UNIVERSITY OF CALIFORNIA, SAN DIEGO

Extremal Graphs and Additive Combinatorics

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Craig Michael Timmons

Committee in charge:

Professor Jacques Verstraëte, Chair
Professor Fan Chung Graham
Professor Ronald Graham
Professor Ramamohan Paturi
Professor Jeffrey Remmel
Professor Alex Vardy

2014

The dissertation of Craig Michael Timmons is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2014

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to start by thanking my advisor Jacques Verstraëte. His encouragement and support has helped tremendously during my time at UCSD. I am grateful to my committee members for their time and support as well. I would like to thank Sheila, Grace, and Lexi who keep me motivated to work hard. Fan Chung Graham, André Kündgen, Amber Puha, and Todd Kemp have played an important role in my development as a mathematician and instructor. Hooman, Janine, Mike, Jay, Rob, Mark, Andy, and many other graduate students at UCSD have made working there a lot fun. Lastly I would like to thank my parents: Mom, Dad, Deedee, Dan, and Mary.

Chapter 4 is a reprint of the material as it appears in "Sidon sets and graphs without 4-cycles" coauthored with Michael Tait. This paper has been submitted for publication. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 5 is a reprint of the material as it appears in "An ordered Turán problem for bipartite graphs," Timmons, Craig. *Electron. J. Combin.*, 19 (4), P43, 2012. The dissertation author was the primary investigator and author of this paper.

Chapter 6 is a reprint of the material as it appears in "A counterexample to sparse removal" coauthored with Jacques Verstraëte. This paper has been submitted for publication. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 7 is a reprint of the material as it appears in "k-fold Sidon sets" coauthored with Javier Cilleruelo. This paper has been submitted for publication. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 8 is a reprint of the material as it appears in "Upper and lower bounds on B_k^+-sets," Timmons, Craig. *Integers*, A14:1, 1-27, 2014. The dissertation author was the primary investigator and author of this paper.

VITA

2004	B. A. in Liberal Studies, California State University, San Marcos
2007	M. S. in Mathematics, California State University, San Marcos
2009–2014	Graduate Teaching Assistant, Department of Mathematics, University of California, San Diego
2014	Ph. D. in Mathematics, University of California, San Diego

PUBLICATIONS

C. Timmons, "Star coloring high girth planar graphs," *Electron. J. Combin.* 15(1) #R124, 2008.

H. A. Kierstead, A. Kündgen, C. Timmons, "Star coloring planar bipartite graphs," *J. Graph Theory*, 60(1) 1-10, 2009.

A. Kündgen, C. Timmons, "Star coloring planar graphs from small lists," *J. Graph Theory*, 63(4) 324-337, 2010.

S. M. Cioabă, A. Kündgen, C. Timmons, V. V. Vysotsky, "Covering complete r-graphs with spanning complete r-partite r-graphs," *Combin. Probab. Comput.*, 20, 519-527, 2011.

C. Timmons, "An ordered Turán problem for bipartite graphs," *Electron. J. Combin.* 19(4) #P43, 2012.

C. Timmons, "Upper and lower bounds on B_k^+-sets," *Integers*, 14:A1, 1-27, 2014.

X. Peng, C. Timmons, "Infinite Turán problems for bipartite graphs," submitted.

J. Cilleruelo, C. Timmons, "k-fold Sidon sets," submitted.

M. Tait, C. Timmons, "Sidon sets and graphs without 4-cycles," submitted.

C. Timmons, J. Verstraëte, "A counterexample to sparse removal," submitted.

X. Peng, R. Tesoro, C. Timmons, "Bounds for generalized Sidon sets," submitted.

ABSTRACT OF THE DISSERTATION

Extremal Graphs and Additive Combinatorics

by

Craig Michael Timmons

Doctor of Philosophy in Mathematics

University of California, San Diego, 2014

Professor Jacques Verstraëte, Chair

In this thesis we investigate graphs that are constructed using objects from additive number theory. Sidon sets are used to construct C_4-free graphs that show $\mathrm{ex}(q^2-q-2, C_4) \geq \frac{1}{2}q^3 - q^2 - O(q^{3/4})$ whenever q is an odd prime power. This disproves a conjecture of Abreu, Balbuena, and Labbate and improves the current lower bound by $\frac{q}{2}$. Sidon sets are also used to show that there are bipartite graphs F for which the following holds: for infinitely many n there is an n-vertex graph with $\Omega(\mathrm{ex}(n, F))$ edges and has the property that every edge lies in exactly one copy of F. This disproves a conjecture of Solymosi concerning a possible extension of the removal lemma to sparse graphs. An ordered Turán problem for bipartite graphs is introduced and studied. We prove upper bounds on this problem and use B_k-sets to obtain lower bounds. Generalizations of Sidon sets, such as k-fold Sidon sets

and B_k^*-sets, are studied. We prove that for any integer $k \geq 3$, if $A \subset \{1, 2, \ldots, N\}$ is a B_k^*-set, then $|A| \leq (\frac{1}{4} + o_k(1))N^{1/k}$. This improves the previously best known upper bound by a factor of $\frac{1}{4}$. For odd $k \geq 3$, we construct B_k^*-sets that have more elements than the Bose-Chowla B_k-sets. A consequence is that there are B_3^*-sets in $\{1, 2, \ldots, N\}$ that are asymptotically larger than any B_3-set contained in $\{1, 2, \ldots, N\}$. For any integer $k \geq 3$, we prove that if $A \subset \{1, 2, \ldots, N\}$ is a k-fold Sidon set, then $|A| \leq (N/k)^{1/2} + O((N/k)^{1/4})$. This improves the current upper bound by a factor of $(1/k)^{1/2}$.

Chapter 1

Introduction

In this thesis we investigate graphs that are constructed using Sidon sets, k-fold Sidon sets, and B_k^*-sets. These constructions generate graphs with interesting properties from an extremal point of view. Algebraic properties of these sets from additive number theory are used to deduce structural properties about the graph. For instance it is known that Sidon sets in finite abelian groups can be used to construct graphs that do not contain a 4-cycle. In addition, we investigate k-fold Sidon sets and B_k^*-sets as these are interesting objects in their own right. In fact, there are similarities between bipartite Turán problems and the theory of invariant linear equations which was initiated by Ruzsa [71]. Sidon sets and their generalizations play an important role in this theory. We begin by giving an outline of this thesis followed by a summary of most of our main results.

1.1 Outline of Thesis

Our main results are stated in Section 1.2. Section 1.3 contains some frequently used notation. Chapter 2 is an introduction to the bipartite Turán problem and Chapter 3 is an introduction to Sidon sets. In these chapters we state some of the important results in these two areas and introduce techniques that will be used in later chapters.

Our contributions are contained in Chapters 4, 5, 6, 7, and 8. In Chapter 4 we use Sidon sets to construct graphs that are C_4-free and have many edges. These

graphs disprove a conjecture of Abreu, Balbuena, and Labbate [1] concerning the Turán number of C_4. In Chapter 5 we investigate an ordered version of the Turán problem for bipartite graphs that was motivated by the work of Erdős, Czipszer, and Hajnal [23], and Dudek and Rödl [24]. For this problem, B_k-sets are used to construct graphs that do not contain certain ordered cycles of length $2k$. In Chapter 6 we use Sidon sets to disprove a conjecture of Solymosi concerning a possible extension of the removal lemma to sparse graphs. We also show how k-fold Sidon sets can be used to construct other counterexamples to Solymosi's conjecture. These k-fold Sidon sets were first introduced by Lazebnik and Verstraëte [55] in their work on a hypergraph Turán problem. We study k-fold Sidon sets in Chapter 7. In Chapter 8 we study another generalization of Sidon sets called B_k^*-sets and prove new upper bounds on the maximum size of a B_k^*-set contained in $\{1, 2, \ldots, N\}$.

1.2 Main Results

In this section we summarize several of our main results and it is written so that it is self contained. We have made an attempt to minimize the number of references and technical details while still presenting an accurate picture of our work.

1.2.1 Sidon Sets and Graphs Without 4-cycles

Let F be a graph. The *Turán number of* F is the maximum number of edges in an n-vertex graph that does not contain F as a subgraph. Write $\mathrm{ex}(n, F)$ for this maximum. Estimating Turán numbers for different graphs F is one of the most well studied problems in extremal graph theory. A case of particular interest is the Turán number of C_4, the cycle with four vertices. A well known double counting argument of Kővári, Sós, and Turán [51] (see also Reiman [68]) implies $\mathrm{ex}(n, C_4) \leq \frac{1}{2}n^{3/2} + \frac{n}{2}$. Using polarities of projective planes, Brown [14], Erdős, Rényi, and Sós [32] constructed for each prime power q, a graph with $q^2 + q + 1$

vertices, $\frac{1}{2}q(q+1)^2$ edges, and no C_4. This implies

$$\mathrm{ex}(q^2+q+1, C_4) \geq \frac{1}{2}q(q+1)^2$$

for any prime power q. Erdős conjectured that this lower bound is best possible. Füredi [43] proved this conjecture and so

$$\mathrm{ex}(q^2+q+1, C_4) = \frac{1}{2}q(q+1)^2$$

for any prime power $q \geq 15$.

In 2010, Abreu, Balbuena, and Labbate [1] used the polarity graphs of Brown, Erdős, Rényi, and Sós to obtain new lower bounds on $\mathrm{ex}(n, C_4)$. In particular, they showed that for any odd prime power q,

$$\mathrm{ex}(q^2-q-2, C_4) \geq \tfrac{1}{2}q^3 - q^2 - \tfrac{1}{2}q + 1.$$

They conjectured that this lower bound is best possible. Using Sidon sets, we show [79] that their conjecture is false and improve this lower bound by $\frac{q}{2}$.

Theorem 1.2.1 (Tait, Timmons, 2013)**.** *If q is an odd prime power, then*

$$\mathrm{ex}(q^2-q-2, C_4) \geq \tfrac{1}{2}q^3 - q^2 - O(q^{3/4}).$$

It is well known that Sidon sets can be used to construct graphs that do not contain C_4. To obtain our lower bound, we prove that Bose-Chowla Sidon sets have an arithmetic property that, to the best of our knowledge, has not appeared in the literature.

1.2.2 Ordered Turán Problems

Given the large amount of research on the bipartite Turán problem, it is natural to considered generalizations and extensions. Czipszer, Erdős, Hajnal [23], and Dudek, Rödl [24] considered a Turán type problem in which the vertices of the host graph and of the forbidden graph are ordered. Taking their work as motivation, we introduced the following ordered Turán problem. Let G be an n-vertex graph with $V(G) = \{1, 2, \ldots, n\}$. View the vertices of G as being

ordered in the obvious way. Let Z_4 be a 4-cycle whose vertices are positive integers $n_1 < n_2 < n_3 < n_4$, and whose edges are n_1n_3, n_1n_4, n_2n_3, and n_2n_4. Let $\mathrm{ex}(n, Z_4)$ be the maximum number of edges in an n-vertex graph with vertex set $\{1, 2, \ldots, n\}$ that does not contain a Z_4. If G is a C_4-free graph with $V(G) = \{1, 2, \ldots, n\}$, then G is also Z_4-free thus

$$\mathrm{ex}(n, C_4) \leq \mathrm{ex}(n, Z_4).$$

From the asymptotic formula $\mathrm{ex}(n, C_4) = \frac{1}{2}n^{3/2} + o(n^{3/2})$, we obtain the lower bound $\frac{1}{2}n^{3/2} + o(n^{3/2}) \leq \mathrm{ex}(n, Z_4)$. Using a counting argument inspired by Füredi [40], we prove the following result.

Theorem 1.2.2 (Timmons, 2012 [80])**.** *For any integer $n \geq 1$,*

$$\mathrm{ex}(n, Z_4) \leq \tfrac{2}{3}n^{3/2} + O(n^{5/4}).$$

We conjecture that $\mathrm{ex}(n, Z_4) = \frac{1}{2}n^{3/2} + o(n^{3/2})$, but this is still open.

Using projective planes we showed that the difference between $\mathrm{ex}(n, C_4)$ and $\mathrm{ex}(n, Z_4)$ can be made arbitrarily large.

Theorem 1.2.3 (Timmons, 2012 [80])**.** *For any prime p,*

$$\mathrm{ex}(p^2 + p + 1, Z_4) \geq p^2 + \mathrm{ex}(p^2 + p + 1, C_4).$$

We considered other types of ordered Turán problems for complete bipartite graphs and even cycles. For more on this topic, we refer the reader to Chapter 5.

1.2.3 A Counterexample to Sparse Removal

One of the most important theorems in extremal graph theory is the triangle removal lemma first proved by Ruzsa and Szemerédi in 1976 [73]. The triangle removal lemma states that any n-vertex graph with $o(n^3)$ triangles can be made triangle free by removing $o(n^2)$ edges. This was later generalized to the graph removal lemma by Erdős, Frankl, and Rödl [29], although they did not state the graph removal lemma as it is usually stated today. Let H be a graph with h vertices. The graph removal lemma states that if G is an n-vertex graph with $o(n^h)$ copies of H, then G can be made H-free by removing $o(n^2)$ edges.

It was observed by Ruzsa and Szemerédi [73] that the triangle removal lemma implies a weak form of Roth's Theorem [69] on 3-term arithmetic progressions in $\mathbb{Z}$. This was the starting point of a line of research in which removal lemmas for graphs and hypergraphs were used to prove results in combinatorial number theory such as Szemerédi's Theorem [76] on k-term arithmetic progressions and the Corner's Theorem of Atjai and Szemerédi [4]. This approach has been considerably successful. The graph removal lemma has been generalized in several directions such as removal in hypergraphs, random graphs, and groups. For a more thorough discussion, we refer the reader to the survey of Conlon and Fox [22]. Despite these advances, there is no known removal lemma for sparse graphs. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *sparse* if G_n has n vertices and $e(G_n)/n^2 \to 0$ as $n \to \infty$. As a possible removal lemma for sparse graphs, Solymosi [75] made the following conjecture.

Conjecture 1.2.4. [Solymosi, 2011] Let F be a graph with $\mathrm{ex}(n, F) = O(n^\alpha)$ where $1 < \alpha < 2$. If G is an n-vertex graph with the property that every edge of G lies in one copy of F, then G has $o(n^\alpha)$ edges.

Using Sidon sets, we construct infinitely many bipartite graphs F for which Conjecture 1.2.4 is false.

Theorem 1.2.5 (Timmons, Verstraëte, 2013 [82])**.** *There are infinitely many bipartite graphs F with $\mathrm{ex}(n, F) = \Theta(n^{3/2})$ for which the following holds: for infinitely many n, there is an n-vertex graph G with $\Omega(n^{3/2})$ edges such that every edge of G lies in exactly one copy of F.*

The smallest graph for which we know Conjecture 1.2.4 is false has 10 vertices and 16 edges. Conjecture 1.2.4 is still open for C_4. Lazebnik and Verstraëte [55] introduced objects called k-fold Sidon sets which are a generalization of Sidon sets. It is known that k-fold Sidon sets can be used to construct graphs with the property that every edge is in exactly one 4-cycle (see [82] for details). Therefore it is of interest to investigate these objects and this leads us into our next topic.

1.2.4 k-fold Sidon Sets

Let $k \geq 2$ be an integer. Let $N > k$ be an integer that is coprime to each integer in the set $\{1, 2, \ldots, k\}$. We say that a set $A \subset \mathbb{Z}_N$ is a *k-fold Sidon set* if A has only trivial solutions to each equation of the form

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 = 0$$

where $1 \leq |c_i| \leq k$, and $c_1 + c_2 + c_3 + c_4 = 0$. We define k-fold Sidon sets in $\{1, 2, \ldots, N\} \subset \mathbb{Z}$ in the same way, except that we do not require N to be relatively prime to each integer in $\{1, 2, \ldots, k\}$.

A short counting argument shows that a k-fold Sidon set $A \subset \mathbb{Z}_N$ has at most $(N/k)^{1/2} + 1$ elements. Lazebnik and Verstraëte [55] showed that for any $t \in \mathbb{N}$, there is a 2-fold Sidon set $A \subset \mathbb{Z}_N$ such that $|A| \geq \frac{1}{2}N^{1/2} - 3$ where $N = 2^{2^{t+1}} + 2^{2^t} + 1$. This shows that for $k = 2$, $N^{1/2}$ is the correct order of magnitude. Obtaining upper bounds in $\mathbb{Z}_N$ is often easier than obtaining upper bounds in $\{1, 2, \ldots, N\}$. For example, a short counting argument shows that a Sidon set $A \subset \mathbb{Z}_N$ has at most $N^{1/2} + 1$ elements, whereas it is a famous result of Erdős and Turán [37] that a Sidon set $A \subset \{1, 2, \ldots, N\}$ has at most $N^{1/2}+O(N^{1/4})$ elements. We generalize the Erdős-Turán upper bound and show that one can match the upper bound of $(N/k)^{1/2} + 1$ in $\mathbb{Z}_N$ on the size of a k-fold Sidon set in $\{1, 2, \ldots, N\}$, up to an error term.

Theorem 1.2.6 (Cilleruelo, Timmons, 2013 [20]). *Let $k \geq 1$ be an integer and let $1 \leq c_1 < c_2 < \cdots < c_k$ be k distinct integers. If $A \subset \{1, 2, \ldots, N\}$ is a set with only trivial solutions to $c_i(x_1 - x_2) = c_j(x_3 - x_4)$ for each $1 \leq i \leq j \leq k$, then*

$$|A| \leq \left(\tfrac{N}{k}\right)^{1/2} + O\left(\left(\tfrac{c_k^2 N}{k}\right)^{1/4}\right).$$

If we take $c_j = j$ for $1 \leq j \leq k$, then we immediately obtain that a k-fold Sidon set $A \subset \{1, 2, \ldots, N\}$ has at most $(N/k)^{1/2} + O((kN)^{1/4})$ elements.

Using the methods of Lazebnik and Verstraëte [55], we construct sets which show that Theorem 1.2.6 is close to best possible.

Theorem 1.2.7 (Cilleruelo, Timmons, 2013 [20]). *There exists k distinct integers $c_1, \ldots, c_k$ and infinitely many N such that there is a set $A \subset \mathbb{Z}_N$ with*

$$|A| \geq (1 - o(1))\frac{N^{1/2}}{k},$$

and A has only trivial solutions to $c_i(x_1 - x_2) = c_j(x_3 - x_4)$ for each $1 \leq i \leq j \leq k$.

1.2.5 B_k^+-sets

A well studied generalization of Sidon sets are B_k-sets. Let Γ be an abelian group. A set $A \subset \Gamma$ is called a *B_k-set* if

$$a_1 + \cdots + a_k = b_1 + \cdots + b_k \text{ with } a_i, b_j \in A,$$

implies $(a_1, \ldots, a_k)$ is a permutation of $(b_1, \ldots, b_k)$. A B_2-set is the same as a Sidon set. The problem concerning B_k-sets that has attracted the most attention is that of determining the maximum size of a B_k-set in $\{1, 2, \ldots, N\}$. A construction of Bose and Chowla [13] shows that for any integers $k \geq 2$ and $N \geq 1$, there is a B_k-set $A \subset \{1, 2, \ldots, N\}$ with $|A| = (1 + o(1))N^{1/k}$. Erdős [27], Bose and Chowla [13] have conjectured that this is best possible, but a matching upper bound is well out of a reach at the moment. The best known upper bounds are due to Green [47] and are roughly of the form $\frac{k}{e}N^{1/k}$ for the maximum size of a B_k-set contained in $\{1, 2, \ldots, N\}$. The main obstacle in improving the upper bound is obtaining a good upper bound in $\mathbb{Z}_N$. In $\mathbb{Z}_N$, a short counting argument gives the best upper bound (see Chapter 8) and improving this bound is a difficult open problem.

Ruzsa considered a generalization of B_k-sets that are sometimes called B_k^*-sets. A set A is a *B_k^*-set* if

$$a_1 + \cdots + a_k = a_{k+1} + \cdots + a_{2k} \text{ with } a_i \in A,$$

implies $a_i = a_j$ for some $1 \leq i < j \leq 2k$. If A is a B_k-set, then it is also a B_k^*-set but the converse is not true in general. Ruzsa [71] proved that a B_k^*-set in $\{1, 2, \ldots, N\}$ has at most $(1 + o(1))k^{2-1/k}N^{1/k}$ elements. Let $F_k(N)$ and $F_k^*(N)$ denote the maximum size of a B_k and B_k^*-set, respectively, contained in the interval $\{1, 2, \ldots, N\}$. The results of Green [47] and Ruzsa [71] do not rule out the possibility that $F_k(N)$ is asymptotic to $F_k^*(N)$. Using Bose-Chowla B_k-sets, we proved that $F_3(N)$ is not asymptotic to $F_3^*(N)$.

Theorem 1.2.8 (Timmons, 2012 [81]). *Let $k \geq 3$ be an odd integer. For any integer $N \geq 1$,*

$$F_k^*(N) \geq (1 + o(1))2^{1-1/k}N^{1/k}.$$

If $F_k(N) = (1+o(1))N^{1/k}$, as Erdős conjectured, then Theorem 1.2.8 shows that $F_k(N)$ is not asymptotic to $F_k^*(N)$ for all odd $k \geq 3$. It is still unknown however if $F_4(N)$ is asymptotic to $F_4^*(N)$.

Using arguments of Ruzsa [71], we improve the upper bound on $F_k^*(N)$ by a factor of $\frac{1}{4}$ for large k.

Theorem 1.2.9 (Timmons, 2012 [81]). *If $k \geq 3$ is an integer, then*

$$F_k^*(N) \leq \left(\tfrac{1}{4} + \epsilon(k)\right) k^2 N^{1/k}$$

where $\epsilon(k) \to 0$ as $k \to \infty$.

Theorems 1.2.8 and 1.2.9 do not encompass all of our results on this topic. We introduce objects that we call B_k^+-sets, which are another generalization of B_k-sets, and investigate the corresponding function $F_k^+(N)$. For more on this topic we refer the reader to Chapter 8.

1.3 Frequently Used Notation

- $[N] = \{1, 2, \ldots, N\}$.
- $\mathbb{F}_q$ is the finite field with q elements.
- $\mathbb{F}_q^*$ is the multiplicative group of nonzero elements of the finite field $\mathbb{F}_q$.
- $\mathbb{Z}_N$ is the group of integers modulo N.
- $\mathrm{ex}(n, F)$ is the Turán number of F.
- K_t is the complete graph on t vertices.
- $K_{n,m}$ is the complete bipartite graph with n vertices in one part and m vertices in the other.

- C_k is a cycle of length k.
- $e(G)$ is the number of edges of the graph G.

Chapter 2

Bipartite Turán Problems

In this chapter we discuss the Turán problem for graphs. We start with the two most important results in this area: Turán's Theorem and the Erdős-Stone-Simonovits Theorem. We then briefly discuss Turán numbers of trees. Section 2.3 forms the bulk of the chapter. In this section we discuss the bipartite Turán problem with an emphasis on even cycles and complete bipartite graphs. We introduce some of the techniques and methods, and present some of the more important theorems such as the Even Circuit Theorem of Bondy and Simonovits, and the Kővari-Sós-Turán upper bound on the maximum number of edges in a graph with no $K_{s,t}$.

2.1 Non-Bipartite Graphs

Let $\mathcal{F}$ be a family of graphs. We say that a graph G is *$\mathcal{F}$-free* if no subgraph of G is isomorphic to a graph in $\mathcal{F}$. The maximum number of edges in an n-vertex graph that is $\mathcal{F}$-free is called the *Turán number* of $\mathcal{F}$ and is denoted by $\mathrm{ex}(n, \mathcal{F})$. When the family $\mathcal{F}$ consists of a single graph, say $\mathcal{F} = \{F\}$, then we write $\mathrm{ex}(n, F)$ instead of $\mathrm{ex}(n, \{F\})$.

Determining Turán numbers of different families of graphs is one of the most well-studied problems in extremal graph theory. The earliest result is due to Mantel [61] who, in 1907, determined the Turán number of K_3. Mantel proved that $\mathrm{ex}(n, K_3) = \lfloor n/2 \rfloor \lceil n/2 \rceil$ for all $n \geq 1$. Furthermore, he showed that if G is an

n-vertex K_3-free graph with $\lfloor n/2 \rfloor \lceil n/2 \rceil$ edges, then G is isomorphic to $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. Mantel's Theorem was generalized by Turán in 1941. Turán's Theorem is often considered to be the first theorem in extremal graph theory. In order to state this result we require the definition of the Turán graph. Given integers $n \geq 1$ and $r \geq 2$, let $T_{n,r}$ be the r-partite graph whose part sizes are as equal as possible, and all pairs of vertices that belong to different parts are adjacent. Observe that $T_{n,r}$ is K_{r+1}-free.

Theorem 2.1.1 (Turán, 1941 [83]). *Let $r \geq 3$ be an integer. For any integer $n \geq 1$,*

$$\mathrm{ex}(n, K_r) = e(T_{n,r-1}).$$

Furthermore, any n-vertex K_r-free graph with $\mathrm{ex}(n, K_r)$ edges is isomorphic to $T_{n,r-1}$.

Shortly after Turán's Theorem appeared, Erdős and Stone determined an asymptotic formula for the Turán number of blow ups of complete graphs. Given integers $r \geq 2$ and $t \geq 1$, let $K_r(t)$ be the complete r-partite graph with t vertices in each part.

Theorem 2.1.2 (Erdős, Stone, 1946 [36]). *Let $r \geq 2$ and $t \geq 1$ be integers. Given $\epsilon > 0$, there exists an $n_0 = n_0(r, t, \epsilon)$ such that the following holds: if $n \geq n_0$ and G is an n-vertex graph with*

$$e(G) \geq \left(1 - \frac{1}{r-1} + \epsilon\right) \frac{n^2}{2},$$

then G contains a subgraph isomorphic to $K_r(t)$. In particular,

$$\mathrm{ex}(n, K_r(t)) = \left(1 - \frac{1}{r-1}\right) \frac{n^2}{2} + o(n^2)$$

for any integer $t \geq 1$.

The Erdős-Stone Theorem is sometimes refered to as the fundamental theorem of extremal graph theory. It was later observed by Erdős and Simonovits [33] that the Erdős-Stone Theorem gives an asymptotic formula for the Turán number of any family of non-bipartite graphs.

Theorem 2.1.3 (Erdős-Stone, Erdős-Simonovits, 1966)**.** *Let $\mathcal{F}$ be a family of graphs and let $r = \min_{F\in\mathcal{F}} \chi(F)$. If $r \geq 2$, then*

$$\mathrm{ex}(n, \mathcal{F}) = \left(1 - \frac{1}{r-1}\right)\frac{n^2}{2} + o(n^2).$$

The special case when $\mathcal{F}$ consists of a single non-bipartite graph is worth mentioning on its own. If F is a graph with $\chi(F) \geq 3$, then by Theorem 2.1.3,

$$\mathrm{ex}(n, F) = \left(1 - \frac{1}{\chi(F)-1}\right)\frac{n^2}{2} + o(n^2).$$

From this we see that it is the chromatic number of a non-bipartite graph that controls its Turán number. More precise results on the error term are known (see [9]) and they rely estimates on bipartite Turán numbers.

Having an asymptotic formula for the Turán number of any non-bipartite graph, it is natural to look for a similar formula for bipartite graphs.

2.2 Turán Numbers of Trees

When F is a bipartite graph, Theorem 2.1.3 implies $\mathrm{ex}(n, F) = o(n^2)$, but it is certainly desirable to look for better estimates. For instance a classical result of Erdős and Gallai [30] from 1959 gives the Turán number of the path. If $r \geq 1$ is an integer and P_r is the path of length r, then $\mathrm{ex}(n, P_r) \leq \frac{r-1}{2}n$. Equality occurs if and only if r divides n, in which case the unique extremal graph is a disjoint union of $\frac{n}{r}$ copies of the complete graph K_r. In general, one can show that for any tree T with $r+1$ vertices,

$$\frac{r-1}{2}n \leq \mathrm{ex}(n, T) \leq rn \tag{2.1}$$

(see Proposition 2.1.8 [86]). The famous Erdős-Sós conjecture is that if $n \geq r+1$, then $\mathrm{ex}(n, T) \leq \frac{r-1}{2}n$ where T is any tree on $r+1$ vertices. Recent progress on this conjecture has been made by Atjai, Komlós, Simonovits, and Szemerédi (see [45], Theorem 6.2). They proved that there is an integer r_0 so that if T is any tree with $r+1 > r_0$ vertices, then $\mathrm{ex}(n, T) \leq \frac{r-1}{2}n$.

By (2.1) and Theorem 2.1.3, we can say that Turán numbers of trees are linear in n, and Turán numbers of non-bipartite graphs are quadratic in n. What

remains then is bipartite graphs that contain a cycle. As we shall see shortly, the picture here is more complicated.

2.3 Bipartite Turán Numbers

Let F be a bipartite graph that contains a cycle. By the Erdős-Stone-Simonovits Theorem, $\mathrm{ex}(n, F) = o(n^2)$. On the other hand, using random graphs one can construct an F-free n-vertex graph with n^{1+c} edges where $c > 0$ is a positive constant depending only on F. From this we deduce that the Turán number of any bipartite graph that contains a cycle is not linear and also not quadratic. The first question that one would like to answer then is given F, can we determine the order of magnitude of $\mathrm{ex}(n, F)$? Erdős and Simonovits [34] conjectured that for any bipartite graph F, there exists a constant $c = c_F$ such that

$$\lim_{n\to\infty} \frac{\mathrm{ex}(n, F)}{n^c}$$

exists and is positive. This conjecture is still open for many bipartite graphs such as C_8 and $K_{4,4}$.

An early result of Kővári, Sós, and Turán [51] provides an upper bound on the Turán number of $K_{s,t}$. We include the proof as it is a good example as how double counting arguments can be used to obtain upper bounds on bipartite Turán numbers.

Theorem 2.3.1 (Kővári, Sós, Turán, 1954)**.** *Let $s \geq t \geq 2$ be integers. For any integer $n \geq 1$,*

$$\mathrm{ex}(n, K_{s,t}) \leq \frac{1}{2}(s-1)^{1/t} n^{2-1/t} + \frac{tn}{2}.$$

Proof. Let G be an n-vertex $K_{s,t}$-free graph. The sum

$$\sum_{v\in V(G)} \binom{d(v)}{t}$$

counts sets of t distinct vertices that are contained in the neighborhood of some vertex in G. Since G is $K_{s,t}$-free, no subset of vertices of size t can be counted

more than $s-1$ times by this sum. Therefore,

$$(s-1)\binom{n}{t} \geq \sum_{v \in V(G)} \binom{d(v)}{t}.$$

By convexity,

$$\sum_{v \in V(G)} \binom{d(v)}{t} \geq n\binom{2e(G)/n}{t}$$

from which we deduce that

$$(s-1)\binom{n}{t} \geq n\binom{2e(G)/n}{t}.$$

To complete the proof we use the trivial estimates $(s-1)\binom{n}{t} \leq (s-1)\frac{n^t}{t!}$ and $n\binom{2e(G)/n}{t} \geq \frac{n}{t!}\left(\frac{2e(G)}{n} - t\right)^t$. Some simple algebra then shows that the inequality

$$(s-1)\frac{n^t}{t!} \geq \frac{n}{t!}\left(\frac{2e(G)}{n} - t\right)^t$$

implies the desired upper bound on $e(G)$. □

It is believed by many that the exponent $2-\frac{1}{t}$ is correct; that is for $s \geq t \geq 2$, $\mathrm{ex}(n, K_{s,t})$ is of order $n^{2-1/t}$. This is still open for most values of s and t. Using a more sophisticated double counting argument, Füredi [40] proved

$$\mathrm{ex}(n, K_{s,t}) \leq \frac{1}{2}(s-t+1)^{1/t} n^{2-1/t} + O(n^{2-2/t}) \tag{2.2}$$

for $s \geq t \geq 2$. This is essentially the best known upper bound on the Turán number for $K_{s,t}$. Recently Nikiforov [64] improved the error term.

Obtaining good lower bounds on $\mathrm{ex}(n, K_{s,t})$ is a difficult problem. Incidence graphs of projective planes imply $\mathrm{ex}(n, K_{2,2}) \geq \frac{1}{2\sqrt{2}} n^{3/2} + o(n^{3/2})$ which matches the exponent of (2.2). Graphs constructed independently by Erdős, Rényi [31] and Brown [14] improve the constant $\frac{1}{2\sqrt{2}}$ to $\frac{1}{2}$. This gives one of the rare instances in which we have an asymptotic formula for the Turán number of a bipartite graph that contains a cycle:

$$\mathrm{ex}(n, K_{2,2}) = \frac{1}{2} n^{3/2} + o(n^{3/2}).$$

Füredi [42] generalized some of the results of [31], [14] and constructed $K_{2,s}$-free n-vertex graphs with $\frac{1}{2}\sqrt{s-1}n^{3/2}+o(n^{3/2})$ edges for any $s \geq 2$. Therefore, if $s \geq 2$ then

$$\mathrm{ex}(n, K_{2,s}) = \frac{1}{2}\sqrt{s-1}n^{3/2} + o(n^{3/2}).$$

A construction of Brown [14] and (2.2) give

$$\mathrm{ex}(n, K_{3,3}) = \frac{1}{2}n^{5/3} + o(n^{5/3}).$$

To give the reader an idea of what these constructions look like, we take a moment to present the $K_{2,s}$-free graphs constructed by Füredi [42], and the $K_{3,3}$-free graphs constructed by Brown [14].

$K_{2,s}$-free graphs. Let $s \geq 2$ be an integer. Let q be a prime power chosen so that s divides $q-1$. Let $\mathbb{F}_q$ be the finite field with q elements. The multiplicative group of nonzero elements of $\mathbb{F}_q$ is isomorphic to $\mathbb{Z}_{q-1}$. Choose an element $h \in \mathbb{F}_q^*$ that generates an s element subgroup in $\mathbb{F}_q^*$. Let $H = \{1, h, h^2, \ldots, h^{s-1}\}$. Define an equivalence relation on $\mathbb{F}_q \times \mathbb{F}_q \backslash \{(0,0)\}$ by saying that (a,b) and (a',b') are equivalent if $h^\alpha(a,b) = (a',b')$ for some $\alpha \in \{0,1,\ldots,s-1\}$. Write $\langle a,b\rangle$ for the equivalence class that contains (a,b). Let $G_{q,s}$ be the graph whose vertices are the set of equivalence classes. We join $\langle a,b\rangle$ to $\langle x,y\rangle$ if $ax+by \in H$. The graph $G_{q,s}$ has $\frac{q^2-1}{s}$ vertices and at least $\frac{1}{2}(q-1)(\frac{q^2-1}{s})$ edges. For the remaining details, including a proof that $G_{q,s}$ is $K_{2,s}$-free, see [42].

$K_{3,3}$-free graphs. Let p be an odd prime. If $p \equiv 3 (\mathrm{mod}\ 4)$, let α be a fixed nonzero quadratic residue in $\mathbb{F}_p$, otherwise let α be a fixed non-quadratic residue in $\mathbb{F}_p$. Let G_p have vertex set $\mathbb{F}_p^3$. Given $x = (x_1, x_2, x_3) \in \mathbb{F}_p^3$, let

$$S(x) = \{(y_1, y_2, y_3) : (x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 = \alpha\}.$$

Vertices x and y are adjacent if $y \in S(x)$. The graph G_p has p^3 vertices, at least $\frac{1}{2}p^3(p^2-p)$ edges, and is $K_{3,3}$-free (see [14]).

When $s \geq (t-1)!+1$, the projective norm graphs constructed by Kollár, Rónyai, and Szabó [50] (see also the paper of Alon, Rónyai, Szabó [3]) show that $\mathrm{ex}(n, K_{s,t}) \geq c_{s,t}n^{2-1/t}$. For other values of s and t, there are no lower bounds that match (2.2) in order of magnitude. The graph $K_{4,4}$ is the simplest complete bipartite graph for which the order of magnitude of $\mathrm{ex}(n, K_{4,4})$ is not known.

Another family of bipartite graphs that has been studied is even cycles. Of course a 4-cycle is the same as $K_{2,2}$ so we will discuss even cycles of length at least 6. Here the picture is similar to that of complete bipartite graphs. A general upper bound of Bondy and Simonovits [11] gives the correct order of magnitude of $\mathrm{ex}(n, C_{2k})$ for $k \in \{2, 3, 5\}$. If $k \notin \{2, 3, 5\}$, then the order of magnitude of $\mathrm{ex}(n, C_{2k})$ is not known.

Upper bounds on $\mathrm{ex}(n, C_{2k})$ are not as easy to obtain as upper bounds on $\mathrm{ex}(n, K_{s,t})$, unless of course $k = 2$. The first good upper bound on $\mathrm{ex}(n, C_{2k})$ is due to Bondy and Simonovits [11] who proved $\mathrm{ex}(n, C_{2k}) \leq 100n^{1+1/k}$ for any $k \geq 2$. This is sometimes called the Even Circuit Theorem. The best known upper bound on $\mathrm{ex}(n, C_{2k})$ for general k is due to Pikhurko [67] who, using ideas of Verstraëte [84], showed that

$$\mathrm{ex}(n, C_{2k}) \leq (k-1)n^{1+1/k} + 16(k-1)n. \tag{2.3}$$

A proof that $\mathrm{ex}(n, C_{2k}) = O(n^{1+1/k})$ would be too long to reproduce here. The simplest proof is perhaps the one found in [84]. It is unlikely that (2.3) is asymptotically best possible for any k. Constructions of C_6-free and C_{10}-free graphs given by Benson [7] and Wenger [85] show that the exponent $1 + 1/k$ is correct in these cases. Füredi, Naor, and Verstraëte [44] proved that for sufficiently large n, $\mathrm{ex}(n, C_6) \leq 0.6272n^{4/3}$. They also gave a construction which shows $\mathrm{ex}(n, C_6) > 0.5338n^{4/3}$ for infinitely many n. Lazebnik, Ustimenko, and Woldar [54], using the construction of Wenger [85], proved $\mathrm{ex}(n, C_{10}) > 4/5^{6/5}n^{6/5} + o(n^{6/5})$. To summarize, $\mathrm{ex}(n, C_4)$ is known asymptotically. For $k \in \{3, 5\}$, the order of $\mathrm{ex}(n, C_{2k})$ is $n^{1+1/k}$. For $k \notin \{2, 3, 5\}$, $\mathrm{ex}(n, C_{2k})$ is $O(n^{1+1/k})$ but there is no matching lower bound.

For more on bipartite Turán problems, we refer the reader to the excellent survey of Füredi and Simonovits [45].

Chapter 3

Sidon Sets

In this chapter we introduce Sidon sets which are objects from additive number theory. Sidon sets and their generalizations play a central role in this thesis. In Section 3.1 we state some well known upper bounds on the sizes of finite Sidon sets. This section includes a proof of the famous Erdős-Turán upper bound on the size of a Sidon set contained in $\{1, 2, \ldots, N\}$. In Section 3.2 we introduce several constructions of Sidon sets in various groups. Some of these constructions will be used in later chapters.

3.1 Upper Bounds

Let Γ be an abelian group. A set $A \subset \Gamma$ is a *Sidon set* if

$$a + b = c + d \text{ with } a, b, c, d \in A,$$

implies (a, b) is a permutation of (c, d).

The study of combinatorial properties of Sidon sets has its origins in a paper of Erdős and Turán [37] from 1941. According to Erdős [26], Hungarian mathematician Simon Sidon introduced Sidon sets to Erdős in 1932 or 1933. Sidon was interested in Sidon sets because of a connection to problems in analysis (see [72] for a simple example how Sidon sets arise naturally in an analytic context). Erdős viewed these sets from a combinatorial point of view and what emerged was a beautiful collection of problems that have drawn a large amount of interest

from the mathematical community. Sidon sets and problems related to Sidon sets appear frequently in the problem papers of Erdős. A survey of O'Bryant [65] contains over 100 references to papers on Sidon and Sidon type sets.

According to Erdős and Turán [37], it was Sidon who was the first to ask for the maximum size of a Sidon set contained in the interval $\{1, 2, \ldots, N\}$. Let

$$F_2(N) = \max_{A \subset [N]} \{|A| : A \text{ is a Sidon set}\}.$$

We begin with a simple proposition that provides an upper bound on $F_2(N)$.

Proposition 3.1.1. *For any integer $N \geq 1$,*

$$F_2(N) \leq (2N)^{1/2} + 1.$$

Proof. Let $A \subset [N]$ be a Sidon set. Each ordered pair of distinct elements of A determines a unique difference that is contained in the interval

$$\{-N+1, -N+2, \ldots, N-2, N-1\}.$$

This implies $|A|(|A|-1) \leq 2N$ so that $|A| \leq (2N)^{1/2} + 1$. □

In 1941, Erdős and Turán [37] improved the upper bound of Propsition 3.1.1. The following result is one of the most well-known theorems on Sidon sets.

Theorem 3.1.2 (Erdős, Turán, 1941)**.** *For any integer $N \geq 1$,*

$$F_2(N) \leq N^{1/2} + O(N^{1/4}). \tag{3.1}$$

Proof. Let $A \subset [N]$ be a Sidon set. Let $m \geq 1$ be an integer and let $I = [m]$. For any integer j, let $I_j = \{1+j, 2+j, \ldots, m+j\}$. The sum $\sum_{j=-m+1}^{N} |A \cap I_j|$ counts each element of A exactly m times. By Cauchy-Schwarz,

$$\begin{aligned} |A|^2 m^2 &= \left(\sum_{j=-m+1}^{N} |A \cap I_j| \right)^2 \leq (N+m-1) \sum_{j=-m+1}^{N} |A \cap I_j|^2 \\ &= (N+m-1) \left(2 \sum_{j=-m+1}^{N} \binom{|A \cap I_j|}{2} + m|A| \right). \end{aligned}$$

Given $d \in [m-1]$, there is at most one pair $\{a,b\}$ with $a,b \in A$ and $b-a=d$. The sum $\sum_{j=-m+1}^{N} \binom{|A\cap I_j|}{2}$ counts such pairs and if $\{a,b\}$ is a pair with $b-a=d$ for some $d \in [m-1]$, then this pair is counted exactly $m-d$ times. Therefore,

$$\sum_{j=-m+1}^{N} \binom{|A \cap I_j|}{2} \leq \sum_{d=1}^{m-1}(m-d) \leq m^2.$$

Combining the previous two estimates, we get

$$|A|^2 m^2 \leq (N+m-1)(m^2+m|A|). \tag{3.2}$$

The rest of the proof involves optimizing the choice of m and simplifying.

Rewrite (3.2) as $|A|^2 \leq (\frac{N+m}{m})(m+|A|)$. We can complete the square to get

$$\begin{aligned} |A| &\leq \left(N+m+\frac{(N+m)^2}{4m^2}\right)^{1/2} + \frac{N+m}{2m} \\ &= N^{1/2}\left(1+\frac{m}{N}+\frac{(N+m)^2}{4Nm^2}\right)^{1/2} + \frac{N}{2m} + \frac{1}{2} \\ &< N^{1/2}\left(1+\frac{m}{2N}+\frac{(N+m)^2}{8Nm^2}\right) + \frac{N}{2m} + \frac{1}{2} \\ &= N^{1/2} + \frac{m}{2N^{1/2}} + \frac{N^{3/2}}{8m^2} + \frac{N^{1/2}}{4m} + \frac{1}{8N^{1/2}} + \frac{N}{2m} + \frac{1}{2}. \end{aligned}$$

If we choose $m = \lfloor N^{3/4} \rfloor$, then we get after simplifying $|A| < N^{1/2} + N^{1/4} + 1$ as required. □

Erdős and Turán conjectured $F_2(N) \leq N^{1/2} + C$ where C is an absolute constant. This is still open and Erdős [26] offered $500 for a proof or disproof of this conjecture. Currently the best known upper bound on $F_2(N)$ is due to Cilleruelo [19] who obtained $F_2(N) < N^{1/2} + N^{1/4} + 1/2$ as a consequence of a more general result. This improved the previous upper bound of $F_2(N) < N^{1/2} + N^{1/4} + 1$ which had been established in 1969 by Lindström [57]. His argument was essentially the same as that of Erdős and Turán and it is this bound that we proved above. It seems safe to say that new ideas will be required to prove or disprove the conjecture of Erdős and Turán.

The leading term $N^{1/2}$ in (3.1) is best possible. It was observed by Erdős [25] that a construction of Singer [74] could be used to give a lower bound on $F_2(N)$.

Theorem 3.1.3 (Singer, 1938)**.** *If q is a prime power, then there is a Sidon set $A \subset \mathbb{Z}_{q^2+q+1}$ with $|A| = q+1$.*

Given an integer $N \geq 1$, let

$$C_2(N) = \max_{A \subset \mathbb{Z}_N} \{|A| : A \text{ is a Sidon set}\}.$$

If $A \subset \mathbb{Z}_N$ is a Sidon set, then A is also a Sidon set when viewed as a subset of $\mathbb{Z}$ so $C_2(N) \leq F_2(N)$ for all $N \geq 1$. Theorem 3.1.3 gives the lower bound

$$q+1 \leq C_2(q^2+q+1) \leq F_2(q^2+q+1) \tag{3.3}$$

for any prime power q. The function $F_2(N)$ is nondecreasing in N so using estimates on gaps between primes (see [6]), one obtains

$$F_2(N) \geq N^{1/2} - O(N^{0.2625}).$$

Singer's construction and (3.1) determine the asymptotic behaviour of $F_2(N)$:

$$\lim_{N \to \infty} \frac{F_2(N)}{N^{1/2}} = 1.$$

Unfortunately it is not known if $C_2(N)$ is nondecreasing in N so we do not obtain the same lower bound for $C_2(N)$. Since $C_2(N) \leq F_2(N)$, (3.1) implies $C_2(N) \leq N^{1/2} + O(N^{1/4})$ although the following simple argument gives a better upper bound.

Proposition 3.1.4. *If $N \geq 1$ is any integer, then*

$$C_2(N) \leq \sqrt{N - 3/4} + 1/2.$$

Proof. Let $A \subset \mathbb{Z}_N$ be a Sidon set. Each pair of distinct elements of A determine a unique nonzero element of $\mathbb{Z}_N$ so $|A|(|A|-1) \leq N-1$. After some simplifying we find that $|A| \leq \sqrt{N - 3/4} + 1/2$ as required. □

By Theorem 3.1.3,

$$\limsup_{N \to \infty} \frac{C_2(N)}{N^{1/2}} = 1. \tag{3.4}$$

There are other constructions of B_2-sets with many elements but they all use primes in some way. There is no known general construction that works for arbitrary N. This leads us into our next section which is lower bounds on $F_2(N)$ and $C_2(N)$.

3.2 Lower Bounds

In this section we present several constructions of Sidon sets in a variety of groups.

Proposition 3.2.1. *The set $A = \{(a, a^2) : a \in \mathbb{N}\}$ is a Sidon set in the group $\mathbb{Z}^2$.*

Proof. Suppose $(a, a^2) + (b, b^2) = (c, c^2) + (d, d^2)$. Then

$$a - c = d - b \text{ and } a^2 - c^2 = d^2 - b^2.$$

If $a - c = 0$, then $(a, a^2) = (c, c^2)$ and $(b, b^2) = (d, d^2)$. Othwerwise, we can divide $a^2 - c^2 = d^2 - b^2$ through by $a - c = d - b$ to get $a + c = b + d$. Add this last equation to $a - c = d - b$ to get $2a = 2d$. This implies that $a = d$ and $b = c$. □

The same proof can be used to prove the next proposition.

Proposition 3.2.2. *Let q be an odd prime power. The set $A = \{(a, a^2) : a \in \mathbb{F}_q\}$ is a Sidon set in the group $\mathbb{F}_q^2$.*

The Sidon sets of Proposition 3.2.2 will be used in Chapter 6. The next example is due to Erdős and Turán [37]. Given an integer y, let $(y)_p$ be the unique integer between 1 and p that satisfies $y \equiv (y)_p (\text{mod } p)$.

Theorem 3.2.3. *Let p be an odd prime. The set*

$$A = \{x + 2p(x^2)_p : 1 \leq x \leq p\}$$

is a Sidon set in $\mathbb{Z}$ that is contained in the interval $\{1, 2, \ldots, 2p^2 + p\}$.

Proof. Suppose

$$x + 2p(x^2)_p + y + 2p(y^2)_p = u + 2p(u^2)_p + v + 2p(v^2)_p \tag{3.5}$$

for some $1 \leq x, y, u, v \leq p$. Since $|x + y - u - v| \leq 2p - 2$, (3.5) implies

$$x + y \equiv u + v (\text{mod } p) \text{ and } x^2 + y^2 \equiv u^2 + v^2 (\text{mod } p).$$

As in the proof of Proposition 3.2.1, we have either $x \equiv u (\text{mod } p)$ or $x \equiv v (\text{mod } p)$. Since $1 \leq x, y, u, v \leq p$, these congruences become equalities so $\{x, y\} = \{u, v\}$. □

The Sidon sets of Theorem 3.2.3 show that $F_2(N) \geq (\frac{1}{\sqrt{2}} - o(1))N^{1/2}$. As noted by Erdős [25], the following result of Singer [74] gives a better lower bound.

Theorem 3.2.4 (Singer, 1938)**.** *If q is an odd prime power, then there is a Sidon set $A \subset \mathbb{Z}_{q^2+q+1}$ with $|A| = q + 1$.*

The construction of A in Theorem 3.2.4 is more complicated than the previous examples. Since we will not need these Sidon sets, we omit the details and refer the reader to [13].

The next construction is due to Bose and Chowla [13]. We will use these Sidon sets in Chapters 4 and 7.

Theorem 3.2.5 (Bose, Chowla, 1962)**.** *If q is a prime power, then there is a Sidon set $A \subset \mathbb{Z}_{q^2-1}$ with $|A| = q$.*

Proof. Let θ be a generator of $\mathbb{F}_{q^2}^*$. Let $\mathbb{F}_q = \{b_1, \ldots, b_q\}$. Define $a_i \in \mathbb{Z}_{q^2-1}$ by

$$\theta^{a_i} = \theta + b_i \text{ for } i = 1, 2, \ldots, q.$$

Let $A = \{a_1, \ldots, a_q\}$. We will now show that this is a Sidon set in $\mathbb{Z}_{q^2-1}$.

Suppose $a_i + a_j \equiv a_k + a_l (\text{mod } q^2 - 1)$. Then $\theta^{a_i}\theta^{a_j} = \theta^{a_k}\theta^{a_l}$. We can rewrite this equation as $(\theta + b_i)(\theta + b_j) = (\theta + b_k)(\theta + b_l)$ and then cancel θ^2 from both sides. What remains is a degree one equation with coefficients in $\mathbb{F}_q$ which is impossible, unless the equation is trivial. Therefore, $\{b_i, b_j\} = \{b_k, b_l\}$ which implies $\{a_i, a_j\} = \{a_k, a_l\}$. □

In 1993, Ruzsa [71] gave one of the simplest constructions of a Sidon set that has many elements.

Theorem 3.2.6 (Ruzsa, 1993)**.** *If p is a prime, then there is a Sidon set $A \subset \mathbb{Z}_{p^2-p}$ with $|A| = p - 1$.*

Proof. Let p be a prime and let θ be a generator of the multiplicative group $\mathbb{F}_p^*$. Define $p - 1$ residues a_i by

$$a_i \equiv i (\text{mod } p - 1) \quad \text{and} \quad a_i \equiv \theta^i (\text{mod } p)$$

for $i = 1, 2, \ldots, p - 1$. Let $A = \{a_1, \ldots, a_{p-1}\}$. We now show that A is a Sidon set in $\mathbb{Z}_{p(p-1)}$. Suppose $a_i + a_j \equiv b (\text{mod } p(p - 1))$. Then

$$i + j \equiv b (\text{mod } p - 1) \quad \text{and} \quad \theta^i + \theta^j \equiv b (\text{mod } p).$$

Working over $\mathbb{F}_p$, we see that

$$x^2 - bx + \theta^b = (x - \theta^i)(x - \theta^j)$$

is the unique factorization of $x^2 - bx + \theta^b$. Therefore, b determines i and j uniquely up to ordering. □

Theorems 3.2.4, 3.2.5, and 3.2.6 give the following lower bounds on $C_2(N)$:

1. $C_2(q^2 + q + 1) \geq q + 1$ for any prime power q.
2. $C_2(q^2 - 1) \geq q$ for any prime power q.
3. $C_2(p^2 - p) \geq p - 1$ for any prime p.

Recall that if $A \subset \mathbb{Z}_N$ is a Sidon set, then $|A|(|A| - 1) \leq N - 1$. This simple inequality shows that each of the three lower bounds given above is best possible. Despite this, determining the asymptotic behavior of $C_2(N)$ is still a challenging open problem. It seems likely that $C_2(N)$ is asymptotic to $F_2(N)$. The best bounds that we currently have are

$$\frac{1}{\sqrt{2}} \leq \liminf_{N \to \infty} \frac{C_2(N)}{N^{1/2}} \leq \limsup_{N \to \infty} \frac{C_2(N)}{N^{1/2}} \leq 1.$$

The lower bound follows from the observation that a Sidon set $A \subset \{1, \ldots, \lfloor N/2 \rfloor\}$ is also a Sidon set when viewed as a subset of $\mathbb{Z}_N$. A consequence of this observation is that $F_2(\lfloor N/2 \rfloor) \leq C_2(N)$ for all $N \geq 1$.

For more on Sidon sets, the standard reference is the book of Halberstam and Roth [48].

Chapter 4

Sidon Sets and Graphs Without 4-cycles

The problem of determining the maximum number of edges in an n-vertex graph that does not contain a 4-cycle has a rich history in extremal graph theory. Using Sidon sets, for each odd prime power q we construct a graph with q^2-q-2 vertices, $\frac{1}{2}q^3-q^2-O(q^{3/4})$ edges, and no 4-cycle [79]. This disproves a conjecture of Abreu, Balbuena, and Labbate [1] on the Turán number $\mathrm{ex}(q^2-q-2, C_4)$ where q is an odd prime power.

4.1 Introduction

Let F be a graph. Recall that the *Turán number* of F, denoted $\mathrm{ex}(n, F)$, is the maximum number of edges in an n-vertex graph that does not contain F as a subgraph. As mentioned in Chapter 2, one of the most studied cases is when $F = C_4$, the cycle on four vertices. Before stating our main result, we recall some of the known results concerning the Turán number of C_4. Some of these results were mentioned in Chapter 2.

The Kővári-Sós-Turán bound [51] implies $\mathrm{ex}(n, C_4) \leq \frac{1}{2}n^{3/2} + \frac{n}{2}$ for every $n \geq 1$. Using polarity graphs of projective planes Brown [14], Erdős, Rényi, and Sós [32] independently proved that for each prime power q, $\mathrm{ex}(q^2+q+1, C_4) \geq \frac{1}{2}q(q+1)^2$. Therefore, $\mathrm{ex}(n, C_4) = \frac{1}{2}n^{3/2} + o(n^{3/2})$ is an asymptotic formula for the

Turán number of C_4.

The exact value of $\mathrm{ex}(n, C_4)$ was determined using computer searches ([21], [70]) for all $n \leq 31$. Füredi [43] proved that whenever $q \geq 15$ is a prime power, $\mathrm{ex}(q^2+q+1, C_4) \leq \frac{1}{2}q(q+1)^2$ thus we have the exact result $\mathrm{ex}(q^2+q+1, C_4) = \frac{1}{2}q(q+1)^2$ for all prime powers $q \geq 15$. It was also shown in [43] that the only graphs with q^2+q+1 vertices and $\frac{1}{2}q(q+1)^2$ edges that do not contain a 4-cycle are orthogonal polarity graphs of finite projective planes. Along with the constructions of [14] and [32], the results of Füredi are the most important contributions to the 4-cycle Turán problem. Recently Firke, Kosek, Nash, and Williford [38] proved that for even q, $\mathrm{ex}(q^2+q, C_4) \leq \frac{1}{2}q(q+1)^2 - q$. If q is a power of two, then we have the exact result $\mathrm{ex}(q^2+q, C_4) = \frac{1}{2}q(q+1)^2 - q$. The lower bound in this case comes from taking an orthogonal polarity graph of a projective plane of order q and removing a vertex of degree q. In [38], it is stated that if q is a power of two and G is a C_4-free graph with q^2+q vertices and $\frac{1}{2}q(q+1)^2 - q$ edges, then G must be obtained by deleting a vertex of minimum degree from a polarity graph. Hence we have characterizations of the extremal graphs when $n = q^2+q+1$ and q is a prime power, and when $n = q^2+q$ and q is a power of two. The results we have mentioned so far describe all of the cases in which an exact formula for $\mathrm{ex}(n, C_4)$ is known.

While investigating adjacency matrices of polarity graphs, Abreu, Balbuena, and Labbate [1] were able to find subgraphs of a polarity graph that have many edges. By deleting such a subgraph, they proved that for any prime power q,

$$\mathrm{ex}(q^2-q-2, C_4) \geq \begin{cases} \frac{1}{2}q^3 - q^2 - \frac{q}{2} + 1 & \text{if } q \text{ is odd,} \\ \frac{1}{2}q^3 - q^2 & \text{if } q \text{ is even.} \end{cases}$$

They conjectured that these bounds are best possible. Our main result shows that when q is an odd prime power, this lower bound can be improved by $\frac{q}{2} - O(q^{3/4})$.

Theorem 4.1.1 (Tait, Timmons, 2013)**.** *If q is an odd prime power, then*

$$\mathrm{ex}(q^2-q-2, C_4) \geq \frac{1}{2}q^3 - q^2 - O(q^{3/4}).$$

We will construct graphs without 4-cycles using the Sidon sets constructed by Bose and Chowla [13]. Let Γ be an abelian group. Recall a set $A \subset \Gamma$ is a *Sidon*

set if whenever $a+b=c+d$ with $a,b,c,d \in A$, the pair (a,b) is a permutation of (c,d).

Let q be a prime power and θ be a generator of the multiplicative group $\mathbb{F}_{q^2}^*$ where $\mathbb{F}_{q^2}^*$ is the nonzero elements of the finite field $\mathbb{F}_{q^2}$. Bose and Chowla proved [13] that

$$A(q,\theta) := \{a \in \mathbb{Z}_{q^2-1} : \theta^a - \theta \in \mathbb{F}_q\}$$

is a Sidon set in the group $\mathbb{Z}_{q^2-1}$ (see also Chapter 3, Section 3.2).

Definition 4.1.2. Let q be a prime power and θ be a generator of the multiplicative group $\mathbb{F}_{q^2}^*$. The graph $G_{q,\theta}$ is the graph with vertex set $\mathbb{Z}_{q^2-1}$ and two distinct vertices i and j are adjacent if and only if $i+j=a$ for some $a \in A(q,\theta)$.

It is known that Sidon sets can be used to construct C_4-free graphs. We will prove a result about the Bose-Chowla Sidon sets, Lemma 4.2.6, that will help us find a subgraph of $G_{q,\theta}$ with $q+1$ vertices and has many edges. We remove this subgraph to obtain a graph with q^2-q-2 vertices and at least $\frac{1}{2}q^3 - q^2 - O(q^{3/4})$ edges. In addition to providing examples of graphs with no 4-cycle, the graphs $G_{q,\theta}$ have been used to solve other extremal problems [16].

We would like to remark that we could have defined $G_{q,\theta}$ as a polarity graph in the following way. Let $\mathcal{P} = \mathbb{Z}_{q^2-1}$ and let $\mathcal{L}$ be the set of q^2-1 translates of $A(q,\theta)$. That is, $\mathcal{L} = \{A_1, A_2, \ldots, A_{q^2-1}\}$ where $A_i := A(q,\theta) + i$. This defines a geometry in the obvious way: $i \in \mathcal{P}$ is incident to $A_j \in \mathcal{L}$ if and only if $i \in A_j$. We define a polarity by $\pi(i) = A_{q^2-1-i}$ for all $i \in \mathcal{P}$, and $\pi(A_i) = q^2-1-i$ for all $A_i \in \mathcal{L}$. The fact that π is a polarity can be checked directly. We choose to use Definition 4.1.2 as it is more convenient for our argument.

4.2 Proof of Theorem 4.1.1

Fix an odd prime power q and a generator θ of the multiplicative group $\mathbb{F}_{q^2}^*$. We write A for the Sidon set $A(q,\theta)$ in $\mathbb{Z}_{q^2-1}$ and observe that $|A| = q$. All of our manipulations will be done in the group $\mathbb{Z}_{q^2-1}$ or in the finite field $\mathbb{F}_{q^2}$. If it is not clear from the context we will state which algebraic structure we are working in.

The first two lemmas are known. We present proofs for completeness.

Lemma 4.2.1. *The graph $G_{q,\theta}$ does not contain a 4-cycle.*

Proof. Suppose $ijkl$ is a 4-cycle in $G_{q,\theta}$. There are elements $a, b, c, d \in A$ such that $i + j = a$, $j + k = b$, $k + l = c$, and $l + i = d$. This implies

$$a + c = b + d.$$

Since A is a Sidon set, (a, c) is a permutation of (b, d). If $a = b$ then $i + j = j + k$ so $i = k$. If $a = d$ then $i + j = l + i$ so $j = l$. In either case we have a contradiction thus $G_{q,\theta}$ does not contain a 4-cycle. □

Lemma 4.2.2. *If $A - A := \{a - b : a, b \in A\}$ then*

$$A - A = \mathbb{Z}_{q^2-1} \backslash \{q + 1, 2(q + 1), 3(q + 1), \ldots, (q - 2)(q + 1)\}.$$

Proof. Suppose $s(q + 1) \in A - A$ for some $1 \leq s \leq q - 2$. Write $s(q + 1) = a - b$ where $a, b \in A$ and $a \neq b$. We have for some $\alpha, \beta \in \mathbb{F}_q$,

$$\theta^{s(q+1)} = \theta^{a-b} = \theta^a \theta^{-b} = (\theta + \alpha)(\theta + \beta)^{-1}.$$

From this we obtain

$$\theta + \alpha = (\theta + \beta)(\theta^{q+1})^s$$

but $\theta^{q+1} \in \mathbb{F}_q$ so $\theta + \alpha = (\theta + \beta)\gamma$ for some $\gamma \in \mathbb{F}_q$. Since θ does not satisfy a nontrivial linear relation over $\mathbb{F}_q$ we must have $\gamma = 1$ hence $\alpha = \beta$ (in $\mathbb{F}_{q^2}$) so $a = b$ (in $\mathbb{Z}_{q^2-1}$). From this we get $s(q + 1) = 0$ which contradicts the fact that $1 \leq s \leq q - 2$. This shows that

$$(A - A) \cap \{q + 1, 2(q + 1), \ldots, (q - 2)(q + 1)\} = \emptyset.$$

Since A is a Sidon set, $|A - A| = q(q - 1) + 1$ which is precisely the number of elements in the set

$$\mathbb{Z}_{q^2-1} \backslash \{q + 1, 2(q + 1), \ldots, (q - 2)(q + 1)\}$$

and this completes the proof of the lemma. □

Let i be a vertex in $G_{q,\theta}$. If $i + i \in A$ then the degree of i is $q - 1$. If $i + i \notin A$ then the degree of i is q. We call a vertex of degree $q - 1$ an *absolute point*.

Lemma 4.2.3. *Distinct vertices i and j in $G_{q,\theta}$ have a common neighbor if and only if $i - j \in (A - A)\backslash\{0\}$.*

Proof. First suppose i and j are distinct vertices that have a common neighbor k. Then $i + k = a$ and $k + j = b$ for some $a, b \in A$ so $i - j = (a - k) - (b - k) = a - b$. Since $i \neq j$, we get that $a - b \neq 0$.

Now suppose $i - j = a - b$ for some $a, b \in A$ with $a \neq b$. Let $k = a - i$. Then $k + i = a$ so k is adjacent to i. Also, $k = a - i = b - j$ so $k + j = b$ and k is adjacent to j. □

Lemma 4.2.4. *If i is an absolute point then $i + \frac{q^2-1}{2}$ is also an absolute point.*

Proof. If $2i = a$ for some $a \in A$ then $2(i + \frac{q^2-1}{2}) = 2i = a$. □

Lemma 4.2.5. *Let i and j be two distinct absolute points of $G_{q,\theta}$. If $i \neq j + \frac{q^2-1}{2}$ then i and j have a common neighbor and if $i = j + \frac{q^2-1}{2}$ then i and j do not have a common neighbor.*

Proof. By Lemmas 4.2.2 and 4.2.3, i and j have a common neighbor unless $i - j = s(q + 1)$ for some $1 \leq s \leq q - 2$. Since i and j are absolute points, there exists elements $a, b \in A$ such that $2i = a$ and $2j = b$ thus $a - b = 2s(q + 1)$. By Lemma 4.2.2, it must be the case that $a = b$ so $2i = 2j$. The solutions to $2x \equiv 2y (\text{mod } q^2 - 1)$ are $x = y$ and $x = y + \frac{q^2-1}{2}$ hence $i = j$ or $i = j + \frac{q^2-1}{2}$. Thus i and j will have a common neighbor whenever they are distinct absolute points with $i \neq j + \frac{q^2-1}{2}$ and will not have a common neighbor when $i = j + \frac{q^2-1}{2}$. □

Lemma 4.2.6. *Let $\{a_1, a_2, a_3\}$ and $\{b_1, b_2, b_3\}$ be subsets of A with a_1, a_2, and a_3 all distinct and b_1, b_2, and b_3 all distinct. If*

$$2b_1 - a_1 = 2b_2 - a_2 = 2b_3 - a_3$$

then two of the ordered pairs (a_1, b_1), (a_2, b_2), (a_3, b_3) are equal.

The proof of Lemma 4.2.6 is simple but it is not short. For this reason we postpone the proof until after the proof of Theorem 4.1.1.

Lemma 4.2.7. *Any vertex j is adjacent to at most two absolute points.*

Proof. Suppose j is a vertex of $G_{q,\theta}$ that is adjacent to three distinct absolute points i_i, i_2, and i_3. There exists elements $a_1, a_2, a_3, b_1, b_2, b_3 \in A$ such that

$$2i_k = a_k \text{ and } i_k + j = b_k$$

for $k = 1, 2, 3$. Since i_1, i_2, i_3 are all distinct, b_1, b_2, and b_3 must all be distinct. If $a_k = a_l$ for some $1 \leq k < l \leq 3$ then $i_k = i_l + \frac{q^2-1}{2}$. In this case, the vertices i_k and i_l are absolute points with a common neighbor but this is impossible by Lemma 4.2.5. We conclude that a_1, a_2, and a_3 are all distinct. For each k, we can write $i_k + j = b_k$ as $2j = 2b_k - a_k$ so that

$$2b_1 - a_2 = 2b_2 - a_2 = 2b_3 - a_3.$$

By Lemma 4.2.6, $(a_k, b_k) = (a_l, b_l)$ for some $1 \leq k < l \leq 3$ but we have already argued that a_k and a_l are distinct. This gives the needed contradiction and completes the proof of the lemma. □

Proof of Theorem 4.1.1. Let P be the absolute points of $G_{q,\theta}$. By Lemma 4.2.4, the absolute points come in pairs so we can write

$$P = \{i_1, i_1 + \frac{q^2-1}{2}, i_2, i_2 + \frac{q^2-1}{2}, \ldots, i_t, i_t + \frac{q^2-1}{2}\}$$

where $2t$ is the number of absolute points of $G_{q,\theta}$. When q is odd, $q^2 - 1$ is even and we can write $q^2 - 1 = 2^r m$ where $r \geq 1$ is an integer and m is odd. If $a \in A$ then the congruence

$$2x \equiv a (\text{mod } 2^r m)$$

has no solution when a is odd and two solutions if a is even. Therefore t is exactly the number of even elements of A when we view A as a subset of $\mathbb{Z}$. Lindström [60] proved that dense Sidon sets are close to evenly distributed among residue classes. In particular, the results of [60] imply that

$$t = \frac{q}{2} + O(q^{3/4}) \tag{4.1}$$

so we know that we have $q+O(q^{3/4})$ absolute points in $G_{q,\theta}$. The number of vertices of $G_{q,\theta}$ is q^2-1 and the number of edges of $G_{q,\theta}$ is

$$e(G) = \frac{1}{2}\left(q(q^2-1-2t)+(q-1)(2t)\right) = \frac{1}{2}q^3 - \frac{1}{2}q - t.$$

Let $S \subset V(G_{q,\theta})$ with $|S| = q+1$ and let t_S be the number of absolute points in S. The graph $G_{q,\theta}\backslash S$ has q^2-q-2 vertices and

$$\frac{1}{2}q^3 - \frac{1}{2}q - t - e(S) - e(S,\overline{S}) \tag{4.2}$$

edges. Here $e(S,\overline{S})$ is the number of edges of $G_{q,\theta}$ with exactly one endpoint in S. We can rewrite $e(S)+e(S,\overline{S})$ as

$$e(S)+e(S,\overline{S}) = \sum_{i\in S} d(i) - e(S) = (q+1-t_S)q + t_S(q-1) - e(S) = q^2+q-t_S-e(S).$$

By (4.2) we can write the number of edges of $G_{q,\theta}\backslash S$ as

$$\frac{1}{2}q^3 - \frac{1}{2}q - t - (q^2+q-t_S-e(S)) = \frac{1}{2}q^3 - q^2 - \frac{3}{2}q - t + t_S + e(S). \tag{4.3}$$

For any $1 \le j_1 < j_2 \le t$, the pair i_{j_1} and i_{j_2} of absolute points have a unique common neighbor by Lemmas 4.2.5 and 4.2.1. Set $k = \lfloor \frac{1}{2}\sqrt{8q+9} - \frac{1}{2} \rfloor$ and note that for large enough q we have $k \le t$. The integer k is chosen so that it is as large as possible and still satisfies the inequality $\binom{k}{2} + k \le q+1$. Let $S_1 = \{i_1, \ldots, i_k\}$. For each pair $1 \le j_1 < j_2 \le k$, let x_{j_1,j_2} be the unique common neighbor of the absolute points i_{j_1} and i_{j_2}. Let $S_2 = \{x_{j_1,j_2} : 1 \le j_1 < j_2 \le k\}$. By Lemma 4.2.7, S_2 consists of $\binom{k}{2}$ distinct vertices. A short calculation shows that $\binom{k}{2} + k \ge q - O(\sqrt{q})$. Let S_3 be a set of $q+1-\binom{k}{2}-k$ vertices chosen arbitrarily from $V(G_{q,\theta})\backslash(S_1 \cup S_2)$. Let S be the subgraph of $G_{q,\theta}$ induced by the vertices $S_1 \cup S_2 \cup S_3$. By construction, S has $q+1$ vertices and at least $2\binom{k}{2}$ edges so

$$t_S + e(S) \ge k + 2\binom{k}{2} \ge 2q - O(\sqrt{q}).$$

By (4.1) and (4.3), removing the vertices of S from $G_{q,\theta}$ leaves a graph with q^2-q-2 vertices and at least

$$\frac{1}{2}q^3 - q^2 - 2q + 2q - O(q^{3/4}) = \frac{1}{2}q^3 - q^2 - O(q^{3/4}).$$

edges. □

Now we return to the proof of Lemma 4.2.6.

Proof of Lemma 4.2.6. Let $\{a_1, a_2, a_3\}, \{b_1, b_2, b_3\} \subset A$ with a_1, a_2, and a_3 all distinct, and b_1, b_2, and b_3 all distinct. Since $a_k, b_k \in A$, there exists elements $c_k, d_k \in \mathbb{F}_q$ such that

$$\theta^{a_k} = \theta + c_k \text{ and } \theta^{b_k} = \theta + d_k$$

for $k = 1, 2, 3$. Observe that c_1, c_2, and c_3 are all distinct and so are d_1, d_2, and d_3.

The generator θ satisfies a degree two polynomial over $\mathbb{F}_q$, say $\theta^2 = \alpha\theta + \beta$ where $\alpha, \beta \in \mathbb{F}_q$. Since θ generates $\mathbb{F}_{q^2}^*$, it cannot be the case that $\alpha = 0$ and if $\beta = 0$, then $\theta(\theta - \alpha) = 0$ which is impossible since $\theta \notin \mathbb{F}_q$. The polynomial $X^2 - 3X + 3\beta \in \mathbb{F}_q[X]$ has at most two roots in $\mathbb{F}_q$. Without loss of generality, we may assume that

$$c_1^2 - 3c_1\alpha + 3\beta \neq 0 \tag{4.4}$$

since c_1, c_2, and c_3 are all distinct. This fact will be important towards the end of the proof.

Consider the equation $2b_1 + a_2 = 2b_2 + a_1$. We can rewrite this as

$$(\theta + d_1)^2(\theta + c_2) = (\theta + d_2)^2(\theta + c_1).$$

If we expand, use $\theta^2 = \alpha\theta + \beta$, and regroup we obtain

$$\begin{aligned} & \theta(2d_1\alpha + c_2\alpha + d_1^2 + 2d_1c_2) + (2d_1\beta + c_2\beta + d_1^2c_2) \\ = \ & \theta(2d_2\alpha + c_1\alpha + d_2^2 + 2d_2c_1) + (2d_2\beta + c_1\beta + d_2^2c_1). \end{aligned}$$

These coefficients are all in $\mathbb{F}_q$ so we must have

$$2d_1\alpha + c_2\alpha + d_1^2 + 2d_1c_2 = 2d_2\alpha + c_1\alpha + d_2^2 + 2d_2c_1 \tag{4.5}$$

and

$$2d_1\beta + c_2\beta + d_1^2c_2 = 2d_2\beta + c_1\beta + d_2^2c_1. \tag{4.6}$$

Similar arguments show that both (4.5) and (4.6) hold with c_3 replacing c_2 and d_3 replacing d_2. We view c_1 and d_1 as begin fixed and (c_2, d_2) and (c_3, d_3) as solutions to the system

$$2d_1\alpha + X\alpha + d_1^2 + 2d_1X = 2Y\alpha + c_1\alpha + Y^2 + 2Yc_1, \tag{4.7}$$

$$2d_1\beta + X\beta + d_1^2X = 2Y\beta + c_1\beta + Y^2c_1. \tag{4.8}$$

One solution is $(X, Y) = (c_1, d_1)$. If we can show that the system (4.7), (4.8) has at most two solutions then we are done as this forces two of the pairs (c_1, d_1), (c_2, d_2), (c_3, d_3) to be the same and the pair (c_k, d_k) uniquely determines the pair (a_k, b_k). Multiply (4.7) by c_1 and then subtract (4.8) to eliminate Y^2 and obtain

$$(2c_1d_1\alpha + c_1d_1^2 + c_1\beta - 2d_1\beta - c_1^2\alpha) + X(\alpha c_1 + 2c_1d_1 - \beta - d_1^2) = Y(2c_1^2 + 2c_1\alpha - 2\beta). \tag{4.9}$$

Next we subtract α times (4.8) from β times (4.7) to get

$$d_1^2\beta + X(2d_1\beta - d_1^2\alpha) = Y^2(\beta - \alpha c_1) + Y(2c_1\beta). \tag{4.10}$$

If we knew that the coefficient of X was nonzero in (4.9) and $\beta - \alpha c_1 \neq 0$ then we could easily deduce that there are at most two solutions (X, Y). Unfortunately we do not know this and so we have to work to overcome this obstacle.

Suppose (4.9) is an equation where the coefficients of X and Y are both 0. Then

$$2c_1^2 + 2c_1\alpha - 2\beta = 0 \quad \text{and} \quad \alpha c_1 + 2c_1d_1 - \beta - d_1^2 = 0.$$

Since q is odd, the first equation can be rewritten as $c_1^2 + c_1\alpha - \beta$. Subtracting the second equation $c_1^2 + c_1\alpha - \beta$ gives $c_1^2 - 2c_1d_1 + d_1^2 = 0$ hence $(c_1 - d_1)(c_1 + d_1) = 0$.

If $c_1 = d_1$ then $\theta^{a_1} = \theta + c_1 = \theta + d_1 = \theta^{b_1}$ so $a_1 = b_1$ (in $\mathbb{Z}_{q^2-1}$). Using $2b_1 - a_1 = 2b_2 - a_2$ we get $b_1 + a_2 = b_2 + b_2$ so $b_1 = b_2$, a contradiction. Assume $c_1 = -d_1$. Then $c_1 \neq 0$ and $d_1 \neq 0$ otherwise $c_1 = d_1$ which we already know does not occur. Since both coefficients of X and Y are 0 in (4.9) the constant term must also be 0 so, using $c_1 = -d_1$,

$$\begin{aligned} 0 &= 2c_1d_1\alpha + c_1d_1^2 + c_1\beta - 2d_1\beta - c_1^2\alpha \\ &= -3c_1^2\alpha + c_1^3 + 3c_1\beta \\ &= c_1(c_1^2 - 3c_1\alpha + 3\beta). \end{aligned}$$

By (4.4) this is impossible. We conclude that at least one of the coefficients of X or Y in (4.9) must be nonzero.

If the coefficient of X in (4.9) is nonzero then we can write $X = \gamma_1 Y + \gamma_2$ for some $\gamma_1, \gamma_2 \in \mathbb{F}_q$. Substituting this equation into (4.7) gives a quadratic equation in

Y which has at most two solutions and Y uniquely determines X since $X = \gamma_1 Y + \gamma_2$ and we are done.

Assume now that $\alpha c_1 + 2c_1 d_1 - \beta - d_1^2 = 0$. Then (4.9) gives a unique solution for Y. Since $(X, Y) = (c_1, d_1)$ is a solution we must have that all solutions to the system (4.7), (4.8) have $Y = d_1$. Substituting into (4.7) and (4.8) we get

$$X(\alpha + 2d_1) = c_1(\alpha + 2d_1)$$
$$X(\beta + d_1^2) = c_1(\beta + d_1^2).$$

If $d_1 = 0$ then $X\alpha = c_1\alpha$ and since $\alpha \neq 0$ we get $X = c_1$ are we are done.

Assume $d_1 \neq 0$. If either $\alpha + 2d_1$ or $\beta + d_1^2$ are nonzero then we are done. Assume $\alpha + 2d_1 = \beta + d_1^2 = 0$. If we substitute $Y = d_1$ into (4.10) then we get

$$Xd_1(2\beta - d_1\alpha) = d_1 c_1(2\beta - d_1\alpha).$$

Again, if $2\beta - d_1\alpha$ is nonzero we are done so assume $2\beta - d_1\alpha = 0$. Using the three equations

$$\alpha + 2d_1 = 0,\ \beta + d_1^2 = 0\ , 2\beta - d_1\alpha = 0$$

we have

$$0 = 2\beta - d_1\alpha = 2(-d_1^2) - d_1(-4d_1) = 2d_1^2$$

so $d_1 = 0$ giving the needed contradiction.

□

Chapter 4 is a reprint of "Sidon sets and graphs without 4-cycles," 2013. Tait, Michael; Timmons, Craig. This paper has been submitted for publication. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 5

Ordered Turán Problems

In this Chapter we introduce an ordered version of the Turán problem for bipartite graphs. Let G be a graph with $V(G) = \{1, 2, \ldots, n\}$ and view the vertices of G as being ordered in the natural way. A *zig-zag* $K_{s,t}$, denoted $Z_{s,t}$, is a complete bipartite graph $K_{s,t}$ whose parts $A = \{n_1 < n_2 < \cdots < n_s\}$ and $B = \{m_1 < m_2 < \cdots < m_t\}$ satisfy the condition $n_s < m_1$. A *zig-zag* C_{2k} is an even cycle C_{2k} whose vertices in one part precede all of those in the other part. Write $\mathcal{Z}_{2k}$ for the family of zig-zag $2k$-cycles. We investigate the Turán numbers $\mathrm{ex}(n, Z_{s,t})$ and $\mathrm{ex}(n, \mathcal{Z}_{2k})$. In particular, we show that $\mathrm{ex}(n, Z_{2,2}) \leq \frac{2}{3}n^{3/2} + O(n^{5/4})$. For infinitely many n we construct a $Z_{2,2}$-free n-vertex graph with at least $(n - \sqrt{n} - 1) + \mathrm{ex}(n, K_{2,2})$ edges.

Our motivation for introducing ordered Turán problems comes from a connection between Sidon sets and Z_4-free graphs. As discussed in Chapter 2, obtaining good lower bounds on $\mathrm{ex}(n, C_{2k})$ is a difficult problem. Using generalizations of Sidon sets called B_k-sets, we can construct graphs that are $\mathcal{Z}'_{2k}$-free, where $\mathcal{Z}'_{2k}$ is a particular subfamily of $\mathcal{Z}_{2k}$. This construction works for every $k \geq 2$, and gives lower bounds of order $n^{1+1/k}$ on the ordered Turán number $\mathrm{ex}(n, \mathcal{Z}'_{2k})$.

5.1 Introduction

Given an n-vertex graph G, label its vertices with the integers in $[n]$ using each integer exactly once. This induces a natural ordering of the vertices of G and

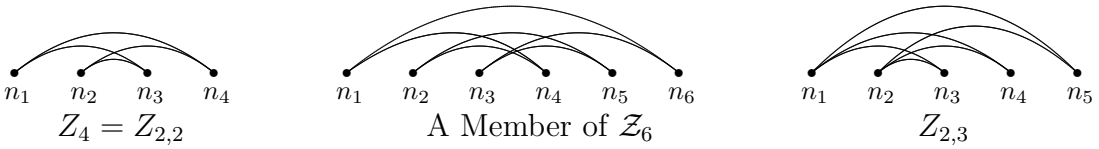

Figure 5.1: $Z_4 = Z_{2,2}$, a Member of $\mathcal{Z}_6$, and $Z_{2,3}$

we use this ordering to distinguish between different types of a fixed subgraph. This idea is not new to Turán theory. Czipszer, Erdős, and Hajnal [23] and Dudek and Rödl [24] investigated Turán problems for increasing paths of length k. An increasing path of length k is a sequence of k edges $n_1n_2, n_2n_3, \ldots, n_kn_{k+1}$ such that $n_1 < n_2 < \cdots < n_{k+1}$.

Let H be a bipartite graph with parts A and B. Let $f : \{1, 2\} \to \{A, B\}$ be a bijection and call f an *ordering* of the parts. A *zig-zag* H relative to f and bipartition $\{A, B\}$ is a copy of H in G such that all of the vertices in $f(1)$ precede all of the vertices in $f(2)$. One of the reason we consider zig-zag bipartite graphs as opposed to bipartite graphs that do not zig-zag is because there exist graphs with $\frac{1}{8}n^2 + o(n^2)$ edges that do not contain increasing paths of length 2. One such graph is obtained by joining each even vertex to all of the odd vertices that come after it in the ordering. If a complete bipartite graph does not zig-zag, then it will contain an increasing path of length 2. In contrast, if a zig-zag bipartite graph is forbidden, then the number of edges will not be quadratic in n. Our focus will be on zig-zag complete bipartite graphs and zig-zag even cycles so we specialize the notation.

As before let G be an n-vertex graph with $V(G) = [n]$ and consider the vertices of G as being ordered. A *zig-zag* $K_{s,t}$, which will be denoted by $Z_{s,t}$, is a $K_{s,t}$ whose parts $A = \{n_1 < n_2 < \cdots < n_s\}$ and $B = \{m_1 < m_2 < \cdots < m_t\}$ satisfy the condition $n_s < m_1$. A *zig-zag* $2k$*-cycle*, denoted Z_{2k}, is a $2k$-cycle whose vertices are $\{n_1 < n_2 < \cdots < n_{2k}\}$ and $A = \{n_1, \ldots, n_k\}$, $B = \{n_{k+1}, \ldots, n_{2k}\}$ is the bipartition. Let $\mathcal{Z}_{2k}$ be the family of all zig-zag $2k$-cycles. Observe that for $k = 2$, $\mathcal{Z}_{2k}$ consists of a single graph and we simply write Z_4 for this family.

Any n-vertex $K_{s,t}$-free graph G can be used to define a $Z_{s,t}$-free graph so $\mathrm{ex}(n, K_{s,t}) \leq \mathrm{ex}(n, Z_{s,t})$. A non-trivial relation between $\mathrm{ex}(n, K_{s,t})$ and $\mathrm{ex}(n, Z_{s,t})$ can be viewed as a compactness result (see [34]) since one is forbidding a special

type of $K_{s,t}$ rather than forbidding all $K_{s,t}$'s. The same remark applies to zig-zag even cycles as well.

Our original motivation for investigating zig-zag bipartite graphs comes from a problem in additive number theory. A set $A \subset \mathbb{Z}$ is a *B_2-set* if $a_1 + a_2 = b_1 + b_2$ with $a_i, b_j \in A$ implies $\{a_1, a_2\} = \{b_1, b_2\}$. B_2-sets, also called *Sidon* sets, were introduced in the early 1930's and since then they have attracted the attention of many researchers. These sets were discussed in Chapter 3, but we briefly recall a few results. Let $F_2(n)$ be the maximum size of a B_2-set contained in $[n]$. Erdős and Turán [37] proved $F_2(n) < n^{1/2} + O(n^{1/4})$. In 1968 Lindström [57], refining the argument of Erdős and Turán, proved $F_2(n) < n^{1/2} + n^{1/4} + 1$. Cilleruelo [19] obtained $F_2(n) < n^{1/2} + n^{1/4} + 1/2$ as a consequence of a more general result. Erdős conjectured $F_2(n) < n^{1/2} + O(1)$ and offered \$500 for a proof or disproof of this conjecture [26]. The error term of $n^{1/4}$ has not been improved since the original argument of Erdős and Turán.

The connection between B_2-sets and ordered Turán theory is given by the following construction. Let $A \subset [n]$ be a B_2-set. Define the graph G_A by $V(G_A) = [n]$ and

$$E(G_A) = \{ij : i = j + a \text{ for some } a \in A\}.$$

It is easily checked that G_A is Z_4-free and so bounds on $\mathrm{ex}(n, Z_4)$ translate to bounds on $F_2(n)$.

Our first theorem gives an upper bound on $\mathrm{ex}(n, \mathcal{Z}_{2k})$.

Theorem 5.1.1. *Let $k \geq 2$ be an integer. For any n,*

$$\mathrm{ex}(n, \mathcal{Z}_{2k}) \leq \frac{k - 3/2}{2^{1/k} - 1} n^{1+1/k} + (2k - 3)n \log_2 n.$$

For comparison, recall that the best known upper bound on $\mathrm{ex}(n, C_{2k})$ is due to Pikhurko [67] who showed that

$$\mathrm{ex}(n, C_{2k}) \leq (k - 1)n^{1+1/k} + 16(k - 1).$$

The next theorem gives an upper bound on $\mathrm{ex}(n, Z_{s,t})$ corresponding to Füredi's upper bound on $\mathrm{ex}(n, K_{s,t})$.

Theorem 5.1.2. *Let $t \geq s \geq 2$ be integers. For any n,*

$$\mathrm{ex}(n, Z_{s,t}) \leq \frac{(t-1)^{1/s}}{2-1/s} n^{2-1/s} + \left((t-1)^{1/s} + \frac{1}{2}(s-1)\right) n^{3/2-1/2s} + (s-1)n.$$

For small values of s and t there is a gap between the upper bounds on $\mathrm{ex}(n, K_{s,t})$ and $\mathrm{ex}(n, Z_{s,t})$. When $s = t = 2$, the upper bound of Theorem 5.1.2 gives $\mathrm{ex}(n, Z_{2,2}) \leq \frac{2}{3}n^{3/2} + O(n^{5/4})$ whereas (2.2) gives $\mathrm{ex}(n, K_{2,2,}) \leq \frac{1}{2}n^{3/2} + O(n)$.

We are able to give a construction using projective planes which shows

$$\limsup_{n\to\infty} (\mathrm{ex}(n, Z_4) - \mathrm{ex}(n, C_4)) = \infty.$$

Unfortunately we were unable to determine whether or not $\mathrm{ex}(n, Z_4) \sim \mathrm{ex}(n, C_4)$, but we do have the following theorem.

Theorem 5.1.3. *For any prime p,*

$$\mathrm{ex}(p^2 + p + 1, Z_4) \geq p^2 + \mathrm{ex}(p^2 + p + 1, C_4).$$

The upper bound $\mathrm{ex}(n, Z_4) \leq \frac{2}{3}n^{3/2} + o(n^{3/2})$ given by Theorem 5.1.2 can probably be improved. We believe that the constructions are best possible.

Conjecture 5.1.4. The zig-zag Turán number $\mathrm{ex}(n, Z_4)$ satisfies

$$\mathrm{ex}(n, Z_4) \leq \frac{1}{2}n^{3/2} + o(n^{3/2}).$$

The notion of compactness [34] has produced several interesting problems concerning Turán numbers for bipartite graphs. Similar questions can be asked for our ordered version of the problem.

Problem 5.1.5. Is it true that for any bipartite graph H and any zig-zag ZH we have

$$\mathrm{ex}(n, ZH) = O(\mathrm{ex}(n, H))?$$

A positive answer to Problem 5.1.5 is supported by Theorem 5.1.2.

Another interesting problem related to compactness is the following. Let $\mathcal{Z}_{2k}^{\times}$ be the sub-family of $\mathcal{Z}_{2k}$ that consists of all Z_{2k}'s with a longest or shortest edge.

Problem 5.1.6. Is $\mathrm{ex}(n, \mathcal{Z}_{2k}^{\times}) = O(n^{1+1/k})$ for $k \geq 3$?

We will discuss Problem 5.1.6 in more detail in Chapter 9. For now we remark that it is not difficult to show $\mathrm{ex}(n, \mathcal{Z}_{2k}^{\times}) > cn^{1+1/k}$ for all $k \geq 3$. This may come as a surprise considering the difficulty in finding good lower bounds on $\mathrm{ex}(n, C_{2k})$ for $k \notin \{2, 3, 5\}$.

In the next three sections we prove Theorems 5.1.1, 5.1.2, and 5.1.3. In this chapter all floor and ceiling symbols are omitted whenever they do not affect the results.

5.2 Proof of Theorem 5.1.1

Let $k \geq 2$ be an integer. Let G be a $\mathcal{Z}_{2k}$-free graph with $V(G) = [n]$. Given two subsets $A, B \subset V(G)$, write $A < B$ if all of the elements of A are less than the smallest element of B. For disjoint subsets $A, B \subset V(G)$, let $G(A, B)$ be the subgraph of G with $V(G(A, B)) = A \cup B$ and

$$E(G(A, B)) = \{ij \in E(G) : i \in A, j \in B\}.$$

For any pair of subsets $A, B \subset V(G)$ with $A < B$, the graph $G(A, B)$ is C_{2k}-free since G is $\mathcal{Z}_{2k}$-free. Given integers m_1 and m_2, let $\mathrm{ex}(m_1, m_2, C_{2k})$ be the maximum number of edges in a C_{2k}-free bipartite graph with m_1 vertices in one part and m_2 vertices in the other. Naor and Verstraëte [63] proved an upper bound on $\mathrm{ex}(m_1, m_2, C_{2k})$ that implies a C_{2k}-free bipartite graph with m vertices in each part has at most $(2k-3)(m^{1+1/k} + 2m)$ edges. Applying this bound to $G(A_1, B_1)$ where $A_1 = \{1, 2, \ldots, n/2\}$ and $B_1 = \{n/2+1, n/2+2, \ldots, n\}$ gives

$$e(G(A_1, B_1)) \leq (2k-3)((n/2)^{1+1/k} + n).$$

We repeat the argument on the sets $A_{2,1} = \{1, 2, \ldots, n/4\}$, $B_{2,1} = \{n/4+1, n/4+2, \ldots, n/2\}$ and on the sets $A_{2,2} = \{n/2+1, n/2+2, \ldots, 3n/4\}$, $B_{2,2} = \{3n/4+$

$1, 3n/4+2, \ldots, n\}$. Continuing in this fashion gives

$$\begin{aligned} e(G) &\leq \sum_{l=1}^{\log_2 n} 2^{l-1}\mathrm{ex}(n2^{-l}, n2^{-l}, C_{2k}) \\ &\leq \sum_{l=1}^{\log_2 n} 2^{l-1}(2k-3)\left((n/2^l)^{1+1/k} + 2n/2^l\right) \\ &\leq \frac{(k-3/2)n^{1+1/k}}{2^{1/k}} \sum_{l=0}^{\infty} \left(\frac{1}{2^{1/k}}\right)^l + (2k-3)n\log_2 n \\ &= \frac{k-3/2}{2^{1/k}-1} n^{1+1/k} + (2k-3)n\log_2 n. \end{aligned}$$

5.3 Proof of Theorem 5.1.2

The following lemma was used by Füredi [40] in proving (2.2). The proof of the lemma is an easy application of Jensen's Inequality. For $k \geq 1$ and $x \geq k-1$ define $\binom{x}{k} = \frac{1}{k!}x(x-1)\cdots(x-k+1)$. If $k-1 > x \geq 0$ define $\binom{x}{k} = 0$. For fixed k each of these functions is convex.

Lemma 5.3.1 (Füredi, [40]). *If $n, k \geq 1$ are integers and $c, y, x_1, \ldots, x_k$ are non-negative real numbers and $\sum_{i=1}^n \binom{x_i}{k} \leq c\binom{y}{k}$ then*

$$\sum_{i=1}^n x_i \leq yc^{1/k}n^{1-1/k} + (k-1)n. \tag{5.1}$$

Proof. Let $s = \sum_{i=1}^n x_i$. If $s \leq n(k-1)$ then the inequality holds so assume $\frac{s}{n} - k + 1 > 0$. Apply Jensen's Inequality to get $\sum_{i=1}^n \binom{x_i}{k} \geq n\binom{s/n}{k}$ which implies $c\binom{y}{k} \geq n\binom{s/n}{k}$. Rearranging this inequality gives

$$\frac{y(y-1)(y-2)\cdots(y-k+1)}{(s/n)(s/n-1)\cdots(s/n-k+1)} \geq \frac{n}{c}.$$

The left hand side can be bounded above by $\left(\frac{y}{s/n-k+1}\right)^k$ to get $\left(\frac{y}{s/n-k+1}\right)^k \geq \frac{n}{c}$. Solving this inequality for s gives (5.1). □

Define the *back neighborhood* of a vertex $i \in V(G)$ to be the set

$$\Gamma^-(i) = \{j < i : ji \in E(G)\}.$$

Let G be a $Z_{s,t}$-free graph with $V(G) = \{1, 2, \ldots, n\}$. Define an n by n bipartite graph H with parts $L = \{b_1, b_2, \ldots, b_n\}$, $P = \{1, 2, \ldots, n\}$, and edge set

$$E(H) = \{\{i, b_j\} : i \in \Gamma^-(j)\}.$$

H is the incidence graph of the back neighborhoods $\{\Gamma^-(i)\}_{i=1}^n$ of G.

It is easy to check that $e(H) = e(G)$ and H has no complete bipartite subgraph with t vertices in L and s vertices in P. Let $k = n^{1/2-1/2s}$. For $j = 1, 2, \ldots, k$ let

$$P_j = \{1 + (j-1)\tfrac{n}{k}, 2 + (j-1)\tfrac{n}{k}, \ldots, \tfrac{jn}{k}\}.$$

Any back neighborhood $\Gamma^-(i)$ is a subset of $\{1, 2, \ldots, i-1\}$ and so the neighbors of b_i in H are contained in the set $\{1, 2, \ldots, i-1\}$. If $i < (j-1)\frac{n}{k} + 1$ then $d_{P_j}(b_i) = 0$ hence

$$e(L, P_j) = \sum_{i=1}^{n} d_{P_j}(b_i) = \sum_{i=1+(j-1)\frac{n}{k}}^{n} d_{P_j}(b_i). \tag{5.2}$$

Recall $\binom{x}{s} = 0$ if $0 \le x < s$ and so

$$\sum_{i=1}^{n} \binom{d_{P_j}(b_i)}{s} = \sum_{i=(j-1)\frac{n}{k}+1}^{n} \binom{d_{P_j}(b_i)}{s}.$$

Each subset of size s in P_j can be counted at most $t-1$ times in the sum $\sum_{i=1}^n \binom{d_{P_j}(b_i)}{s}$ therefore

$$(t-1)\binom{n/k}{s} \ge \sum_{i=1}^{n} \binom{d_{P_j}(b_i)}{s} = \sum_{i=1+(j-1)\frac{n}{k}}^{n} \binom{d_{P_j}(b_i)}{s}.$$

By Lemma 5.3.1,

$$\sum_{i=1+(j-1)\frac{n}{k}}^{n} d_{P_j}(b_i) \le \frac{n}{k}(t-1)^{1/s}\left(n\left(1 - \frac{j-1}{k}\right)\right)^{1-1/s} + (s-1)\left(n\left(1 - \frac{j-1}{k}\right)\right).$$

Using this inequality and (5.2), we obtain

$$\begin{aligned} e(H) &= \sum_{j=1}^{k} e(L, P_j) \\ &\leq \frac{(t-1)^{1/s} n^{2-1/s}}{k} \sum_{j=1}^{k} (1-(j-1)/k)^{1-1/s} + (s-1)n \sum_{j=1}^{k} (1-(j-1)/k) \\ &\leq \frac{(t-1)^{1/s} n^{2-1/s}}{k} (1 + \int_0^k (1-\frac{x}{k})^{1-1/s} dx) + (s-1)n(1 + \int_0^k (1-\frac{x}{k}) dx) \\ &= \frac{(t-1)^{1/s}}{2-1/s} n^{2-1/s} + \frac{(t-1)^{1/s} n^{2-1/s}}{k} + \frac{(s-1)nk}{2} + (s-1)n \\ &= \frac{(t-1)^{1/s}}{2-1/s} n^{2-1/s} + \left((t-1)^{1/s} + \frac{1}{2}(s-1) \right) n^{3/2-1/2s} + (s-1)n. \end{aligned}$$

Since $e(G) = e(H)$, this completes the proof.

5.4 A Lower Bound

Graphs constructed by Erdős, Rényi [31] and Brown [14] show that $\mathrm{ex}(q^2+q+1, C_4) \geq \frac{1}{2}q(q+1)^2$ where q is any odd prime power. Since $\mathrm{ex}(n, Z_4) \geq \mathrm{ex}(n, C_4)$, this implies

$$\mathrm{ex}(n, Z_4) \geq \frac{1}{2} n^{3/2} - o(n^{3/2}).$$

The construction we present improves this lower bound in the error term. Recall Füredi [43] proved $\mathrm{ex}(q^2+q+1, C_4) \leq \frac{1}{2}q(q+1)^2 = \frac{1}{2}q^3 + q^2 + \frac{1}{2}q$ for $q \geq 15$. Using the constructions of Erdős, Rényi, and Brown we have the exact result $\mathrm{ex}(q^2+q+1, C_4) = \frac{1}{2}q(q+1)^2$ for prime power q. For each prime p, we construct a Z_4-free graph with p^2+p+2 vertices, maximum degree $p+1$, and $\frac{1}{2}p^3+2p^2+\frac{3}{2}p+1$ edges. It follows that there exists a Z_4-free graph on p^2+p+1 vertices with at least $\frac{1}{2}p^3+2p^2+\frac{1}{2}p$ edges and so for prime $p \geq 15$,

$$\mathrm{ex}(p^2+p+1, C_4) + p^2 \leq \mathrm{ex}(p^2+p+1, Z_4).$$

Before giving the construction we point out a connection between Z_4-free graphs on n-vertices and collections subsets of $[n]$. Let G be a Z_4-free graph on $[n]$. The back neighborhoods $\{\Gamma^-(i)\}_{i=2}^n$ form a collection of subsets of $[n]$ that satisfy

1. $\Gamma^-(i) \subseteq [i-1]$ for $2 \leq i \leq n$.

2. For any $i \neq j$, $|\Gamma^-(i) \cap \Gamma^-(j)| \leq 1$.

Conversely any collection of sets $\{A_i\}_{i=1}^{n-1}$ that satisfy $A_i \subset [i]$ and $|A_i \cap A_j| \leq 1$ for $i \neq j$ can be used to define a Z_4-free graph G by setting $\Gamma^-(i+1) = A_i$. We will construct a family of sets $A_1, \ldots, A_{p^2+p+1}$ that satisfies these conditions by labeling the points of a projective plane using the numbers $1, 2, \ldots, p^2+p+1$ and by labeling the lines of the plane using the numbers $1, 2, \ldots, p^2+p+1$. Suppose l is a line that is assigned label i and we denote this by l_i. Our goal is to make the sum

$$\sum_{i=1}^{p^2+p+1} |l_i \cap [i]|$$

as large as possible for if G is defined by setting $\Gamma^-(i+1) = l_i \cap [i]$ for $1 \leq i \leq p^2+p+1$, then

$$e(G) = \sum_{i=2}^{p^2+p+2} |\Gamma^-(i)| = \sum_{i=1}^{p^2+p+1} |l_i \cap [i]|.$$

Now we proceed with the construction. Fix a prime p. The number p^2+p+1 and the numbers in the array

$$\begin{array}{lllll} p^2+p & p^2+p-1 & p^2+p-2 & \ldots & p^2+1 \\ p^2 & p^2-1 & p^2-2 & \ldots & p^2-p+1 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ p & p-1 & p-2 & \ldots & 1 \end{array}$$

are the points of the projective plane. To define the lines of the plane it is convenient to let $a = p^2+p+1$ and let $a_{i-1,j}$ be the (i,j)-entry in the array above. The lines of the plane are $l(r) = \{a, a_{r,1}, a_{r,2}, \ldots, a_{r,p}\}$ where $0 \leq r \leq p$, and $l(r,c) = \{a_{0,r}, a_{1,r+c}, a_{2,r+2c}, \ldots, a_{p,r+pc}\}$ where $1 \leq r, c \leq p$. The second subscript $r+jc$ is reduced modulo p so that its value is in $\{1, 2, \ldots, p\}$. These names are used only to define the lines and at this point we are ready to assign the labels $1, 2, \ldots, p^2+p+1$ to the lines. Each label will be used exactly once and there are p^2+p+1 lines so to label the lines we just drop the name after the label has been assigned. Give $l(0)$ label p^2+p+1 and for $1 \leq r \leq p$, give $l(r)$ label $p^2-p(r-1)$.

$$l(0) \to l_{p^2+p+1} \qquad l(r) \to l_{p^2-p(r-1)} \text{ for } 1 \le r \le p.$$

To determine the label assigned to $l(r,c)$, we look at which point $l(r,c)$ contains from the $(c+1)$-st row of the array. This point, or more precisely its label, will be the label assigned to $l(r,c)$ unless this element is a multiple of p. Specifically for $1 \le r, c \le p$,

$$l(r,c) \to \begin{cases} l_{a_{c,r+c^2}} & \text{if } a_{c,r+c^2} < p^2-(c-1)p, \\ l_{p^2+p-(c-1)} & \text{if } a_{c,r+c^2} = p^2-(c-1)p. \end{cases}$$

Before going further an example is needed. For $p = 3$ form the array

$$\begin{array}{ccc} 12 & 11 & 10 \\ 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{array}$$

In this case $a = 13$, $a_{0,1} = 12$, $a_{0,2} = 11$, $a_{0,3} = 10, a_{1,1} = 9, \ldots, a_{3,3} = 1$. Below we show the lines before the label assignments are given and the new labels. For lines of the form $l(r,c)$ we have underlined the point of the line used to determine its label.

$l(0) = \{13, 12, 11, 10\} \to l_{13}$, $l(1) = \{13, 9, 8, 7\} \to l_9$, $l(2) = \{13, 6, 5, 4\} \to l_6$

$l(3) = \{13, 3, 2, 1\} \to l_3$,

$l(1,1) = \{12, \underline{8}, 4, 3\} \to l_8$, $l(1,2) = \{11, 7, \underline{5}, 3\} \to l_5$, $l(1,3) = \{10, 9, 6, \underline{3}\} \to l_{10}$

$l(2,1) = \{12, \underline{7}, 6, 2\} \to l_7$, $l(2,2) = \{11, 9, \underline{4}, 2\} \to l_4$, $l(2,3) = \{10, 8, 5, \underline{2}\} \to l_2$

$l(3,1) = \{12, \underline{9}, 5, 1\} \to l_{12}$, $l(3,2) = \{11, 8, \underline{6}, 1\} \to l_{11}$, $l(3,3) = \{10, 7, 4, \underline{1}\} \to l_1$

To compute $\sum_{i=1}^{p^2+p+1} |l_i \cap [i]|$, we divide it into three smaller sums. It is easy to check

$$\sum_{i=1}^{p} |l_{ip} \cap [ip]| = p \cdot p \tag{5.3}$$

and

$$\sum_{i=p^2+1}^{p^2+p+1} |l_i \cap [i]| = (p+1)(p+1). \tag{5.4}$$

For fixed j with $0 \le j \le p-1$,

$$\sum_{i=jp+1}^{(j+1)p-1} |l_i \cap [i]| = (p-1)(j+1). \tag{5.5}$$

Putting (5.3), (5.4), and (5.5) together gives

$$\sum_{i=1}^{p^2+p+1} |l_i \cap [i]| = p^2 + (p+1)^2 + (p-1)\sum_{j=1}^{p} j = \frac{1}{2}p^3 + 2p^2 + \frac{3}{2}p + 1.$$

We summarize this construction as a result on labeling points and lines of a projective plane. Define a *labeling* of a projective plane $(\mathcal{P}, \mathcal{L})$ of order q to be a pair of bijections $L_{\mathcal{P}} : \mathcal{P} \to \{1, 2, \ldots, q^2+q+1\}$ and $L_{\mathcal{L}} : \mathcal{L} \to \{l_1, l_2, \ldots, l_{q^2+q+1}\}$.

Proposition 5.4.1. *Let $(\mathcal{P}, \mathcal{L})$ be a projective plane of order p where p is prime. There is a labeling of the points $L_{\mathcal{P}} : \mathcal{P} \to \{1, 2, \ldots, p^2+p+1\}$ and the lines $L_{\mathcal{L}} : \mathcal{L} \to \{l_1, l_2, \ldots, l_{p^2+p+1}\}$ such that*

$$\sum_{i=1}^{p^2+p+1} |l_i \cap [i]| = \frac{1}{2}p^3 + 2p^2 + \frac{3}{2}p + 1.$$

An easier way to obtain a labeling of a projective plane of order q is to label the points and lines randomly. The sum $X = \sum_{i=1}^{q^2+q+1} |l_i \cap [i]|$ is a random variable whose expectation and variance can be computed exactly as $\mathbb{E}X = \frac{1}{2}q^3 + q^2 + \frac{3}{2}q + 1$ and $\mathrm{Var}X = \frac{n}{12}\sqrt{n - 3/4} - \frac{n}{24} + \frac{1}{12}\sqrt{n - 3/4} - \frac{1}{24}$ where $n = q^2 + q + 1$. Furthermore X has a nice symmetry property that allows one to prove that there are outcomes where $X \geq \mathbb{E}X + \frac{1}{2}\sqrt{\mathrm{Var}X}$. This method produces a labeling with $X \geq \frac{1}{2}q^3 + q^2 + O(q^{1.5})$. When q is prime this matches our construction in the leading term but is not as good in second term. On the other hand, we do not know of any other method to label projective planes whose order is not a prime.

One may suspect that with some clever labeling we can find a sequence of projective planes of order $q_1 < q_2 < \ldots$ such that

$$\sum_{i=1}^{q_k^2+q_k+1} |l_i \cap [i]| > \left(\frac{1}{2} + \epsilon\right) q_k^3$$

for a fixed $\epsilon > 0$. The next result shows that this cannot be done.

Proposition 5.4.2. *Let $(\mathcal{P}, \mathcal{L})$ be a projective plane of order q. If $L_{\mathcal{P}}$, $L_{\mathcal{L}}$ is a labeling of $(\mathcal{P}, \mathcal{L})$ then*

$$\sum_{i=1}^{q^2+q+1} |l_i \cap [i]| = \frac{1}{2}q^3 + o(q^3).$$

To prove Proposition 5.4.2 we need the following lemma (see [49]).

Lemma 5.4.3. *Let G be a d-regular, n-vertex bipartite graph with parts X, Y and set $\lambda = \max_{i \neq 1,n} |\lambda_i|$ where $\lambda_1 \geq \cdots \geq \lambda_n$ are the eigenvalues of the adjacency matrix of G. For any $S \subset X$, $T \subset Y$,*

$$\left| e(S,T) - \frac{2d|S||T|}{n} \right| \leq \frac{\lambda}{2}(|S| + |T|).$$

Proof of Proposition 5.4.2. Let $(\mathcal{P}, \mathcal{L})$ be a projective plane of order q and let $L_{\mathcal{P}}$, $L_{\mathcal{L}}$ be a labeling of $(\mathcal{P}, \mathcal{L})$. Let $t = q^{1/4}$ and for $1 \leq i \leq t$, let $S_i = \{1, 2, \ldots, \frac{iq^2}{t}\}$ and

$$T_i = \{l_{1+(i-1)\frac{q^2}{t}}, l_{2+(i-1)\frac{q^2}{t}}, \ldots, l_{\frac{q^2}{t}+(i-1)\frac{q^2}{t}}\}.$$

Let A be the adjacency matrix of the incidence graph of $(\mathcal{P}, \mathcal{L})$. The eigenvalues of A are $q+1$, $-(q+1)$, each with multiplicity 1, and all other eigenvalues are $\pm\sqrt{q}$. This can be seen by considering the matrix A^2 which has the rather simple form $A^2 = \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix}$ where $B = I + qJ$ and J is the all 1's matrix. Using Lemma 5.4.3,

$$\begin{aligned}
\sum_{i=1}^{q^2+q+1} |l_i \cap [i]| &\leq (q+1)^2 + \sum_{i=1}^{q^2} |l_i \cap [i]| \leq (q+1)^2 + \sum_{i=1}^{t} e(S_i, T_i) \\
&\leq (q+1)^2 + \sum_{i=1}^{t} \left(\frac{2(q+1)q^4 i}{2(q^2+q+1)t^2} + \frac{\sqrt{q}}{2}\left(\frac{iq^2}{t} + \frac{q^2}{t}\right)\right) \\
&\leq (q+1)^2 + \frac{q^4(t+1)}{2(q-1)t} + \frac{q^{5/2}}{2} + \frac{q^{5/2}(t+1)}{4} \\
&= \frac{1}{2}q^3 + o(q^3).
\end{aligned}$$

A similar argument gives the lower bound $\sum_{i=1}^{q^2+q+1} |l_i \cap [i]| \geq \frac{1}{2}q^3 + o(q^3)$. □

Proof of Theorem 5.1.3. By Proposition 5.4.1, there exists a Z_4-free graph G_p with p^2+p+2 vertices and $\frac{1}{2}p^3+2p^2+\frac{3}{2}p+1$ edges for prime p. Furthermore this graph

has maximum degree $p+1$. Let G'_p be a subgraph of G_p with p^2+p+1 vertices and at least $\frac{1}{2}p^3+2p^2+\frac{1}{2}p$ edges. Then

$$\begin{aligned}\mathrm{ex}(p^2+p+1, Z_4) &\geq e(G'_p) \geq \frac{1}{2}p^3+2p^2+\frac{1}{2}p \\ &= \mathrm{ex}(p^2+p+1, C_4)+p^2.\end{aligned}$$

□

5.5 Sidon Sets and Z_4-free Graphs

In this section we show how B_2-sets can be used to construct Z_4-free graphs. Let $A \subset [n]$ be a B_2-set. Define the graph G_A by $V(G_A)=[n]$ and

$$E(G_A)=\{ij : i=j+a, a\in A\}.$$

Perhaps the most interesting feature of this construction is that, in general, G_A will contain many C_4's but will still be Z_4-free. For each vertex i and pair $\{a,b\}\subset A$ with $i+a+b\leq n$, the vertices $\{i, i+a, i+b, i+a+b\}$ form a C_4 in G.

Lemma 5.5.1. *If $A\subset[n]$ is a B_2-set, then G_A is a Z_4-free graph with $\sum_{i=1}^{n-1}|A\cap[i]|$ edges.*

Proof. Suppose $n_1<n_2<n_3<n_4$ are the vertices of a Z_4 in G. There exists $a,b,c,d\in A$ such that $n_3=a+n_1$, $n_4=b+n_1$, $n_3=c+n_2$, and $n_4=d+n_2$. This implies $a-c+d-b=0$ so that $\{a,d\}=\{b,c\}$. If $a=b$ then $n_3=n_4$ and if $a=c$ then $n_1=n_2$. In either case we have a contradiction.

The number of edges of G_A is $\sum_{i=1}^{n}|\Gamma^-(i)|$ and $|\Gamma^-(i)|=|A\cap[i-1]|$. □

Observe that for any such G_A,

$$\mathrm{ex}(n, Z_4)\geq e(G_A).$$

If we can accurately estimate $\sum_{i=1}^{n-1}|A\cap[i]|$ then upper bounds on $\mathrm{ex}(n,Z_4)$ can imply upper bounds on $|A|$. We can use the following result of Cilleruelo to show $e(G_A)=\frac{1}{2}n^{3/2}-o(n^{3/2})$ provided A is chosen appropriately.

Theorem 5.5.2 (Cilleruelo, [17]). *If $A \subset [n]$ is a B_2-set with $n^{1/2} - L$ elements then any interval of length cn contains $c|A| + e_I$ elements of A where*

$$|e_I| \leq 52n^{1/4}(1 + c^{1/2}n^{1/8})(1 + L_+^{1/2}n^{-1/8})$$

and $L_+ = \max\{0, L\}$.

Theorem 5.5.3. *For each B_2-set A with $|A| = n^{1/2}$ there exists an n-vertex Z_4-free graph G_A with*

$$e(G_A) \geq \frac{1}{2}n^{3/2} - O(n^{5/4}).$$

Furthermore G_A has at least $\frac{n^2}{18} - O(n^{15/8})$ 4-cycles.

Proof. Suppose $A \subset [n]$ is a B_2-set with $|A| = n^{1/2}$. Let $k = n^{1/4}$ and for $1 \leq j \leq k$, let

$$P_j = \{1 + \frac{(j-1)n}{k}, 2 + \frac{(j-1)n}{k}, \ldots, \frac{jn}{k}\}.$$

Using Theorem 5.5.2,

$$\sum_{i=1}^{n} |A \cap [i]| \geq \sum_{j=2}^{k} \sum_{i \in P_j} |A \cap [i]| \geq \sum_{j=2}^{k} \frac{n}{k} \left| A \cap \left[\frac{(j-1)n}{k} \right] \right| = \frac{n}{k} \sum_{j=2}^{k} \left(\frac{j-1}{k}|A| + e_{I_j} \right)$$

where e_{I_j} satisfies the inequality

$$e_{I_j} \geq -52n^{1/4} \left(1 + \sqrt{\frac{j-1}{k}} n^{1/8} \right).$$

Now $\frac{n}{k} \sum_{j=2}^{k} \frac{j-1}{k}|A| = \frac{1}{2}n^{3/2} - \frac{n^{3/2}}{2k}$ and so it remains to find a lower bound on the sum $-\frac{52n^{5/4}}{k} \sum_{j=1}^{k-1}(1 + \sqrt{j/k}n^{1/8})$. Estimating the sum with an integral gives

$$-\frac{52n^{5/4}}{k} \sum_{j=1}^{k-1} \left(1 + \sqrt{j/k} n^{1/8} \right) \geq -\frac{52n^{5/4}}{k} \left(k + \frac{2n^{1/8}k}{3} \right) = -52n^{5/4} - \frac{104n^{11/8}}{3}.$$

Thus

$$\sum_{i=1}^{n} |A \cap [i]| \geq \frac{1}{2}n^{3/2} - \frac{n^{3/2}}{2k} - 52n^{5/4} - \frac{104n^{11/8}}{3} \geq \frac{1}{2}n^{3/2} - O(n^{5/4}).$$

To prove the statement concerning 4-cycles, observe Theorem 5.5.2 implies

$$|A \cap [1, n/3]| \geq \frac{1}{3}n^{1/2} - 104n^{3/8}.$$

For any vertex $i \in [1, n/3]$ and pair $\{a, b\} \subset A \cap [1, n/3]$, the vertices $\{i, i+a, i+b, i+a+b\}$ form a 4-cycle. If $\alpha = \frac{1}{3}n^{1/3} - 104n^{3/8}$, then there are $\frac{n}{3}\binom{\alpha}{2} = \frac{n^2}{18} - O(n^{15/8})$ such 4-cycles.

□

For q a prime power, there exists B_2-sets $A \subset [q^2-1]$ with $|A| = q$ (see [13]). Applying Theorem 5.5.3 to such a B_2-set gives a Z_4-free graph with $\frac{1}{2}n^{3/2} + o(n^{3/2})$ edges where $n = q^2 - 1$.

Chapter 5 is a reprint of "An ordered Turán problem for bipartite graphs," Timmons, Craig. *Electr. J. Combin.* 19 (4) #P43, 2012.

Chapter 6

A Counterexample to Sparse Removal

The graph removal lemma is one of the most important results in extremal graph theory. It states that if H is a graph with h vertices, then any n-vertex graph G with $o(n^h)$ copies of H can be made H-free by removing $o(n^2)$ edges. The special case when H is a triangle was proved by Ruzsa and Szemerédi [73]. This special case, often called the triangle removal lemma, is known to imply a weak form of Roth's Theorem [69] on three term arithmetic progressions in subsets of the integers. There is a large amount of research aimed at obtaining graph removal type lemmas in different settings such as hypergraphs, directed graphs, and finite groups. One area where we do not yet have an analogue of the graph removal lemma is sparse graphs.

Solymosi [75] conjectured that if H is a graph and $\mathrm{ex}(n, H) = O(n^\alpha)$ where $\alpha > 1$, then any n-vertex graph with the property that each edge lies in exactly one copy of H has $o(n^\alpha)$ edges. This can be viewed as a possible extension of the removal lemma to sparse graphs. Using the removal lemma, it is easy to show that the conjecture is true when H is a non-bipartite graph and so the conjecture concerns bipartite graphs. In this chapter we use Sidon sets to exhibit infinitely many bipartite graphs H for which the conjecture is false.

6.1 Introduction

One of the most important consequences of the Regularity Lemma [77] is the graph removal lemma, also known as the removal lemma.

Theorem 6.1.1 (Graph Removal Lemma)**.** *Let H be a graph with h vertices. Given $\epsilon > 0$, there exists a $\delta = \delta(H, \epsilon) > 0$ and an integer $n_0 = n_0(H, \epsilon)$ such that the following holds: if $n \geq n_0$ and G is an n-vertex graph with at most δn^h copies of H, then G can be made H-free by removing at most ϵn^2 edges.*

This is the modern formulation of the graph removal lemma. Ruzsa and Szemerédi [73] proved Theorem 6.1.1 in that case that H is a triangle. Erdős, Frankl, and Rödl [29] proved a generalization of the triangle removal lemma and, while they did not state the graph removal lemma as it is stated above, all of the ideas needed to prove Theorem 6.1.1 are present in [29]. The removal lemma is a central tool in extremal combinatorics with many applications [22]. Because of its importance, removal type lemmas have been extended to other settings such as directed graphs, random graphs, and hypergraphs. One area in which we still do not have a removal lemma is in the realm of sparse graphs. A sequence $\{G_n\}_{n\in\mathbb{N}}$ of graphs is *sparse* if G_n has n vertices and $e(G_n)/n^2 \to 0$ as $n \to \infty$. There are many natural families of sparse graphs. For example, any sequence of $K_{s,t}$-free graphs is sparse graph by the Kővári-Sós-Turán Theorem (see Chapter 2). Theorem 6.1.1 is trivial for sparse graphs. To see this, let H be any graph and let $\epsilon > 0$. Suppose $\{G_n\}$ is a sequence of sparse graphs. Choose n_0 so that $e(G_n) < \epsilon n^2$ for all $n \geq n_0$. Then for every $n \geq n_0$, we can make the graph G_n H-free by removing all of the edges of G_n. Since $e(G_n) < \epsilon n^2$, this removes at most ϵn^2 edges and we did not require any assumption on how many copies of H were contained in G_n.

Solymosi [75] conjectured an analogue of the removal lemma that would apply to sparse graphs which involved the Turán number of H. The *exponent* of a graph H, when it exists, is a real number α such that $\mathrm{ex}(n, H) = \Theta(n^\alpha)$ as $n \to \infty$. Erdős and Simonovits [34] conjectured that every graph has an exponent, but this conjecture remains open. In the case that H is not bipartite the exponent is 2. In the bipartite case the exponent is generally not known. Solymosi's conjecture is

as follows.

Conjecture 6.1.2. *If H is a graph with exponent α, then any n-vertex graph in which every edge is in exactly one copy of H has $o(n^\alpha)$ edges.*

In the case $H = C_4$ i.e. H is a quadrilateral, Solymósi [75] conjectured that if an n-vertex graph is a union of $\Theta(n^{3/2})$ edge-disjoint quadrilaterals, then the graph contains $\Omega(n^2)$ quadrilaterals. By the removal lemma, Conjecture 6.1.2 is true for non-bipartite graphs H, so the conjecture is interesting for bipartite graphs.

Our main result shows that there are infinitely many bipartite graphs H for which Conjecture 6.1.2 is false. Let H_k be the graph with vertex set $V(H_k) = \{1, 2, \ldots, 2k\}$ and edge set $E(H_k) = \{1i, 2i, 3j, 4j : 3 \leq i \leq k+2 < j \leq 2k\}$ – see Figure 6.1.

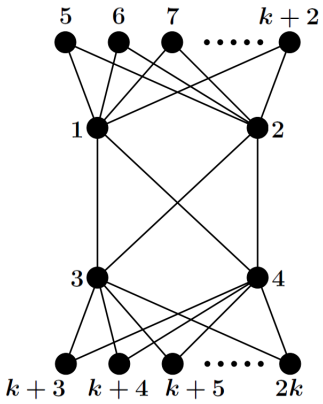

Figure 6.1: The Bipartite Graph H_k

We shall prove that H_k has exponent $\frac{3}{2}$ for all $k \geq 3$ (Section 6.2.4), and then the following theorem shows that the graphs H_k for $k \geq 5$ give counterexamples to Conjecture 6.1.2.

Theorem 6.1.3. *Let $k \geq 5$ and let P be the set of primes $p \equiv 1$ mod 4. There exists a sequence of graphs $(G_p)_{p \in P}$ such that G_p has $n = 2kp^2$ vertices, $\Theta(n^{3/2})$ edges, and every edge of every G_p is contained in exactly one copy of H_k.*

The proof of Theorem 6.1.3 is based on a construction described in Section 6.2 involving simple field arithmetic and the quadratic character of $\mathbb{F}_p$. The proof

is given in Sections 6.2 – 6.2.4. The case that H is a quadrilateral in Conjecture 6.1.2 remains open.

6.2 Outline of the Proof of Theorem 6.1.3

The proof of Theorem 6.1.3 is achieved in three steps. Before we describe these steps, we introduce some notation. Throughout this section, p is a prime, k is a positive integer, and $\mathbb{F}_p$ denotes the finite field of order p. Let χ denote the quadratic character of $\mathbb{F}_p$, namely $\chi(x) = 1$ if x is a non-zero quadratic residue, $\chi(x) = -1$ if x is not a quadratic residue, and $\chi(0) = 0$. The graph H_k has vertex set $[2k] := \{1, 2, \ldots, 2k\}$ and edge set

$$E(H_k) := \{1i, 2i, 3j, 4j : 3 \leq i \leq k+2 < j \leq 2k\}.$$

Let $V(G)$ and $E(G)$ denote the vertex set and edge set of a graph G. If G_1 and G_2 are edge-disjoint graphs, then we write $G = G_1 \oplus G_2$ if $V(G) = V(G_1) \cup V(G_2)$ and $E(G) = E(G_1) \cup E(G_2)$.

6.2.1 Step I. Construction of Graphs $G_{\Gamma,\Lambda,S}(H)$.

Let Γ be a finite abelian group and $S \subseteq \Gamma$. Let H be an arbitrary graph with vertex set $[k]$, edge set E, and let $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_k\} \subset \mathbb{Z}$. For $i \in [k]$ let $X_i = \Gamma \times \{i\}$. For $ij \in E$ with $i < j$, let $G_{\Gamma,\Lambda,S}(ij)$ be the bipartite graph with parts X_i and X_j, where $x \in X_i$ is adjacent to $y \in X_j$ whenever there exists $a \in S$ such that $y = x + (\lambda_j - \lambda_i)a$. Finally, define the following k-partite graph with parts $X_1, X_2, \ldots, X_k$:

$$G_{\Gamma,\Lambda,S}(H) = \bigcup_{ij \in E} G_{\Gamma,\Lambda,S}(ij).$$

A key observation is that $G_{\Gamma,\Lambda,S}(H)$ is built up from edge-disjoint copies of H. The following is proved in Section 6.3:

Proposition 6.2.1. *Let $k \in \mathbb{N}$ and let H be a k-vertex graph. For any finite abelian group Γ, $S \subseteq \Gamma$, and $\Lambda \subset \mathbb{Z}$, the graph $G_{\Gamma,\Lambda,S}(H)$ has $k|\Gamma|$ vertices, $|E||\Gamma||S|$*

edges, and there exist induced subgraphs $H(1), H(2), \ldots, H(|\Gamma||S|)$ *of* $G_{\Gamma,\Lambda,S}(H)$*, each isomorphic to* H*, such that*

$$G_{\Gamma,\Lambda,S}(H) = \bigoplus_{i=1}^{|\Gamma||S|} H(i).$$

The proof of Theorem 6.1.3 involves a suitable choice of the group Γ, the set S, and the set $\Lambda \subset \mathbb{Z}$. The graphs $G_{\Gamma,\Lambda,S}(H)$ may be useful as constructions for other extremal graph theory problems.

6.2.2 Step II. The Choice of Γ, S, and Λ.

Let $\Gamma = \mathbb{F}_p^2$ and $S = \{(a, a^2) : a \in \mathbb{F}_p^*\}$ where $p \equiv 1 \bmod 4$ is prime. We aim to show that there exists a choice of $\Lambda = \{\lambda_1, \ldots, \lambda_k\}$ such that every edge of $G_p = G_{\Gamma,\Lambda,S}(H_k)$ is in exactly one copy of H_k, which is the heart of Theorem 6.1.3. It is convenient to let $\lambda_{ij} = \lambda_j - \lambda_i$. After some arithmetic preparations in Section 6.4.1, we prove the following in Section 6.4.2.

Proposition 6.2.2. *Let* $p \equiv 1$ *mod 4 be a prime,* $\Gamma = \mathbb{F}_p^2$*, and* $S = \{(a, a^2) : a \in \mathbb{F}_p^*\}$*. Suppose* $\Lambda = \{\lambda_i : i \in [2k]\}$ *and*

1. $\lambda_1 = 0$*,* $\lambda_2 = 1$*.*
2. *For* $i \in \{3, 4\}$ *and* $j \in \{1, 2\}$*,* $\chi(\lambda_{1i}\lambda_{i2}) = \chi(\lambda_{3j}\lambda_{j4}) = -1$*.*
3. *For* $5 \leq i \leq k + 2 < j \leq 2k$*, we have* $\chi(\lambda_{1i}\lambda_{i2}) = \chi(\lambda_{3j}\lambda_{j4}) = 1$*.*

Then every edge of $G_p = G_{\Gamma,\Lambda,S}(H_k)$ *is contained in exactly one copy of* H_k*.*

Note that $\chi(\lambda_{ij}) = \chi(-\lambda_{ji}) = \chi(\lambda_{ji})$ since $p \equiv 1 \bmod 4$. A set Λ satisfying the conditions in Proposition 6.2.2 is called *suitable*. We prove the following in Section 6.5:

Proposition 6.2.3. *If* $k \geq 5$ *and* $p \geq 4k+3$*, then there exist a suitable set* $\Lambda \subset \mathbb{F}_p$*.*

If $\Gamma = \mathbb{F}_p^2$ and $S = \{(a, a^2) : a \in \mathbb{F}_p^*\}$, then it is not hard to show $G_{\Gamma,S,\Lambda}(C_4)$ contains $\Theta(n^2)$ quadrilaterals for any set Λ, so $G_{\Gamma,S,\Lambda}(C_4)$ cannot be used as a counterexample to Conjecture 6.1.2 when $H = C_4$.

6.2.3 Step III. The Turán Number for H_k.

It remains to show $\mathrm{ex}(n, H_k) = \Theta(n^{3/2})$, which is the final step in the proof of Theorem 6.1.3. A counting argument will be used to determine the order of magnitude of $\mathrm{ex}(n, H_k)$.

Proposition 6.2.4. *For any integer $k \geq 3$, $ex(n, H_k) = \Theta(n^{3/2})$.*

This proposition is proved in Section 6.6. Since $K_{2,k} \subset H_k$, the results of Füredi [42] show

$$\mathrm{ex}(n, H_k) \geq \mathrm{ex}(n, K_{2,k}) \geq \frac{1}{2}\sqrt{k-1}n^{3/2} + o(n^{3/2})$$

and this establishes the lower bound in Proposition 6.2.4.

6.3 Proof of Proposition 6.2.1

It follows from the definition of $G_{\Gamma,\Lambda,S}(H)$ that $G_{\Gamma,\Lambda,S}(H)$ has $k|\Gamma|$ vertices. We now prove that $G_{\Gamma,\Lambda,S}(H)$ is a union of edge-disjoint copies of H. We observe that H appears naturally as a subgraph of $G_{\Gamma,\Lambda,S}(H)$ in the following manner: for $v \in \Gamma$ and $a \in S$, the set $\{v_i = v + \lambda_i a : i \in [k]\}$ induces a subgraph $H(v, a)$ isomorphic to H, since $v_i v_j \in G_{\Gamma,\Lambda,S}(H)$ if and only if $ij \in E(H)$. Also, if $(x, i), (y, j) \in H(v, a)$ and $(x, i), (y, j) \in H(w, b)$, then

$$x = v + \lambda_i a = w + \lambda_i b \quad y = v + \lambda_j a = w + \lambda_j b$$

and therefore $\lambda_i(a - b) = \lambda_j(a - b)$ which means $a = b$ and so $v = w$. Therefore, no two subgraphs $H(v, a)$ share any pair of vertices. We conclude

$$G_{\Gamma,\Lambda,S}(H) = \bigoplus_{(v,a)\in\Gamma\times S} H(v, a).$$

In particular, $G_{\Gamma,\Lambda,S}(H)$ has $|E||\Gamma||S|$ edges. This proves Proposition 6.2.1.

6.4 Proof of Proposition 6.2.2

This section is split into two parts. First we describe the interaction between quadrilaterals in $G_p(H) := G_{\Gamma,\Lambda,S}(H)$ for general H, $\Gamma = \mathbb{F}_p^2$, $\Lambda \subset \mathbb{F}_p$ and $S =$

$\{(a, a^2) : a \in \mathbb{F}_p^*\}$. We then show that if the conditions of Proposition 6.2.2 are satisfied, then every edge of $G_p = G_p(H_k)$ is in exactly one copy of H_k.

6.4.1 Quadrilaterals in $G_{\Gamma,\Lambda,S}(H)$.

Throughout this subsection, H is a graph, $G_p(H) = G_{\Gamma,\Lambda,S}(H)$ where $\Gamma = \mathbb{F}_p^2$, $\Lambda \subset \mathbb{F}_p$ and $S = \{(a, a^2) : a \in \mathbb{F}_p^*\}$. The following simple but key arithmetic lemma is due to Ruzsa [71].

Lemma 6.4.1. *Let $\alpha, \beta, \gamma, \delta, a, b, c, d \in \mathbb{F}_p^*$ where $\alpha + \beta = \gamma + \delta \neq 0$. If*

$$\alpha(a, a^2) + \beta(b, b^2) = \gamma(c, c^2) + \delta(d, d^2),$$

then $\alpha\beta(a - b)^2 = \gamma\delta(c - d)^2$.

Proof. Multiply $\alpha a^2 + \beta b^2 = \gamma c^2 + \delta d^2$ by $\alpha + \beta = \gamma + \delta$ to get

$$\alpha^2 a^2 + \alpha\beta(a^2 + b^2) + \beta^2 b^2 = \gamma^2 c^2 + \gamma\delta(c^2 + d^2) + \delta^2 d^2.$$

Subtracting the square of $\alpha a + \beta b = \gamma c + \delta d$ gives $\alpha\beta(a - b)^2 = \gamma\delta(c - d)^2$. □

This lemma has a number of consequences relative to the distribution of quadrilaterals in $G(ij) := G_{\Gamma,\Lambda,S}(ij)$ for $ij \in H$. For $hi, ij \in E(H)$, let $G(hij) = G(hi) \cup G(ij)$.

Lemma 6.4.2. *For any edge $ij \in E(H)$, the graph $G(ij)$ is C_4-free.*

Proof. Consider a quadrilateral $(x, i)(y, j)(z, i)(w, j) \subset G(ij)$. For some $a, b, c, d \in \mathbb{F}_p^*$,

$$y = x + \lambda_{ij}(a, a^2) = z + \lambda_{ij}(b, b^2) \quad w = x + \lambda_{ij}(c, c^2) = z + \lambda_{ij}(d, d^2).$$

Canceling out x, y, z and w and dividing by λ_{ij}, we obtain in each component $a + b = c + d$ and $a^2 + b^2 = c^2 + d^2$. By Lemma 6.4.1, $(a - b)^2 = (c - d)^2$ and so $a - b = c - d$ or $a - b = d - c$. In either case, together with $a + b = c + d$ we find $\{a, b\} = \{c, d\}$. But then $(x, i) = (z, i)$ or $(y, j) = (w, j)$, a contradiction. Therefore $G(ij)$ has no C_4. □

Lemma 6.4.3. *For any edges $hi, ij \in E(H)$, the graph $G(hij)$ is $K_{2,3}$-free.*

Proof. Suppose $G(hij)$ contains a $K_{2,3}$. By Lemma 6.4.2, $G(hi)$ and $G(ij)$ are C_4-free, and so the three vertices of degree two in the $K_{2,3}$ must be in X_i, say $(z_1, i), (z_2, i), (z_3, i)$, and the other two vertices are $(x, h) \in X_h$ and $(y, j) \in X_j$. By definition, for some $a_1, a_2, a_3 \in \mathbb{F}_p^*$ and $b_1, b_2, b_3 \in \mathbb{F}_p^*$,

$$y = x + \lambda_{hi}(a_r, a_r^2) + \lambda_{ij}(b_r, b_r^2)$$

for $r \in \{1, 2, 3\}$. Note that the a_r are distinct and the b_r are distinct, for if $a_r = a_s$ or $b_r = b_s$ for some $r \neq s$, then $z_r = z_s$ and so $(z_r, i) = (z_s, i)$, a contradiction. On the other hand, by Lemma 6.4.1, $(a_r - b_r)^2 = (a_s - b_s)^2$ for all $r, s \in \{1, 2, 3\}$. Taking square roots, some pair of square roots has the same sign, namely for some $r \neq s$, we have $a_r - b_r = a_s - b_s$. For all r, s, we also have

$$\lambda_{hi}a_r + \lambda_{ij}b_r = \lambda_{hi}a_s + \lambda_{ij}b_s.$$

Subtracting $\lambda_{hi}(a_r - b_r) = \lambda_{hi}(a_s - b_s)$ from this equation, we obtain $\lambda_{hj}b_r = \lambda_{hj}b_s$. Therefore, $b_r = b_s$ and $(z_r, i) = (z_s, i)$ which is a contradiction. □

Lemma 6.4.4. *If $C = (x, i)(y, j)(z, k)(w, l)$ is a quadrilateral in $G_p(H)$ and $(x, i)(y, j)(z, k) \subset H(v, a)$, then $C \subset H(v, a)$.*

Proof. By definition of $G_p(H)$, there exist $a, c, d \in \mathbb{F}_p^*$ such that

$$y = x + \lambda_{ij}(a, a^2) \quad z = y + \lambda_{jk}(a, a^2) \quad w = z + \lambda_{kl}(c, c^2) \quad x = w + \lambda_{li}(d, d^2).$$

This implies

$$\lambda_{ij}(a, a^2) + \lambda_{jk}(a, a^2) = \lambda_{lk}(c, c^2) + \lambda_{il}(d, d^2).$$

By Lemma 6.4.1 with $\alpha = \lambda_{ij}$, $\beta = \lambda_{jk}$, $\gamma = \lambda_{lk}$, and $\delta = \lambda_{il}$, noting $\alpha + \beta = \gamma + \delta$, we have

$$\gamma\delta(c - d)^2 = 0$$

and therefore $c = d$. Now

$$\lambda_{ik}a = \lambda_{ij}a + \lambda_{jk}a = \lambda_{lk}c + \lambda_{il}d = \lambda_{ik}c$$

so we conclude $a = c = d$. In particular, since $z = v + \lambda_{0k}(a, a^2)$,

$$w = z + \lambda_{kl}(a, a^2) = v + \lambda_{0k}(a, a^2) + \lambda_{kl}(a, a^2) = v + \lambda_{0l}(a, a^2)$$

and we conclude $C \subset H(v, a)$. □

Lemma 6.4.5. *If $C = (x, i)(y, j)(z, k)(w, l)$ is a quadrilateral in $G_p(H)$ and $\chi(\lambda_{ij}\lambda_{jk}\lambda_{kl}\lambda_{li}) = -1$ then $C \subset H(v, a)$ for some $(v, a) \in \mathbb{F}_p^2 \times S$.*

Proof. By definition of $G_p(H)$, there exist $a, b, c, d \in \mathbb{F}_p^*$ such that

$$\lambda_{ij}(a, a^2) + \lambda_{jk}(c, c^2) = \lambda_{il}(d, d^2) + \lambda_{lk}(b, b^2).$$

By Lemma 1,

$$\lambda_{ij}\lambda_{jk}(a - c)^2 = \lambda_{il}\lambda_{lk}(d - b)^2.$$

Since $\chi(\lambda_{ij}\lambda_{jk}\lambda_{il}\lambda_{lk}) = \chi(\lambda_{ij}\lambda_{jk}\lambda_{kl}\lambda_{li}) = -1$, we conclude $a = c$ and $b = d$. The equation $\lambda_{ij}a + \lambda_{jk}c = \lambda_{il}d + \lambda_{lk}b$ reduces to $\lambda_{ik}a = \lambda_{ik}d$ and we conclude $a = b = c = d$. Letting $v = x - \lambda_{0i}(a, a^2)$, we have $C \subset H(v, a)$. □

6.4.2 Proof of Proposition 6.2.2

Suppose $F \subset G_p = G_{\Gamma,\Lambda,S}(H_k)$ is isomorphic to H_k and let $\phi : V(H_k) \to V(F)$ be an isomorphism. We aim to show that $F = H(v, a)$ for some $v \in \mathbb{F}_p$ and $a \in S$ by finding, via Lemma 6.4.5, a quadrilateral $C^* \subset H_k$ with $\phi(C^*) \subset H(v, a)$. This is the point where we make heavy use of the last two conditions in Proposition 6.2.2. Let $\mathfrak{c} : V(H_k) \to \Lambda$ be the proper vertex-coloring of H_k given by $\mathfrak{c}(x) = \lambda_i$ if and only if $\phi(x) \in X_i$. Let $[m, n] := \{m, m + 1, m + 2, \ldots, n\}$.

Claim 1. *Each quadrilateral in H_k is colored with at least three colors, and each $K_{2,3}$ in H_k is colored with at least four colors. Furthermore,*

$$\{\mathfrak{c}(1), \mathfrak{c}(2), \mathfrak{c}(3), \mathfrak{c}(4)\} \subset \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}.$$

Proof of Claim 1. The first statement is an immediate consequence of Lemmas 6.4.2 and 6.4.3. Consider the vertex $1 \in V(H_k)$ (see Figure 1).

If $\mathfrak{c}(1) \in \{\lambda_5, \lambda_6, \ldots, \lambda_{k+2}\}$ then each neighbor of 1 is assigned color λ_1 or color λ_2. This follows from the fact that if $xy \in E(H_k)$ and $\mathfrak{c}(x) = \lambda_i$ and $\mathfrak{c}(y) = \lambda_j$,

then $ij \in E(H_k)$. The neighbors of 1 are also neighbors of 2 so that $\mathfrak{c}$ assigns one color to at least three common neighbors of 1 and 2. This is impossible by the first part of the claim. A similar argument shows that $\mathfrak{c}(1) \notin \{\lambda_{k+3}, \lambda_{k+4}, \ldots, \lambda_{2k}\}$ thus $\mathfrak{c}(1) \notin \{\lambda_5, \lambda_6, \ldots, \lambda_{2k}\}$. By symmetry, we must have $\mathfrak{c}(i) \notin \{\lambda_5, \lambda_6, \ldots, \lambda_{2k}\}$ for $1 \leq i \leq 4$ and so

$$\{\mathfrak{c}(1), \mathfrak{c}(2), \mathfrak{c}(3), \mathfrak{c}(4)\} \subset \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}.$$

This proves Claim 1. □

Claim 2. *There is a quadrilateral $C^* = ghij$ in H_k such that*

$$\chi\left((\mathfrak{c}(g) - \mathfrak{c}(h))(\mathfrak{c}(h) - \mathfrak{c}(i))(\mathfrak{c}(i) - \mathfrak{c}(j))(\mathfrak{c}(j) - \mathfrak{c}(g))\right) = -1.$$

Proof of Claim 2. By Claim 1 there is a path $ghi \subset 1324$ such that $\mathfrak{c}(g), \mathfrak{c}(h), \mathfrak{c}(i)$ are distinct. Without loss of generality, assume that $g = 1$, $h = 3$, and $i = 2$. If we can find a $j \in [5, k+2]$ such that $\mathfrak{c}(j) \notin \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ then by the conditions of Proposition 6.2.2,

$$\chi(\lambda_{13}\lambda_{32}\lambda_{2j}\lambda_{j1}) = \chi(\lambda_{13}\lambda_{32}) \cdot \chi(\lambda_{2j}\lambda_{j1}) = -1$$

and so $132j$ is the required cycle. To find j, note that no two vertices in $[5, k+2]$ have color $\mathfrak{c}(3)$ otherwise those two vertices together with $1, 3$ and 2 form a 3-colored $K_{2,3}$, contradicting Claim 1. Similarly, no two vertices in $[5, k+2]$ have color $\mathfrak{c}(4)$. Since $k \geq 5$, there are at least three vertices in $[5, k+2]$ and so one of them, say j, has a color not in $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$, as required. □

We now complete the proof of Proposition 6.2.2. By Lemma 6.4.5, there exists $v \in \mathbb{F}_p^2$ and $a \in S$ such that $\phi(C^*) \subset H(v, a)$. Given any edge e of $F = \phi(H_k)$, there exist quadrilaterals $C^* = C^1, C^2, \ldots, C^r \subset H_k$ such that $e \in E(\phi(C^r))$, and $E(\phi(C^i)) \cap E(\phi(C^{i+1})) \neq \emptyset$ for $i < r$. By Lemma 6.4.4, we inductively have $\phi(C^i) \subset H(v, a)$ for all $i \leq r$. In particular, $e \in E(H(v, a))$ and we conclude $F \subset H(v, a)$. Since $H(v, a)$ is an induced subgraph of $G_p(H_k)$ isomorphic to F, we conclude $F = H(v, a)$. The completes the proof of Proposition 6.2.2.

6.5 Proof of Proposition 6.2.3

To build a suitable set $\Lambda \subset \mathbb{F}_p$, we use the following identity (see Theorem 5.48, [56]):

Proposition 6.5.1. *Let* $f(x) = a_2x^2 + a_1x + a_0 \in \mathbb{F}_p[x]$ *where* $a_2 \neq 0$. *If* $a_1^2 - 4a_0a_2 \neq 0$ *then*

$$\sum_{c\in\mathbb{F}_p} \chi(f(c)) = -\chi(a_2).$$

Let $k \geq 5$, $p \geq 4k+3$ be prime with $p \equiv 1 (\text{mod } 4)$, and $\lambda_1 = 0$, $\lambda_2 = 1$. By Proposition 6.5.1,

$$\sum_{c\in\mathbb{F}_p} \chi(c^2 - c) = -1 \tag{6.1}$$

so there are least $\frac{p-3}{2} \geq 2k$ elements of $\mathbb{F}_p$ for which $\chi(c^2-c) = -1$. Let λ_3 be any one of them and observe that since $\chi(\lambda_3^2 - \lambda_3) = -1$, we have $\lambda_3 \neq 0$ and $\lambda_3 \neq 1$. Using the fact that $\chi(-1) = 1$,

$$-1 = \chi(\lambda_3^2 - \lambda_3) = \chi((0-\lambda_3)(\lambda_3 - 1)) = \chi((\lambda_1 - \lambda_3)(\lambda_3 - \lambda_2)).$$

Next we choose λ_4. Let $g(x) = (1 - \chi(x^2 - x))(1 - \chi(\lambda_3 x))$ and $X = \{c \in \mathbb{F}_p : g(c) = 4\}$. If $c \in X$ then $\chi(c^2 - c) = \chi(\lambda_3 c) = -1$. Suppose $c \in \mathbb{F}_p$ and $0 < g(c) \leq 2$. Then either $\chi(c^2 - c) = 0$ or $\chi(\lambda_3 c) = 0$ which means $c = 0$ or $c = 1$ thus

$$\Big|\sum_{c\in\mathbb{F}_p} g(c) - 4|X|\Big| \leq 2 \cdot 2. \tag{6.2}$$

Expanding $g(c)$ and using Proposition 6.5.1,

$$\sum_{c\in\mathbb{F}_p} g(c) = p + \sum_{c\in\mathbb{F}_p} \big(-\chi(c^2-c) - \chi(\lambda_3)\chi(c) + \chi(c^2)\chi(\lambda_3)\chi(c-1)\big) = p+1. \tag{6.3}$$

Here we have used the well known fact that $\sum_{c\in\mathbb{F}_p} \chi(c) = 0$. By (6.2) and (6.3),

$$|X| \geq \frac{p+1}{4} - 1$$

and since $\frac{p-3}{4} \geq 1$, the set X is not empty. Choose λ_4 so that $-1 = \chi(\lambda_4(\lambda_4 - 1)) = \chi(\lambda_3\lambda_4)$. Then λ_4 is not equal to any of λ_1, λ_2, or λ_3. The relation $-1 =$

$\chi(\lambda_4(\lambda_4-1))$ implies $-1=\chi((\lambda_1-\lambda_4)(\lambda_4-\lambda_2))$ and the relation $-1=\chi(\lambda_3\lambda_4)$ implies $-1=\chi((\lambda_3-\lambda_1)(\lambda_1-\lambda_4))$. Furthermore

$$\begin{aligned}\chi((\lambda_3-\lambda_2)(\lambda_2-\lambda_4)) &= \chi((\lambda_3-1)(\lambda_4-1))\\ &= \chi(\lambda_3(\lambda_3-1))\chi(\lambda_3\lambda_4)\chi(\lambda_4(\lambda_4-1))\\ &= (-1)(-1)(-1)=-1\end{aligned}$$

so that all of the requirements of Condition 2 are satisfied.

Choosing the remaining λ_i's will be straightforward. By (6.1), there are at least $\frac{p-3}{2}$ elements $c\in\mathbb{F}_p$ for which $\chi(c^2-c)=1$. Since $\frac{p-3}{2}\geq 2k$ we can choose $\lambda_5,\lambda_6,\ldots,\lambda_{k+2}$ so that none of these are equal to λ_1, λ_2, λ_3 or λ_4, and $\chi((\lambda_1-\lambda_i)(\lambda_i-\lambda_2))=1$ for all $4<i\leq k+2$.

Let $f(x)=(\lambda_3-x)(x-\lambda_4)=-x^2+(\lambda_3+\lambda_4)x-\lambda_3\lambda_4$. If $a_2=-1$, $a_1=\lambda_3+\lambda_4$, and $a_0=-\lambda_3\lambda_4$, then $a_1^2-4a_0a_2=(\lambda_3-\lambda_4)^2\neq 0$. By Proposition 6.5.1,

$$\sum_{c\in\mathbb{F}_p}\chi(f(c))=-\chi(-1)=-1$$

and there are at least $\frac{p-3}{2}$ elements c for which $\chi(f(c))=1$. Since $\frac{p-3}{2}\geq 2k$ we can choose $\lambda_{k+3},\lambda_{k+4},\ldots,\lambda_{2k}$ so that all of $\lambda_1,\lambda_2,\ldots,\lambda_{2k}$ are distinct and for any $4<i\leq k+2$,

$$\chi((\lambda_3-\lambda_{i+k-2})(\lambda_{i+k-2}-\lambda_4))=1.$$

This completes the proof of Proposition 6.5.

6.6 Proof of Proposition 6.2.4

The claim $\mathrm{ex}(n,H_k)=\Theta(n^{3/2})$ follows by proving $\mathrm{ex}(n,H_k)\leq kn^{3/2}$ for all $k\geq 3$ and large enough n. We follow the method of dependent random choice [39], but we do not optimize the constants in the upper bound we obtain for $\mathrm{ex}(n,H_k)$.

Let G be an H_k-free n-vertex graph with average degree $d:=2|E(G)|/n$. It is sufficient to show $d\leq 2k\sqrt{n}$. Choose uniformly at random a pair of vertices $\{x_1,x_2\}$ in G and let $S=N(x_1,x_2)$, the common neighborhood of x_1 and x_2. Given a pair of vertices $\{y_1,y_2\}$, let $d(y_1,y_2)=|N(y_1,y_2)|$. Let $X=|S|$ and let Y be the

number of $\{y_1, y_2\} \subset S$ with $d(y_1, y_2) \leq 2k$. If $X - Y \geq k+2$, then there exist two vertices $\{x_1, x_2\} \subset V(G)$ with $d(x_1, x_2) > k+2$ and some pair $\{y_1, y_2\} \subset N(x_1, x_2)$ with $d(y_1, y_2) > 2k$, and we easily find a copy of H_k by mapping $\{1, 2\}$ to $\{x_1, x_2\}$ and $\{3, 4\}$ to $\{y_1, y_2\}$, a contradiction. So $X - Y \leq k + 1$.

On the other hand, using convexity of binomial coefficients,

$$\begin{aligned} k+1 &\geq \mathbb{E}(X - Y) \\ &\geq \sum_{y \in V(G)} \frac{\binom{d(y)}{2}}{\binom{n}{2}} - \binom{n}{2} \cdot \frac{\binom{2k}{2}}{\binom{n}{2}} \\ &\geq \frac{1}{\binom{n}{2}} n \binom{d}{2} - \binom{2k}{2} \\ &\geq \frac{d(d-1)}{n} - \binom{2k}{2}. \end{aligned}$$

It follows that if n is large enough, then $d \leq 2k\sqrt{n}$. This proves Proposition 6.2.4.

6.7 Concluding Remarks

A generalization of Sidon sets was given in [55]. Consider all equations $\alpha a + \beta b = \gamma c + \delta d$ where $\alpha + \beta = \gamma + \delta$ and $1 \leq \alpha, \beta, |\gamma|, |\delta| \leq k$. If $A \subset \Gamma$ has no solution to any of these equations other than $\{a, b\} = \{c, d\}$, then A is called a *k-fold Sidon set*. It is easy to see that a k-fold Sidon set in a finite abelian group Γ has size $O(|\Gamma|^{1/2})$. Bounds on k-fold Sidon sets will be discussed in more detail in Chapter 7 which is based on [20]. A 2-fold Sidon set of size roughly $\sqrt{N}/2$ in $\mathbb{Z}_N$ is constructed for infinitely many N in [55], but it is an open question to construct a k-fold Sidon set of size $\Theta(\sqrt{N})$ in any abelian group Γ_N of order N for any $k \geq 3$. In fact the following is conjectured in [55].

Conjecture 6.7.1. *Let $k \in \mathbb{N}$. Then there exists $c_k > 0$ such that for any $N \in \mathbb{N}$, there exists a k-fold Sidon set $A \subset [N]$ of size at least $c_k\sqrt{N}$.*

The densest current construction available is due to Ruzsa [71], who showed that for each k there exists a k-fold Sidon set of size $N^{1/2-o(1)}$ in $\mathbb{Z}_N$. The construction used for Theorem 6.1.3 is easily adapted to give counterexamples to Conjecture

6.1.2 if there exists a k-fold Sidon set of size $\Theta(\sqrt{N})$ in an abelian group Γ_N of order N. The following theorem is proved in the same way as Theorem 6.1.3, by taking $G_N = G_{\Gamma_N,\Lambda,S}(C)$ where C is a quadrilateral, $\Lambda = \{0,3,1,2\}$ and S is a 3-fold Sidon set:

Theorem 6.7.2. *If there exists a 3-fold Sidon set of size $\Theta(\sqrt{N})$ in an abelian group Γ of order N, then there exists a $4N$-vertex graph G_N with $\Theta(N^{3/2})$ edges such that every edge of G_N is contained in exactly one quadrilateral.*

A small step in answering Conjecture 6.7.1 would be a solution to the following problem.

Problem 6.7.3. Determine if for infinitely many N, there is a set $A \subset [N]$ with $|A| \geq cN^{1/2}$ where A has only trivial solutions to each of the equations

$$x_1 - x_2 = x_3 - x_4, \;\; x_1 - x_2 = 2(x_3 - x_4), \;\; \text{and } x_1 - x_2 = 3(x_3 - x_4).$$

A more general setting is given in [55] in the language of hypergraphs [8]. An r-uniform hypergraph has *girth five* if whenever one pair of vertices is selected from each hyperedge, the resulting graph has no cycles of length at most four (and in particular no double edges). In [55] it is conjectured that for every $r \geq 2$, there exists an r-uniform hypergraph on n vertices with girth five and $\Theta(n^{3/2})$ hyperedges. This is settled for $r = 2$ by Erdős and Rényi [31], and for $r = 3$ in [55]. In the case $r = 4$, if this conjecture is true then we may place a quadrilateral in each hyperedge to obtain a graph with $\Theta(n^{3/2})$ edges and n vertices, in which every edge is in exactly one quadrilateral. Using Ruzsa's construction [71], one finds for each $r \geq 3$ an n-vertex r-uniform hypergraph of girth five with $n^{3/2-o(1)}$ edges. In particular, if $f_r(n)$ is the maximum number of hyperedges in an r-uniform n-vertex hypergraph of girth five, then $f_2(n) = \Theta(n^{3/2})$, $f_3(n) = \frac{1}{6}n^{3/2} + o(n^{3/2})$, and for some constant $c_r > 0$,

$$\frac{n^{3/2}}{\exp(c_r\sqrt{\log n})} \leq f_r(n) \leq \frac{1}{r(r-1)}n^{3/2} + O(n)$$

for all $r \geq 4$. We conclude this chapter with the following problem concerning this inequality.

Problem 6.7.4. Determine if $f_r(n) = \Omega(n^{3/2})$ for $r \geq 4$.

Chapter 6 is a reprint of "On a counterexample to sparse removal," 2013. Timmons, Craig; Verstraëte, Jacques. This paper has been submitted for publication. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 7

k-fold Sidon Sets

Let $k \geq 1$ be an integer. A set $A \subset \mathbb{Z}$ is a *k-fold Sidon set* if A has only trivial solutions to each equation of the form $c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 = 0$ where $0 \leq |c_i| \leq k$, and $c_1 + c_2 + c_3 + c_4 = 0$. We prove [20] that for any integer $k \geq 1$, a k-fold Sidon set $A \subset [N]$ has at most $(N/k)^{1/2} + O((Nk)^{1/4})$ elements. Indeed we prove that given any k positive integers $c_1 < \cdots < c_k$, any set $A \subset [N]$ that contains only trivial solutions to $c_i(x_1 - x_2) = c_j(x_3 - x_4)$ for each $1 \leq i \leq j \leq k$, has at most $(N/k)^{1/2} + O((c_k^2N/k)^{1/4})$ elements. On the other hand, for any $k \geq 2$ we can exhibit k positive integers $c_1, \ldots, c_k$ and a set $A \subset [N]$ with $|A| \geq (\frac{1}{k} + o(1))N^{1/2}$, such that A has only trivial solutions to $c_i(x_1 - x_2) = c_j(x_3 - x_4)$ for each $1 \leq i \leq j \leq k$.

7.1 Introduction

Let

$$c_1x_1 + \cdots + c_rx_r = 0 \tag{7.1}$$

be an integer equation where $c_i \in \mathbb{Z}\backslash\{0\}$, and $c_1 + \cdots + c_r = 0$. Call such an equation an *invariant equation*. A solution $(x_1, \ldots, x_r) \in \mathbb{Z}^r$ to (7.1) is *trivial* if there is a partition of $\{1, \ldots, r\}$ into nonempty sets $T_1, \ldots, T_m$ such that for every $1 \leq i \leq m$, we have $\sum_{j \in T_i} c_j = 0$, and $x_{j_1} = x_{j_2}$ whenever $j_1, j_2 \in T_i$. A natural extremal problem is to determine the maximum size of a set $A \subset [N]$

with only trivial solutions to (7.1). This problem was investigated in detail by Ruzsa [71]. One of the important open problems from [71] is the genus problem. Given an invariant equation $E : c_1x_1 + \cdots + c_rx_r = 0$, the *genus* $g(E)$ is the largest integer m such that there is a partition of $\{1, \ldots, r\}$ into nonempty sets $T_1, \ldots, T_m$, such that $\sum_{j \in T_i} c_j = 0$ for $1 \leq i \leq m$. Ruzsa proved that if E is an invariant equation and $A \subset [N]$ has only trivial solutions to E, then $|A| \leq c_E N^{1/g(E)}$. Here c_E is a positive constant depending only on the equation E. Determining if there are sets $A \subset [N]$ with $|A| = N^{1/g(E)-o(1)}$ and having only trivial solutions to E is open for most equations. In particular, the genus problem is open for the equation $2x_1 + 2x_2 = 3x_3 + x_4$. This equation has genus 1 but the best known construction [71] gives a set $A \subset [N]$ with $|A| \geq cN^{1/2}$ where $c > 0$ is a positive constant. More generally, Ruzsa showed that for any four variable equation $E : c_1x_1 + c_2x_2 = c_3x_3 + c_4x_4$ with $c_1 + c_2 = c_3 + c_4$ and $c_i \in \mathbb{N}$, there is a set $A \subset [N]$ with only trivial solutions to E and $|A| \geq c_E N^{1/2-o(1)}$. In this chapter we consider special types of four variable invariant equations.

Let $k \geq 1$ be an integer. A set $A \subset \mathbb{Z}$ is a *k-fold Sidon* set if A has only trivial solutions to each equation of the form

$$c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 = 0$$

where $0 \leq |c_i| \leq k$, and $c_1 + c_2 + c_3 + c_4 = 0$. A 1-fold Sidon set is a Sidon set. A 2-fold Sidon set has only trivial solutions to each of the equations

$$x_1 + x_2 - x_3 - x_4 = 0, \qquad 2x_1 + x_2 - 2x_3 - x_4 = 0, \qquad 2x_1 - x_2 - x_3 = 0.$$

One can also define k-fold Sidon sets in $\mathbb{Z}_N$. We must add the condition that N is relatively prime to all integers in the set $\{1, 2, \ldots, k\}$. The reason for this is that if a coefficient $c_i \in \{1, 2, \ldots, k\}$ has a common factor with N, then in $\mathbb{Z}_N$ one could have $c_i(a_1 - a_2) = 0$ with $a_1 \neq a_2$. In this case, if $|A| \geq 3$, we can choose $a_3 \in A \backslash \{a_1, a_2\}$, and obtain the nontrivial solution $(x_1, x_2, x_3, x_4) = (a_1, a_2, a_3, a_3)$ to the equation $c_i(x_1 - x_2) + x_3 - x_4 = 0$.

Lazebnik and Verstraëte [55] were the first to define k-fold Sidon sets. They conjectured the following.

Conjecture 7.1.1 (Lazebnik, Verstraëte [55])**.** For any integer $k \geq 3$, there is a positive constant $c_k > 0$ such that for all integers $N \geq 1$, there is a k-fold Sidon set $A \subset [N]$ with $|A| \geq c_k N^{1/2}$.

This conjecture is still open. Lazebnik and Verstraëte proved that for infinitely many N, there is a 2-fold Sidon set $A \subset \mathbb{Z}_N$ with $|A| \geq \frac{1}{2}N^{1/2} - 3$. Axenovich [5] and Verstraëte (unpublished) observed that one can adapt Ruzsa's construction for four variable equations (Theorem 7.3, [71]) to construct k-fold Sidon sets $A \subset [N]$ or $A \subset \mathbb{Z}_N$ with $|A| \geq c_k N^{1/2} e^{-c_k\sqrt{\log N}}$ for any $k \geq 3$. An affirmative answer to Conjecture 7.1.1, even in the case when $k = 3$, would have applications to hypergraph Turán problems [55] and extremal graph theory [82].

Since any k-fold Sidon set is a Sidon set, the trivial upper bound $|A| \leq \sqrt{N - 3/4} + 1/2$ for a Sidon set $A \subset \mathbb{Z}_N$, and the Erdős-Turán bound $|A| \leq N^{1/2} + O(N^{1/4})$ for any Sidon set $A \subset [N]$, also hold for k-fold Sidon sets. We will obtain better upper bounds for k-fold Sidon sets. Instead of considering all the possible equations $c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 = 0$ with $c_1 + c_2 + c_3 + c_4 = 0$, we will take advantage only of the equations of the form

$$c_1(x_1 - x_2) = c_2(x_3 - x_4).$$

For any $c_1, \ldots, c_k$ with $(c_i, N) = 1$, if $A \subset \mathbb{Z}_N$ contains only trivial solutions to $c_i(x_1 - x_2) = c_j(x_3 - x_4)$ for each $1 \leq i \leq j \leq k$, then

$$|A| \leq \sqrt{\frac{N-1}{k} + \frac{1}{4}} + \frac{1}{2}. \tag{7.2}$$

To see this, consider all elements of the form $c_i(x - y)$ where $1 \leq i \leq k$, and $x \neq y$ are elements of A. All of these elements are distinct and nonzero. Therefore, $k|A|(|A| - 1) \leq N - 1$ which is equivalent to (7.2).

The short counting argument used to obtain (7.2) does not work in $\mathbb{Z}$. Using a more sophisticated argument, we can show that a bound similar to (7.2) does hold in $\mathbb{Z}$.

Theorem 7.1.2. *Let $k \geq 1$ be an integer and $1 \leq c_1 < c_2 < \cdots < c_k$ be a set of k distinct integers. If $A \subset [N]$ is a set with only trivial solutions to $c_i(x_1 - x_2) =$*

$c_j(x_3 - x_4)$ *for each* $1 \le i \le j \le k$, *then*

$$|A| \le \left(\frac{N}{k}\right)^{1/2} + O\left(\left(\frac{c_k^2 N}{k}\right)^{1/4}\right).$$

Taking $c_j = j$ for $1 \le j \le k$, we have the following corollary.

Corollary 7.1.3. *If* $k \ge 1$ *is an integer and* $A \subset [N]$ *is a* k*-fold Sidon set, then*

$$|A| \le \left(\frac{N}{k}\right)^{1/2} + O((kN)^{1/4}).$$

It is natural to ask if we can improve Corollary 7.1.3 if we make full use of the assumption that A is a k-fold Sidon set. For example, the bound $|A| \le (N/3)^{1/2} + O(N^{1/4})$ holds under the assumption that $A \subset [N]$ has only trivial solutions to $c_1(x_1 - x_2) = c_2(x_3 - x_4)$ for each $1 \le c_1 \le c_2 \le 3$. A 3-fold Sidon set additionally has only trivial solutions to $2x_1 + 2x_2 = 3x_3 + x_4$. Our argument does not capture this property. It is not known if this additional assumption would improve the upper bound $|A| \le (N/3)^{1/2} + O(N^{1/4})$.

The method used by Lazebnik and Verstraëte to construct 2-fold Sidon sets is rather robust. Using this method, we prove the following theorem.

Theorem 7.1.4. *There exist* k *distinct integers* $c_1, \ldots, c_k$ *and infinitely many* N*, such that there is a set* $A \subset \mathbb{Z}_N$ *with*

$$|A| \ge \frac{N^{1/2}}{k}(1 - o(1))$$

and having only trivial solutions to $c_i(x_1 - x_2) = c_j(x_3 - x_4)$ *for each* $1 \le i \le j \le k$.

The next section contains the proof of Theorem 7.1.2. Section 7.3 contains the proof of Theorem 7.1.4.

7.2 Proof of Theorem 7.1.2

For finite sets $B, C \subset \mathbb{Z}$, define

$$r_{B-C}(x) = |\{(b, c) : b - c = x, b \in B, c \in C\}|.$$

The following useful lemma has appeared in the literature (see [19] or [71]).

Lemma 7.2.1. *For any finite sets $B, C \subset \mathbb{Z}$,*

$$\frac{(|B||C|)^2}{|B+C|} \leq |B||C| + \sum_{x \neq 0} r_{B-B}(x) r_{C-C}(x). \tag{7.3}$$

Proof. By Cauchy-Schwarz,

$$\begin{aligned} \frac{(|B||C|)^2}{|B+C|} &= \frac{\left(\sum_{x \in B+C} r_{B+C}(x)\right)^2}{|B+C|} \leq \sum_x r_{B+C}^2(x) \\ &= \sum_x r_{B-B}(x) r_{C-C}(x) = |B||C| + \sum_{x \neq 0} r_{B-B}(x) r_{C-C}(x). \end{aligned}$$

□

Proof of Theorem 7.1.2. Let $1 \leq c_1 < c_2 < \cdots < c_k$ be k distinct integers. Let $A \subset [N]$ be a set with only trivial solutions to $c_i(x_1 - x_2) = c_j(x_3 - x_4)$ for each $1 \leq i \leq j \leq k$. Let

$$B_{r,i} = \{x : c_r x + i \in A\}$$

for $1 \leq r \leq k$ and $0 \leq i \leq c_r - 1$. Therefore,

$$|A| = \sum_{i=0}^{c_r - 1} |\{a \in A : \ a \equiv i \pmod{c_r}\}| = \sum_{i=0}^{c_r - 1} |B_{r,i}|$$

so by Cauchy-Schwarz,

$$|A|^2 = \left(\sum_{i=0}^{c_r - 1} |B_{r,i}| \right)^2 \leq c_r \sum_{i=0}^{c_r - 1} |B_{r,i}|^2. \tag{7.4}$$

For any $y \neq 0$,

$$\sum_{r=1}^{k} \sum_{i=0}^{c_r - 1} r_{B_{r,i} - B_{r,i}}(y) \leq 1. \tag{7.5}$$

To see this, suppose

$$y = x_1 - x_2 = x_3 - x_4 \tag{7.6}$$

where $x_1, x_2 \in B_{r,i}$ and $x_3, x_4 \in B_{r',i'}$ for some $1 \leq r, r' \leq k$, $1 \leq i \leq c_r - 1$, and $1 \leq i' \leq c_{r'} - 1$. There are elements $a_1, a_2, a_3, a_4 \in A$ such that

$$c_r x_1 + i = a_1, \quad c_r x_2 + i = a_2, \quad c_{r'} x_3 + i' = a_3, \quad \text{and} \quad c_{r'} x_4 + i' = a_4.$$

Then (7.6) implies

$$\frac{1}{c_r}(a_1 - i) - \frac{1}{c_r}(a_2 - i) = \frac{1}{c_{r'}}(a_3 - i') - \frac{1}{c_{r'}}(a_4 - i'),$$

thus $c_{r'}(a_1 - a_2) = c_r(a_3 - a_4)$. Since $y \neq 0$, we have $a_1 \neq a_2$ and $a_3 \neq a_4$ and the we would have a non trivial solution of the equation.

Let $C = \{0, 1, \ldots, m-1\}$. For any $1 \leq r \leq k$ and $0 \leq i \leq c_r - 1$, the set $B_{r,i} + C$ is contained in the interval $\{0, 1, \ldots, N/c_r + m - 1\}$. This gives the trivial estimate $|B_{r,i} + C| \leq N/c_r + m$. By Lemma 7.2.1,

$$\frac{|B_{r,i}|^2 m^2}{N/c_r + m} \leq |B_{r,i}|m + \sum_{y \neq 0} r_{B_{r,i} - B_{r,i}}(y) r_{C-C}(y).$$

We sum this inequality over all $1 \leq r \leq k$ and $0 \leq i \leq c_r - 1$ to get

$$\begin{aligned} m^2 \sum_{r=1}^{k} \frac{1}{N/c_r + m} \sum_{i=0}^{c_r - 1} |B_{r,i}|^2 &\leq \sum_{r=1}^{k} \sum_{i=0}^{c_r - 1} |B_{r,i}|m \\ &+ \sum_{y \neq 0} \sum_{r=1}^{k} \sum_{i=0}^{c_r - 1} r_{B_{r,i} - B_{r,i}}(y) r_{C-C}(y) \\ &\leq k|A|m + \sum_{y \neq 0} r_{C-C}(y) \\ &\leq m(k|A| + m). \end{aligned}$$

From (7.4) we deduce

$$m^2 |A|^2 \sum_{r=1}^{k} \frac{1}{N + c_r m} \leq m(k|A| + m). \tag{7.7}$$

The left hand side of (7.7) is at least $\frac{|A|^2 k m^2}{N + c_k m}$. Therefore, $\frac{|A|^2 k m}{N + c_k m} \leq k|A| + m$, and

$$|A|^2 k m \leq (N + c_k m)(m + k|A|).$$

From this inequality, we obtain

$$\begin{aligned} \left(|A| - \left(\frac{N}{2m} + \frac{c_k}{2} \right) \right)^2 &\leq \frac{N}{k} + \frac{c_k m}{k} + \left(\frac{N}{2m} + \frac{c_k}{2} \right)^2 \\ &\leq \frac{N}{k} + \frac{c_k m}{k} + \frac{N^2}{2m^2} + \frac{c_k^2}{2} \\ &= \frac{N}{k} \left(1 + \frac{c_k m}{N} + \frac{Nk}{2m^2} + \frac{k c_k^2}{2N} \right). \end{aligned}$$

Upon solving for $|A|$, we get

$$\begin{aligned} |A| &\leq \left(\frac{N}{k}\right)^{1/2}\left(1+\frac{c_k m}{N}+\frac{Nk}{2m^2}+\frac{kc_k^2}{2N}\right)+\frac{N}{2m}+\frac{c_k}{2} \\ &\leq \left(\frac{N}{k}\right)^{1/2}+\frac{c_k m}{k^{1/2}N^{1/2}}+\frac{N^{3/2}k^{1/2}}{2m^2}+\frac{k^{1/2}c_k^2}{2N^{1/2}}+\frac{N}{2m}+\frac{c_k}{2}. \end{aligned}$$

Take $m = \lceil (N^{3/4}k^{1/4})/c_k^{1/2} \rceil$ to get $|A| \leq \left(\frac{N}{k}\right)^{1/2} + O((c_k^2 N/k)^{1/4})$. This completes the proof of Theorem 7.1.2.

□

7.3 Proof of Theorem 7.1.4

Let $k \geq 2$ be an integer. Let p be a prime, and let $M \geq 1$ be a large integer. Let r be any prime with $r > Mk$. Let $i \geq 1$ be an integer, and set $t = r^i$ and $q = p^t$.

We will prove that for $c_j = p^{j-1}$ for $j = 1, \ldots k$ there exists a set $A \subset \mathbb{Z}_{q^2-1}$ with $|A| \geq \frac{q}{k}\left(1-\frac{1}{M}\right) - (p^4-1)(M-1)$ and having only trivial solutions to

$$x_1 - x_2 = p^{j-1}(x_3 - x_4)$$

for $1 \leq j \leq k$. This proves Theorem 7.1.4 because as i tends to infinity, the term $\frac{q}{k}\left(1-\frac{1}{M}\right)$ is the dominant term. M can be taken as large as we want, and $(p^4-1)(M-1)$ is constant with respect to i.

Let θ be a generator of the cyclic group $\mathbb{F}_{q^2}^*$. Bose and Chowla [13] proved that the set

$$C(q,\theta) = \{a \in \mathbb{Z}_{q^2-1} : \theta^a - \theta \in \mathbb{F}_q\}$$

is a Sidon set in $\mathbb{Z}_{q^2-1}$. Lindström [59] proved

$$B(q,\theta) = \{b \in \mathbb{Z}_{q^2-1} : \theta^b + \theta^{qb} = 1\}$$

is a translate of $C(q,\theta)$ and is therefore a Sidon set.

Lemma 7.3.1. *The map $x \mapsto px$ is an injection from $\mathbb{Z}_{q^2-1}$ to $\mathbb{Z}_{q^2-1}$ that maps $B(q,\theta)$ to $B(q,\theta)$.*

Proof. The map $x \mapsto px$ is 1-to-1 since p is relatively prime to q^2-1. If $b \in B(q,\theta)$, then

$$1 = (\theta^b + \theta^{qb})^p = \theta^{pb} + \theta^{q(pb)}$$

so $pb \in B(q,\theta)$. □

Let $\pi : B(q,\theta) \to B(q,\theta)$ be the permutation $\pi(b) = pb$. As in [55], we use the cycles of π to define A. Let $\sigma = (b_1, \ldots, b_m)$ be a cycle of π. If $m < k$, then remove all elements of σ from $B(q,\theta)$. If $m \geq k$, then remove all b_j in σ for which j is not divisible by k. Do this for each cycle of π. Let A be the resulting subset of $B(q,\theta)$.

Lemma 7.3.2. *For each $c \in \{1, p, p^2, \ldots, p^{k-1}\}$, A has only trivial solutions to*

$$x_1 - x_2 = c(x_3 - x_4).$$

Proof. Suppose $a_1, a_2, a_3, a_4 \in A$ and $a_1 - a_2 = p^j(a_3 - a_4)$ for some $0 \leq j \leq k-1$. By Lemma 7.3.1, there are elements $b_3, b_4 \in B(q,\theta)$ such that $p^j a_3 = b_3$ and $p^j a_4 = b_4$. This gives $a_1 - a_2 = b_3 - b_4$. Since $B(q,\theta)$ is a Sidon set, either $a_1 = a_2$, $b_3 = b_4$ or $a_1 = b_3$, $a_2 = b_4$.

If $a_1 = a_2$ and $b_3 = b_4$, then $a_3 = a_4$ and the solution (a_1, a_2, a_3, a_4) is trivial. Suppose $a_1 = b_3$ and $a_2 = b_4$. This implies $b_3 \in A$, so both $p^j a_3$ and a_3 are in A. This contradicts the way in which A was constructed. □

Lemma 7.3.3. $|A| \geq \frac{q}{k}\left(1 - \frac{1}{M}\right) - (p^4 - 1)(M - 1)$.

Proof. In order to obtain a lower bound on $|A|$, we need to estimate the number of cycles of π that are short. For instance, if all cycles of π have length less than k, then $|A| = 0$. For a cycle σ of π with length $mk \geq Mk$, we delete at most $m(k-1)$ elements from $B(q,\theta)$ and keep at least $m-1$ elements.

We estimate the number of cycles of length at most $Mk - 1$. Let $\sigma = (b, pb, \ldots, p^{e-1}b)$ be a cycle of π of length e where $e \leq Mk - 1$. The integer e is the smallest positive integer such that $p^e b \equiv b \pmod{q^2-1}$. This is the same as saying that the order of p in the multiplicative group of units $\mathbb{Z}_n^*$ is e where $n = \frac{q^2-1}{\gcd(b, q^2-1)}$. Since

$$p^{4t} - 1 = (p^{2t} - 1)(p^{2t} + 1) = (q^2 - 1)(p^{2t} + 1)$$

we have $p^{4t} \equiv 1 (\text{mod } q^2 - 1)$, so e must divide $4t = 4r^i$. Since r is prime and $r \geq Mk$, e cannot divide r, so e must divide 4. To count the number the number of cycles of π with length at most $Mk - 1$, it is enough to count the elements $x \in \mathbb{Z}_{q^2-1} \backslash \{0\}$ such that $p^4 x \equiv x (\text{mod } q^2 - 1)$. This follows from the fact that if $e \in \{1, 2\}$ and $p^e x \equiv x (\text{mod } q^2 - 1)$, then $p^4 x \equiv x (\text{mod } q^2 - 1)$. The number of solutions to this congruence is $\gcd(p^4 - 1, q^2 - 1) \leq p^4 - 1$. Therefore, there are at most $p^4 - 1$ cycles of π of length at most $Mk - 1$. For a cycle of length at least Mk, the proportion of elements of the cycle that are put into A is at least $\frac{M-1}{Mk}$ (the function $f(x) = \frac{x-1}{xk}$ is increasing provided $k > 0$). Since $|B(q, \theta)| = q$,

$$|A| \geq \left(q - (p^4 - 1)Mk\right) \left(\frac{M-1}{Mk}\right) = \frac{q}{k} \left(1 - \frac{1}{M}\right) - (p^4 - 1)(M - 1).$$

□

Theorem 7.1.4 follows from Lemmas 7.3.2 and 7.3.3.

Chapter 7 is a reprint of "k-fold Sidon sets," 2013. Cilleruelo, Javier; Timmons, Craig. This paper has been submitted for publication. The dissertation author was one of the primary investigators and authors of this paper.

Chapter 8

B_k^+-sets

Let Γ be an abelian group. A set $A \subset \Gamma$ is a *B_k^+-set* if whenever

$$a_1 + \cdots + a_k = b_1 + \cdots + b_k \text{ with } a_i, b_j \in A,$$

there is an i and a j such that $a_i = b_j$. If A is a B_k-set then it is also a B_k^+-set, but the converse is not true in general. Determining the largest size of a B_k-set in the interval $\{1, 2, \ldots, N\} \subset \mathbb{Z}$ or in the cyclic group $\mathbb{Z}_N$ is a well-studied problem. In this chapter we investigate the corresponding problem for B_k^+-sets. We prove nontrivial upper bounds on the maximum size of a B_k^+-set contained in the interval $\{1, 2, \ldots, N\}$. For odd $k \geq 3$, we construct B_k^+-sets that have more elements than the B_k-sets constructed by Bose and Chowla. We prove that any B_3^+-set $A \subset \mathbb{Z}_N$ has at most $(1 + o(1))(8N)^{1/3}$ elements. A set A is a *B_k^*-set* if whenever $a_1 + \cdots + a_k = a_{k+1} + \cdots + a_{2k}$ with $a_i \in A$, there is an $i \neq j$ such that $a_i = a_j$. We obtain new upper bounds on the maximum size of a B_k^*-set $A \subset \{1, 2, \ldots, N\}$, a problem first investigated by Ruzsa.

8.1 Introduction

Let Γ be an abelian group. A set $A \subset \Gamma$ is a *B_k^+-set* if

$$a_1 + \cdots + a_k = b_1 + \cdots + b_k \text{ with } a_1, \ldots, a_k, b_1, \ldots, b_k \in A \tag{8.1}$$

implies $a_i = b_j$ for some i and j. A set A is a *B_k-set* if (8.1) implies $(a_1, \ldots, a_k)$ is a permutation of $(b_1, \ldots, b_k)$. If A is a B_k-set then A is also a B_k^+-set, but in general

the converse is not true. Often B_2-sets are called *Sidon sets* and have received much attention since they were first studied by Erdős and Turán [37] in 1941. Let $F_k(N)$ be the maximum size of a B_k-set $A \subset [N]$ and let $C_k(N)$ be the maximum size of a B_k-set $A \subset \mathbb{Z}_N$. If $A \subset \mathbb{Z}_N$ is a B_k-set, then A is also a B_k-set when viewed as a subset of $\mathbb{Z}$. Thus for any $k \geq 2$, $C_k(N) \leq F_k(N)$.

Before discussing the functions $C_k(N)$ and $F_k(N)$ in more detail, we recall some of the results from Chapter 3. Erdős and Turán [37] proved that for any $N \geq 1$, $F_2(N) \leq N^{1/2} + O(N^{1/4})$. Their argument was used by Lindström [59] to show $F_2(N) \leq N^{1/2} + N^{1/4} + 1$. In 2010, Cilleruelo [19] obtained $F_2(N) \leq N^{1/2} + N^{1/4} + \frac{1}{2}$ as a consequence of a more general result. This is the best known upper bound on $F_2(N)$. By counting differences $a - b$ with $a \neq b$, it is easy to prove $C_2(N) \leq \sqrt{N} + 1$. There are several constructions of dense B_2-sets (see [74], [13], [71]) that show $C_2(N) \geq N^{1/2}$ for infinitely many N. These constructions imply $F_2(N) \sim \sqrt{N}$ and $\limsup \frac{C_2(N)}{\sqrt{N}} = 1$.

For $k \geq 3$, bounds on $F_k(N)$ and $C_k(N)$ are not as precise. For each $k \geq 2$ and prime power q, Bose and Chowla [13] constructed a B_k-set $A \subset \mathbb{Z}_{q^k-1}$ with $|A| = q$ so that

$$(1 + o(1))N^{1/k} \leq F_k(N).$$

The current upper bounds on $F_k(N)$ and $C_k(N)$ do not match this lower bound for any $k \geq 3$. If $A \subset [N]$ is a B_k-set then each k-multiset in A gives rise to a unique sum in $\{1, \ldots, kN\}$. Therefore, $\binom{|A|+k-1}{k} \leq kN$ which implies $F_k(N) \leq (k! \cdot kN)^{1/k}$. Similar counting shows $C_k(N) \leq (k!N)^{1/k}$. By considering differences one can improve these bounds. We illustrate this idea with an example that is relevant to our results. Let $A \subset \mathbb{Z}_N$ be a B_3-set. There are $\binom{|A|}{2}(|A| - 2)$ sums of the form $a_1 + a_2 - a_3$ where a_1, a_2, and a_3 are distinct elements of A. Let $A^{(2)} = \{\{x, y\} : x, y \in A, x \neq y\}$. It is not hard to check that each $n \in \mathbb{Z}_N$ has at most one representation as $n = a_1 + a_2 - a_3$ with $\{a_1, a_2\} \in A^{(2)}$ and $a_3 \in A \backslash \{a_1, a_2\}$. This implies $\binom{|A|}{2}(|A| - 2) \leq N$ so $|A| \leq (2N)^{1/3} + 2$. In general, for any $k \geq 2$

$$C_k(N) \leq \left(\left\lfloor \frac{k}{2} \right\rfloor ! \left\lceil \frac{k}{2} \right\rceil ! N \right)^{1/k} + O_k(1), \tag{8.2}$$

and

$$F_k(N) \leq \left(\left\lfloor \frac{k}{2} \right\rfloor ! \left\lceil \frac{k}{2} \right\rceil ! \cdot \left\lceil \frac{k}{2} \right\rceil N \right)^{1/k} + O_k(N^{1/2k}). \tag{8.3}$$

These bounds were first obtained by Jia [52] in the even case, and Chen [15] in the odd case. The best upper bounds on $F_k(N)$ are to due to Green [47]. For every $k \geq 2$, (8.3) has been improved (see for example [47] or [18]), but there is no value of $k \geq 3$ for which (8.2) has been improved. This is interesting since all of the constructions take place in cyclic groups and provide lower bounds on $C_k(N)$. For other bounds on B_k-sets the interested reader is referred to Green [47], Cilleruelo [18], O'Bryant's survey [65], or the book of Halberstam and Roth [48].

Now we discuss B_k^+-sets. Write $F_k^+(N)$ for the maximum size of a B_k^+-set $A \subset [N]$, and $C_k^+(N)$ for the maximum size of a B_k^+-set $A \subset \mathbb{Z}_N$. Ruzsa [71] proved that a set $A \subset [N]$ with no solution to the equation $x_1 + \cdots + x_k = y_1 + \cdots + y_k$ in $2k$ distinct integers has at most $(1 + o(1))k^{2-1/k}N^{1/k}$ elements. Call such a set a B_k^*-set and define $F_k^*(N)$ in the obvious way. Any B_k^+-set is also a B_k^*-set so that $F_k^+(N) \leq F_k^*(N)$. Using the constructions of Bose and Chowla [13] and Ruzsa's Theorem 5.1 of [71], we get for every $k \geq 3$,

$$(1 + o(1))N^{1/k} \leq F_k(N) \leq F_k^+(N) \leq F_k^*(N) \leq (1 + o(1))k^{2-1/k}N^{1/k}.$$

In this chapter we improve this upper bound on $F_k^+(N)$ and $F_k^*(N)$. We also improve this lower bound on $F_k^+(N)$ for all odd $k \geq 3$, and we prove a nontrivial upper bound on $C_3^+(N)$. We do not consider the case when $k = 2$. The reason for this is that Ruzsa [71] proved $F_2^*(N) \leq N^{1/2} + 4N^{1/4} + 11$, and thus $F_2(N) \sim F_2^+(N) \sim F_2^*(N) \sim N^{1/2}$. In fact, a B_2-set is the same as a B_2^+-set.

Our first result is a construction which shows that for any odd $k \geq 3$, there is a B_k^+-set in $[N]$ that has more elements than any known B_k-set contained in $[N]$.

Theorem 8.1.1. *For any prime power q and odd integer $k \geq 3$, there is a B_k^+-set $A \subset \mathbb{Z}_{2(q^k-1)}$ with $|A| = 2q$.*

Using known results on densities of primes (see [6] for example), Theorem 8.1.1 implies

Corollary 8.1.2. *For any integer $N \geq 1$ and any odd integer $k \geq 3$,*

$$F_k^+(N) \geq (1 + o(1))2^{1-1/k}N^{1/k}.$$

Green proved $F_3(N) \leq (1 + o(1))(3.5N)^{1/3}$. We will use a Bose-Chowla B_3-set to construct a B_3^+-set $A \subset [2q^3]$ with $|A| = 2q = (4 \cdot 2q^3)^{1/3}$. Putting the two results together we see that A is denser than any B_3-set in $[2q^3]$ for sufficiently large prime powers q. Our construction and Green's upper bound show that $F_3(N)$ and $F_3^*(N)$ are not asymptotically the same.

The proof of Theorem 8.1.1 is based on a simple lemma, Lemma 8.2.1, which implies

$$2C_k(N) \leq C_k^+(2N) \text{ for any odd } k \geq 3. \tag{8.4}$$

This inequality provides us with a method of estimating $C_k(N)$ by proving upper bounds on $C_k^+(N)$ for odd k. Our next theorem provides such an estimate when $k = 3$.

Theorem 8.1.3. *If $A \subset \mathbb{Z}_N$ is a B_3^+-set, then*

$$|A| \leq (1 + o(1))(8N)^{1/3}.$$

Theorem 8.1.3 and (8.4) imply

Corollary 8.1.4. *If $A \subset \mathbb{Z}_N$ is a B_3-set, then*

$$|A| \leq (1 + o(1))(2N)^{1/3}.$$

As shown above, there is a simpler argument that implies this bound. The novelty here is that our results imply (8.2) for $k = 3$. It is important to mention that the error term we obtain is larger than the error term in the bound $C_3(N) \leq (2N)^{1/3} + 2$. We feel that any improvement in the leading term of Theorem 8.1.3 or (8.2) would be significant.

In $\mathbb{Z}$ we obtain the following bounds for small k.

Theorem 8.1.5. *(i) If $A \subset [N]$ is a B_3^+-set, then*

$$|A| \leq (1 + o(1))(18N)^{1/3}.$$

(ii) If $A \subset [N]$ is a B_4^+-set, then

$$|A| \leq (1+o(1))(272N)^{1/4}.$$

Recall that Ruzsa [71] proved $F_k^*(N) \leq (1+o(1))k^{2-1/k}N^{1/k}$ which implies $F_k^+(N) \leq (1+o(1))k^{2-1/k}N^{1/k}$. For $k \geq 5$, we were able to improve this upper bound on $F_k^+(N)$ by modifying arguments of Ruzsa. Our method also applies to B_k^*-sets. As a consequence, we improve the upper bound on $F_k^*(N)$ for all $k \geq 3$. We state our result only for $k = 3$ and for large k. For other small values of k the reader is referred to Table 1 in Section 8.6.

Theorem 8.1.6. *If $A \subset [N]$ is a B_3^*-set, then*

$$|A| \leq (1+o(1))(162N)^{1/3}.$$

If $A \subset [N]$ is a B_k^*-set, then

$$|A| \leq \left(\frac{1}{4} + \epsilon(k)\right) k^2 N^{1/k}$$

where $\epsilon(k) \to 0$ as $k \to \infty$.

We remark that Ruzsa's upper bound on $F_k^*(N)$ is asymptotic to $k^2N^{1/k}$. Our results do not rule out the possibility of $F_k^+(N)$ being asymptotic to $F_k^*(N)$.

Problem 8.1.7. Determine whether or not $F_k^+(N)$ is asymptotic to $F_k^*(N)$ for $k \geq 3$.

If $A \subset [N]$ is a B_k^*-set, then the number of solutions to $2x_1+x_2+\cdots+x_{k-1} = y_1+\cdots+y_k$ with $x_i, y_j \in A$ is $o(|A|^k)$ (see [71]). A B_k^*-set allows solutions to this equation with $x_1, \ldots, x_{k-1}, y_1, \ldots, y_k$ all distinct, but such a solution cannot occur in a B_k^+-set. If it were true that $F_k^+(N)$ is asymptotic to $F_k^*(N)$, then this would confirm the belief that it is the sums of k distinct elements of A that control the size of A and the lower order sums should not matter. Jia [52] defines a *semi-B_k-set* to be a set A with the property that all sums consisting of k distinct elements of A are distinct. He states that Erdős conjectured [27] that a semi-B_k-set $A \subset [N]$ should satisfy $|A| \leq (1+o(1))N^{1/k}$. A positive answer to Problem 8.1.7 would be evidence in favor of this conjecture.

At this time we do not know how to construct B_{2k}^+-sets or B_{2k}^*-sets for any $k \geq 2$ that are bigger than the corresponding Bose-Chowla B_{2k}-sets. We were able to construct interesting B_4^+-sets in the non-abelian setting.

Let G be a non-abelian group. A set $A \subset G$ is a *non-abelian B_k-set* if

$$a_1 a_2 \cdots a_k = b_1 b_2 \cdots b_k \text{ with } a_i, b_j \in A \tag{8.5}$$

implies $a_i = b_i$ for $1 \leq i \leq k$. If $A \subset G$ is a non-abelian B_k-set, then every k-letter word in $|A|$ is different so $|A|^k \leq |G|$. Odlyzko and Smith [66] proved that there exist infinitely many groups G such that G has a non-abelian B_4-set $A \subset G$ with $|A| = (1 + o(1)) \left(\frac{|G|}{1024}\right)^{1/4}$. They actually proved a more general result that gives constructions of non-abelian B_k-sets for all $k \geq 2$. The case when $k = 4$ is the only result that we will need. Define a *non-abelian B_k^+-set* to be a set $A \subset G$ such that (8.5) implies $a_i = b_i$ for some $i \in \{1, 2, \ldots, k\}$. As in the abelian setting, a non-abelian B_k-set is also a non-abelian B_k^+-set but the converse is not true in general. Using a construction of [66], we prove

Theorem 8.1.8. *For any prime p with $p - 1$ divisible by 4, there is a non-abelian group G of order $48(p^4 - 1)$ that contains a non-abelian B_4^+-set $A \subset G$ with*

$$|A| = \frac{1}{2}(p - 1).$$

Our result shows that there are infinitely many groups G such that G has a non-abelian B_4^+-set A with $|A| = \left(\frac{|G|}{768}\right)^{1/4} + o(|G|^{1/4})$. We conclude our introduction with the following conjecture concerning B_{2k}^+-sets.

Conjecture 8.1.9. If $k \geq 4$ is any even integer, then there exists a positive constant c_k such that for infinitely many N,

$$F_k^+(N) \geq (1 + c_k + o(1))N^{1/k}.$$

If Conjecture 8.1.9 is true with $c_k = 2^{1-1/k} - 1$ as in the odd case, then using Green's upper bound $F_4(N) \leq (1 + o(1))(7N)^{1/4}$, we can conclude that $F_4(N)$ and $F_4^*(N)$ are not asymptotically the same just as in the case when $k = 3$. Our hope is that a positive answer to Conjecture 8.1.9 will either provide an analogue of (8.4) for even $k \geq 4$, or a construction of a B_k^+-set that does not use Bose-Chowla B_k-sets.

8.2 Proof of Theorem 8.1.1

In this section we show how to construct B_k^+-sets for odd $k \geq 3$. Our idea is to take a dense B_k-set A and a translate of A.

Lemma 8.2.1. *If $A \subset \mathbb{Z}_N$ is a B_k-set where $k \geq 3$ is odd, then*

$$A^+ := \{a + bN : a \in A, b \in \{0, 1\}\}$$

is a B_k^+-set in $\mathbb{Z}_{2N}$.

Proof. Let $k \geq 3$ be odd and suppose

$$\sum_{i=1}^{k} a_i + b_i N \equiv \sum_{i=1}^{k} c_i + d_i N \pmod{2N} \tag{8.6}$$

where $a_i, c_i \in A$, and $b_i, d_i \in \{0, 1\}$. Taking (8.6) modulo N gives

$$\sum_{i=1}^{k} a_i \equiv \sum_{i=1}^{k} c_i \pmod{N}.$$

Since A is a B_k-set in $\mathbb{Z}_N$, $(a_1, \ldots, a_k)$ must be a permutation of $(c_1, \ldots, c_k)$. If we label the a_i's and c_i's so that $a_1 \leq a_2 \leq \cdots \leq a_k$ and $c_1 \leq c_2 \leq \cdots \leq c_k$, then $a_i = c_i$ for $1 \leq i \leq k$. Rewrite (8.6) as

$$\sum_{i=1}^{k} b_i N \equiv \sum_{i=1}^{k} d_i N \pmod{2N}.$$

The sums $\sum_{i=1}^{k} b_i$ and $\sum_{i=1}^{k} d_i$ have the same parity. Since k is odd and $b_i, d_i \in \{0, 1\}$, there must be a j such that $b_j = d_j$, so $a_j + b_j N \equiv c_j + d_j N \pmod{2N}$. □

Let q be a prime power, $k \geq 3$ be an odd integer, and A_k be a Bose-Chowla B_k-set with $A_k \subset \mathbb{Z}_{q^k-1}$ (see [13] for a description of A_k). Let

$$A_k^+ = \{a + b(q^k - 1) : a \in A_k, b \in \{0, 1\}\}.$$

By Lemma 8.2.1, A_k^+ is a B_k^+-set in $\mathbb{Z}_{2(q^k-1)}$ and $|A_k^+| = 2|A_k| = 2q$. This proves Theorem 8.1.1.

8.3 Proof of Theorem 8.1.3

Let $A \subset \mathbb{Z}_N$ be a B_3^+-set. If N is odd, then $2x \equiv 2y \pmod N$ implies $x \equiv y \pmod N$. If N is even, then $2x \equiv 2y \pmod N$ implies $x \equiv y \pmod N$ or $x \equiv y+N/2 \pmod N$. Because of this, the odd case is quite a bit easier to deal with and so we present the more difficult case. In this section N is assumed to be even. If N is odd, then the proof of Theorem 8.1.5(i) given in the next section works in $\mathbb{Z}_N$. The only modification needed is to divide by N instead of $3N$ when applying Cauchy-Schwarz. For simplicity of notation, we write $x = y$ rather than $x \equiv y \pmod N$.

For $n \in \mathbb{Z}_N$, define

$$f(n) = \#\left\{(\{a,c\},b) \in A^{(2)} \times A : n = a - b + c, \{a,c\} \cap \{b\} = \emptyset\right\}.$$

Recall that $A^{(2)} = \{\{x,y\} : x,y \in A, x \neq y\}$. The sum $\sum f(n)(f(n)-1)$ counts the number of ordered pairs $((\{a,c\},b),(\{x,z\},y))$ such that the tuples $(\{a,c\},b)$ and $(\{x,z\},y)$ are distinct, and both are counted by $f(n)$. For each such pair we cannot have $\{a,c\} = \{x,z\}$. Otherwise, the tuples would be equal. If $((\{a,c\},b),(\{x,z\},y))$ is counted by $\sum f(n)(f(n)-1)$, then $a+y+c = x+b+z$. By the B_3^+ property, $\{a,y,c\} \cap \{x,b,z\} \neq \emptyset$ so that $\{a,c\} \cap \{x,z\} \neq \emptyset$ or $b = y$. The tuples are distinct, so both of these cases cannot occur at the same time.

Case 1: $\{a,c\} \cap \{x,z\} \neq \emptyset$ and $b \neq y$.

Without loss of generality, assume $a = x$. Cancel a from both sides of the equation $a - b + c = x - y + z$ and solve for c to get $c = b - y + z$. Here we are using the ordering of the tuples $((\{a,c\},b),(\{x,z\},y))$ to designate which element is solved for after the cancellation of the common term.

If $z = b$, then $c + y = 2b$ and we have a 3-term arithmetic progression (a.p. for short). The number of trivial 3-term a.p.'s in A is at most $2|A|$ since for any $a \in A$,

$$a + a = 2a = 2(a + N/2).$$

Next we count the number of nontrivial 3-terms a.p.'s. By nontrivial, we mean that all terms involved in the a.p. are distinct, and $a+a = 2(a+N/2)$ is considered to be trivial.

If $p+q=2r$ is a 3-term a.p., then call p and q *outer terms*. Let p be an outer term of the 3-term a.p. $p+q=2r$ where $p,q,r\in A$. We will show that p is an outer term of at most one other nontrivial a.p. Let $p+q'=2r'$ be another a.p. with $q',r'\in A$ and $(q,r,)\neq(q',r')$.

If $r=r'$, then $p+q=2r=2r'=p+q'$ so $q=q'$. This is a contradiction and so we can assume that $r\neq r'$.

If $q=q'$, then $2r=p+q=p+q'=2r'$ so $r'=r$ or $r'=r+N/2$. Thus, $p+q=2r$ or $p+q=2(r+N/2)$.

Now suppose $r\neq r'$ and $q\neq q'$. Since $2r-q=p=2r'-q'$ we have by the B_3^+ property,

$$\{r,q'\}\cap\{r',q\}\neq\emptyset.$$

The only two possibilities are $r=q$ or $r'=q'$, but in either of these cases we get a trivial 3-term a.p. Putting everything together proves the following lemma.

Lemma 8.3.1. *If $A\subset\mathbb{Z}_N$ is a B_3^+-set, then the number of 3-term arithmetic progressions in A is at most $4|A|$.*

Given a fixed element $a\in A$ and a fixed 3-term a.p. $c+y=2b$ in A, there are at most 4! ways to form an ordered tuple of the form $((\{a,c\},b),(\{a,b\},y))$. The number of ordered tuples counted by $\sum f(n)(f(n)-1)$ when $\{a,c\}\cap\{x,z\}\neq\emptyset$ and $z=b$ is at most $4!|A|\cdot 4|A|=96|A|^2$. The first factor of $|A|$ in the expression $4!|A|\cdot 3|A|$ comes from the number of ways to choose the element a.

Assume now that $z\neq b$. Recall that we have solved for c to get $c=b-y+z$. If $b=y$, then $c=z$ which implies $\{a,c\}=\{x,z\}$, a contradiction as the tuples are distinct. By definition $y\neq z$, so $c=b-y+z$ where $\{b,z\}\in A^{(2)}$ and $\{y\}\cap\{b,z\}=\emptyset$. The number of ways to write c in this form is $f(c)$. Given such a solution $\{b,z\},y$ counted by $f(c)$, there are two ways to order b and z, and $|A|$ ways to choose $a=x$. The number of ordered tuples we obtain when $\{a,c\}\cap\{x,z\}\neq\emptyset$ and $z\neq b$ is at most $|A|\cdot 2\sum_{c\in A}f(c)$. This completes the analysis in Case 1.

Before addressing Case 2, the case when $b=y$ and $\{a,c\}\cap\{x,z\}=\emptyset$, some additional notation is needed. For $d\in A+A$, define

$$S(d)=\big\{\{a,b\}\in A^{(2)}:a+b=d \text{ and there is a pair } \{a',b'\}\in A^{(2)}$$

with $\{a,b\} \cap \{a',b'\} = \emptyset$ and $a'+b' = d\}$.

Let $d_1, d_2, \ldots, d_M$ be the integers for which $S(d_i) \neq \emptyset$. Write S_i^2 for $S(d_i)$ and define

$$T_i^1 = \{a : a \in \{a,b\} \text{ for some } \{a,b\} \in S_i^2\}.$$

Let $s_i = |S_i^2|$ and $d_1, d_2, \ldots, d_m$ be the integers for which $s_i = 2$. Let $d_{m+1}, \ldots, d_M$ be the integers for which $s_i \geq 3$. For $1 \leq i \leq M$, we will use the notation $S_i^2 = \{\{a_1^i, b_1^i\}, \{a_2^i, b_2^i\}, \ldots, \{a_{s_i}^i, b_{s_i}^i\}\}$. A simple, but important, observation is that for any fixed $i \in \{1, \ldots, M\}$, any element of A appears in at most one pair in S_i^2.

If A was a B_3-set, then there would be no d_i's. This suggests that a B_3^+-set or a B_3^*-set that is denser than a B_3-set should have many d_i's. The B_3^+-set A_3^+ constructed in Theorem 8.1.1 has $m \approx \frac{1}{2}\binom{|A_3^+|}{2}$. However, if A_3^+ is viewed as a subset of $\mathbb{Z}$, then $m \approx \frac{1}{4}\binom{|A_3^+|}{2}$ (see Lemma 8.4.3 which also holds in $\mathbb{Z}_N$ if N is odd).

Case 2: $b = y$ and $\{a,c\} \cap \{x,z\} = \emptyset$.

If $b = y$, then $a + c = x + z$. There are $|A|$ choices for $b = y$ and

$$\sum_{i=1}^{M} |S_i^2|(|S_i^2| - 1)$$

ways to choose an ordered pair of different sets $\{a,c\}, \{x,z\} \in A^{(2)}$ with $a + c = x + z$, and $\{a,c\} \cap \{x,z\} = \emptyset$.

Putting Cases 1 and 2 together gives the estimate

$$\sum f(n)(f(n)-1) \leq |A| \left(2\sum_{c \in A} f(c) + \sum_{i=1}^{M} |S_i^2|(|S_i^2| - 1) \right) + 96|A|^2. \qquad (8.7)$$

Our goal is to find upper bounds on the sums

$$\textstyle\sum_{c \in A} f(c) \text{ and } \sum_{i=1}^{M} |S_i^2|(|S_i^2| - 1).$$

Lemma 8.3.2. *If $x \in T_i^1 \cap T_j^1$ for some $i \neq j$, then (i) $\max\{s_i, s_j\} \leq 3$ and (ii) if $s_i = s_j = 3$, then for some $x_1, y, z \in A$ depending on i and j, we have $d_j = d_i + N/2$ and $S_i^2 = \{\{x, x_1\}, \{y, z\}, \{y + \frac{N}{2}, z + \frac{N}{2}\}\}$, $S_j^2 = \{\{x, x_1 + \frac{N}{2}\}, \{y + \frac{N}{2}, z\}, \{y, z + \frac{N}{2}\}\}$.*

Proof. If $s_i = 2$ and $s_j = 2$ then we are done. Assume $s_j > 2$. Let $S_i^2 = \{\{a_1^i, b_1^i\}, \ldots, \{a_{s_i}^i, b_{s_i}^i\}\}$ and $S_j^2 = \{\{a_1^j, b_1^j\}, \ldots, \{a_{s_j}^j, b_{s_j}^j\}\}$. Without loss of generality, suppose $x = a_i^1$ and $x = a_j^1$. By definition, $s_i \geq 2$ so we can write $d_i = x + b_1^i = a_2^i + b_2^i$ and $d_j = x + b_1^j = a_2^j + b_2^j = a_3^j + b_3^j$.

Solve for x to get $x = a_2^i + b_2^i - b_1^i = a_2^j + b_2^j - b_1^j$. This can be rewritten as

$$a_2^i + b_2^i + b_1^j = a_2^j + b_2^j + b_1^i. \tag{8.8}$$

Since $d_i \neq d_j$, b_1^i cannot be b_1^j therefore b_1^j is not on the right hand side of (8.8), and b_1^i is not on the left hand side of (8.8). By the B_3^+ property, $\{a_2^i, b_2^i\} \cap \{a_2^j, b_2^j\} \neq \emptyset$. The same argument can be repeated with a_3^j in place of a_2^j and b_3^j in place of b_2^j to get

$$\{a_2^i, b_2^i\} \cap \{a_3^j, b_3^j\} \neq \emptyset.$$

Recall any element of A can occur at most once in the list $a_1^j, b_1^j, a_2^j, b_2^j, \ldots a_{s_j}^j, b_{s_j}^j$ thus $s_j \leq 3$. By symmetry, $s_i \leq 3$.

Now suppose $s_i = s_j = 3$. Repeating the argument above, we have for each $2 \leq k \leq 3$ and $2 \leq l \leq 3$,

$$|\{a_l^i, b_l^i\} \cap \{a_k^j, b_k^j\}| = 1.$$

This intersection cannot have size 2 since $d_i \neq d_j$. Without loss of generality, let $y = a_2^i = a_2^j$, $z = b_2^i = a_3^j$, $u = a_3^i = b_2^j$, and $v = b_3^i = b_3^j$. We represent these equalities between T_i^1 and T_j^1 using a bipartite graph with parts T_i^1 and T_j^1 where $w \in T_i^1$ is adjacent to $w' \in T_j^1$ if and only if $w = w'$ (see Figure 8.1).

The equalities $d_i = y+z = u+v$ and $d_j = y+u = z+v$ imply $d_i - d_j = z-u$ and $d_i - d_j = u - z$. Therefore $2z = 2u$. If $z = u$, then this is a contradiction since the elements in the list x, b_1^i, y, z, u, v are all distinct. It is in this step that the parity of N plays an important role. We conclude $u = z + N/2$ and

$$d_j = y + u = y + (z + N/2) = y + z + N/2 = d_i + N/2.$$

Let $b_1^i = x_1$ so $b_1^j = x_1 + N/2$. Since $d_i = y + z = u + v$ and $u = z + N/2$,

$$v = y + z - u = y + z - (z + N/2) = y - N/2 = y + N/2.$$

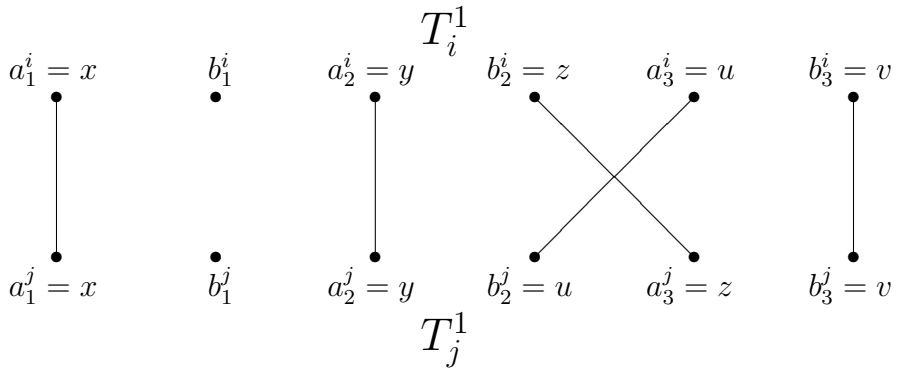

Figure 8.1: Equality Graph for Lemma 8.3.2

Substituting $u = z + N/2$ and $v = y + N/2$ gives the assertion about the pairs in S_i^2 and S_j^2 when $s_i = s_j = 3$. □

Corollary 8.3.3. *If $s_i \geq 4$, then for any $j \neq i$, $T_i^1 \cap T_j^1 = \emptyset$. Furthermore, any $x \in A$ is in at most two T_i^1's with $s_i = 3$.*

Proof. The first statement follows immediately from Lemma 8.3.2. For the second statement, suppose $x \in T_i^1 \cap T_j^1$ with $s_i = s_j = 3$ and $i \neq j$. By Lemma 8.3.2, $\{x, x_1\} \in S_i^2$ and $\{x, x_1 + N/2\} \in S_j^2$ for some $x_1 \in A$. If $x \in T_k^1$ with $k \neq i$, then $\{x, x_1 + N/2\} \in S_k^2$ so $d_j = x + (x_1 + N/2) = d_k$ and $j = k$. □

Lemma 8.3.4. *If $A \subset \mathbb{Z}_N$ is a B_3^+-set, then*

$$\sum_{c \in A} f(c) \leq |A|^2 + 7|A|.$$

Proof. For $c \in A$, let

$$g_1(c) = \#\left\{(\{x, z\}, y) \in A^{(2)} \times A : c = x - y + z, c \neq y, \{x, z\} \cap \{y\} = \emptyset\right\}$$

and

$$g_2(c) = \#\left\{(\{x, z\}, y) \in A^{(2)} \times A : c = x - y + z, c = y, \{x, z\} \cap \{y\} = \emptyset\right\}.$$

For each $c \in A$, $f(c) = g_1(c) + g_2(c)$. The sum $\sum_{c \in A} g_2(c)$ is exactly the number of nontrivial 3-term a.p.'s in A. By Lemma 8.3.1, $\sum_{c \in A} g_1(c) \leq 4|A|$. Estimating $\sum_{c \in A} g_1(c)$ takes more work. To compute $g_1(c)$ with $c \in A$, we first choose an

i with $c \in T_i^1$, and then choose one of the pairs $\{x, z\} \in S_i^2 \backslash \{c, y\}$ to obtain a solution $c = x - y + z$ with $c \neq y$ and $\{x, z\} \cap \{y\} = \emptyset$.

If $c \notin T_1^1 \cup \cdots \cup T_M^1$, then the equation $c + y = x + z$ with c, y, x, and z all distinct has no solutions in A so $g_1(c) = 0$. Assume $c \in T_1^1 \cup \cdots \cup T_M^1$.

Case 1: $c \notin T_1^1 \cup \cdots \cup T_m^1$.

By Corollary 8.3.3, there are two possibilities. One is that there is a unique j with $c \in T_j^1$ and $s_j \geq 3$. In this case, $|S_j^2| \leq \frac{|A|}{2}$ so $g_1(c) \leq \frac{|A|}{2}$. The other possibility is that $c \in T_i^1 \cap T_j^1$ with $s_i = s_j = 3$ and $i \neq j$. In this case, $g_1(c) \leq 4$ because we can choose either i or j, and then one of the two pairs in S_i^2 or S_j^2 that does not contain c.

Case 2: $c \in T_1^1 \cup \cdots \cup T_m^1$.

By Lemma 8.3.2, c is not in any T_j^1 with $s_j \geq 4$ and c is in at most two T_j^1's with $s_j = 3$. There are at most $|A|$ T_i^1's with $c \in T_i^1$ since there are at most $|A|$ pairs $\{c, y\}$ that contain c so $g_1(c) \leq |A| + 4$.

In all cases, $g_1(c) \leq |A| + 4$ and

$$\sum_{c \in A} f(c) = \sum_{c \in A} (g_1(c) + g_2(c)) \leq |A|(|A| + 4) + 4|A|$$

which proves the lemma. □

Lemma 8.3.5. *If $g_1(c)$ is the function of Lemma 8.3.4, then*

$$2 \sum_{i=1}^{M} |S_i^2|(|S_i^2| - 1) = \sum_{c \in A} g_1(c).$$

Proof. Define an edge colored graph G with vertex set A, edge set $\cup_{i=1}^{M} S_i^2$, and the color of edge $\{a, b\}$ is $a + b$. The sum $\sum_{i=1}^{M} |S_i^2|(|S_i^2| - 1)$ counts ordered pairs $(\{c, y\}, \{x, z\})$ of distinct edges of G where $\{c, y\}$ and $\{x, z\}$ have the same color, i.e. $c + y = x + z$, and c, y, x, and z are all distinct elements of A. The sum $\sum_{c \in A} g_1(c)$ counts each such ordered pair $(\{c, y\}, \{x, z\})$ exactly two times, one contribution coming from $g_1(c)$ and the other from $g_1(y)$. □

By Lemma 8.3.5,

$$\sum_{i=1}^{M} |S_i^2|(|S_i^2| - 1) \leq \frac{1}{2} \sum_{c \in A} f(c). \tag{8.9}$$

Next we use the following version of the Cauchy-Schwarz inequality.

Lemma 8.3.6 (Cauchy-Schwarz)**.** *If $x_1, \ldots, x_n$ are real numbers, $t \in \{1, 2, \ldots, n-1\}$, and $\Delta = \frac{1}{t}\sum_{i=1}^{t} x_i - \frac{1}{n}\sum_{i=1}^{n} x_i$, then*

$$\sum_{i=1}^{n} x_i^2 \geq \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 + \frac{tn\Delta^2}{n-t}.$$

A simple counting argument shows that $\sum f(n) = \binom{|A|}{2}(|A|-2)$. Let $\sum_{c\in A} f(c) = \delta|A|^2$. If

$$\Delta := \frac{1}{|A|}\sum_{c\in A} f(c) - \frac{1}{N}\sum_{n} f(n) = \delta|A| - \frac{1}{N}\sum_{n} f(n)$$

then, using Ruzsa's bound $|A| = O(N^{1/3})$ and $C_3^+(N) \leq F_3^+(N)$, we get

$$\Delta = \delta|A| - \frac{\binom{|A|}{2}(|A|-2)}{N} \geq \delta|A| - C$$

where C is some absolute constant. By Lemma 8.3.6,

$$\begin{aligned}\sum f(n)^2 &\geq \frac{\binom{|A|}{2}^2(|A|-2)^2}{N} + \frac{|A|\cdot N(\delta|A|-C)^2}{N-|A|} \\ &= \frac{\binom{|A|}{2}^2(|A|-2)^2}{N} + \delta^2|A|^3\frac{\left(1-\frac{C}{\delta|A|}\right)^2}{1-\frac{|A|}{N}}.\end{aligned}$$

By (8.7) and (8.9),

$$\begin{aligned}\sum f(n)^2 &\leq \sum f(n) + |A|\left(2\sum_{c\in A} f(c) + \sum_{i=1}^{M}|S_i^2|(|S_i^2|-1)\right) + 96|A|^2 \\ &\leq \frac{|A|^3}{2} + \frac{5|A|}{2}\sum_{c\in A} f(c) + 96|A|^2 \\ &= |A|^3\left(\frac{1+5\delta}{2}\right) + 96|A|^2.\end{aligned}$$

Combining the two estimates on $\sum f(n)^2$ gives the inequality

$$\frac{\binom{|A|}{2}^2(|A|-2)^2}{N} + \delta^2|A|^3\frac{\left(1-\frac{C}{\delta|A|}\right)^2}{1-\frac{|A|}{N}} \leq |A|^3\left(\frac{1+5\delta}{2}\right) + 96|A|^2. \tag{8.10}$$

If $\delta = 0$, then (8.10) is not valid but we still get

$$\frac{\binom{|A|}{2}^2(|A|-2)^2}{N} \leq \frac{|A|^3}{2} + 96|A|^2.$$

This inequality implies $|A| \leq (1+o(1))(2N)^{1/3}$. Assume that $\delta > 0$. In this case, (8.10) simplifies to

$$|A| \leq (1+o(1))\left(2+10\delta-4\delta^2\right)^{1/3} N^{1/3}. \tag{8.11}$$

At this point we find the maximum of the right hand side of (8.11) using the fact that $0 \leq \delta \leq 1+\frac{7}{|A|}$, which follows from Lemma 8.3.4. For $|A| \geq 28$, the maximum occurs when $\delta = 1+\frac{7}{|A|}$ therefore, after some simplifying, we find

$$|A| \leq (1+o(1))(8N)^{1/3}.$$

8.4 Proof of Theorem 8.1.5(i)

The proof of Theorem 8.1.5(i) follows along the same lines as the proof of Theorem 8.1.3. We will use the same notation as in the previous section. The derivation of (8.7) is very similar except in $\mathbb{Z}$, or in $\mathbb{Z}_N$ with N odd, there are fewer 3-term a.p.'s in A. Regardless, (8.7) still holds under the assumption that $A \subset [N]$ is a B_3^+-set, or $A \subset \mathbb{Z}_N$ is a B_3^+-set and N odd.

Next we prove a lemma that corresponds to Lemma 8.3.2.

Lemma 8.4.1. *If $x \in T_i^1 \cap T_j^1$ for distinct i and j, then either $s_i = s_j = 2$, or if $s_j > 2$, then $s_i = 2$, $s_j = 3$, and $|T_i^1 \cap T_j^1| \geq 3$.*

Proof. The proof of this lemma is exactly the same as the proof of Lemma 8.3.2 up until the point where we write the equation $2z = 2u$. In $\mathbb{Z}$ (or $\mathbb{Z}_N$ with N odd), this implies $z = u$ which is a contradiction since the elements x, b_1^i, y, z, u, v are all distinct. This allows us to conclude that $T_i^1 \cap T_j^1 = \emptyset$ for any $i \neq j$ with $s_i \geq 3$ and $s_j \geq 3$.

The assertion $|T_i^1 \cap T_j^1| \geq 3$ can be verified with some easy computations. Alternatively, one can just ignore $a_3^i = u$ and $b_3^i = v$ in Figure 1 to see $|T_i^1 \cap T_j^1| \geq 3$. □

Corollary 8.4.2. *If $m+1 \leq i < j \leq M$, then $T_i^1 \cap T_j^1 = \emptyset$.*

Proof. If $x \in T_i \cap T_j$ with $i \neq j$, then by Lemma 8.4.1, one of s_i or s_j must be equal to 2. □

The next lemma has no corresponding lemma from the previous section. It will be used to estimate $\sum_{c \in A} f(c)$.

Lemma 8.4.3. *If $A \subset [N]$ is a B_3^+-set or if $A \subset \mathbb{Z}_N$ is a B_3^+-set and N is odd, then for any $a \in A$, the number of distinct $i \in \{1, 2, \ldots, m\}$ such that $a \in T_i^1$ is at most $\frac{|A|}{2}$.*

Proof. To make the notation simpler, we suppose $a \in T_i^1$ for $1 \leq i \leq k$ and we will show $k \leq \frac{|A|}{2}$. The case when $a \in T_{i_1}^1 \cap \cdots \cap T_{i_k}^1$ for some sequence $1 \leq i_1 < \cdots < i_k \leq m$ is the same. For this lemma we deviate from the notation $S_i^2 = \{\{a_1^i, b_1^i\}, \ldots \{a_{s_i}^i, b_{s_i}^i\}\}$. Write $S_i^2 = \{\{a, a_i\}, \{b_i, c_i\}\}$ and $a + a_i = b_i + c_i$ where $1 \leq i \leq k$, and for fixed i, the elements a, a_i, b_i, and c_i are all distinct. Observe $a_1, \ldots, a_k$ are all distinct since the sums $a + a_i$ are all distinct. For $1 \leq i \leq k$, $a = b_i + c_i - a_i$. Therefore,

$$b_i + c_i + a_j = b_j + c_j + a_i$$

for any $1 \leq i, j \leq k$. These two sums must intersect and they cannot intersect at a_j or a_i, unless $i = j$, so for $2 \leq j \leq k$,

$$\{b_1, c_1\} \cap \{b_j, c_j\} \neq \emptyset.$$

Let $2 \leq j \leq l$ be the indices for which the sums intersect at b_1. Let $l+1 \leq j \leq k$ be the indices for which the sums intersect at c_1. Let $b = b_1$ and $c = c_1$. We have

the k equations

$$\begin{aligned} a + a_1 &= b + c, \\ a + a_2 &= b + c_2, \\ \vdots &\quad \vdots \\ a + a_l &= b + c_l, \\ a + a_{l+1} &= b_{l+1} + c, \\ \vdots &\quad \vdots \\ a + a_k &= b_k + c. \end{aligned}$$

We will show that $a_1, \ldots, a_k, c_1, \ldots, c_l, b_{l+1}, \ldots, b_k$ are all distinct which implies $2k \leq |A|$.

Suppose $a_i = b_j$ for some $2 \leq i \leq l$ and $l+1 \leq j \leq k$. Then $a + b_j = a + a_i = b + c_i$, but $a = b_j + c - a_j$. Therefore, $b + c_i = a + b_j = 2b_j + c - a_j$ which implies $2b_j + c = b + c_i + a_j$. The elements a_j, b_j, and c are all distinct so these sums cannot intersect at a_j. Similarly they cannot intersect at c. The only remaining possibility is $b_j = c_i$, but then $a_i = b_j = c_i$, which is a contradiction. We conclude that a_i and b_j are distinct for $2 \leq i \leq l$ and $l+1 \leq j \leq k$. A similar argument shows that a_j and c_i are distinct for $l+1 \leq j \leq k$ and $2 \leq i \leq l$.

Suppose now that $a_i = c_{i'}$ for some $2 \leq i \neq i' \leq l$. Then $b + c_i = a + a_i = a + c_{i'} = a + (a + a_{i'} - b)$, so that $2b + c_i = 2a + a_{i'}$. Since $2 \leq i' \leq l$, these sums cannot intersect at b and they cannot intersect at a. If $c_i = a_{i'}$, then $a = b$ which is impossible. The equation $2b + c_i = 2a + a_{i'}$ contradicts the B_3^+ property. Note that $2b = 2a$ need not imply $a = b$ if $A \subset \mathbb{Z}_N$ with N even. We conclude that $a_i \neq c_{i'}$ for each $2 \leq i \neq i' \leq l$. Similarly, $a_j \neq b_{j'}$ for $l+1 \leq j \neq j' \leq k$.

The previous two paragraphs imply

$$\{a_1, a_2, \ldots, a_k\} \cap \{c_2, c_3, \ldots, c_l, b_{l+1}, b_{l+2}, \ldots, b_k\} = \emptyset.$$

To finish the proof we show $\{c_2, c_3, \ldots, c_l\} \cap \{b_{l+1}, b_{l+2}, \ldots, b_k\} = \emptyset$. Suppose $c_i = b_j$ for some $2 \leq i \leq l$ and $l+1 \leq j \leq k$. Then

$$a + a_i = b + c_i = b + b_j = b + (a + a_j - c) = b + a + a_j - (a + a_1 - b) = a_j + 2b - a_1$$

which implies $a+a_i+a_1 = a_j+2b$. Since $i < l+1 \leq j$, these sums cannot intersect at a_j. They cannot intersect at b either since a, a_i, b, and c_i are all distinct whenever $1 \leq i \leq l$. This is a contradiction. Therefore, $c_i \neq b_j$ for all $2 \leq i \leq l$ and $l+1 \leq j \leq k$. □

Lemma 8.4.4. *If $A \subset [N]$ is a B_3^+-set, then*

$$\sum_{c \in A} f(c) \leq \frac{|A|^2}{2} + 4|A|.$$

Proof. Again we write f as a sum of the simpler functions g_1 and g_2. Recall that for $c \in A$,

$$g_1(c) = \#\left\{(\{x,z\},y) \in A^{(2)} \times A : c = x-y+z, c \neq y, \{x,z\} \cap \{y\} = \emptyset\right\},$$

and

$$g_2(c) = \#\left\{(\{x,z\},y) \in A^{(2)} \times A : c = x-y+z, c = y, \{x,z\} \cap \{y\} = \emptyset\right\}.$$

For each $c \in A$, $f(c) = g_1(c) + g_2(c)$. The sum $\sum_{c \in A} g_2(c)$ is exactly the number of nontrivial 3-term a.p.'s in A. By Lemma 8.3.1, this is at most $4|A|$.

If $c \notin T_1^1 \cup \cdots \cup T_M^1$, then the equation $c+y = x+z$ with c, y, x, and z all distinct has no solutions in A so $g_1(c) = 0$. Assume $c \in T_1^1 \cup \cdots \cup T_M^1$.

Case 1: $c \notin T_1^1 \cup \cdots \cup T_m^1$.

By Corollary 8.4.2, there is a unique j with $c \in T_j^1$ and $m+1 \leq j \leq M$. For such a j we have $|S_j^2| \leq \frac{|A|}{2}$ by Corollary 8.4.2. There is a unique pair in S_j^2 that contains c so y is determined. There are at most $\frac{|A|}{2}$ choices for the pair $\{x,z\} \in S_j^2 \backslash \{c,y\}$ so $g_1(c) \leq \frac{|A|}{2}$.

Case 2: $c \in T_1^1 \cup \cdots \cup T_m^1$.

First assume $c \notin T_{m+1}^1 \cup \cdots \cup T_M^1$. A solution to $c+y = x+z$ with c, y, x, and z all distinct corresponds to a choice of an S_i^2 with $1 \leq i \leq m$ and $c \in T_i^1$. By Lemma 8.4.3, c is in at most $\frac{|A|}{2}$ T_i^1's and so $g_1(c) \leq \frac{|A|}{2}$.

Lastly suppose $c \in T_{m+1}^1 \cup \cdots \cup T_M^1$. There is a unique j with $c \in T_j^1$ and $m+1 \leq j \leq M$. Furthermore, for this j we have $|T_j^1| = 6$ by Lemma 8.4.1. If $c \in T_i^1$ with $1 \leq i \leq m$ then, again by Lemma 8.4.1, $|T_i^1 \cap T_j^1| \geq 3$. There are

$\binom{6}{3}$ 3-subsets of T_j^1 and given such a 3-subset, there are $\binom{3}{1}$ ways to pair up an element in the 3-subset with c in S_i^2. This implies c is in at most $3\binom{6}{3}$ S_i^2's with $1 \le i \le m$, so $g_1(c) \le 2 + 3\binom{6}{3} \le \frac{|A|}{2}$. The 2 comes from choosing one of the two pairs in $S_j^2 \backslash \{c, y\}$. □

The rest of the proof of Theorem 8.1.5(i) is almost identical to that of Theorem 8.1.3. If $\sum_{c \in A} f(c) = \delta |A|^2$, then by (8.7) and (8.9),

$$\sum f(n)^2 \le |A|^3 \left(\frac{1 + 5\delta}{2} \right) + O(|A|^2).$$

We use the same version of the Cauchy-Schwarz inequality to get

$$\frac{\binom{|A|}{2}^2 (|A| - 2)^2}{3N} + \delta^2 |A|^3 \frac{\left(1 - \frac{C}{\delta |A|}\right)}{1 - \frac{|A|}{3N}} \le |A|^3 \left(\frac{1 + 5\delta}{2} \right) + O(|A|^2). \tag{8.12}$$

If $\delta = 0$, then

$$\frac{\binom{|A|}{2}^2 (|A| - 2)^2}{3N} \le \frac{|A|^3}{2} + O(|A|^2)$$

which implies $|A| \le (1 + o(1))(6N)^{1/3}$. Assume $\delta > 0$. Then (8.12) simplifies to

$$|A| \le (1 + o(1))(6 + 30\delta - 12\delta^2)^{1/3} N^{1/3}.$$

By Lemma 8.4.4, $0 \le \delta \le \frac{1}{2} + \frac{3}{|A|}$. The maximum occurs when $\delta = \frac{1}{2} + \frac{3}{|A|}$ and we get

$$|A| \le (1 + o(1))(18N)^{1/3}.$$

If we were working in $\mathbb{Z}_N$ with N odd, then in (8.12) the $3N$ can be replaced by N. Some simple calculations show that we get Theorem 8.1.3 in the odd case. We actually obtain the upper bound $|A| \le (1 + o(1))(6N)^{1/3}$ when $A \subset \mathbb{Z}_N$ is a B_3^+-set and N is odd.

8.5 Proof of Theorem 8.1.5(ii)

Let $A \subset [N]$ be a B_4^+-set. For $n \in [-2N, 2N]$, define

$$f(n) = \#\{(\{a_1, a_2\}, \{b_1, b_2\}) \in A^{(2)} \times A^{(2)} : a_1 + a_2 - b_1 - b_2 = n, \{a_1, a_2\} \cap \{b_1, b_2\} = \emptyset\}.$$

Recall that $A^{(2)} = \{\{x, y\} : x, y \in A, x \neq y\}$.

Lemma 8.5.1. *If $A \subset [N]$ is a B_4^+-set, then A is a B_2-set.*

Proof. Suppose $a + b = c + d$ with $a, b, c, d \in A$. If $\{a, b\} \cap \{c, d\} = \emptyset$, then the equation $2(a + b) = 2(c + d)$ contradicts the B_4^+ property so $\{a, b\} \cap \{c, d\} \neq \emptyset$. Since $a + b = c + d$ and $\{a, b\} \cap \{c, d\} \neq \emptyset$, we have $\{a, b\} = \{c, d\}$. □

Lemma 8.5.2. *If $A \subset [N]$ is a B_4^+-set, then for any integer n, $f(n) \leq 2|A|$.*

Proof. Suppose $f(n) \geq 1$. Fix a tuple $(\{a_1, a_2\}, \{b_1, b_2\})$ counted by $f(n)$. Let $(\{c_1, c_2\}, \{d_1, d_2\})$ be another tuple counted by $f(n)$, not necessarily different from $(\{a_1, a_2\}, \{b_1, b_2\})$. Then $a_1 + a_2 - b_1 - b_2 = c_1 + c_2 - d_1 - d_2$ so

$$a_1 + a_2 + d_1 + d_2 = c_1 + c_2 + b_1 + b_2. \tag{8.13}$$

By the B_4^+ property, $\{a_1, a_2, d_1, d_2\} \cap \{c_1, c_2, b_1, b_2\} \neq \emptyset$. In order for this intersection to be non-empty, it must be the case that $\{a_1, a_2\} \cap \{c_1, c_2\} \neq \emptyset$ or $\{b_1, b_2\} \cap \{d_1, d_2\} \neq \emptyset$.

Case 1: $\{a_1, a_2\} \cap \{c_1, c_2\} \neq \emptyset$.

Assume $a_1 = c_1$. There are at most $|A|$ choices for c_2 so we fix one. The equality $a_1 = c_1$ and (8.13) imply

$$d_1 + d_2 = b_1 + b_2 + c_2 - a_2. \tag{8.14}$$

The right hand side of (8.14) is determined. By Lemma 8.5.1, there is at most one pair $\{d_1, d_2\}$ such that (8.14) holds.

Case 2: $\{a_1, a_2\} \cap \{c_1, c_2\} = \emptyset$ and $\{b_1, b_2\} \cap \{d_1, d_2\} \neq \emptyset$.

Again there is no loss in assuming $b_1 = d_1$. There are at most $|A|$ choices for d_2 so fix one. The equality $b_1 = d_1$ and (8.13) imply

$$c_1 + c_2 = a_1 + a_2 - b_2 + d_2. \tag{8.15}$$

The right hand side of (8.15) is determined and there is at most one pair $\{c_1, c_2\}$ satisfying (8.15) as before.

Putting the two possibilities together we get at most $2|A|$ solutions $(\{c_1, c_2\}, \{d_1, d_2\})$. We have also accounted for the solution $(\{a_1, a_2\}, \{b_1, b_2\})$ in our count so $f(n) \leq 2|A|$. □

Lemma 8.5.3. *If $A \subset [N]$ is a B_4^+-set, then*

$$\sum f(n)(f(n)-1) \leq 2|A| \sum_{n \in A-A} f(n). \tag{8.16}$$

Proof. The left hand side of (8.16) counts the number of ordered tuples

$$((\{a_1, a_2\}, \{b_1, b_2\}), (\{c_1, c_2\}, \{d_1, d_2\}))$$

such that $(\{a_1, a_2\}, \{b_1, b_2\}) \neq (\{c_1, c_2\}, \{d_1, d_2\})$, and both tuples are counted by $f(n)$. Equation (8.13) holds for these tuples. As before we consider two cases.

Case 1: $\{a_1, a_2\} \cap \{c_1, c_2\} \neq \emptyset$.

Assume $a_1 = c_1$ so that $a_2 - c_2 = b_1 + b_2 - d_1 - d_2$.

If $\{b_1, b_2\} \cap \{d_1, d_2\} \neq \emptyset$, say $b_1 = d_1$, then $a_2 - c_2 = b_2 - d_2$. We can rewrite this equation as $a_2 + d_2 = b_2 + c_2$ so that $\{a_2, d_2\} = \{b_2, c_2\}$. Since $\{a_1, a_2\} \cap \{b_1, b_2\} = \emptyset$, it must be the case that $a_2 = c_2$ and $d_2 = b_2$. This contradicts the fact that the tuples are distinct. We conclude $\{b_1, b_2\} \cap \{d_1, d_2\} = \emptyset$.

There are $|A|$ choices for the element $a_1 = c_1$ and we fix one. Since $a_2 - c_2 = b_1 + b_2 - d_1 - d_2$ and $\{b_1, b_2\} \cap \{d_1, d_2\} = \emptyset$, there are $f(a_2 - c_2)$ ways to choose $\{b_1, b_2\}$ and $\{d_1, d_2\}$. Also observe that each $n \in A - A$ with $n \neq 0$ has a unique representation as $n = a_2 - c_2$ with $a_2, c_2 \in A$. This follows from the fact that A is a B_2-set.

Case 2: $\{a_1, a_2\} \cap \{c_1, c_2\} = \emptyset$ and $\{b_1, b_2\} \cap \{d_1, d_2\} \neq \emptyset$.

The argument in this case is essentially the same as that of Case 1.

Putting the two cases together proves the lemma.

□

Observe $\sum f(n) = \binom{|A|}{2}\binom{|A|-2}{2}$. Using Cauchy-Schwarz, and Lemmas 8.5.3 and 8.5.2,

$$\begin{aligned}
\frac{\left(\binom{|A|}{2}\binom{|A|-2}{2}\right)^2}{4N} &\leq \sum f(n)^2 \leq \binom{|A|}{2}\binom{|A|-2}{2} + 2|A| \sum_{n \in A-A} f(n) \\
&\leq \frac{|A|^4}{4} + 2|A||A-A| \cdot 2|A| \\
&\leq \frac{|A|^4}{4} + 4|A|^4 = \frac{17|A|^4}{4}.
\end{aligned}$$

After rearranging we get

$$|A| \leq (1+o(1))(16 \cdot 17N)^{1/4} = (1+o(1))(272N)^{1/4}.$$

8.6 Proof of Theorem 8.1.6

Lemma 8.6.1. *Let A be a B_k^+-set with $k \geq 4$. If $k = 2l$, then there is a subset $A' \subset A$ such that A' is a B_l^+-set and $|A'| \geq |A| - 2l$. If $k = 2l+1$, then there is a subset $A' \subset A$ such that $|A'| \geq |A| - 2k$ and A' is either a B_l^+-set or a B_{l+1}^+-set.*

Proof. Suppose $k = 2l$ with $l \geq 2$. If A is not a B_l^+-set, then there is a set of $2l$ (not necessarily distinct) elements $a_1, \ldots, a_{2l} \in A$, such that

$$a_1 + \cdots + a_l = a_{l+1} + \cdots + a_{2l}$$

and $\{a_1, \ldots, a_l\} \cap \{a_{l+1}, \ldots, a_{2l}\} = \emptyset$. Let $A' = A \backslash \{a_1, a_2, \ldots, a_{2l}\}$. If A' is not a B_l^+-set, then there is another set of $2l$ elements of A', say $b_1, \ldots, b_{2l}$, such that

$$b_1 + \cdots + b_l = b_{l+1} + \cdots + b_{2l}$$

and $\{b_1, \ldots, b_l\} \cap \{b_{l+1}, \ldots, b_{2l}\} = \emptyset$. Adding these two equations together gives

$$a_1 + \cdots + a_l + b_1 + \cdots + b_l = a_{l+1} + \cdots + a_{2l} + b_{l+1} + \cdots + b_{2l}$$

with $\{a_1, \ldots, a_l, b_1, \ldots, b_l\} \cap \{a_{l+1}, \ldots, a_{2l}, b_{l+1}, \ldots, b_{2l}\} = \emptyset$. This is a contradiction.

The case when $k = 2l+1 \geq 5$ can be handled in a similar way.

□

It is easy to modify the proof of Lemma 8.6.1 to obtain a version for B_k^*-sets.

Lemma 8.6.2. *Let A be a B_k^*-set with $k \geq 4$. If $k = 2l$, then there is a subset $A' \subset A$ such that A' is a B_l^*-set and $|A'| \geq |A| - 2l$. If $k = 2l+1$, then there is a subset $A' \subset A$ such that $|A'| \geq |A| - 2k$ and A' is either a B_l^*-set or a B_{l+1}^*-set.*

For $A \subset [N]$ and $j \geq 2$, let

$$\sigma_j(n) = \#\left\{(a_1, \ldots, a_j) \in A^j : a_1 + \cdots + a_j = n\right\}.$$

Let $e(x) = e^{2\pi ix}$ and $f(t) = \sum_{a\in A} e(at)$. For any $j \geq 1$, $f(t)^j = \sum \sigma_j(n)e(nt)$ so by Parseval's Identity, $\sum \sigma_j(n)^2 = \int_0^1 |f(t)|^{2j}dt$. The next lemma is (5.9) of [71].

Lemma 8.6.3. *If $A \subset [N]$ is a B_k^*-set, then*

$$\sum \sigma_k(n)^2 \leq (1+o(1))k^2|A| \sum \sigma_{k-1}(n)^2. \tag{8.17}$$

In [71], Ruzsa estimates the right hand side of (8.17) using Hölder's Inequality and shows

$$\sum \sigma_{k-1}(n)^2 \leq \left(\sum \sigma_k(n)\right)^{\frac{k-2}{k-1}} |A|^{\frac{1}{k-1}}.$$

Our next lemma uses Hölder's Inequality in a different way.

Lemma 8.6.4. *Let $A \subset [N]$ be a B_k^*-set. If $k \geq 4$ is even, then*

$$\sum \sigma_k(n)^2 \leq (1+o(1))k^k|A|^{k/2} \sum \sigma_{k/2}(n)^2.$$

If $k = 2l+1 \geq 5$, then

$$\sum \sigma_k(n)^2 \leq (1+o(1)) \max\left\{k^{k+1}|A|^{l+1} \sum \sigma_l(n)^2, k^{k-1}|A|^l \sum \sigma_{l+1}(n)^2\right\}.$$

Proof. First assume that $k = 2l \geq 4$. By Lemma 8.6.2, we may assume that A is a B_l^*-set. Otherwise, we pass to a subset of A that is a B_l^* set and has at least $|A| - 2k$ elements. Applying Hölder's Inequality with $p = \frac{k}{k-2}$ and $q = \frac{k}{2}$, we get

$$\begin{aligned}
\sum \sigma_{k-1}(n)^2 &= \int_0^1 |f(t)|^{2(k-1)}dt = \int_0^1 |f(t)|^{\frac{2k}{p}} |f(t)|^{\frac{2l}{q}} dt \\
&\leq \left(\int_0^1 |f(t)|^{2k}dt\right)^{1/p} \left(\int_0^1 |f(t)|^{2l}dt\right)^{1/q} \\
&= \left(\sum \sigma_k(n)^2\right)^{(k-2)/k} \left(\sum \sigma_l(n)^2\right)^{2/k}.
\end{aligned}$$

Substituting this estimate into (8.17) and solving for $\sum \sigma_k(n)^2$ gives the first part of the lemma.

Now assume $k = 2l+1 \geq 5$. Again by Lemma 8.6.2, we can assume that A is either a B_l^*-set or a B_{l+1}^*-set.

Suppose A is a B_l^*-set. Applying Hölder's Inequality with $p = \frac{k+1}{k-1}$ and $q = \frac{k+1}{2}$, we get

$$\sum \sigma_{k-1}(n)^2 \leq \left(\sum \sigma_k(n)^2\right)^{\frac{k-1}{k+1}} \left(\sum \sigma_l(n)^2\right)^{\frac{2}{k+1}}.$$

This inequality and (8.17) imply

$$\sum \sigma_k(n)^2 \leq (1+o(1))k^{k+1}|A|^{\frac{k+1}{2}} \sum \sigma_l(n)^2.$$

If A is a B_{l+1}^*-set instead, then apply Hölder's Inequality with $p = \frac{l}{l-1}$ and $q = \frac{1}{l}$ and proceed as above. It is in this step that we must assume $k = 2l+1 \geq 5$ otherwise if $k = 3$, then $l = 1$ and p is not defined. □

For $k \geq 2$ let c_k^+ be the smallest constant such that for any B_k^+-set A,

$$\sum \sigma_k(n)^2 \leq (1+o(1))c_k^+|A|^k.$$

Define c_k^* similarly. The techniques of [71] can be used to show that $c_k^* \leq k^{2k}$ so c_k^+ and c_k^* are well defined. Observe that for any $k \geq 2$, $c_k^+ \leq c_k^*$. Using Lemma 8.6.4, it is not difficult to show that for even $k \geq 4$,

$$c_k^+ \leq k^k c_{k/2}^+ \text{ and } c_k^* \leq k^k c_{k/2}^*, \tag{8.18}$$

Similarly, one can show that for odd $k = 2l+1 \geq 5$,

$$c_k^+ \leq \max\left\{k^{k+1}c_l^+, k^{k-1}c_{l+1}^+\right\} \text{ and } c_k^* \leq \max\left\{k^{k+1}c_l^*, k^{k-1}c_{l+1}^*\right\}. \tag{8.19}$$

Lemma 8.6.5. *Let $A \subset [N]$ be a B_k^+-set. If $k \geq 4$ is even, then*

$$|A| \leq (1+o(1))\left(k^{k+1}c_{k/2}^+ N\right)^{1/k}. \tag{8.20}$$

If $k = 2l+1 \geq 5$, then

$$|A| \leq (1+o(1))\left(k^k \cdot \max\{k^2 c_l^+, c_{l+1}^+\}N\right)^{1/k}. \tag{8.21}$$

The same inequalities hold under the assumption that $A \subset [N]$ is a B_k^*-set provided that the c_k^+'s are replaced with c_k^*'s.

Proof. By Cauchy-Schwarz,

$$\frac{|A|^{2k}}{kN} \le \sum \sigma_k(n)^2 \tag{8.22}$$

for any $k \ge 2$.

First suppose $k \ge 4$ is even. By (8.22) and Lemma 8.6.4,

$$\frac{|A|^{2k}}{kN} \le \sum \sigma_k(n)^2 \le (1+o(1))k^k|A|^{k/2}\sum \sigma_{k/2}(n)^2 \le (1+o(1))k^k c_{k/2}^+|A|^k.$$

Solving this inequality for $|A|$ proves (8.20).

Now suppose $k = 2l+1 \ge 5$. By (8.22) and Lemma 8.6.4,

$$\begin{aligned}\frac{|A|^{2k}}{kN} &\le \sum \sigma_k(n)^2 \le (1+o(1))\max\left\{k^{k+1}c_l^+|A|^k, k^{k-1}c_{l+1}^+|A|^k\right\} \\ &= (1+o(1))|A|^k k^{k-1}\max\{k^2c_l^+, c_{l+1}^+\}.\end{aligned}$$

□

Lemma 8.6.5 shows that we can obtain upper bounds on B_k^+-sets and B_k^*-sets recursively. To start the recursion we need estimates on c_2^+, c_2^*, c_3^+, and c_3^*.

Lemma 8.6.6. *If A is a B_2^*-set, then*

$$\sum \sigma_2(n)^2 \le 2|A|^2 + 32|A|$$

and therefore $c_2^* \le 2$.

Proof. Let $\delta(n) = \#\{(a_1,a_2) \in A^2 : a_1 - a_2 = n\}$. Observe $\sum \sigma_2(n)^2 = \sum \delta(n)^2$. In [71] (see Theorem 4.7) it is shown that $\delta(n) \le 1$ for any $n \ne 0$, and $\delta(n) = 2$ for at most $8|A|$ integers n. We conclude

$$\sum \delta(n)^2 \le \delta(0)^2 + 8|A|\cdot 4 + |A-A| \le 2|A|^2 + 32|A|.$$

□

Lemma 8.6.7. *If $A \subset [N]$ is a B_3^+-set, then*

$$\sum \sigma_3(n)^2 \le (1+o(1))18|A|^3$$

and therefore $c_3^+ \le 18$.

Proof. Let $A \subset [N]$ be a B_3^+-set and let

$$r_2(n) = \#\left\{\{a,b\} \in A^{(2)} : a+b=n\right\}.$$

Define $2 \cdot A := \{2a : a \in A\}$. For $n \in 2 \cdot A$, $\sigma_2(n) = 2r_2(n)+1$ and $\sigma_2(n) = 2r_2(n)$ otherwise. The sum $\sum_{n \in 2\cdot A} r_2(n)$ counts the number of 3-term a.p.'s in A so by Lemma 8.3.1,

$$\begin{aligned}\sum \sigma_2(n)^2 &= 4\sum r_2(n)^2 + 4\sum_{n\in 2\cdot A} r_2(n) + |2\cdot A| \\ &\leq 4\sum r_2(n)^2 + 4\cdot 4|A| + |A| = 4\sum r_2(n)^2 + 17|A|.\end{aligned}$$

Using the notation and results of Section 3, and the inequality $x^2 \leq 2x(x-1)$ for $x \geq 2$, we have

$$\sum r_2(n)^2 = \sum_{i=1}^{M} |S_i^2|^2 \leq 2\sum_{i=1}^{M} |S_i^2|(|S_i^2|-1) \leq \sum_{c\in A} f(c) \leq \frac{|A|^2}{2} + 3|A|.$$

Combining this inequality with (8.17) gives

$$\begin{aligned}\sum \sigma_3(n)^2 &\leq (1+o(1))3^2|A|\sum \sigma_2(n)^2 \leq (1+o(1))9|A|(4\sum r_2(n)^2 + 17|A|) \\ &\leq (1+o(1))9|A|(2|A|^2 + 29|A|) \leq (1+o(1))(18|A|^3 + 261|A|^2).\end{aligned}$$

□

Lemma 8.6.8. *If $A \subset [N]$ is a B_3^*-set, then*

$$\sum \sigma_3(n)^2 \leq (1+o(1))54|A|^3$$

and therefore $c_3^* \leq 54$.

Proof. Let $A \subset [N]$ be a B_3^*-set. The idea of the proof is motivated by the same arguments that we used for B_3^+-sets. For $d \in A+A$, let

$$P^2(d) = \{\{a,b\} \in A^{(2)} : a+b=d\}.$$

Define $m_0 = 0$ and for $1 \leq j \leq 4$, let $d_{m_{j-1}+1}, d_{m_{j-1}+2}, \ldots, d_{m_j}$ be the integers for which $|P^2(d_i)| = j$. Let $d_{m_4+1}, d_{m_4+2}, \ldots, d_M$ be the integers for which $|P^2(d_i)| \geq 5$. Write P_i^2 for $P^2(d_i)$, p_i for $|P_i^2|$, and for $1 \leq i \leq M$, let

$$Q_i^1 = \{a : a \in \{a,b\} \text{ for some } \{a,b\} \in P_i^2\}.$$

We will use the notation $P_i^2 = \{\{a_1^i, b_1^i\}, \ldots, \{a_{p_i}^i, b_{p_i}^i\}\}$. A difference between the P_i^2's of this section and the S_i^2's of earlier sections is that we allow for a P_i^2 to contain only one pair.

Lemma 8.6.9. *If $x \in Q_i^1 \cap Q_j^1$ for some $i \neq j$ where $p_i \geq 3$ and $p_j \geq 3$, then $p_i + p_j \leq 7$.*

Proof. Without loss of generality, assume $x = a_1^i$ and $x = a_1^j$ where

$$P_i^2 = \{\{a_1^i, b_1^i\}, \{a_2^i, b_2^i\}, \ldots, \{a_{p_i}^i, b_{p_i}^i\}\} \text{ and } P_j^2 = \{\{a_1^j, b_1^j\}, \{a_2^j, b_2^j\}, \ldots, \{a_{p_j}^j, b_{p_j}^j\}\}.$$

For $2 \leq l \leq p_i$ we have $d_i = x + b_1^i = a_l^i + b_l^i$. Similarly, for $2 \leq k \leq p_j$ we have $d_j = x + b_1^j = a_k^j + b_k^j$. Then $a_l^i + b_l^i - b_1^i = x = a_k^j + b_k^j - b_1^j$, so

$$a_l^i + b_l^i + b_1^j = a_k^j + b_k^j + b_1^i \text{ for any } 2 \leq l \leq p_i \text{ and } 2 \leq k \leq p_j. \tag{8.23}$$

If $b_1^j \in T_i^1$, then there is no loss in assuming $b_1^j \in \{a_2^i, b_2^i\}$. The same assumption may be made with i and j interchanged. This means that for $l \geq 3$, b_1^j is not a term in the sum $a_l^i + b_l^i$ and for $k \geq 3$, b_1^i is not a term in the sum $a_k^j + b_k^j$. The B_3^* property and (8.23) imply

$$|\{a_l^i, b_l^i\} \cap \{a_k^j, b_k^j\}| = 1 \text{ for any } 3 \leq l \leq p_i \text{ and } 3 \leq k \leq p_j. \tag{8.24}$$

In particular, $\{a_3^i, b_3^i\} \cap \{a_3^j, b_3^j\} \neq \emptyset$ and $\{a_3^i, b_3^i\} \cap \{a_4^j, b_4^j\} \neq \emptyset$ so that $p_j \leq 4$. Here we are using the fact that any element of A can occur at most once in the list $a_1^i, b_1^i, \ldots, a_{p_i}^i, b_{p_i}^i$. By symmetry, $p_i \leq 4$.

If $p_i = p_j = 4$, then by (8.24), $\{a_3^i, b_3^i, a_4^i, b_4^i\} = \{a_3^j, b_3^j, a_4^j, b_4^j\}$ but then $2d_i = a_3^i + b_3^i + a_4^i + b_4^i = 2d_j$ implying $d_i = d_j$, a contradiction. □

Corollary 8.6.10. *If $p_i \geq 4$ and $p_j \geq 4$ with $i \neq j$, then $Q_i^1 \cap Q_j^1 = \emptyset$.*

Using the definition of the P_i^2's, we can write

$$\sum r_2(n)^2 = \sum_{i=1}^{M} |P_i^2|^2 = m_1 + 4(m_2 - m_1) + 9(m_3 - m_2) + 16(m_4 - m_3) + \sum_{i=m_4+1}^{M} |P_i^2|^2.$$

If $p_i = p_j = 4$ for some $i \neq j$, then $Q_i^1 \cap Q_j^1 = \emptyset$ by Corollary 8.6.10 so $m_4 - m_3 \leq \frac{|A|}{8}$. For $1 \leq i \leq 3$, let $\delta_i |A|^2 = m_i - m_{i-1}$. Then

$$\sum r_2(n)^2 \leq |A|^2(\delta_1 + 4\delta_2 + 9\delta_3) + \sum_{i=m_4+1}^{M} |P_i^2|^2 + 2|A|. \tag{8.25}$$

Define a graph H with vertex set $Q_{m_2+1}^1 \cup \cdots \cup Q_{m_3}^1$ and edge set $P_{m_2+1}^2 \cup \cdots \cup P_{m_3}^2$. Let $n = |V(H)|$. The graph H has $3(m_3 - m_2) = 3\delta_3|A|^2$ edges so $3\delta_3|A|^2 \leq \frac{n}{2}$ which can be rewritten as

$$\sqrt{6\delta_3}|A| \leq |Q_{m_2+1}^1 \cup \cdots \cup Q_{m_3}^1|. \tag{8.26}$$

For any i and j with $m_2 + 1 \leq i \leq m_3$ and $m_4 + 1 \leq j \leq M$, $Q_i^1 \cap Q_j^1 = \emptyset$ by Lemma 8.6.9. Thus (8.26) implies

$$\sum_{i=m_4+1}^{M} |P_i^2| = \frac{1}{2} \sum_{i=m_4+1}^{M} |Q_i^1| = \frac{1}{2}|Q_{m_4+1}^1 \cup \cdots \cup Q_M^1| \leq \frac{1}{2}(1 - \sqrt{6\delta_3})|A|.$$

We conclude $\sum_{i=m_4+1}^{M} |P_i^2|^2 \leq \left(\frac{1-\sqrt{6\delta_3}}{2}\right)^2 |A|^2$. This estimate and (8.25) give

$$\sum r_2(n)^2 \leq |A|^2 \left(\delta_1 + 4\delta_2 + 9\delta_3 + \frac{1}{4}(1 - \sqrt{6\delta_3})^2\right) + 2|A|. \tag{8.27}$$

Each pair $\{a, b\} \in A^{(2)}$ is in at most one P_i^2 so

$$|A|^2(\delta_1 + 2\delta_2 + 3\delta_3) = m_1 + 2(m_2 - m_1) + 3(m_3 - m_2) \leq \binom{|A|}{2} \leq \frac{|A|^2}{2}.$$

The maximum of $\delta_1 + 4\delta_2 + 9\delta_3 + \frac{1}{4}(1 - \sqrt{6\delta_3})^2$ subject to the conditions $\delta_1 + 2\delta_2 + 3\delta_3 \leq \frac{1}{2}$, $\delta_1 \geq 0$, $\delta_2 \geq 0$, and $\delta_3 \geq 0$ is $\frac{3}{2}$, achieved when $\delta_1 = \delta_2 = 0$ and $\delta_3 = \frac{1}{6}$. By (8.27),

$$\sum r_2(n)^2 \leq \frac{3|A|^2}{2} + 2|A|. \tag{8.28}$$

An immediate consequence is that

$$\sum_{n \in 2 \cdot A} r_2(n) = \sum 1_{2 \cdot A}(n) r_2(n) \leq |A|^{1/2} \left(\sum r_2(n)^2\right)^{1/2} \leq 2|A|^{3/2}. \tag{8.29}$$

Next we proceed as in Lemma 8.6.7. Using (8.29) and (8.28),

$$\begin{aligned} \sum \sigma_2(n)^2 &= 4 \sum r_2(n)^2 + 4 \sum_{n \in 2 \cdot A} r_2(n) + |2 \cdot A| \\ &\leq 6|A|^2 + 8|A|^{3/2} + 9|A|. \end{aligned}$$

By Lemma 8.6.3,

$$\sum \sigma_3(n)^2 \leq (1+o(1))3^2|A| \sum \sigma_2(n)^2.$$

The previous two estimates show that $\sum \sigma_3(n)^2 \leq (1+o(1))54|A|^3$. This completes the proof of Lemma 8.6.8. □

Corollary 8.6.11. *If $A \subset [N]$ is a B_3^*-set, then*

$$|A| \leq (1+o(1))(162N)^{1/3}.$$

Proof. Let $A \subset [N]$ be a B_3^*-set. By Cauchy-Schwarz and Lemma 8.6.8,

$$\frac{|A|^6}{3N} \leq \sum \sigma_3(n)^2 \leq (1+o(1))54|A|^3.$$

□

So far we have shown $c_2^+ \leq c_2^* \leq 2$, $c_3^+ \leq 18$, and $c_3^* \leq 54$. Now we describe our method for obtaining upper bounds on $F_k^+(N)$ and $F_k^*(N)$. Assume we have upper bounds on $c_2^+, c_3^+, \ldots, c_{k-1}^+$. Lemma 8.6.5 gives an upper bound on $|A|$ in terms of $c_{k/2}^+$ when k is even, and in terms of c_l^+ and c_{l+1}^+ when $k = 2l+1 \geq 5$. An upper bound on c_k^+ is obtained from (8.18) and (8.19). We can also apply this method to B_k^*-sets. The upper bounds we obtain are given in Table 1 below. They have been rounded up to the nearest tenth. They hold for large enough N without error terms.

Table 8.1: Upper Bounds on B_k^+-sets and B_k^*-sets

k	U.b. of [71] on F_k^*	Our U.b. on F_k^*	Our U.b. on F_k^+
3	$6.3N^{1/3}$	$5.5N^{1/3}$	$2.7N^{1/3}$
4	$11.4N^{1/4}$	$6.8N^{1/4}$	$4.1N^{1/4}$
5	$18.2N^{1/5}$	$11.2N^{1/5}$	$11N^{1/5}$
6	$26.8N^{1/6}$	$15.8N^{1/6}$	$13.1N^{1/6}$
7	$37.2N^{1/7}$	$21.6N^{1/7}$	$18.5N^{1/7}$
8	$49.4N^{1/8}$	$22.7N^{1/8}$	$22.7N^{1/8}$

We conclude this section with our proof of the second statement of Theorem 8.1.6. Recall that (8.18) states $c_k^* \leq k^k c_{k/2}^*$ for any even $k \geq 4$. For $k = 2l+1 \geq 5$, (8.19) gives $c_k^* \leq k^{k+1} \max\{c_l^*, c_{l+1}^*\}$. For $x \geq 0$, let $\lceil x \rceil$ be the smallest integer greater than or equal to x. Let $\lfloor x \rfloor$ be the greatest integer less than or equal to x. For $k \geq 0$, define $\phi_1(k) = \lceil \frac{k}{2} \rceil$ and $\phi_i(k) := \phi_1(\phi_{i-1}(k))$ for $i \geq 2$. A simple induction argument can be used to show that for all $i \geq 1$, $\phi_i(k) \leq k2^{-i} + \sum_{t=0}^{i-1} 2^{-t}$. The conclusion is that for every $i \geq 1$, $\phi_i(k) \leq k2^{-i} + 2$. For any $k \geq 5$,

$$c_k^* \leq k^{k+1} \prod_{i=1}^{\lfloor \log_2 k \rfloor} \phi_i(k)^{\phi_i(k)+1} \leq k^{k+1} \prod_{i=1}^{\lfloor \log_2 k \rfloor} \left(k2^{-i} + 2\right)^{k2^{-i}+3}.$$

Taking k-th roots,

$$\begin{aligned}
(c_k^*)^{1/k} &\leq k^{1+1/k} \prod_{i=1}^{\lfloor \log_2 k \rfloor} (k2^{-i} + 2)^{2^{-i}+3/k} \\
&\leq k^{1+1/k} \left(\frac{k}{2} + 2\right)^{\frac{3\log_2 k}{k}} \prod_{i=1}^{\lfloor \log_2 k \rfloor} (k2^{-i} + 2)^{2^{-i}} \\
&\leq k^{1+1/k} k^{\frac{3\log_2 k}{k}} k^{\sum_{i=1}^{\lfloor \log_2 k \rfloor} 2^{-i}} \prod_{i=1}^{\lfloor \log_2 k \rfloor} \left(2^{-i} + \frac{2}{k}\right)^{2^{-i}} \\
&\leq k^2 k^{\frac{4\log_2 k}{k}} \prod_{i=1}^{\lfloor \log_2 k \rfloor} \left(2^{-i} + \frac{2}{k}\right)^{2^{-i}}.
\end{aligned}$$

We claim the sequence $(c_k^*)^{1/k}$ is bounded above by a function $F(k)$ that tends to $\frac{k^2}{4}$ as $k \to \infty$. With this in mind, we rewrite the previous inequality as

$$\frac{4(c_k^*)^{1/k}}{k^2} \leq 4k^{\frac{4\log_2 k}{k}} \prod_{i=1}^{\lfloor \log_2 k \rfloor} \left(2^{-i} + \frac{2}{k}\right)^{2^{-i}}. \tag{8.30}$$

It is easy to check $k^{\frac{4\log_2 k}{k}} \to 1$ as $k \to \infty$. Using $\sum_{n=0}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}$ from elementary calculus, we obtain

$$\prod_{i=1}^{\lfloor \log_2 k \rfloor} (2^{-i})^{2^{-i}} = \left(\frac{1}{2}\right)^{\sum_{i=1}^{\lfloor \log_2 k \rfloor} i2^{-i}} \to \frac{1}{4}$$

as $k \to \infty$. Using the inequality $1 + x \le e^x$ for $x \ge 0$, we have

$$\begin{aligned} 1 &\le \frac{\prod_{i=1}^{\lfloor \log_2 k \rfloor} (2^{-i} + 2/k)^{2^{-i}}}{\prod_{i=1}^{\lfloor \log_2 k \rfloor} (2^{-i})^{2^{-i}}} = \prod_{i=1}^{\lfloor \log_2 k \rfloor} \left(1 + \frac{2^{i+1}}{k}\right)^{2^{-i}} \\ &\le \prod_{i=1}^{\lfloor \log_2 k \rfloor} e^{2^{i+1}/k} \le e^{\frac{1}{k}\sum_{i=2}^{\lfloor \log_2 k \rfloor} 2^i} \le e^{1/k}. \end{aligned}$$

As $k \to \infty$, $e^{1/k} \to 1$ so

$$\prod_{i=1}^{\lfloor \log_2 k \rfloor} \left(2^{-i} + \frac{2}{k}\right)^{2^{-i}} \to \frac{1}{4}.$$

This shows that the right hand side of (8.30) tends to 1 as $k \to \infty$ which proves the claim.

Given $\epsilon > 0$, we can choose k large enough so that $k^{1/k}(c_k^*)^{1/k} \le (1 + \epsilon)\frac{k^2}{4}$. The theorem now follows from the definition of c_k^* and the estimate $\frac{|A|^{2k}}{kN} \le \sum \sigma_k(n)^2$.

8.7 Proof of Theorem 8.1.8

Lemma 8.7.1. *If $A \subset G$ is a non-abelian B_k-set and $B \subset H$ is a non-abelian B_k^+-set, then $A \times B$ is a non-abelian B_k^+-set in $G \times H$.*

Proof. Suppose $a_1, \dots, a_k, a_1', \dots, a_k' \in A$, $b_1, \dots, b_k, b_1', \dots, b_k' \in B$ and

$$(a_1, b_1) \cdots (a_k, b_k) = (a_1', b_1') \cdots (a_k', b_k').$$

Then $a_1 \cdots a_k = a_1' \cdots a_k'$ and $b_1 \cdots b_k = b_1' \cdots b_k'$ so that $a_i = a_i'$ for every i and $b_j = b_j'$ for some j. Thus, $(a_j, b_j) = (a_j', b_j')$. □

Let $\mathbb{F}_4 = \{0, 1, a, b\}$ be the finite field with four elements. Let

$$H = \left\{ \begin{pmatrix} x & y \\ 0 & x^{-1} \end{pmatrix} : x \in \mathbb{F}_4^*, y \in \mathbb{F}_4 \right\}.$$

H is a group under matrix multiplication and $|H| = 12$. Let

$$\alpha = \begin{pmatrix} a & 1 \\ 0 & b \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} a & a \\ 0 & b \end{pmatrix}.$$

Simple computations show that α and β satisfy $\alpha^3 = \beta^3 = \mathrm{id}$ and $\alpha^2\beta = \beta^2\alpha$.

Lemma 8.7.2. *The set $\{\alpha, \beta\}$ is a B_4^+-set in H.*

Proof. Suppose there is a solution to the equation $x_1x_2x_3x_4 = y_1y_2y_3y_4$ with $x_i \neq y_i$ for $1 \leq i \leq 4$, and $x_i, y_j \in \{\alpha, \beta\}$ for all i, j. Without loss of generality, assume $x_1 = \alpha$ and $y_1 = \beta$. There are eight cases which we can deal with using the relations $\alpha^3 = \beta^3 = \mathrm{id}$ and $\alpha^2\beta = \beta^2\alpha$. Instead of considering each individually, we handle several cases at the same time.

Case 1: $\alpha^4 = \beta^4$ or $\alpha^3\beta = \beta^3\alpha$ or $\alpha\beta^3 = \beta\alpha^3$.

If any of these equations hold, then the relation $\alpha^3 = \beta^3 = \mathrm{id}$ implies $\alpha = \beta$, a contradiction.

Case 2: $\alpha^2\beta\alpha = \beta^2\alpha\beta$ or $\alpha^2\beta^2 = \beta^2\alpha^2$.

If either of these equations hold, then the relation $\alpha^2\beta = \beta^2\alpha$ implies $\alpha = \beta$.

Case 3: $\alpha\beta\alpha^2 = \beta\alpha\beta^2$.

Multiplying the equation on the right by β and using $\beta^3 = \mathrm{id}$, we get $\alpha\beta\alpha^2\beta = \beta\alpha$. On the other hand, $\alpha\beta\alpha^2\beta = \alpha\beta^3\alpha = \alpha^2$ so combining the two equations we get $\beta\alpha = \alpha^2$. This implies $\alpha = \beta$, a contradiction.

Case 4: $\alpha\beta\alpha\beta = \beta\alpha\beta\alpha$.

Multiply the equation on the left by β^2 to get $\beta^2\alpha\beta\alpha\beta = \alpha\beta\alpha$. This can be rewritten as $\alpha^2\beta^2\alpha\beta = \alpha\beta\alpha$ using $\beta^2\alpha = \alpha^2\beta$. Replace $\beta^2\alpha$ with $\alpha^2\beta$ on the left hand side of $\alpha^2\beta^2\alpha\beta = \alpha\beta\alpha$ and cancel α to get $\beta^2 = \beta\alpha$. This implies $\beta = \alpha$.

Case 5: $\alpha\beta^2\alpha = \beta\alpha^2\beta$.

Using the relation $\beta^2\alpha = \alpha^2\beta$, we can rewrite this equation as $\alpha^3\beta = \beta^3\alpha$ which implies $\alpha = \beta$ since $\alpha^3 = \beta^3 = \mathrm{id}$.

□

The set $\{\alpha, \beta\}$ is not a non-abelian B_4-set since $\alpha^2\beta\beta = \beta^2\alpha\beta$. The next theorem is a special case of a result of Odlyzko and Smith. We will use it in our construction.

Theorem 8.7.3 (Odlyzko, Smith, [66])**.** *For each prime p with $p-1$ divisible by 4, there is a non-abelian group G of order $4(p^4-1)$ and a non-abelian B_4-set $A \subset G$*

with

$$|A| = \frac{1}{4}(p-1).$$

Armed with Lemma 8.7.1, Lemma 8.7.2, and Theorem 8.7.3, we now prove Theorem 8.1.8.

Let p be any prime with $p-1$ divisible by 4. By Theorem 8.7.3, there is a group G_1 of order $4(p^4-1)$ and a non-abelian B_4-set $A_1 \subset G_1$ with $|A_1| = \frac{1}{4}(p-1)$. Define the group G to be the product group $G = G_1 \times H$. Let $A = A_1 \times \{\alpha, \beta\}$. Clearly $|G| = 12 \cdot 4(p^4-1)$, $|A| = \frac{1}{2}(p-1)$, and by Lemma 8.7.1, A is a non-abelian B_4^+-set in G.

Chapter 8 is a reprint of "Upper and lower bounds on B_k^+-sets," Timmons, Craig. Integers, 14A1 (2014), 1-27. The dissertation author was the primary investigator and author of this paper.

Chapter 9

Future Work

In this final chapter we discuss some future work.

9.1 The Turán Number of C_4

Abreu, Balbuena, and Labbate [1] proved for that any prime power q,

$$\mathrm{ex}(q^2-q-2,C_4) \geq \begin{cases} \frac{1}{2}q^3-q^2-\frac{q}{2}+1 & \text{if } q \text{ is odd,} \\ \frac{1}{2}q^3-q^2 & \text{if } q \text{ is even.} \end{cases}$$

They conjectured that these lower bounds are best possible. In Chapter 4 we showed

$$\mathrm{ex}(q^2-q-2,C_4) \geq \frac{1}{2}q^3-q^2-O(q^{3/4})$$

for each odd prime power q which disproves the conjecture of Abreu et al. in the case when q is odd. The case when q is even is still open.

Problem 9.1.1. Determine if the lower bound

$$\mathrm{ex}(q^2-q-2,C_4) \geq \frac{1}{2}q^3-q^2, q \text{ a power of } 2$$

is best possible.

Recall the graph $G_{q,\theta}$ defined in Chapter 4. This graph has vertex set $\mathbb{Z}_{q^2-1}$ where q is a prime power. Distinct vertices x and y are joined in $x+y \in A(q,\theta)$ where

$$A(q,\theta) = \{a \in \mathbb{Z}_{q^2-1} : \theta^a - \theta \in \mathbb{F}_q\}$$

is a Bose-Chowla Sidon set.

Recently, it was shown that the graph $G_{q,\theta}$ is a large induced subgraph of the graphs constructed by Brown [14], Erdős, Rényi, and Sós [32]. Let us introduce some notation for these graphs. Let q be a prime power and let ER_q be the graph with vertex set the points of $PG(2,q)$. Two distinct vertices (x_0,x_1,x_2) and (y_0,y_1,y_2) are adjacent if $x_0y_0+x_1y_1+x_2y_2=0$. The graph ER_q has q^2+q+1 vertices, $\frac{1}{2}q(q+1)^2$ edges, and is C_4-free. These are precisely the graphs constructed in [14] and [32].

Theorem 9.1.2 (Tait, Timmons, 2014 [78])**.** *Let $q \geq 15$ be a prime power. The graph $G_{q,\theta}$ is isomorphic to an induced subgraph of ER_q.*

Theorem 9.1.2 allows one to study ER_q by studying Bose-Chowla Sidon sets. In [62], Mubayi and Williford investigated the independence number of ER_q. An open problem mentioned in [62] is to construct an independent set I in ER_q for any q that is not an even power of 2 with $|I| = q^{3/2} + O(q)$, or show that no such set exists. The graphs $G_{q,\theta}$ and ER_q differ by only $q+2$ vertices and so the independence number of $G_{q,\theta}$ differs from the independence number of ER_q by at most $q+2$. Using the definition of adjacency, we find that an independent set I in $G_{q,\theta}$ is a set $I \subset \mathbb{Z}_{q^2-1}$ such that $x+y \notin A(q,\theta)$ for all $x \neq y$, $x,y \in I$; that is

$$\alpha(G_{q,\theta}) = \max_{I \subset \mathbb{Z}_{q^2-1}} \{|I| : ((I+I)\backslash(2\cdot I)) \cap A(q,\theta) = \emptyset\}$$

where $2 \cdot I = \{2x : x \in I\}$.

Problem 9.1.3. Let $A(q,\theta)$ be a Bose-Chowla Sidon set. Determine

$$\max_{I \subset \mathbb{Z}_{q^2-1}} \{|I| : ((I+I)\backslash(2\cdot I)) \cap A(q,\theta) = \emptyset\}.$$

Another open problem from [62] is to find an induced subgraph of ER_q, q a power of 2, that is triangle free and has at least $\frac{q^2}{2} + O(q^{3/2})$ vertices. Again, since this problem concerns induced subgraphs, finding such a subgraph in $G_{q,\theta}$ would solve this problem in ER_q.

Problem 9.1.4. Find the largest set $J \subset \mathbb{Z}_{q^2-1}$ such that for any $x,y,z \in J$ with x,y,z all distinct, at least one of the sums $x+y$, $x+z$, or $y+z$ is not contained in $A(q,\theta)$.

9.2 Ordered Turán for Even Cycles

Determining the order of magnitude of $\text{ex}(n, C_{2k})$ for $k \notin \{2, 3, 5\}$ is one of the most important open problems concerning bipartite Turán numbers. The upper bound $\text{ex}(n, C_{2k}) = O(n^{1+1/k})$ holds for all $k \geq 2$ [11]. Algebraic methods have been used to prove lower bounds on $\text{ex}(n, C_{2k})$. For example Lazebnik, Ustimenko, and Woldar [53] proved that $\text{ex}(n, C_{2k}) = \Omega(n^{1+\frac{2}{3k-3+t}})$ for all $k \geq 2$ where $t = 0$ if k is odd, and $t = 1$ if k is even. Despite these efforts, it is still unknown if $\text{ex}(n, C_{2k}) = \Omega(n^{1+1/k})$ for $k \notin \{2, 3, 5\}$. Since $\text{ex}(n, C_{2k}) \leq \text{ex}(n, \mathcal{Z}_{2k})$, it may be easier to solve the following problem.

Problem 9.2.1. Determine if $\text{ex}(n, \mathcal{Z}_{2k}) = \Omega(n^{1+1/k})$ for $k \notin \{2, 3, 5\}$.

The family $\mathcal{Z}_6$ contains six different ordered 6-cycles (see Figure 9.1). The family $\mathcal{Z}_8$ is even larger and in general, the number of ordered graphs in the family $\mathcal{Z}_{2k}$ increases with k. If instead of forbidding all ordered $2k$-cycles in the family $\mathcal{Z}_{2k}$ we forbid a particular subfamily, then B_k-sets can be used to obtain lower bounds of the form $cn^{1+1/k}$. The construction is as follows. Let $A \subset [n]$ be a B_k-set. Let G_A be the graph with vertex set $[n]$ and vertices i and j are adjacent if $i = j + a$ for some $a \in A$. Consider a Z_{2k} in G_A. Suppose this Z_{2k} is $x_1y_1x_2y_2 \dots x_ky_kx_1$ where $x_i < y_j$ for all i, j. Since A is a B_k-set,

$$\{y_1 - x_1, y_2 - x_2, \dots, y_k - x_k\} = \{y_1 - x_2, y_2 - x_3, \dots, y_k - x_1\}$$

and so we cannot have an edge that is longer or shorter than all of the other edges. Let $\mathcal{Z}_{2k}^{\times}$ be the family of Z_{2k}'s with a longest or shortest edge. For example

$$\mathcal{Z}_6^{\times} = \{Z_6^2, Z_6^3, Z_6^4, Z_6^5, Z_6^6\}.$$

The ordered 6-cycle Z_6^2 has an edge that is shorter than all of the others and the ordered 6-cycles Z_6^3, Z_6^4, Z_6^5, and Z_6^6 have an edge that is longer than all of the others. Using the B_k-sets constructed by Bose and Chowla, [13], we get the lower bound

$$\text{ex}(n, \mathcal{Z}_{2k}^{\times}) \geq cn^{1+1/k}$$

for all $k \geq 3$. Here $c > 0$ is a constant independent of k. We would like to know if there is an upper bound that matches the order of magnitude of this lower bound.

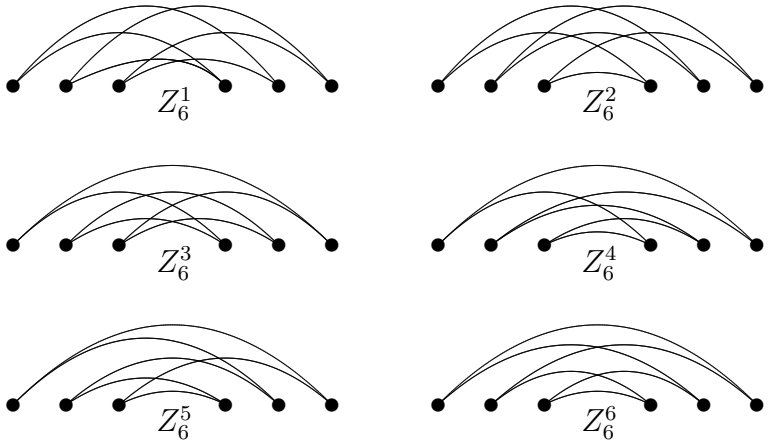

Figure 9.1: The Family $\mathcal{Z}_6$

Problem 9.2.2. Let $k \geq 3$ be an integer. Determine if

$$\mathrm{ex}(n, \mathcal{Z}_{2k}^{\times}) = O(n^{1+1/k}).$$

A much more difficult problem is the following.

Problem 9.2.3. Let $k \geq 3$ be an integer. Determine if for any $Z_{2k} \in \mathcal{Z}_{2k}$, we have

$$\mathrm{ex}(n, Z_{2k}) = O(n^{1+1/k}).$$

We can also use B_3^+-sets to give interesting constructions. Let $A \subset [n]$ be a B_3^+-set with $|A| = (1 + o(1))(4n)^{1/3}$. The existence of such a set was proved in Chapter 8. Define G_A as before. One can check that G_A is Z_6^6-free and, using an inequality of Katz and an argument of Graham [16], we can show that G_A has at least $.79n^{4/3}$ edges for large enough n. Using the upper bound on $\mathrm{ex}(n, C_6)$ of Füredi, Naor, and Verstraëte [44], we get

$$\mathrm{ex}(n, C_6) + .15n^{4/3} \leq \mathrm{ex}(n, Z_6^6).$$

Since $\mathrm{ex}(n, C_6) = \Theta(n^{4/3})$, we see that forbidding just one ordered 6-cycle is not the same as forbidding all 6-cycles. It would be very nice to prove a corresponding upper bound of the form $\mathrm{ex}(n, Z_6^6) = O(n^{4/3})$ which is a special case of Problem 9.2.3.

9.3 3-fold Sidon Sets

In Chapter 7 we showed if $A \subset [N]$ is a k-fold Sidon set, then

$$|A| \leq \left(\frac{N}{k}\right)^{1/2} + O((kN)^{1/4}).$$

Lazebnik and Verstraéte [55] constructed 2-fold Sidon sets $A \subset [N]$ for infinitely many N with $|A| \geq (1/2 + o(1))N^{1/2}$. The next step is of course constructing 3-fold Sidon sets.

Problem 9.3.1. Construct a 3-fold Sidon set $A \subset [N]$ for infinitely many N with $|A| \geq cN^{1/2}$, or show that the maximum size of a 3-fold Sidon set in $[N]$ is $o(N^{1/2})$.

A 3-fold Sidon set $A \subset [N]$ with $|A| \geq cN^{1/2}$ is known to imply the existence of a graph with $c_1 N$ vertices, $c_2 N^{3/2}$ edges, and every edge is in exactly one cycle of length four [82]. A problem that may be easier is the following.

Problem 9.3.2. Construct a set $A \subset [N]$ for infinitely many N with $|A| \geq cN^{1/2}$ and A has only trivial solutions to

$$c_1(x_1 - x_2) = c_2(x_3 - x_4)$$

for $1 \leq c_1 \leq c_2 \leq 3$, or show that no such sequence of sets exists.

Another problem is to determine the maximum size of a 2-fold Sidon set in $\mathbb{Z}_N$ or $[N]$. Let $S_k(N)$ be the maximum size of a k-fold Sidon set in $\mathbb{Z}_N$. For any integer $t \geq 1$, there are 2-fold Sidon sets $A \subset \mathbb{Z}_N$, $N = 2^{2^{t+1}} + 2^{2^t} + 1$, with $|A| \geq \frac{1}{2}N^{1/2} - 3$ (see [55]). Theorem 7.1.2 gives an upper bound of $(N/2)^{1/2} + O(N^{1/4})$ and so

$$\frac{1}{2} \leq \limsup_{N\to\infty} \frac{S_2(N)}{N^{1/2}} \leq \frac{1}{2^{1/2}}.$$

Problem 9.3.3. Determine $\limsup_{N\to\infty} \frac{S_2(N)}{N^{1/2}}$.

In the case of Sidon sets, we have $\limsup_{N\to\infty} \frac{S_1(N)}{N^{1/2}} = 1$ by [37] and [74].

9.4 B_k^*-sets

In Chapter 8 we used Bose-Chowla B_k-sets to prove that

$$F_k^*(N) \geq (1 + o(1))2^{1-1/k}N^{1/k} \tag{9.1}$$

for all odd $k \geq 3$. It would be interesting to determine if a bound similar to (9.1) holds when $k \geq 4$ is even.

Problem 9.4.1. Determine if there is a positive constant $c > 0$ such that

$$F_k^*(N) \geq (1 + c + o(1))N^{1/k}$$

for even $k \geq 4$.

The results of Chapter 8 imply that $F_3(N)$ is not asymptotic to $F_3^+(N)$. We do not know if $F_3^+(N)$ and $F_3^*(N)$ are asymptotic to one another.

Problem 9.4.2. Determine if $F_3^+(N)$ is asymptotic to $F_3^*(N)$.

Bibliography

[1] M. Abreu, C. Balbuena, D. Labbate, *Adjacency matrices of polarity graphs and other C_4-free graphs of large size*, Des. Codes Cryptogr. (2010), 55, 221-233.

[2] N. Alon, M. Krivelevich, B. Sudakov, *Turán numbers of bipartite graphs and related Ramsey-type questions*, Special issue on Ramsey theory. Combin. Probab. Comput. 12 (2003), no. 5-6, 477-494.

[3] N. Alon, L. Rónyai, T. Szabó, *Norm-graphs: variations and applications*, J. Combin.łTheory B, **76** (2) (1999), p. 280-290.

[4] M. Ajtai, E. Szemerédi, *Sets of lattice points that form no squares*, Stud. Sci. Math. Hungar. **9** (1974), 9-11.

[5] M. Axenovich, personal communication

[6] R. C. Baker, G. Harman, J. Pintz, *The difference between consecutive primes II*, Proc. London Math. Soc. **83** (3) (2001), 532-562.

[7] C. T. Benson, *Minimal regular graphs of girths eight and twelve*, Canad. J. Math. **18** (1966), 1091-1094.

[8] C. Berge, *Hypergraphes*, Combinatoire des ensembles finis, Dunod, Paris, (1987) xii+240 pp.

[9] B. Bollobás, *Extremal Graph Theory*, Academic Press Inc. (London) Ltd., 1978.

[10] B. Bollobás, *Modern Graph Theory*, Graduate texts in mathematics; 184, Springer 1998.

[11] J. A. Bondy, M. Simonovits, *Cycles of even length in graphs*, J. Combin. Theory B, **16** (1974), p. 97-105.

[12] R. C. Bose, *An affine analogue of Singer's theorem*, J. Ind. Math. Soc. 6 (1942), 1-15.

[13] R. C. Bose, S. Chowla, *Theorems in the additive theory of numbers*, Comment. Math. Helv. **37** (1962/1963), 141-147.

[14] W. G. Brown, *On graphs that do not contain a Thomsen graph*, Canada Math. Bull. **9** (1966), 281-289.

[15] S. Chen, *On the size of finite Sidon sequences*, Proc. Amer. Math. Soc. Vol. 121 **2** (1994), 353-356.

[16] F. R. K. Chung, *Diameters and eigenvalues*, J. American Mathematical Society, Vol. 2, No. 2, 187-196, 1989.

[17] J. Cilleruelo, *Gaps in dense Sidon sets*, Integers, 11:A11, 2001.

[18] J. Cilleruelo, *New upper bound for finite B_h Sequences*, Advances in Mathematics, **159** (2001), 1-17.

[19] J. Cilleruelo, *Sidon sets in* $\mathbb{N}^d$, J. Combin. Theory, Series A **117** (2010), 857-871.

[20] J. Cilleruelo, C. Timmons, *k-fold Sidon sets*, submitted.

[21] C. R. J. Clapham, A. Flockhart, J. Sheehan, *Graphs without four-cycles*, J. Graph Theory 13 (1989), no. 1, 29-47.

[22] D. Conlon, J. Fox, *Graph removal lemmas*, Surveys in Combinatorics, Cambridge University Press, 2013, 1-50.

[23] J. Czipszer, P. Erdős, A. Hajnal, *Some extremal problems on infinite graphs*, Publications of the Math. Inst. of the Hungarian Academy of Sci. Ser. A **7** (1962), p. 441-456.

[24] A. Dudek, V. Rödl, *On the Turán properties of infinite graphs*, Electr. J. Combin. **15** (2008) #R47.

[25] P. Erdős, *On a problem of Sidon in additive number theory, and on some related results - addendum*,

[26] P. Erdős, *A survey of problems in combinatorial number theory*, Annals of Discrete Mathematics 6 (1980), 89-115.

[27] P. Erdős, *Some problems and results on combinatorial number theory*, Graph theory and its applications, Ann. New York Acad. Sci. 576 (1989), 132-145.

[28] P. Erdős, *Some problems in additive number theory*, Amer. Math. Monthly **77** (1970), 619-621.

[29] P. Erdős, P. Frankl, V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), 113-121.

[30] P. Erdős, T. Gallai, *On maximal paths and circuits of graphs*, Acta Math. Acad. Sci. Hungar. **10** (1959) 337-356.

[31] P. Erdős, A. Rényi, *On a problem in the theory of graphs*, Publ. Math. Inst. Hungar. Acad. Sci. 7A (1962), p. 623-641.

[32] P. Erdős, A. Rényi, V. T. Sós, *On a problem of graph theory*, Studia Sci. Math. Hungar. 1 (1966), 215-235.

[33] P. Erdős, M. Simonovits, *A limit theorem in graphs theory*, Studia Sci. Math. Hung. **1**, (1966), 51-57.

[34] P. Erdős, M. Simonovits, *Compactness results in extremal graph theory*, Combinatorica **2** (3) (1982), p. 275-288.

[35] P. Erdős, M. Simonovits, *Cube-saturated graphs and related problems*, Progress in graph theory (Waterloo, Ont., 1982), 203-218, Academic Press, Toronto, 1984.

[36] P. Erdős, A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946) 1087-1091.

[37] P. Erdős, P. Turán, *On a problem of Sidon in additive number theory, and on some related results*, Journal of the London Mathematical Society, 16 (1941).

[38] F. Firke, P. Kosek, E. Nash, J. Williford, *Extremal graphs without 4-cycles*, J. Combin. Theory, Series B **103** (2013) 327-336.

[39] J. Fox, B. Sudakov, *Dependent random choice*, Random Structures Algorithms 38 (2011), no. 1-2, 68–99.

[40] Z. Füredi, *An upper bound on Zarankiewicz' Problem*, Combin. Probab. Comput. **5** (1996), p. 29-33.

[41] Z. Füredi, *Graphs without quadrilaterals*, J. Combin. Theory B, **34** (2) (1983), p. 187-190.

[42] Z. Füredi, *New asymptotics for bipartite Turán numbers*, J. Combin. Theory, Series A, **75** (1) (1996), p. 141-144.

[43] Z. Füredi, *On the number of edges of quadrilateral-free graphs*, J. Combin. Theory, Series B **68**, 1-6 (1996).

[44] Z. Füredi, A. Naor, J. Verstraëte, *On the Turán number for the hexagon*, Adv. Math. 203 (2006), no. 2, p. 476-496.

[45] Z. Füredi, M. Simonovits, *The history of degenerate (bipartite) extremal graph problems*, arXiv:1306.5167

[46] S. W. Graham, B_h *sequences*, in "Proceedings, Conference in Honor of Heini Halberstam" (B. C. Berndt, H. G. Diamond, and A. J. Hildebrand, Eds.), 337-355, Birkhäuser, Basel, 1996.

[47] B. Green, *The number of squares and* $B_h[g]$ *sets*, Acta Arith. **100** (2001), 365-390.

[48] H. Halberstam, K. F. Roth, *Sequences*, Vol. I, Clarendon Press, Oxford, 1966.

[49] P. Keevash, B. Sudakov, J. Verstraëte, *On a conjecture of Erdős and Simonovits: Even Cycles*, Combinatorica **33** (6) (2013) 699-732.

[50] J. Kollár, L. Rónyai, T. Szabó, *Norm-graphs and bipartite Turán numbers*, Combinatorica, 16 (3) (1996), p. 399-406.

[51] T. Kővári, V. T. Sós, P. Turán, *On a problem of Zarankiewicz*, Colloq. Math. **3** (1954), p. 50-57.

[52] X. Jia, *On finite Sidon sequences*, J. Number Theory **44** (1993), 84-92.

[53] F. Lazebnik, V. A. Ustimenko, A. J. Woldar, *A new series of dense graphs of large girth,* Bull. Amer. Math. Soc. **32** (1) (1995), 73-79.

[54] F. Lazebnik, V. A. Ustimenko, A. J. Woldar, *Properties of certain families of* $2k$*-cycle-free graphs*, J. Combin. Theory Series B, **60** 2 (1994), p. 293-298.

[55] F. Lazebnik, J. Verstraëte, *On hypergraphs of girth five*, Electr. J. Combin. **10**, (2003), #R25.

[56] R. Lidl, H. Niederreiter, *Finite Fields*, Cambridge University Press, 1997.

[57] B. Lindström, *An inequality for* B_2 *sequences*, J. Combin. Theory **6** (1969), 211-212.

[58] B. Lindström, *A remark on* B_4*-sequences*, J. Combin. Theory **7**, (1969), 276-277.

[59] B. Lindström, *A translate of Bose-Chowla* B_2*-sets*, Studia Sc. Math. Hungar., **36**, (2000), 331-333.

[60] B. Lindström, *Well distribution of Sidon sets in residue classes*, J. Number Theory 69 (2), 197-200 (1998).

[61] W. Mantel, Problem 28, soln. by H. Gouwentak, W. Mantel, J. Teixeira de Mattes, F. Schuh and W. A. Wythoff, *Wiskundige Opgaven* **10** (1907) 60-61.

[62] D. Mubayi, J. Williford, *On the independence number of the Erdős-Rényi and Projective Norm Graphs and a related hypergraph*, J. Graph Theory 56 (2007), no. 2, 113-127.

[63] A. Naor, J. Verstraëte, *A note on bipartite graphs without $2k$-cycles*, Combin. Probab. Comput. **14** (2005), p. 845-849.

[64] V. Nikiforov, *A contribution to the Zarankiewicz problem*, Linear Algebra and its Applications, 432 no. 6 (2010), 1405-1411.

[65] K. O'Bryant, *A complete annotated bibliography of work related to Sidon sequences*, Electr. J. Combin. **DS 11** (2004).

[66] A.M. Odlyzko, W.D. Smith, *Nonabelian sets with distinct k-sums*, Discrete Mathematics 146 (1995), 169-177.

[67] O. Pikhurko, *A note on the Turán function of even cycles*, Proc. Amer. Math. Soc. 140 (2012), 3687-3992.

[68] I. Reiman, *Über ein Problem von K. Zarankiewicz*, (German) Acta. Math. Acad. Sci. Hungar. **9** (1958), 269-273.

[69] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. 28 (1953), 104-109.

[70] P. Rowlinson, Y. Yuansheng, *On extremal graphs without four-cycles*, Utilitas Math. 41 (1992), 204-210.

[71] I. Ruzsa, *Solving a linear equation in a set of integers I*, Acta Arith. **65** 3 (1993), 259-282.

[72] I. Ruzsa, *Erdős and the Integers*, J. Number Theory 79, (1999) 11-163.

[73] I. Ruzsa, E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq. Keszthely, 1976), Vol. II, pp. 939–945, Colloq. Math. Soc. János Bolyai, 18, North-Holland, Amsterdam-New York, 1978.

[74] J. Singer, *A theorem in finite projective geometry and some applications to number theory*, Trans. Amer. Math. Soc. 43 (1938), 377-385.

[75] J. Solymosi, C_4 Removal Lemma for sparse graphs in: Open Problem Session, Mathematisches Forschungsinstitut Oberwolfach, Report No. 01/2011 DOI: 10.4171/OWR/2011/01 Combinatorics (2011).

[76] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199-245.

[77] E. Szemerédi, *Regular partitions of graphs*, Problémes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), pp. 399–401, Colloq. Internat. CNRS, 260, CNRS, Paris, 1978.

[78] M. Tait, C. Timmons, *Orthogonal polarity graphs and Sidon sets*, in preparation.

[79] M. Tait, C. Timmons, *Sidon sets and graphs without 4-cycles*, submitted.

[80] C. Timmons, *An ordered Turan problem for bipartite graphs*, Electr. J. Combin. 19(4) 2012, #43.

[81] C. Timmons, *Upper and lower bounds on B_k^+-sets*, Integers, 14:A1, 1-27, 2014.

[82] C. Timmons, J. Verstraëte, *A counterexample to sparse removal*, submitted.

[83] P. Turán, *On an extremal problem in graph theory* (in Hungarian), Mat. Fiz. Lapok **48** (1941) 436-452.

[84] J. Verstraëte, *Arithmetic progressions of cycle lengths in graphs*, Combin. Probab. Comput. **9** (2000), 369-373.

[85] R. Wenger, *Extremal graphs with no C^4's, C^6's, or C^{10}'s*, J. of Combin. Theory, Series B, **52** (1991), p. 113-116.

[86] D. West, *Introduction to Graph Theory*, 2nd Edition, 2001 Prentice Hall, Inc. Upper Saddle River, NJ 07458.