_\Gamma) < \delta$.*

We do not have examples corresponding to T' for k-product-free sets. It would be interesting to determine whether such examples exist or if Theorem 7.6 can be strengthened to densities below $1/\rho(k)$.

Theorems 7.4 and 7.6 give the structure of extremal (strongly) k-product-free sets in free semigroups. In free groups, Theorems 7.3 and 7.8 solve the density problem, but the structure problem remains open. The simplest open case is the following conjecture.

Conjecture 7.53. *Let $\mathcal{A}$ be a finite set and $\mathcal{G}$ be the free group on alphabet $\mathcal{A}$. If $S \subseteq \mathcal{G}$ is product-free and $d^*(S) = 1/2$, then the following holds. It is possible to label each letter of $\mathcal{A} \cup \mathcal{A}^{-1}$ with a label in $\mathbb{Z}/2\mathbb{Z}$ such that the label of α^{-1} is the negation of the label of α for all $\alpha \in \mathcal{A}$ and S is a subset of*

$$T := \{w \in \mathcal{G} : \textit{the sum of the labels of letters in } w \textit{ is } 1 \bmod 2\}.$$

For strongly k-product-free sets we expect the above conjecture to hold with 2 replaced by k. For k-product-free sets we expect the behaviour to be the same as for free semigroups.

We remark that our methods do give some structure. Similar arguments to Section 7.2.1 show there is a labelling of all words in the subsemigroup $\mathcal{G}^{\alpha\beta}$ (defined in Section 7.4) such that the label of a concatenation is the sum of the individual labels and all words in $S \cap \mathcal{G}^{\alpha\beta}$ have label 1 mod k. What is missing is an understanding of how the labellings interact when letters cancel during concatenation.

Bibliography

[ACFK16] Miklós Abért, Péter Csikvári, Péter E. Frenkel, and Gábor Kun (2016). Matchings in Benjamini–Schramm convergent graph sequences. *Transactions of the American Mathematical Society* **368**(6), 4197–4218. ↑53

[ACJ+26] Maria Axenovich, Wouter Cames van Batenburg, Oliver Janzer, Lukas Michel, and Mathieu Rundström (2026). An improved upper bound for the multicolour Ramsey number of odd cycles. *Journal of Combinatorial Theory, Series B* **179**, 293–298. ↑iii, 14, 68

[ACP+21] Romain Ageron, Paul Casteras, Thibaut Pellerin, Yann Portella, Arpad Rimmel, and Joanna Tomasik (2021). New lower bounds for Schur and weak Schur numbers. arXiv:2112.03175. ↑13, 65

[AEH80] Jin Akiyama, Geoffrey Exoo, and Frank Harary (1980). Covering and packing in graphs. III. Cyclic and acyclic invariants. *Mathematica Slovaca* **30**(4), 405–417. ↑6, 21, 22

[AEH81] Jin Akiyama, Geoffrey Exoo, and Frank Harary (1981). Covering and packing in graphs. IV. Linear arboricity. *Networks* **11**(1), 69–72. ↑22

[Alo88] Noga Alon (1988). The linear arboricity of graphs. *Israel Journal of Mathematics* **62**(3), 311–325. ↑6, 22

[And69] Ian Anderson (1969). A variance method in combinatorial number theory. *Glasgow Mathematical Journal* **10**, 126–129. ↑86

[AS16] Noga Alon and Joel H. Spencer (2016). The probabilistic method. Fourth edn. (John Wiley & Sons). ↑4, 6, 22, 47

[AZ18] Martin Aigner and Günter M. Ziegler (2018). Proofs from THE BOOK. Sixth edn. (Springer, Berlin). ↑49

[Bár78] Imre Bárány (1978). A short proof of Kneser's conjecture. *Journal of Combinatorial Theory, Series A* **25**(3), 325–326. ↑58

[BCC+13] Eli Berger, Krzysztof Choromanski, Maria Chudnovsky, Jacob Fox, Martin Loebl, Alex Scott, Paul Seymour, and Stéphan Thomassé (2013). Tournaments and colouring. *Journal of Combinatorial Theory, Series B* **103**(1), 1–20. ↑10, 55

[BE73] JOHN A. BONDY and PAUL ERDŐS (1973). Ramsey numbers for cycles in graphs. *Journal of Combinatorial Theory, Series B* **14**, 46–54. ↑14, 66

[BEK51] N. G. DE BRUIJN, C. VAN EBBENHORST TENGBERGEN, and D. KRUYSWIJK (1951). On the set of divisors of a number. *Nieuw Archief voor Wiskunde (2)* **23**, 191–193. ↑86

[BKLO16] BEN BARBER, DANIELA KÜHN, ALLAN LO, and DERYK OSTHUS (2016). Edge-decompositions of graphs with high minimum degree. *Advances in Mathematics* **288**, 337–385. ↑4

[BLS19] MATIJA BUCIĆ, SHOHAM LETZTER, and BENNY SUDAKOV (2019). Monochromatic paths in random tournaments. *Random Structures & Algorithms* **54**(1), 69–81. ↑86

[BM24] MATIJA BUCIĆ and RICHARD MONTGOMERY (2024). Towards the Erdős-Gallai cycle decomposition conjecture. *Advances in Mathematics* **437**, Paper No. 109434. ↑5

[BPR16] R. BALASUBRAMANIAN, GYAN PRAKASH, and D. S. RAMANA (2016). Sum-free subsets of finite abelian groups of type III. *European Journal of Combinatorics* **58**, 181–202. ↑17, 90

[Bre73] LEV MEEROVICH BREGMAN (1973). Some properties of nonnegative matrices and their permanents. *Soviet Mathematics-Doklady* **14**, 945–949. ↑47

[BS85] LÁSZLÓ BABAI and VERA T. SÓS (1985). Sidon sets in groups and induced subgraphs of Cayley graphs. *European Journal of Combinatorics* **6**(2), 101–114. ↑17, 90

[BS22] MARCIN BRIAŃSKI and ADAM SZADY (2022). A short note on graphs with long Thomason chains. *Discrete Mathematics* **345**(1), Paper No. 112624. ↑42

[Cau40] AUGUSTIN LOUIS CAUCHY (1840). Exercices d'analyse et de physique mathématique, vol. 1 (Bachelier). ↑44

[CDG+25a] MICHA CHRISTOPH, NEMANJA DRAGANIĆ, ANTÓNIO GIRÃO, EOIN HURLEY, LUKAS MICHEL, and ALP MÜYESSER (2025). Cycle-factors of regular graphs via entropy. *2025 IEEE 66th Annual Symposium on Foundations of Computer Science*, 615–619. ↑iii, 9, 45

[CDG+25b] MICHA CHRISTOPH, NEMANJA DRAGANIĆ, ANTÓNIO GIRÃO, EOIN HURLEY, LUKAS MICHEL, and ALP MÜYESSER (2025). New bounds for linear arboricity and related problems. *Journal of the European Mathematical Society*, to appear. arXiv:2507.20500. ↑iii, 7, 8, 24

[CE96] NEIL J. CALKIN and PAUL ERDŐS (1996). On a class of aperiodic sum-free sets. *Mathematical Proceedings of the Cambridge Philosophical Society* **120**(1), 1–5. ↑18, 89, 90

[CG98] FAN R. K. CHUNG and RONALD L. GRAHAM (1998). Erdős on graphs: His legacy of unsolved problems (AK Peters/CRC Press). ↑13, 65

[CGMS26] Marcelo Campos, Simon Griffiths, Robert Morris, and Julian Sahasrabudhe (2026). An exponential improvement for diagonal Ramsey. *Annals of Mathematics* **203**(3), 869–932. ↑13

[CKL+16] Béla Csaba, Daniela Kühn, Allan Lo, Deryk Osthus, and Andrew Treglown (2016). Proof of the 1-factorization and Hamilton decomposition conjectures. *Memoirs of the American Mathematical Society* **244**(1154), v+164. ↑4

[Des54] Blanche Descartes (1954). *k*-chromatic graphs without triangles. *American Mathematical Monthly* **61**(5), 352–353. ↑10, 55

[DGMS25] Nemanja Draganić, Stefan Glock, David Munhá Correia, and Benny Sudakov (2025). Optimal Hamilton covers and linear arboricity for random graphs. *Proceedings of the American Mathematical Society* **153**(3), 921–935. ↑22

[Dir52] Gabriel Andrew Dirac (1952). Some theorems on abstract graphs. *Proceedings of the London Mathematical Society (3)* **2**, 69–81. ↑25

[DJ17] A. Nicholas Day and J. Robert Johnson (2017). Multicolour Ramsey numbers of odd cycles. *Journal of Combinatorial Theory, Series B* **124**, 56–63. ↑14, 66, 67

[DMM+24] Nemanja Draganić, Richard Montgomery, David Munhá Correia, Alexey Pokrovskiy, and Benny Sudakov (2024). Hamiltonicity of expanders: optimal bounds and applications. arXiv:2402.06603. ↑25

[DMS26] Nemanja Draganić, David Munhá Correia, and Benny Sudakov (2026). Pancyclicity of Hamiltonian graphs. *Journal of the European Mathematical Society* **28**(9), 4047–4062. ↑25

[DP21] Michelle Delcourt and Luke Postle (2021). Progress towards Nash-Williams' conjecture on triangle decompositions. *Journal of Combinatorial Theory, Series B* **146**, 382–416. ↑4

[Ebe16] Sean Eberhard (2016). Product mixing in the alternating group. *Discrete Analysis* Paper No. 2, 18. ↑17, 90

[EFRS78] Paul Erdős, Ralph J. Faudree, Cecil C. Rousseau, and Richard H. Schelp (1978). On cycle-complete graph Ramsey numbers. *Journal of Graph Theory* **2**(1), 53–64. ↑70

[EG75] Paul Erdős and Ronald L. Graham (1975). On partition theorems for finite graphs. *Infinite and finite sets, Colloquia Mathematica Societatis János Bolyai*, vol. 10, 515–527. ↑14, 66, 67

[Ego81] Gregory P. Egorychev (1981). The solution of van der Waerden's problem for permanents. *Advances in Mathematics* **42**(3), 299–305. ↑48

[EGP66] Paul Erdős, A. W. Goodman, and Lajos Pósa (1966). The representation of a graph by set intersections. *Canadian Journal of Mathematics* **18**, 106–112. ↑5

[Eli20] Shalom Eliahou (2020). An adaptive upper bound on the Ramsey numbers $R(3,\dots,3)$. *Integers* **20**, Paper No. A54, 7. ↑65

[Eno81] Hikoe Enomoto (1981). The linear arboricity of 5-regular graphs. Technical report, Department of Information Science, University of Tokyo. ↑22

[EP84] Hikoe Enomoto and Bernard Péroche (1984). The linear arboricity of some regular graphs. *Journal of Graph Theory* **8**(2), 309–324. ↑22

[ER50] Paul Erdős and Richard Rado (1950). A combinatorial theorem. *Journal of the London Mathematical Society* **25**(4), 249–255. ↑15, 72

[Erd47] Paul Erdős (1947). Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society* **53**, 292–294. ↑13

[Erd59] Paul Erdős (1959). Graph theory and probability. *Canadian Journal of Mathematics* **11**, 34–38. ↑10

[Erd63] Paul Erdős (1963). On a problem in graph theory. *The Mathematical Gazette* **47**, 220–223. ↑59

[ERT80] Paul Erdős, Arthur L. Rubin, and Herbert Taylor (1980). Choosability in graphs. *Congressus Numerantium* **26**, 125–157. ↑5

[ES35] Paul Erdős and George Szekeres (1935). A combinatorial problem in geometry. *Compositio Mathematica* **2**, 463–470. ↑13, 71

[Fal81] Dmitry I. Falikman (1981). Proof of the van der Waerden conjecture regarding the permanent of a doubly stochastic matrix. *Matematicheskie Zametki* **29**(6), 931–938. ↑48

[FF22] Uriel Feige and Ella Fuchs (2022). On the path partition number of 6-regular graphs. *Journal of Graph Theory* **101**(3), 345–378. ↑7, 23, 43

[FFJ20] Asaf Ferber, Jacob Fox, and Vishesh Jain (2020). Towards the linear arboricity conjecture. *Journal of Combinatorial Theory, Series B* **142**, 56–79. ↑22

[Fox] Jacob Fox. Personal communication. ↑14, 66

[FR95] David C. Fisher and Jennifer Ryan (1995). Tournament games and positive tournaments. *Journal of Graph Theory* **19**(2), 217–236. ↑61

[FRS14] Uriel Feige, R. Ravi, and Mohit Singh (2014). Short tours through large linear forests. *Integer Programming and Combinatorial Optimization, IPCO 2014, Lecture Notes in Computer Science*, vol. 8494, 273–284. ↑7, 8, 9, 23, 24, 44, 45

[Gaj26] Rishikesh Gajjala (2026). Counterexamples to an extremal conjecture for random cycle-factors. arXiv:2604.26101. ↑53

[GH24] António Girão and Zach Hunter (2024). Monochromatic odd cycles in edge-coloured complete graphs. arXiv:2412.07708. ↑67

[GHI+24] António Girão, Kevin Hendrey, Freddie Illingworth, Florian Lehner, Lukas Michel, Michael Savery, and Raphael Steiner (2024). Chromatic number is not tournament-local. *Journal of Combinatorial Theory, Series B* **168**, 86–95. ↑iii, 11, 57

[GIM+24] António Girão, Freddie Illingworth, Lukas Michel, Michael Savery, and Alex Scott (2024). Flashes and rainbows in tournaments. *Combinatorica* **44**(3), 675–690. ↑iii, 16, 74

[GJ24] Bertille Granet and Felix Joos (2024). Random perfect matchings in regular graphs. *Random Structures & Algorithms* **64**(1), 3–14. ↑53

[GJK+21] Stefan Glock, Felix Joos, Jaehoon Kim, Daniela Kühn, and Deryk Osthus (2021). Resolution of the Oberwolfach problem. *Journal of the European Mathematical Society* **23**(8), 2511–2547. ↑4

[GKLO23] Stefan Glock, Daniela Kühn, Allan Lo, and Deryk Osthus (2023). The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F. *Memoirs of the American Mathematical Society* **284**(1406), v+131. ↑4

[GKO16] Stefan Glock, Daniela Kühn, and Deryk Osthus (2016). Optimal path and cycle decompositions of dense quasirandom graphs. *Journal of Combinatorial Theory, Series B* **118**, 88–108. ↑22

[GL21] Vytautas Gruslys and Shoham Letzter (2021). Cycle partitions of regular graphs. *Combinatorics, Probability and Computing* **30**(4), 526–549. ↑23, 24, 43

[GMS24] Stefan Glock, David Munhá Correia, and Benny Sudakov (2024). Hamilton cycles in pseudorandom graphs. *Advances in Mathematics* **458**, Paper No. 109984. ↑25

[Gow08] W. T. Gowers (2008). Quasirandom groups. *Combinatorics, Probability and Computing* **17**(3), 363–387. ↑17, 90

[GR05] Ben Green and Imre Z. Ruzsa (2005). Sum-free sets in abelian groups. *Israel Journal of Mathematics* **147**, 157–188. ↑17, 90

[Gre02] Joshua E. Greene (2002). A new short proof of Kneser's conjecture. *The American Mathematical Monthly* **109**(10), 918–920. ↑58

[GS71] Ronald L. Graham and Joel H. Spencer (1971). A constructive solution to a tournament problem. *Canadian Mathematical Bulletin* **14**, 45–48. ↑59

[GS06] Charles M. Grinstead and James L. Snell (2006). Grinstead and Snell's Introduction to Probability (Chance Project). ↑102

[GS24] Yuping Gao and Songling Shan (2024). Linear arboricity of robust expanders. arXiv:2405.18494. ↑22

[Gul86a] Filip Guldan (1986). The linear arboricity of 10-regular graphs. *Mathematica Slovaca* **36**(3), 225–228. ↑22

[Gul86b] Filip Guldan (1986). Some results on linear arboricity. *Journal of Graph Theory* **10**(4), 505–509. ↑22

[Hal35] Philip Hall (1935). On Representatives of Subsets. *Journal of the London Mathematical Society* **10**(1), 26–30. ↑41, 43

[Han18] Jie Han (2018). On vertex-disjoint paths in regular graphs. *Electronic Journal of Combinatorics* **25**(2), Paper No. 2.12. ↑23, 24

[Har70] Frank Harary (1970). Covering and packing in graphs. I. *Annals of the New York Academy of Sciences* **175**(1), 198–205. ↑5, 21

[HLTW19] Ararat Harutyunyan, Tien-Nam Le, Stéphan Thomassé, and Hehui Wu (2019). Coloring tournaments: from local to global. *Journal of Combinatorial Theory, Series B* **138**, 166–171. ↑11, 55, 56

[IMS25] Freddie Illingworth, Lukas Michel, and Alex Scott (2025). The structure and density of k-product-free sets in the free semigroup and group. *Journal of the London Mathematical Society* **111**(1), Paper No. e70046, 36. ↑iii, 19, 92

[JM00] Tao Jiang and Dhruv Mubayi (2000). New upper bounds for a canonical Ramsey problem. *Combinatorica* **20**(1), 141–146. ↑16, 72

[JS21] Matthew Jenssen and Jozef Skokan (2021). Exact Ramsey numbers of odd cycles via nonlinear optimisation. *Advances in Mathematics* **376**, Paper No. 107444. ↑14, 66

[JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda (2004). A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM* **51**(4), 671–697. ↑45, 52

[JY26] Oliver Janzer and Fredy Yip (2026). Short monochromatic odd cycles. *Mathematical Proceedings of the Cambridge Philosophical Society*, 1–8. ↑67

[Kah00] Jeff Kahn (2000). Asymptotics of the list-chromatic index for multigraphs. *Random Structures & Algorithms* **17**(2), 117–156. ↑5

[Kar72] Richard M. Karp (1972). Reducibility among combinatorial problems. *Complexity of computer computations*, The IBM Research Symposia Series, 85–103. ↑10

[Ked09] KIRAN S. KEDLAYA (2009). Product-free subsets of groups, then and now. *Communicating mathematics*, *Contemporary Mathematics*, vol. 479, 169–177. ↑17, 90

[Kee14] PETER KEEVASH (2014). The existence of designs. *Annals of Mathematics*, to appear. arXiv:1401.3665. ↑4

[KKK+23] DONG YEAP KANG, TOM KELLY, DANIELA KÜHN, ABHISHEK METHUKU, and DERYK OSTHUS (2023). Graph and hypergraph colouring via nibble methods: A survey. *European Congress of Mathematics*, vol. 21, 771–823. ↑4

[KKOT19] JAEHOON KIM, DANIELA KÜHN, DERYK OSTHUS, and MYKHAYLO TYOMKYN (2019). A blow-up lemma for approximate decompositions. *Transactions of the American Mathematical Society* **371**(7), 4655–4742. ↑4

[KLM24] PETER KEEVASH, NOAM LIFSHITZ, and DOR MINZER (2024). On the largest product-free subsets of the alternating groups. *Inventiones Mathematicae* **237**(3), 1329–1375. ↑17, 90

[Kne55] MARTIN KNESER (1955). Aufgabe 360. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **58**, 27. ↑58

[KS19] CHARLOTTE KNIERIM and PASCAL SU (2019). Improved bounds on the multicolor Ramsey numbers of paths and even cycles. *Electronic Journal of Combinatorics* **26**(1), Paper No. 1.26, 17. ↑66

[KS22] PETER KEEVASH and KATHERINE STADEN (2022). The generalised Oberwolfach problem. *Journal of Combinatorial Theory, Series B* **152**, 281–318. ↑4

[KS25] PETER KEEVASH and KATHERINE STADEN (2025). Ringel's tree packing conjecture in quasirandom graphs. *Journal of the European Mathematical Society* **27**(5), 1769–1826. ↑4

[KSSS24] MATTHEW KWAN, ASHWIN SAH, MEHTAAB SAWHNEY, and MICHAEL SIMKIN (2024). High-girth Steiner triple systems. *Annals of Mathematics* **200**(3), 1059–1156. ↑4

[LC19] QIZHONG LIN and WEIJI CHEN (2019). New upper bound for multicolor Ramsey number of odd cycles. *Discrete Mathematics* **342**(1), 217–220. ↑14, 66

[Lev03] VSEVOLOD F. LEV (2003). Sharp estimates for the number of sum-free sets. *Journal für die Reine und Angewandte Mathematik* **555**, 1–25. ↑106

[Li09] YUSHENG LI (2009). The multi-color Ramsey number of an odd cycle. *Journal of Graph Theory* **62**(4), 324–328. ↑14, 66, 67

[LL09] YUSHENG LI and KO-WEI LIH (2009). Multi-color Ramsey numbers of even cycles. *European Journal of Combinatorics* **30**(1), 114–118. ↑66

[LLNW20] Imre Leader, Shoham Letzter, Bhargav Narayanan, and Mark Walters (2020). Product-free sets in the free semigroup. *European Journal of Combinatorics* **83**, 103003. ↑17, 19, 90

[LMS26] Shoham Letzter, Abhishek Methuku, and Benny Sudakov (2026). Nearly Hamilton cycles in sublinear expanders and applications. *Journal of the London Mathematical Society* **113**(2), Paper No. e70452, 42. ↑23, 24, 44

[Lov78] László Lovász (1978). Kneser's conjecture, chromatic number, and homotopy. *Journal of Combinatorial Theory, Series A* **25**(3), 319–324. ↑58

[LP23] Richard Lang and Luke Postle (2023). An improved bound for the linear arboricity conjecture. *Combinatorica* **43**(3), 547–569. ↑6, 7, 22

[LR95] Hanno Lefmann and Vojtěch Rödl (1995). On Erdős-Rado numbers. *Combinatorica* **15**(1), 85–104. ↑72

[LRT92] Hanno Lefmann, Vojtěch Rödl, and Robin Thomas (1992). Monochromatic vs multicolored paths. *Graphs and Combinatorics* **8**(4), 323–332. ↑15, 71, 72, 76, 80, 81

[ŁS97] Tomasz Łuczak and Tomasz Schoen (1997). On infinite sum-free sets of natural numbers. *Journal of Number Theory* **66**(2), 211–224. ↑18, 89, 90, 91, 95, 130

[ŁS01] Tomasz Łuczak and Tomasz Schoen (2001). Sum-free subsets of right cancellative semigroups. *European Journal of Combinatorics* **22**(7), 999–1002. ↑95

[Łuc95] Tomasz Łuczak (1995). A note on the density of sum-free sets. *Journal of Combinatorial Theory, Series A* **70**(2), 334–336. ↑17, 89, 95, 130

[Mat03] Jiří Matoušek (2003). Using the Borsuk-Ulam theorem. Lectures on topological methods in combinatorics and geometry. Universitext (Springer, Berlin). ↑42

[MM09] Colton Magnant and Daniel M. Martin (2009). A note on the path cover number of regular graphs. *The Australasian Journal of Combinatorics* **43**, 211–217. ↑7, 23, 43

[MMPS25] Richard Montgomery, Alp Müyesser, Alexey Pokrovskiy, and Benny Sudakov (2025). Approximate path decompositions of regular graphs. *Journal of the London Mathematical Society* **112**(2), Paper No. e70269, 40. ↑23, 24

[Mon24] Richard Montgomery (2024). Transversals in Latin squares. *Surveys in combinatorics 2024*, *London Mathematical Society Lecture Note Series*, vol. 493, 131–158. ↑4

[MPS21] Richard Montgomery, Alexey Pokrovskiy, and Benny Sudakov (2021). A proof of Ringel's conjecture. *Geometric and Functional Analysis* **31**(3), 663–720. ↑4

[MR90] Colin McDiarmid and Bruce Reed (1990). Linear arboricity of random regular graphs. *Random Structures & Algorithms* **1**(4), 443–445. ↑22

[MR00] Michael Molloy and Bruce Reed (2000). Near-optimal list colorings. *Random Structures & Algorithms* **17**(3-4), 376–402. ↑5

[MSS15a] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava (2015). Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Annals of Mathematics* **182**(1), 307–325. ↑42

[MSS15b] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava (2015). Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem. *Annals of Mathematics* **182**(1), 327–350. ↑42

[NSS25] Tung Nguyen, Alex Scott, and Paul Seymour (2025). Some results and problems on tournament structure. *Journal of Combinatorial Theory, Series B* **173**, 146–183. ↑11, 56

[NW61] C. St. J. A. Nash-Williams (1961). Edge-disjoint spanning trees of finite graphs. *Journal of the London Mathematical Society* **36**(1), 445–450. ↑5, 21

[NW64] C. St. J. A. Nash-Williams (1964). Decomposition of finite graphs into forests. *Journal of the London Mathematical Society* **39**(1), 12–12. ↑5, 21

[ORS24] Miquel Ortega, Juanjo Rué, and Oriol Serra (2024). Product-free sets in the free group. *Mathematika* **70**(3), e12255. ↑18, 19, 91, 92, 129

[Pér81] Bernard Péroche (1981). On partition of graphs into linear forests and dissections. Rapport de recherche, Centre National de la Recherche Scientifique. ↑22

[Pet91] Julius Petersen (1891). Die Theorie der regulären graphs. *Acta Mathematica* **15**(1), 193–220. ↑38

[Pós76] Lajos Pósa (1976). Hamiltonian circuits in random graphs. *Discrete Mathematics* **14**(4), 359–364. ↑25

[Rad97] Jaikumar Radhakrishnan (1997). An entropy proof of Bregman's theorem. *Journal of Combinatorial Theory, Series A* **77**(1), 161–164. ↑47, 49

[Sch17] Issai Schur (1917). Über die Kongruenz $x^m + y^m \equiv z^m \pmod{p}$. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **25**, 114–117. ↑12, 13, 65

[Sch78] Alexander Schrijver (1978). Vertex-critical subgraphs of Kneser graphs. *Nieuw Archief voor Wiskunde* **26**(3), 454–461. ↑11, 56, 57, 58

[SS16] Alex Scott and Paul Seymour (2016). Open problems for the Barbados Graph Theory Workshop 2016, problem 1. Accessed on 31 May 2026. ↑11, 55, 61

[SYXL06] Yongqi Sun, Yuansheng Yang, Feng Xu, and Bingxi Li (2006). New lower bounds on the multicolor Ramsey numbers $R_r(C_{2m})$. *Graphs and Combinatorics* **22**(2), 283–288. ↑66

[Tom82] PAVEL TOMASTA (1982). Note on linear arboricity. *Mathematica Slovaca* **32**(3), 239–242. ↑22

[TV17] TERENCE TAO and VAN VU (2017). Sum-free sets in groups: a survey. *Journal of Combinatorics* **8**(3), 541–552. ↑17, 90

[Vis12] NISHEETH K. VISHNOI (2012). A permanent approach to the traveling salesman problem. *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, 76–80. ↑9, 24, 44, 47

[Viz64] VADIM G. VIZING (1964). On an estimate of the chromatic class of a p-graph. *Diskret. Analiz* **3**, 25–30. ↑5, 21

[Viz76] VADIM G. VIZING (1976). Coloring the vertices of a graph in prescribed colors. *Diskret. Analiz* **29**, 3–10. ↑5

[Wu99] JIAN-LIANG WU (1999). On the linear arboricity of planar graphs. *Journal of Graph Theory* **31**(2), 129–134. ↑22

[WW08] JIAN-LIANG WU and YU-WEN WU (2008). The linear arboricity of planar graphs of maximum degree seven is four. *Journal of Graph Theory* **58**(3), 210–220. ↑22

[XXC02] XIAO DONG XU, ZHENG XIE, and ZHI CHEN (2002). Upper bounds for Ramsey numbers $R_n(3)$ and Schur numbers. *Mathematics in Economics* **19**(1), 81–84. ↑65

[Zyk49] ALEXANDER A. ZYKOV (1949). On some properties of linear complexes (in Russian). *Matematicheskii Sbornik* **24/66**(2), 163–188. For an English translation see *American Mathematical Society Translations* **79**, 1952. ↑10, 55