New quantitative estimates on the incidence geometry and growth of finite sets

Timothy Gareth Fellgett Jones

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science

School of Mathematics
January 2013

20,000 words
Abstract

This thesis establishes new quantitative records in several problems of incidence geometry and growth. After the necessary background in Chapters 1, 2 and 3, the following results are proven.

Chapter 4 gives new results in the incidence geometry of a plane determined by a finite field of prime order. These comprise a new upper bound on the total number of incidences determined by finitely many points and lines, and a new estimate for the number of distinct lines determined by a finite set of non-collinear points.

Chapter 5 gives new results on expander functions. First, a new bound is established for the two-variable expander $a + ab$ over a finite field of prime order. Second, new expanders in three and four variables are demonstrated over the real and complex numbers with stronger growth properties than any functions previously considered.

Finally, Chapter 6 gives the first bespoke sum-product estimate over function fields, a setting that has so far been largely unexplored for these kinds of problems. This last chapter is joint work with Thomas Bloom.
Acknowledgements

This thesis would not exist, nor would the years leading up to its creation have been as fun, without the support of a great many people.

My supervisor is Misha Rudnev, and without his guidance, support, encouragement and patience nothing here would have been possible. Whilst none of the work here is specifically joint with Misha, everything herein has benefitted from his scrutiny and suggestions. My studentship is funded by the EPSRC, and I have been looked after very well by both the University of Bristol and the School of Mathematics.

The postgraduate community of the School of Mathematics, and my long-running housemates Amanda and Nic, who do proper science, have been an invaluable source of support and distraction. To list everyone from maths would necessitate several pages, so I will limit myself to mentioning my combinatorial colleagues Tom Bloom and Olly Roche-Newton, with whom I have had many fruitful conversations. One of the results in this thesis is joint work with Tom.

I am also grateful to Jo Bryant and everyone in At-Bristol, and to Alastair Iles and the leaders and cubs of the 7th Bristol cub pack, for reminders of the importance of the world outside sumset calculus and covering methods, and for giving me fun, challenging and most especially non-mathematical things to do.

Finally, Helen has made me more cups of tea than I care to recall, and put up with me at the lows (not being able to prove things, and so fed up) and highs (being able to prove something, and so unable to sit still) of research.

I have, overall, been very lucky indeed.
Author’s declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University’s Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate’s own work. Work done in collaboration with, or with the assistance of, others is indicated as such. Any views expressed in the dissertation are those of the author.

Signed:

Date:
**Table of contents**

*Standard notation* ................................................. 13

*Prologue* .......................................................... 15

1 **Incidence geometry** ........................................... 21
    1.1 Incidence bounds in a plane .............................. 21
    1.2 The Szemerédi-Trotter theorem ........................... 25
    1.3 Line counting and Beck’s theorem ........................ 28
    1.4 Incidence geometry in $\mathbb{R}^n$ ....................... 31
    1.5 Other fields .............................................. 34

2 **Growth** .......................................................... 37
    2.1 Sum-product estimates ..................................... 37
    2.2 Expander functions ....................................... 43

3 **A handbook on sumset calculus** ............................ 45
    3.1 Introduction .............................................. 46
    3.2 The Plünnecke-Ruzsa inequalities ........................ 46
    3.3 Partial sumsets ........................................... 51
    3.4 Additive energy .......................................... 56
    3.5 Covering methods ....................................... 59

4 **Incidence theorems over finite fields** ...................... 63
    4.1 Results .................................................. 63
    4.2 Foci and configurations ................................... 66
Table of contents

4.3 How to find foci ............................................. 70
4.4 Bounding partial sum-products ................................. 75
4.5 Proving Theorems 4.1 and 4.2 ................................. 79
4.6 Further work .................................................. 83

5 Expander functions ............................................. 85
5.1 Results ....................................................... 85
5.2 Two-variable finite field expanders ............................ 87
5.3 Three and four-variable real and complex expanders .......... 95
5.4 Further work .................................................. 103

6 A sum-product theorem in function fields ...................... 105
6.1 Results ....................................................... 105
6.2 Function fields .............................................. 107
6.3 Separable sets .............................................. 109
6.4 Finding many separable sets ................................ 110
6.5 Proof of Theorem 6.1 ........................................ 115
6.6 Further work .................................................. 116

Appendices ....................................................... 118
A Pigeonholing ................................................... 119
A.1 Averaging ..................................................... 119
A.2 Dyadic pigeonholing ......................................... 120
A.3 The Cauchy-Schwarz inequality .............................. 120

B Projective geometry ............................................ 123
B.1 Projective space ............................................. 123
B.2 The hyperplane at infinity .................................. 124
B.3 Projective transformations .................................. 124

References ....................................................... 127
List of figures

1.1 Counting incidences ........................................ 22
1.2 $k$ collinear points determine $k - 1$ edges. ............... 26
1.3 The two extremes for counting lines. ........................ 29

4.1 A point set and focus ........................................ 67
4.2 Sum-product configurations .................................. 68
4.3 An illustration of Desargues’ theorem. .................... 84
Standard notation

The following standard conventions are used.

• Write $Y = O(X)$, $Y \ll X$, or $X = \Omega(Y)$ if there is a fixed constant $C$ such that $Y \leq CX$. The constant $C$ is referred to as the ‘implicit constant’.

• Using this notation, we will often use the fact that $X \ll Y + Z$ if and only if $X \ll \max\{Y, Z\}$.

• If $X \ll Y$ and $Y \ll X$ then write $X \approx Y$ or $Y = \Theta(X)$.

• If the implicit constant depends on some parameter $\lambda$ then this is reflected with a subscript, e.g. $Y = O_\lambda(X)$ or $Y \ll_\lambda X$.

• Write $Y \ll X^{1+o(1)}$, $Y \preceq X$, $Y = \tilde{O}(X)$ or $X = \tilde{\Omega}(Y)$ if $Y \ll_\epsilon X^{1+\epsilon}$ for all $\epsilon > 0$. In particular, this notation is used when $Y \ll X (\log X)^\alpha$ for some $\alpha > 0$.

• Write $1$ for an indicator function, that is if $P$ is a proposition then $1(P)$ is equal to 1 if $P$ is true and is zero otherwise.

• The letters $\mathbb{R}, \mathbb{C}, \mathbb{N}, \mathbb{Z}$ have their usual meaning. By $\mathbb{F}_p$ we mean a finite field of prime order $p$, and by $\mathbb{F}_q$ a finite field of order $q$, where $q = p^\alpha$ is a prime power. The letters $p$ and $q$ will also be used to denote points in a plane or higher-dimensional vector space, but the context will ensure that no confusion arises.
For subsets $A, B$ of a field, write $A + B = \{a + b : a \in A, b \in B\}$ and $AB = \{ab : a \in A, b \in B\}$ for the sumset and product set of $A$ and $B$. The difference set $A - B$ and ratio set $A/B$ are defined analogously.

If $x$ is an element and $A$ a subset of a field then write $A + x$ for the translation $A + \{x\}$ and $xA$ for the dilation $\{x\} A$.

If $k$ is a natural number and $A$ is a subset of a field then write $kA$ for the iterated sumset $kA = \underbrace{A + \ldots + A}_k$. Context will ensure that no confusion between a dilation and an iterated sumset of $A$ arises.
Prologue

‘And if you take one from three hundred and sixty-five what remains?’
‘Three hundred and sixty-four, of course.’
Humpty Dumpty looked doubtful, ‘I’d rather see that done on paper,’ he said.
Lewis Carroll, Through the Looking-Glass

This thesis sets new records at the interface of two areas of mathematics: incidence geometry and growth. The first of these, incidence geometry, is about points and lines in a plane and incidences between them; a point is ‘incident’ to a line if it lies on that line. The following sorts of problem are typical:

**Incidence bounds.** Incidences are counted with multiplicity, in the sense that if several lines cross at a single point then we count several incidences. In total, how many incidences could there be between a finite set of points and a finite set of lines?

**Line counting.** A pair of distinct points determines a line. In total, how many distinct lines might be determined by pairs from a finite set of points?

There are many natural generalisations, for example to higher dimensional vector spaces instead of planes, to curves instead of lines, and so forth.

The second area, growth, is about how much bigger a finite subset of a field becomes after passing to its image under some function of two or more variables. Examples are:
**Prologue**

**Sum-product estimates.** For a set $A$ of given cardinality, how much bigger must the quantity $\max\{|A+A|,|AA|\}$ be? Here, $A+A$ and $AA$ are respectively the sets of pairwise sums and products of elements of $A$.

**Expander functions.** Let $F$ be the ambient field and $n \geq 2$ be an integer. What functions $f : F^n \to F$ are there for which the set $f(A) = \{f(a_1, \ldots, a_n) : a_i \in A\}$ is always much bigger than a finite subset $A$ of $F$?

There is a common theme to all of this: estimating how much regimentation can be forced onto a finite set. For example, on the incidence geometry side the existence of too many incidences corresponds to the existence of extremely structured sets of points and lines. Understanding incidence bounds means placing a cap on such structure. As an example on the growth side, the only way in which the sumset $A+A$ or the product set $AA$ can be small is if the elements of $A$ are arranged respectively in some kind of additive or multiplicative ‘conspiracy’, such as an arithmetic or geometric progression. Understanding sum-product estimates therefore means working out how far such conspiracies could coincide.

There are applications of this theme, and of the specific problems above, to theoretical computer science and cryptography, where they can be used to obtain rigorous estimates of pseudorandom behaviour. However these are not pursued here, the motivation instead being for progress on the problems in their own right. The only exception to this approach of not worrying about applications is the interaction between the problems themselves. Incidence results have applications to growth, and growth results have applications to incidences, and this is one of the things that will contribute to progress on both.

When it comes to results the philosophy is one of ‘hard analysis’, placing the emphasis on finding explicit bounds and making them as strong as possible. Quite how strong this might be depends very much on the choice of underlying field. Historically the attention was on real numbers. There are powerful results in this setting, which is still extremely active. With work, many incidence and growth results extend verbatim from the real to complex settings. More recently a lot of research has focused on finite fields. This thesis presents work in all of

---

1Recall that the finite field $\mathbb{F}_p$ of prime order $p$ is simply the set of residues modulo $p$ under addition and multiplication, and that the finite field $\mathbb{F}_q$ of prime power order $q = p^\alpha$ is a degree $\alpha$ algebraic extension of $\mathbb{F}_p$. 16
some remarks on finite fields

Much, but by no means all, of the work in this thesis concerns finite fields, and it is worth pausing to explain a couple of considerations versus the real and complex settings. There is also a slightly different set of considerations in the study of function fields, but these are deferred to Chapter 6 since that is the only place they arise.

There are two main complications in finite fields. The first is the existence of finite subfields, which must be ruled out before anything nontrivial can be said. In the case of the field $\mathbb{F}_p$ of prime order $p$, this collapses into a cardinality condition since there are no proper subfields. For this reason much finite field research focuses on the prime order case; the hoops that one must jump through are similar to the general case, but the mathematics is usually cleaner.

The second complication is an issue of tools and difficulty. Many methods in the real and complex settings depend critically on their topologies and so do not extend to finite fields. This makes it a lot harder to prove things and forces a more combinatorial approach, which is a mixed blessing. The bad news is that quantitative results are usually not as strong. But the good news is that methods and results achieved in finite fields usually extend elsewhere without difficulty. Thus finite field results carry a certain amount of moral authority as they correspond to the worst possible cases.

Although not considered here in detail, it is worth remarking that there is an additional ‘large set’ paradigm for finite field results, with extensive literature. This features strong results for sets satisfying an additional minimum density condition, typically at least a square-root barrier. However the methods have a very different flavour, drawing for example on estimates for exponential sums. On top of this, they cannot be so easily exported to other settings; the interest here is in finite sets, but a minimum density condition in finite fields is usually analogous to a requirement for infinite sets elsewhere.

---

2 The function field $\mathbb{F}_q(t)$ is the field of rational functions over the finite field $\mathbb{F}_q$. 

Structure of the thesis

The first three chapters build up background and preliminaries.

- Chapters 1 and 2 introduce incidence geometry and growth respectively in more detail. They also record particular incidence and growth results for use in later chapters.

- Chapter 3 is a handbook of results in sumset calculus, which is a workhorse of much of the later mathematics.

The subsequent chapters are concerned with original results, developing state of the art incidence geometry and growth in several directions. Between them, there are new results on all four problems from page 15: incidence bounds, line counting, sum-product estimates and expander functions. Moreover they encompass real, complex and finite field settings, and also break new ground in the relatively unexplored function field setting.

- Chapter 4 considers incidence geometry over finite fields and sets two new records.

  * The first is a new incidence bound. If $P$ and $L$ are a set of points and lines respectively in $\mathbb{F}_p^2$ with $|P|, |L| \leq N$ then, writing $I(P, L)$ for the number of incidences between $P$ and $L$, non-trivial bounds are of the form $I(P, L) \ll N^{\frac{3}{2} - \epsilon}$ for $\epsilon > 0$.

  A new bound of $\epsilon \geq \frac{1}{662} - o(1)$ is established, holding whenever $N < p$.

  This improves by an order of magnitude on the previous bound of $\epsilon \geq \frac{1}{10.678}$.

  * The second is a new line counting result. If $P$ is a set of points in $\mathbb{F}_p^2$ with $|P| < p$ then either $\Omega(|P|^{1-o(1)})$ of the points are collinear, or $P$ determines at least $\Omega \left( |P|^{1 + \frac{1}{133} - o(1)} \right)$ distinct lines.

  This improves on previous results in two ways. Quantitatively, the exponent is stronger than the previously best-known $1 + \frac{1}{267}$. And qualitatively, the result applies to all subsets of $\mathbb{F}_p^2$ satisfying the cardinality condition; the previously best-known result applies only for $P$ of the form $P = A \times A$ for $A \subseteq \mathbb{F}_p$. 
• Chapter 5 sets three new records for expander functions.

  * First, there is a result on two-variable expanders in finite fields. If
    \( f(a, b) = a(b + 1) \) then \( |f(A)| \gg |A|^{1 + \frac{1}{106} - o(1)} \) whenever \( A \) is a subset
    of \( \mathbb{F}_p \) with \( |A| < p^{1/2} \). This improves on the previous best-known
    exponent of \( 1 + \frac{1}{106} - o(1) \).

  * Second, there is a result on three-variable expanders for complex num-
    bers. The function \( g(a, b, c) = \frac{a - b}{a-c} \), which has not been previously
    considered in this context, is shown to satisfy \( |g(A)| \gg |A|^{2-o(1)} \) for
    any finite \( A \subseteq \mathbb{C} \). Previously-known functions with this property were
    all of four variables and applied only to sets of real numbers.

  * Third, there is a result on four-variable expanders for real num-
    bers. The function \( h(a, b, c, d) = \frac{(a-b)(c-d)}{(b-c)(a-d)} \), again not previously
    considered in this context, is shown to satisfy \( |h(A)| \gg |A|^2 \) whenever \( A \) is a
    finite set of real numbers. The previously best-obtained bound for a
    function of four variables was \( |A|^{2-o(1)} \).

• Chapter 6 is joint work with Thomas Bloom and breaks new ground by
  establishing a sum-product estimate in the function field \( \mathbb{F}_q(t) \), a relatively
  unexplored setting for this kind of work. Function fields form an interesting
  intermediate case between the finite field and real and complex settings since
  they have an unusually rigid ‘non-archimedean’ topology.

A sum-product estimate of

\[
\max \{|A + A|, |AA|\} \gg_q |A|^{1 + \frac{1}{5} - o(1)}
\]

is established for any finite subset \( A \) of \( \mathbb{F}_q(t) \). The exponent of \( 1 + \frac{1}{5} - o(1) \)
lies between the \( 1 + \frac{1}{11} - o(1) \) known for finite fields and the \( 1 + \frac{1}{3} - o(1) \)
known for real and complex numbers.

There are two appendices, which summarise standard background informa-

• Appendix A covers standard pigeonholing results used throughout the
  thesis. The phrases ‘by averaging’, ‘by Cauchy-Schwarz’ and ‘by dyadic
  pigeonholing’ are deployed frequently and implicitly refer to results from
  here.
Prologue

- **Appendix B** covers material on projective geometry necessary for Chapter 4 and Chapter 5. Additional standard material on the projective theory of cross ratios is covered at an appropriate point in Chapter 5.
Incidence geometry

The prologue mentioned two areas of incidence geometry: incidence bounds and line counting. This chapter introduces them properly, with an emphasis on the real and complex settings, the case of finite fields being deferred to Chapter 4. In so doing it also records results which will be of use in Chapter 5.

Incidence bounds are covered first, establishing a trivial estimate which holds regardless of the underlying field. Non-trivial incidence bounds depend on the setting; the best-understood case is the plane $\mathbb{R}^2$, where the classical Szemerédi-Trotter theorem holds. Line counting is then tackled in $\mathbb{R}^2$, establishing Beck’s theorem as a consequence of Szemerédi-Trotter.

Three ways of generalising to higher-dimensions are also considered, as is a generalisation from the real to complex setting.

Apart from Theorem 1.8, which is a straightforward variation of existing results, all of the mathematics in this chapter is drawn from the literature.

1.1 Incidence bounds in a plane

A plane is simply $F^2$ where $F$ is a field. A point is an element of $F^2$ and a line is the set of points $(x, y) \in F^2$ satisfying an equation

$$ax + by + c = 0$$

for fixed $a, b, c \in F$ that are not all zero. A point $p$ is incident to a line $l$ if $p \in l$. If $P$ is a finite set of points in a plane, and $L$ is a finite set of lines, then write...
Chapter 1. Incidence geometry

$I(P, L)$ for the number of incidences between points in $p$ and lines in $l$, that is

$$I(P, L) = \sum_{p \in P} \sum_{l \in L} \delta_{pl}.$$  

(1.1)

where

$$\delta_{pl} = \begin{cases} 1, & \text{if } p \in l \\ 0, & \text{if } p \notin l. \end{cases}$$

Incidences are therefore counted with multiplicity, as illustrated in Figure 1.1.

![Incidence examples](image)

(a) One incidence  
(b) Two incidences  
(c) Five incidences

Figure 1.1. Counting incidences

It is worth remarking (see Appendix B) that a ‘plane’ could instead be taken to be the projective plane $\mathbb{P}F^2$ containing $F^2$, and ‘lines’ to be projective lines within it. Indeed, all the results in this chapter extend without difficulty to projective space. But since no properties of the projective setting will be used, it makes sense to stay with the affine one. Chapters 4 and 5 make active use of the properties of projective space, and we will cross that bridge when we come to it.

An important topic in incidence geometry is the study of upper bounds on $I(P, L)$ in terms of $|P|$ and $|L|$, referred to as incidence bounds. When consid-
1.1. Incidence bounds in a plane

erating these, the most straightforward observation to make is that each point in $P$ is incident to at most $|L|$ lines in $L$, implying that

$$I(P, L) \leq |P||L|.$$  

However it is immediate from the definition of points and lines that two lines in $L$ are simultaneously incident to at most one point in $P$, and two points in $P$ are simultaneously incident to at most one line in $L$. This fact, combined with the Cauchy-Schwarz inequality, leads to the following somewhat better estimate, which is nonetheless referred to as trivial.

Lemma 1.1 (Trivial incidence bound). If $P$ is a set of points in a plane, and $L$ is a set of lines, then

$$I(P, L) \ll \min \{ |P| + |P|^{1/2}|L|, |L| + |L|^{1/2}|P| \}.$$  

Proof. Cauchy-Schwarz and (1.1) imply that

$$I(P, L)^2 = \left( \sum_{p \in P} \sum_{l \in L} \delta_{pl} \right)^2 \leq |P| \sum_{p \in P} \sum_{l_1, l_2 \in L} \delta_{pl_1} \delta_{pl_2}. \quad (1.2)$$

Split the summation over lines $l_1, l_2 \in L$ according to whether or not $l_1$ and $l_2$ are distinct, so that

$$\sum_{p \in P} \sum_{l_1, l_2 \in L} \delta_{pl_1} \delta_{pl_2} = \sum_{p \in P} \sum_{l \in L} \delta_{pl} + \sum_{p \in P} \sum_{l_1 \neq l_2 \in L} \delta_{pl_1} \delta_{pl_2} = I(P, L) + \sum_{p \in P} \sum_{l_1 \neq l_2 \in L} \delta_{pl_1} \delta_{pl_2}.$$  

Substituting into (1.2) yields

$$I(P, L)^2 \leq |P|I(P, L) + |P| \sum_{l_1 \neq l_2 \in L} \sum_{p \in P} \delta_{pl_1} \delta_{pl_2}. \quad (1.3)$$

Two distinct lines are simultaneously incident to at most one point, and so

$$\sum_{p \in P} \delta_{pl_1} \delta_{pl_2} \leq 1$$
Chapter 1. Incidence geometry

whenever \( l_1 \neq l_2 \), since this is the number of points in \( P \) incident to both \( l_1 \) and \( l_2 \). Hence from (1.3),

\[
I(P, L)^2 \leq |P|I(P, L) + |P||L|^2.
\]

So either \( I(P, L)^2 \ll |P|I(P, L) \) in which case \( I(P, L) \ll |P| \), or \( I(P, L)^2 \ll |P||L|^2 \) in which case \( I(P, L) \ll |P|^{1/2}|L| \). Overall therefore

\[
I(P, L) \ll |P| + |P|^{1/2}|L|.
\]

The proof that \( I(P, L) \ll |L| + |L|^{1/2}|P| \) is the same, except that the roles of points and lines are reversed. \( \square \)

It is often helpful to consider the ‘critical’ case where \( P \) and \( L \) have the same cardinality, say \(|P| = |L| = N\). In this case, the trivial bound from Lemma 1.1 is

\[
I(P, L) \ll N^{3/2}
\]

and so non-trivial bounds will be of the form

\[
I(P, L) \ll N^{3/2-\epsilon}
\]

for \( \epsilon > 0 \).

The following standard result shows that best non-trivial bound that can be hoped for is \( \epsilon = \frac{1}{6} \), that is \( I(P, L) \ll N^{4/3} \).

**Lemma 1.2** (Constraint on incidence bounds). Let \( F \) be a field. If \( F \) has characteristic 0 then for any positive integer \( N \) there is a set \( P \) of points in \( F^2 \) and a set \( L \) of lines, with \(|P|, |L| \approx N\) such that

\[
I(P, L) \approx N^{4/3}.
\]

If the characteristic of \( F \) is \( p > 0 \) then the same result holds, but with the constraint that \( N \) must be less than \( \left( \frac{2}{p} \right)^{3/2} \).

**Proof.** Write \( l_{rs} \) for the line given by \( y = rx + s \). In the zero-characteristic case take

\[
P = \{(x, y) \in \mathbb{Z}^2 : 1 \leq x \leq N^{1/3}, 1 \leq y \leq 2N^{2/3}\}
\]

\[
L = \{l_{rs} : (r, s) \in \mathbb{Z}^2 : 1 \leq r \leq N^{1/3}, 1 \leq s \leq N^{2/3}\}.
\]
There are $\Theta(N)$ distinct points in $P$ and $\Theta(N)$ distinct lines in $L$, and it is easy to check that each line in $L$ is incident to $\Theta(N^{1/3})$ points in $P$, meaning that there are $\Theta(N^{4/3})$ incidences.

The characteristic $p > 0$ case is the same; simply replace $\mathbb{Z}^2$ in the construction with $\mathbb{F}_p^2$. The constraint that $N < \left(\frac{p}{2}\right)^{3/2}$ ensures that all the points and lines are distinct.

1.2 The Szemerédi-Trotter theorem

What can be said about non-trivial incidence bounds? The answer depends on the underlying field over which the plane is defined. This section considers the classical case of the plane $\mathbb{R}^2$ where Szemerédi and Trotter [46] proved the following flagship result of incidence geometry.

**Theorem 1.3** (Szemerédi, Trotter). If $P$ is a set of points in $\mathbb{R}^2$ and $L$ is a set of lines, then

$$I(P, L) \ll |P|^{2/3}|L|^{2/3} + |P| + |L|.$$  

In the critical case this gives $I(P, L) \ll N^{4/3}$, which is sharp up to the implicit constant by Lemma 1.2. Before proving Szemerédi-Trotter it is worth recording an immediate consequence which is often useful in applications to other problems.

**Corollary 1.4.** Let $L$ be a set of lines in $\mathbb{R}^2$. Then the number of points incident to at least $k$ lines in $L$ is $O\left(\frac{|L|^2}{k^3} + \frac{|L|}{k}\right)$. Similarly, the number of lines incident to at least $k$ points in a point set $P$ is $O\left(\frac{|P|^2}{k^3} + \frac{|P|}{k}\right)$.

**Proof.** Let $P_k$ be the set of points incident to at least $k$ lines in $L$. Then

$$|P_k|k \leq I(P_k, L).$$

On the other hand, the Szemerédi-Trotter theorem shows that

$$I(P_k, L) \ll |P_k|^{2/3}|L|^{2/3} + |L|$$

and so comparing upper and lower bounds on $I(P_k, L)$ gives

$$|P_k| \ll \frac{|L|^2}{k^3} + \frac{|L|}{k}$$

as required. The proof for the number of lines incident to at least $k$ points is similar. \qed
Chapter 1. Incidence geometry

Now for a proof of the Szemerédi-Trotter theorem. There have been several proofs since Szemerédi and Trotter’s original. The one here is due to Székely [45]. A third proof, using different techniques again, can be found in a recent paper of Kaplan, Matoušek and Sharir [27].

Proof of Theorem 1.3. Without loss of generality assume that every point is incident to at least one line, and every line is incident to at least one point. Recall that a graph \( G(V, E) \) consists of a set \( V \) of vertices, and a set \( E \) of unordered pairs of vertices, called edges.

Construct a graph \( G(V, E) \) as follows. Take the set \( V \) of vertices to be the set of points \( P \). To construct the set \( E \) of edges, say that \((p_1, p_2)\) is an edge if and only if \( p_1 \) and \( p_2 \) are adjacent along a line in \( L \), in the sense that the line segment connecting them is contained in a line from \( L \) and contains no other points from \( P \).

For each \( l \in L \) write \( k(l) = \sum_{p \in P} \delta_{pl} \) for the number of points \( p \in P \) that are incident to \( L \). Note that \( l \) contains \( k(l) - 1 \) edges from \( E \), as illustrated in Figure 1.2.

\[ |E| = \sum_{l \in L} (k(l) - 1) \]
\[ = \sum_{l \in L} \sum_{p \in P} \delta_{pl} - |L| \]
\[ = I(P, L) - |L|. \quad (1.4) \]

Figure 1.2. \( k \) collinear points determine \( k - 1 \) edges.
1.2. The Szemerédi-Trotter theorem

We now apply a graph-theoretic result. A **drawing** of \( G = G(V, E) \) is an identification of each vertex in \( V \) with a distinct point in \( \mathbb{R}^2 \), and of each edge \((u, v) \in E\) with a curve connecting \( u \) and \( v \). A **crossing** occurs whenever two curves corresponding to edges intersect away from a vertex. Write \( \text{cross}(G) \) for the minimum number of crossings in a drawing of \( G \). We will employ the following lemma; see Chapter 8 of the book of Tao andVu [48] for a proof.

**Lemma 1.5** (Crossing number lemma). Let \( G(V, E) \) be a graph with \(|E| \geq 4|V|\). Then \( \text{cross}(G) \geq \frac{|E|^3}{64|V|^2} \).

Applying the crossing number lemma and (1.4) shows that at least one of the following bounds holds

\[
I(P, L) \ll |P| + |L| \quad (1.5)
\]

\[
\text{cross}(G) \gg \frac{I(P, L)^3}{|P|^2}. \quad (1.6)
\]

Let’s consider what happens when (1.6) holds. It is clear that

\[
\text{cross}(G) \leq |L|^2 \quad (1.7)
\]

since for two edges to cross it is necessary that two lines cross. Comparing the bounds (1.6) and (1.7) yields

\[
\frac{I(P, L)^3}{|P|^2} \ll \text{cross}(G) \leq |L|^2
\]

and hence

\[
I(P, L) \ll |P|^{2/3}|L|^{2/3} \quad (1.8)
\]

whenever (1.6) holds. Thus either (1.5) or (1.8) holds, implying overall that

\[
I(P, L) \ll |P|^{2/3}|L|^{2/3} + |P| + |L|
\]

as required.

The key to the above proof is the crossing number inequality. Beyond this the only properties used are the trivial facts that any two lines are simultaneously incident to at most one point, and any two points are simultaneously incident to at most one line. With a few modifications\(^1\) the same argument goes through

---

\(^1\)See for example Theorem 8.10 of the book of Tao andVu [48].
Chapter 1. Incidence geometry

when \( L \) is a set of curves rather than lines, such that any two points in \( P \) are simultaneously incident to at most \( \alpha \) curves and any two curves in \( L \) are simultaneously incident to at most \( \beta \) points in \( P \). The implicit constant in the statement then depends on \( \alpha \) and \( \beta \) and so the theorem becomes

\[
I(P, L) \ll_{\alpha, \beta} |P|^{2/3} |L|^{2/3} + |P| + |L|.
\]

The required changes are to replace the crossing number lemma for graphs with one for multigraphs, and to replace (1.7) with

\[
\text{cross}(G) \leq \beta |L|^2.
\]

Pach and Sharir \[35\] took this approach further, to encompass the case where any \( k \) points, rather than any two points, are simultaneously incident to at most \( \alpha \) curves. Note that the condition on curves remains the same; any two curves are simultaneously incident to at most \( \beta \) points, rather than any \( k \).

**Theorem 1.6** (Pach, Sharir). Let \( P \) be a set of points in \( \mathbb{R}^2 \) and \( L \) be a set of curves, such that any \( k \) points in \( P \) are simultaneously incident to at most \( \alpha \) curves in \( L \) and any two curves in \( L \) are simultaneously incident to at most \( \beta \) points in \( P \). Then

\[
I(P, L) \ll_{\alpha, \beta} |P|^{\frac{k}{k-1}} |L|^{\frac{\alpha k - 2}{k-1}} + |P| + |L|.
\]

1.3 Line counting and Beck’s theorem

This section gives an application of the Szemerédi-Trotter theorem to the line counting problem. Two distinct points in \( \mathbb{R}^2 \) determine a line, and so for a set \( P \) of points we can consider the set \( L(P) \) of lines determined by pairs of points in \( P \). This could have just one element, which would happen if all the points in \( P \) lie along a single line as in Figure 1.3 (a). Or it could be as large as \( \binom{|P|}{2} \approx |P|^2 \) if the points are in general position as in Figure 1.3 (b).

The following theorem of Beck \[3\] asserts that, up to multiplicative constants, these two extremes are essentially the only two possibilities.

**Theorem 1.7** (Beck). If \( P \) is a set of points in \( \mathbb{R}^2 \) then at least one of the following holds
1.3. Line counting and Beck’s theorem

Figure 1.3. The two extremes for counting lines.

1. At least $\Omega(|P|)$ points in $P$ are incident to a single line.

2. $|L(P)| \approx |P|^2$.

Beck’s theorem was originally published in the same journal edition as the Szemerédi-Trotter theorem. But it is in fact also a consequence of Szemerédi-Trotter. This latter proof, given here, is more straightforward and can be found for example as exercise 8.2.6 in the book of Tao and Vu [48].

Proof of Theorem 1.7. For each $l \in L(P)$ write $\mu(l) = \sum_{p \in P} \delta_{pl}$ for the number of points $p \in P$ that are incident to $l$. It is clear that

$$\mu(l)^2 \approx \sum_{p_1 \neq p_2 \in P} \delta_{p_1l}\delta_{p_2l}$$

since the right-hand side counts pairs of distinct points in $P$ incident to $l$. It follows that

$$\sum_{l \in L(P)} \mu(l)^2 \approx \sum_{p_1 \neq p_2 \in P} \sum_{l \in L(P)} \delta_{p_1l}\delta_{p_2l}.$$  

For fixed $p_1 \neq p_2$ there is precisely one line in $L(P)$ that is incident to $p_1$ and $p_2$, so

$$\sum_{l \in L(P)} \delta_{p_1l}\delta_{p_2l} = 1,$$

and hence by combining the above equations

$$\sum_{l \in L(P)} \mu(l)^2 \approx |P|^2. \quad (1.9)$$
Chapter 1. Incidence geometry

Now for each integer \( j \) let \( L_j \) be the set of \( l \in L(P) \) for which \( 2^j \leq \mu(l) < 2^{j+1} \). Corollary 1.4 implies that

\[
|L_j| \ll \frac{|P|^2}{2^j} + \frac{|P|}{2^j}.
\]

(1.10)

For points \( p, q \in P \) write \( l_{pq} \) for the line determined by \( p \) and \( q \). For any constant \( C \), let \( X_C \) be the set of pairs of distinct points \( p, q \in P \) for which

\[ C \leq \mu(l_{pq}) \leq \frac{|P|}{C}. \]

From (1.10),

\[
|X_C| \approx \sum_{j=\log_2 C}^{\log_2(|P|/C)} |L_j| 2^{2j}
\]

\[
\ll \sum_{j=\log_2 C}^{\log_2(|P|/C)} \left( \frac{|P|^2}{2^j} + \frac{|P|}{2^j} \right) 2^{2j}
\]

(1.11)

Summing geometric series yields

\[
\sum_{j=\log_2 C}^{\log_2(|P|/C)} \frac{1}{2^j} \ll \frac{1}{C}
\]

\[
\sum_{j=\log_2 C}^{\log_2(|P|/C)} 2^j \ll \frac{|P|}{C}
\]

and so (1.11) implies

\[
|X_C| \ll \frac{|P|^2}{C}.
\]

(1.12)

Comparing (1.9) and (1.12) it is possible to pick a constant \( C \) sufficiently large that

\[
|X_C| \leq \frac{1}{2} \sum_{l \in L(P)} \mu(l)^2.
\]

Thus for this fixed \( C \), either

\[
\sum_{l \in L(P): \mu(l) \geq |P|/C} \mu(l)^2 \gg |P|^2.
\]

or

\[
\sum_{l \in L(P): \mu(l) \leq C} \mu(l)^2 \gg |P|^2
\]
1.4 Incidence geometry in $\mathbb{R}^n$

In the former case there are $\Omega(|P|^2)$ pairs $(p, q)$ for which $\mu(l_{pq}) \geq |P|/C$, and so in particular there is at least one line in $L(P)$ that is incident to $\Omega(|P|)$ points. This corresponds to the first case in the conclusion of the theorem.

In the latter case there are $\Omega(|P|^2)$ pairs of points $p, q \in P$ for which $\mu(l_{pq}) \leq C$, meaning that there are $\Omega(|P|^2)$ distinct lines in $L(P)$. This corresponds to the second case in the conclusion of the theorem. \hfill $\Box$

Since it is a consequence of Szemerédi-Trotter, which extends to points and curves, the statement of Beck’s theorem generalises in the same way. This can be further developed by using the Pach-Sharir theorem in place of Szemerédi-Trotter, for example to the following result.

**Theorem 1.8.** For finite $P \subseteq \mathbb{R}^2$, write $C(P)$ for the set of circles determined by non-collinear triples of points from $P$. For any $P$, at least one of the following holds:

1. There are $\Omega(|P|)$ cocircular points in $P$.
2. There are $\Omega(|P|)$ collinear points.
3. $|C(P)| \gg |P|^3$.

The details of the proof are similar to that of Beck’s theorem, using instead the $k = 3$ case of the Pach-Sharir theorem and the fact that three distinct non-collinear points determine a circle.

### 1.4 Incidence geometry in $\mathbb{R}^n$

It is interesting, and useful for applications to other problems, to consider incidence geometry in higher-dimensional vector spaces than the plane. This section generalises the material of the previous two sections from $\mathbb{R}^2$ to $\mathbb{R}^n$. Since Beck’s line-counting theorem follows from the Szemerédi-Trotter incidence bound, the focus is on higher-dimensional incidence bounds, with their application to line counting in $\mathbb{R}^n$ left as an exercise.

There are several ways to formulate higher-dimensional results. One is to simply consider point-line incidences in the higher-dimensional setting. Another
Chapter 1. Incidence geometry

is to consider higher-dimensional objects than lines, for example incidences between points and planes in \( \mathbb{R}^3 \). Somewhere between these two is the study of ‘pseudolines’ which may be higher dimensional than lines but intersect pairwise in at most \( O(1) \) points.

1.4.1 Points and lines

The Szemerédi-Trotter theorem extends by a random projection argument to the setting where the points and lines lie in \( \mathbb{R}^n \) rather than \( \mathbb{R}^2 \). Simply project \( P \) and \( L \) onto a randomly chosen 2-plane in \( \mathbb{R}^n \). With positive probability the number of incidences between the projected points and projected lines will be the same as those between the originals.

Since Szemerédi-Trotter is sharp in \( \mathbb{R}^2 \) it is also sharp in \( \mathbb{R}^n \). However improvements are available subject to constraints on the arrangements of lines so that, for example, not too many of them lie in any one plane. Guth and Katz \[18\] obtained a breakthrough result of this kind in \( \mathbb{R}^3 \) which enabled them to solve the famous ‘distinct distances’ problem of Erdős. Their incidence result is as follows:

**Theorem 1.9** (Guth, Katz). Let \( L \) be a set of lines in \( \mathbb{R}^3 \) of which no more than \( O \left( \frac{|L|^{1/2}}{k^2} \right) \) are simultaneously incident to the same point, plane or regulus\(^2\). Then the number of points in \( \mathbb{R}^3 \) incident to at least \( k \) lines in \( L \) is at most \( O \left( \frac{|L|^{3/2}}{k^2} \right) \).

The bound in Theorem 1.9 should be compared to the Szmerédi-Trotter bound \( O \left( \frac{|L|^2}{k^2} \right) \) arising from Corollary 1.4 when \( k \ll |L|^{1/2} \). The two bounds agree when \( k \approx |L|^{1/2} \) but Theorem 1.9 is stronger for smaller \( k \).

1.4.2 Points and planes in \( \mathbb{R}^3 \)

Some kind of nondegeneracy condition is necessary to say anything interesting about incidences between a set \( P \) of points and a set \( \Pi \) of planes, since if all the planes in \( \Pi \) intersect along a single line that is incident to all the points in \( P \), then \( I(P, \Pi) = |P||\Pi| \).

\(^2\)A regulus is a ‘doubly-ruled’ surface. That is, every point in a regulus \( R \) is incident to at least two lines entirely contained in \( R \).
1.4. Incidence geometry in $\mathbb{R}^n$

Edelsbrunner, Guibas and Sharir [8] proved the following result under the fairly strong condition that no three planes are collinear. There are also plenty of examples [1, 2, 7, 12] of point-plane incidence results under other nondegeneracy conditions.

**Theorem 1.10** (Edelsbrunner, Guibas, Sharir). *Let $P$ and $\Pi$ be a set of points and planes respectively in $\mathbb{R}^3$. If no three planes are collinear then*

\[
I(P, \Pi) \ll |P|^{4/5}|\Pi|^{3/5} + |P| + |\Pi|.
\]

In the original paper [8] this bound is multiplied by a factor of the form $|P|^{o(1)}|\Pi|^{o(1)}$. However Apfelbaum and Sharir [2] showed that this additional factor can be eliminated with more careful analysis and so the refined version is used here.

Like the Szemerédi-Trotter theorem, the Edelsbrunner-Guibas-Sharir theorem has a standard corollary, the derivation of which follows precisely as for Corollary 1.4.

**Corollary 1.11.** *Let $\Pi$ be a set of planes in $\mathbb{R}^3$, no three of which are collinear. Then the number of points incident to at least $k$ planes in $\Pi$ is $O\left(\frac{|\Pi|^3}{k^2} + \frac{|\Pi|}{k}\right)$.*

1.4.3 Points and pseudolines

A set $P$ of points and a collection $V$ of varieties in $\mathbb{R}^n$ form a **pseudoline system** if

1. Any pair of varieties from $V$ intersect in at most $O(1)$ points in $P$.
2. Any pair of points from $P$ are simultaneously incident to at most $O(1)$ varieties in $V$.
3. Points in $P$ can be incident to varieties in $V$ only at their smooth points.
4. If two varieties intersect at a point in $P$ then their tangent spaces intersect *only* at that point.

Solymosi and Tao [44] proved the following result which generalises the Szemerédi-Trotter theorem, up to a loss of $o(1)$ in the exponent of $|P|$, to pseudoline systems in $\mathbb{R}^n$ for which the varieties are of dimension at most $n/2$.  

33
Chapter 1. Incidence geometry

Theorem 1.12 (Solymosi, Tao). Let \( n \geq 2k \) and let \( P \) and \( V \) form a pseudoline system in \( \mathbb{R}^n \). If the varieties in \( V \) are at most \( k \)-dimensional, and all have degree at most \( O(1) \) then

\[
I(P, V) \ll \varepsilon |P|^{2/3 + \varepsilon} |V|^{2/3} + |P| + |L|.
\]

The case with \( n = 4 \) and \( k = 2 \) is particularly important in the next section, as it corresponds to a Szemerédi-Trotter theorem for complex numbers. In this case, Zahl [51] succeeded in eliminating the \( o(1) \) loss to give the following result.

Theorem 1.13 (Zahl). Let \( P \) and \( V \) form a pseudoline system in \( \mathbb{R}^4 \). If the varieties in \( V \) are all \( 2 \)-dimensional, and all have degree at most \( O(1) \) then

\[
I(P, V) \ll |P|^{2/3} |V|^{2/3} + |P| + |L|.
\]

1.5 Other fields

So far we have considered only real vector spaces. But we are also interested in other settings, in particular complex numbers and finite fields. Chapter 4 examines incidence geometry over finite fields and proves several original results, so we keep our powder dry and consider only the complex setting here.

Tóth [49] announced the following generalisation of the Szemerédi-Trotter theorem to \( \mathbb{C}^2 \) in 2003, but nearly ten years later it is still in the review process.

Theorem 1.14 (Tóth). If \( P \) is a set of points in \( \mathbb{C}^2 \) and \( L \) is a set of lines, then

\[
I(P, L) \ll |P|^{2/3} |L|^{2/3} + |P| + |L|.
\]

The usual corollary to Szmerédi-Trotter applies here as well:

Corollary 1.15. Let \( L \) be a set of lines in \( \mathbb{C}^2 \). Then the number of points incident to at least \( k \) lines in \( L \) is \( O \left( \frac{|L|^2}{k^3} + \frac{|L|}{k} \right) \). Similarly, the number of lines incident to at least \( k \) points in \( P \) is \( O \left( \frac{|P|^2}{k^3} + \frac{|P|}{k} \right) \).

More recently, the work of Solymosi, Tao and Zahl on pseudolines yields Tóth’s result as a special case. Indeed, lines in \( \mathbb{C}^2 \) are pseudolines in \( \mathbb{R}^4 \) and so Theorem 1.14 is an immediate corollary of Theorem 1.13.

\[^3\]The 2011 preprint referenced here is a more recent version of the 2003 original.
1.5. Other fields

One can of course then consider complex versions of the higher-dimensional generalisations in Section 1.4. Theorem 1.14 extends to point-line incidences $C^n$, in the same way as Szemerédi-Trotter, although it is not known if the Guth-Katz theorem for $R^3$ generalises to $C^3$. 
2

Growth

Just as the last chapter did for incidence geometry, the purpose here is to introduce the two examples of growth highlighted in the prologue: sum-product estimates and expander functions. The chapter also records particular results that will be useful in Chapter 4.

The idea for sum-product estimates is that, for a subset $A$ of a field, at least one of the sumset $A + A$ and the product set $AA$ must always be large. Chapter 4 will use finite field sum-product estimates as a key ingredient to proving new finite field incidence theorems. Chapter 6 will establish a sum-product theorem in function fields, a relatively unexplored setting for the problem.

Expander functions are slightly different from sum-products. Rather than considering both the sumset and the product set, the concern is with the cardinality of a single image set $f(A)$ of some multivariate function $f$. A medley of new results on expander functions are proved in Chapter 5.

All of the mathematics in this chapter is drawn from the literature.

2.1 Sum-product estimates

Let $A$ be a finite subset of a field $F$, and consider the cardinalities of the sumset

$$A + A = \{a + b : a, b \in A\}$$

and the product set

$$AA = \{ab : a, b \in A\}.$$
Chapter 2. Growth

It is clear that both these sets have at least the same cardinality as $A$. And it is certainly possible to have $|A + A| \approx |A|$ or $|AA| \approx |A|$, for example if $A$ is either an arithmetic or geometric progression. But the idea of sum-product estimates is that it is not possible to have both of these at once. Results are therefore of the form

$$\max \{|A + A|, |AA|\} \gg |A|^{1+\delta}$$

with $\delta > 0$, holding for any finite set $A \subseteq F$ that possibly satisfies some nondegeneracy conditions. Larger values of $\delta$ correspond to stronger results.

Sum-products were first studied by Erdős and Szemerédi [13], who proved the existence of an unquantified $\delta > 0$ holding for any finite set $A$ of integers. They conjectured that $\delta \geq 1 - o(1)$. It is not possible to remove the $o(1)$ from this conjecture, as demonstrated by the case where $A$ is the set of integers between 1 and $N$; in this case $|A + A| \approx |A|$ but $|AA| \approx |A|^{2-o(1)}$.

Sum-product estimates can of course be considered in the more general setting of rings instead of fields, and indeed Erdős and Szemerédi’s initial work concerned the ring of integers. However, when working over general rings the sum-product phenomenon is constrained by the possibility of zero divisors. If $A$ contains too many then the sumset and the product set might both be small. The motivation for sticking with fields in this thesis is therefore that avoiding zero-divisor considerations means working with an integral domain, and when given an integral domain one may as well work with its field of fractions. Tao [47] has recently obtained a general, non-explicit, sum-product result for rings in the absence of too many zero divisors.

2.1.1 Real and complex sum-products

The strongest known sum-product estimate is $\delta \geq \frac{1}{3} - o(1)$, which was proved for any finite set of reals by Solymosi [43] and recently generalised to any finite set of complex numbers by Konyagin and Rudnev [30]. There is a fairly extensive list [9, 14, 34, 41, 42] of previous results in the real and complex settings.

The incidence geometry of Chapter 1 yields a lot of information about sum-products via the Szemerédi-Trotter theorem. An example, not quite as strong as the more recent Solymosi and Konyagin-Rudnev bounds, is the following result of Elekes [9]. A more involved version of this approach can be found in another paper of Solymosi [42].
2.1. Sum-product estimates

**Theorem 2.1** (Elekes). Let \( A \subseteq \mathbb{C} \) be finite. Then

\[
\max \{|A + A|, |AA|\} \gg |A|^{1 + \frac{1}{4}}.
\]

**Proof.** Define a point set \( P \) and a line set \( L \) by

\[
P = (A + A) \times (AA)
\]
\[
L = \{l_{ab} : a, b \in A\}
\]

where \( l_{ab} \) is the line given by \( y = a(x - b) \). It is clear that

\[
|P| = |A + A||AA|
\]
\[
|L| = |A|^2.
\]

Each line \( l_{ab} \in L \) is incident to at least \( |A| \) points in \( P \) since for each \( c \in A \) the point \((b + c, ac)\) is an element of \( P \cap l_{ab} \). Thus

\[
I(P, L) \geq |L||A| = |A|^3.
\]

Applying the the complex Szemerédi-Trotter theorem (Theorem 1.14) shows that

\[
I(P, L) \ll |P|^{2/3}|L|^{2/3} + |P| + |L|
\]
\[
\approx |A + A|^{2/3}|AA|^{2/3}|A|^{1/3} + |A + A||AA| + |A|^2.
\]

Comparing the upper and lower bounds on \( I(P, L) \) yields

\[
|A|^3 \ll |A + A|^{2/3}|AA|^{2/3}|A|^{1/3} + |A + A||AA|.
\]

So either

\[
|A|^3 \ll |A + A|^{2/3}|AA|^{2/3}|A|^{4/3}
\]
or

\[
|A|^3 \ll |A + A||AA|.
\]

In the event of the former, rearranging gives

\[
\max \{|A + A|, |AA|\} \gg |A|^{5/4}.
\]

On the other hand, if the latter holds then

\[
\max \{|A + A|, |AA|\} \gg |A|^{3/2} \geq |A|^{5/4}.
\]

Either way the proof is complete. \( \square \)
2.1.2 Finite field sum-products

Sum-product results for finite fields must be prefaced with a nondegeneracy condition, to prevent the set \( A \) being too close to a copy of a subfield. For example if \( A \) is a field then \( \max\{ |A + A|, |AA| \} = |A| \) and so non-trivial estimates are impossible. To keep this issue as straightforward as possible it is often convenient to concentrate on the finite field \( \mathbb{F}_p \) of prime order \( p \), where it collapses into the requirement for an upper bound on \( |A| \) in terms of \( p \).

Let’s now see what estimates are known in \( \mathbb{F}_p \). The breakthrough result was due to Bourgain, Katz and Tao \([6]\) who proved the existence of an absolute \( \delta > 0 \) dependent on \( \gamma > 0 \) that holds whenever \( |A| < p^{1-\gamma} \).

Garaev \([15]\) made the Bourgain-Katz-Tao result explicit, obtaining \( \delta \geq \frac{1}{14} - o(1) \) whenever \( |A| < p^{1/2} \). This has been subsequently improved by a variety of authors \([5, 20, 31, 40, 37]\). The most recent estimate is \( \delta > \frac{1}{11} - o(1) \) whenever \( |A| < p^{1/2} \), due to Rudnev \([37]\). Li and Roche-Newton \([33]\) built on a technique of Katz and Shen \([28]\) to extend this estimate to any finite field, not necessarily of prime order, so long as \( A \) is not too close to being a subfield.

Chapter 5 will use a particular form of Rudnev’s \( \frac{1}{11} - o(1) \) result. Its statement involves the multiplicative energy \( E_\times(A) \) of \( A \), which is the number of solutions to \( ab = cd \) with \( a, b, c, d \in A \); this is developed more fully in Chapter 3. As Rudnev remarks in his paper, the result is really an upper bound on \( E_\times(A) \) in terms of the sumset \( A + A \), and works just as well when this is replaced with the difference set \( A - A \). In these terms, the estimate can be formulated as follows.

**Theorem 2.2** (Rudnev). Let \( A \subseteq \mathbb{F}_p \). If \( |A| \ll p^{1/2} \) then

\[
E_\times(A)^4 \lesssim |A - A|^7 |A|^4.
\]

The estimate \( \delta \geq \frac{1}{11} - o(1) \) follows from the fact (again, see Chapter 3) that \( E_\times(A) \) is greater than or equal to \( \frac{|A|^4}{|A|} \).

The proofs of all the above finite field results, including Theorem 2.2, take the same overall approach, sketched below.

**Sketch of finite field sum-product proofs.** Given a set \( A \subseteq \mathbb{F}_p \) consider the set

\[
R(A) = \left\{ \frac{a - b}{c - d} : a, b, c, d \in A, a \neq b, c \neq d \right\}.
\]
2.1. Sum-product estimates

Note that if \( \xi \notin R(A) \) then
\[
|A + \xi A| = |A|^2
\]
since different pairs of elements from \( A \) give different elements of \( A + \xi A \). Indeed if there were a solution to to
\[
a + \xi b = c + \xi d.
\]
with \((a, b) \neq (c, d)\) then rearranging would yield the contradiction
\[
\xi = \frac{a - c}{b - d} \in R(A).
\]

Now, assume that \( R(A) \neq \mathbb{F}_p \). This is immediate if \(|A| < p^{1/4}\), and with a little effort we may assume something very similar so long as \(|A| \ll p^{1/2}\). Since \( \mathbb{F}_p \) contains no non-trivial additive subgroups, there must exist
\[
\frac{a - b}{c - d} \in R(A)
\]
such that
\[
\frac{a - b}{c - d} + 1 \notin R(A).
\]
Therefore
\[
|A|^2 = \left| A + \left( \frac{a - b}{c - d} + 1 \right) A \right|
\]
\[
= \left| (c - d) A + (a - b + c - d) A \right|
\]
\[
\leq |AA - AA + AA - AA + AA - AA|.
\]
Thus an iterated sumset of the product set \( AA \) must be large. With some work, and analysing multiplicative energy rather than the product set directly, it turns out that the only way this can happen is if either \( AA \) is large itself or \( A + A \) is large.

As a final remark for this section, it is worth thinking about what a reasonable sum-product conjecture might be for finite fields. Since the conjecture in the real case is \( \delta \geq 1 - o(1) \), it might be natural to suppose that this should hold over \( \mathbb{F}_p \) whenever \(|A| < p^{1/4}\). After all, this condition ensures there is no danger of the sumset or product set filling up the whole field. However the following result of Bourgain [4], refined by Garaev [15], shows that under this constraint the best that can be hoped for is \( \delta \geq \frac{1}{2} \). To obtain \( \delta \geq 1 - o(1) \) will require \(|A| \lesssim p^{1/3}\) as a necessary condition.
Lemma 2.3 (Bourgain, Garaev). For any prime $p$ and any integer $1 \leq N \leq p$ there is a set $A \subseteq \mathbb{F}_p$ with $|A| \approx N$ such that
\[ \max \{|A + A|, |AA|\} \ll N^{1/2}p^{1/2}. \]
In particular there is a set $A \subseteq \mathbb{F}_p$ with $|A| \approx p^{1/2}$ such that
\[ \max \{|A + A|, |AA|\} \ll |A|^{1+\frac{1}{2}}. \]

Proof. Let $g \in \mathbb{F}_p^*$ be a generator of $\mathbb{F}_p^*$. Let $M = p^{1/2}N^{1/2}$. By rearranging orders of summation it is immediate that
\[ \sum_{y \in \mathbb{F}_p} \# (\{g^n : 1 \leq n \leq M\} \cap \{y + j : 1 \leq j \leq M\}) \approx M^2. \]

Therefore there exists $y \in \mathbb{F}_p$ such that
\[ \# (\{g^n : 1 \leq n \leq M\} \cap \{y + j : 1 \leq j \leq M\}) \gg \frac{M^2}{p} = N. \]
It is therefore possible to pick a set $A$ contained in the set on the left hand side such that $|A| \approx N$. Since $A \subseteq \{g^n : 1 \leq n \leq M\}$,
\[ |AA| \ll M. \]
And since $A \subseteq \{y + j : 1 \leq j \leq M\}$,
\[ |A + A| \ll M. \]
So altogether,
\[ \max \{|A + A|, |AA|\} \ll M = N^{1/2}p^{1/2} \]
as required. 

As explained in the prologue, we are not too worried in this thesis about the ‘large-set’ finite field paradigm, where sets are subject to a minimum as well as a maximum density requirement. However it is worth noting that under this regime, with $|A| > p^{2/3}$, Garaev [15] obtained the bound
\[ \max \{|A + A|, |AA|\} \gg N^{1/2}p^{1/2} \]
which is optimal in view of Lemma 2.3.
2.2 Expander functions

Expander functions are another example of growth, and a natural variation on sum-products. An \( n \)-variable expander function is a function \( f : \mathbb{F}^n \to \mathbb{F} \) for which the set

\[
f(A) = \{ f(a_1, \ldots, a_n) : a_i \in A \}
\]

has cardinality at least \( \Omega(|A|^{1+\delta}) \) for some \( \delta > 0 \) and any \( A \subseteq F \), again possibly satisfying some nondegeneracy conditions. Stronger expander results correspond not only to a larger growth exponent \( \delta \) as per sum-product estimates, but also to a smaller number \( n \) of variables.

Sum-product estimates give a trivial method of constructing four-variable expander functions, because it is immediate that

\[
|\{ab + cd : a, b, c, d \in A\}| \geq \max\{|AA|, |A + A|\}.
\]

However it is often possible to do better, both in terms of fewer variables and bigger exponents.

In the two-variable case, Elekes, Nathanson and Ruzsa [10] obtained \( \delta \geq \frac{1}{4} \) for the function

\[
a + \frac{1}{b}
\]

whenever \( A \) is a finite set of reals. This was subsequently extended to \( \delta \geq \frac{5}{19} \) by Li and Roche-Newton [32], and was also considered by Bourgain [4] in the prime finite field setting for an absolute \( \delta > 0 \). The same paper of Bourgain also gives the example of the function

\[
f(a, b) = a^2 + ab
\]

which has the peculiar property of not even requiring \( a \) and \( b \) to be drawn from the same set in order to establish a growth result. Hart, Li and Shen [19] obtained \( \delta \geq \frac{1}{146} \) over finite fields for the function

\[
a + b^2.
\]

However, the strongest-known two-variable expander in the real, complex and finite field setting is the function

\[
a + ab.
\]
This function was first studied by Garaev and Shen [16], and a new result is established in Chapter 5.

In the four-variable case, an expander follows from the recent breakthrough of Guth and Katz [18] on the Erdős distinct distance problem. They showed that a finite point-set $P \subseteq \mathbb{R}^2$ determines at least $\Omega\left(|P|^{1-o(1)}\right)$ distinct pairwise distances. In the particular case $P = A \times A$ this shows that

$$(a - b)^2 + (c - d)^2$$

is a four-variable expander over $\mathbb{R}$ with $\delta > 1 - o(1)$. This methodology was also adapted by Iosevich, Roche-Newton and Rudnev [21] to show that

$$ad - bc$$

is likewise a four-variable expander with $\delta = 1 - o(1)$. New, stronger, results on three and four-variable expanders over $\mathbb{R}$ and $\mathbb{C}$ are proved in Chapter 5.
A handbook on sumset calculus

This chapter comprises a useful arsenal of technical results for analysing sumsets and product sets. Everything is stated in terms of addition and sumsets, but extends without complication to multiplication and product sets.

The basic tools of sumset calculus are the Plünnecke-Ruzsa inequalities, which relate the cardinalities of different sumsets. These will be of use in Chapters 4, 5 and 6.

Partial sumsets are a generalisation of sumsets, where interest is restricted to sums determined by only a subset of possible pairs. Balog-Szemerédi-Gowers type results reduce the study of partial sumsets to that of complete sumsets, but at the price of reduced efficiency. Two such results will come in handy in Chapters 4 and 5.

The additive and multiplicative energy of a set were briefly touched on in Chapter 2 but the treatment here is more developed. Energy is useful for two reasons. First, sets with small sumset have high energy, which is helpful to know when analysing their structure. Second, energy-based arguments go through just as easily for partial as complete sumsets, and so can in some cases be more efficient than a Balog-Szemerédi-Gowers approach. This is useful in Chapters 4 and 5.

Energy considerations also have a particular application in covering methods. When considering complete sumsets, these can sometimes more efficient than the Plünnecke-Ruzsa inequalities. And since they are energy-based, they apply equally well to complete and partial sumsets, and so in some cases are stronger than Balog-Szemerédi-Gowers type results. These will be useful in Chapter 5.
Except for Lemma 3.14 and Lemma 3.15, which are original variations on covering lemmata, all of the mathematics in this chapter is drawn from the literature. However it should be remarked that the very useful Lemma 3.4 does not seem to be widely known or used in the research community.

3.1 Introduction

Chapter 2 defined the sumset $A + A$ and product set $AA$ of a subset $A$ of a field. This extends naturally to the sumset $A + B$ and product set $AB$ of two different sets $A, B$, defined as

$$A + B = \{a + b : a \in A, b \in B\}$$

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

Define also the difference set

$$A - B = A + (-B)$$

and ratio set

$$A/B = A(B \setminus \{0\})^{-1}.$$ 

The essential problem of sumset calculus is that we are given a pair of sets $A, B$ and another related pair $C, D$. Given information about the sum, difference, product or ratio set of $A$ and $B$, what can be said about that of $C$ and $D$?

The results in this chapter are general, in the sense that they depend only on the fact that sets are finite and contained in an abelian group. So they are quoted and proved for sumsets $A + B$, taking $+$ to be an abelian group operation rather than necessarily the additive group of a field. They therefore extend to multiplication and product sets without complication. In some instances it will be convenient to adopt slightly different notation in the multiplicative case to avoid confusion in later chapters; this will be indicated where appropriate.

3.2 The Plünnecke-Ruzsa inequalities

Information on sumsets is often expressed in terms of cardinalities. There are trivial estimates

$$\max(|A|, |B|) \leq |A + B| \leq |A||B|.$$
3.2. The Plünnecke-Ruzsa inequalities

Informally, a sumset set is ‘small’ if its cardinality is close to the lower bound \( \max(|A|,|B|) \) and ‘large’ if it is close to the upper bound \(|A||B|\). The ‘efficiency’ of an estimate corresponds to the number of factors of the form \(|A + B|\) which turn up; the fewer the better since this allows the proof of stronger theorems.

The most basic tool available for analysing sumset cardinalities is the Ruzsa triangle inequality, which applies to difference sets but not immediately to sumsets.

**Lemma 3.1** (Ruzsa triangle inequality). For sets \( A, B, C \) we have

\[
|A - B| \leq \frac{|A - C||B - C|}{|C|}.
\]

**Proof.** For each element \( x \in A - B \), pick \( a_x \in A \) and \( b_x \in B \) such that

\[
x = a_x - b_x.
\]

Consider the map \( f \) from \((A - B) \times C\) to \((A - C) \times (B - C)\) given by

\[
f(x, c) = (a_x - c, b_x - c).
\]

We show that \( f \) is an injection. Suppose that \( f(x, c) = f(y, d) \). Then

\[
a_x - c = a_y - d \tag{3.1}
\]

\[
b_x - c = b_y - d \tag{3.2}
\]

Subtracting \(\text{(3.2)}\) from \(\text{(3.1)}\) gives \(a_x - b_x = a_y - b_y\) and so \(x = y\), which in turn forces \(a_x = a_y\) and \(b_x = b_y\). Plugging this back into \(\text{(3.1)}\) shows that \(c = d\), and so \((x, c) = (y, d)\). Thus \( f \) is an injection and

\[
|A - B||C| \leq |A - C||B - C|.
\]

Rearranging gives the required inequality. \( \square \)

Note that the essential observation used in the above proof is that

\[
a - b = (a - c) - (b - c)
\]

for any \(a, b, c\). Variations on this approach are used several more times in the thesis, and are referred to as ‘Ruzsa-type’.
The Ruzsa triangle inequality is a simple and useful tool, but it applies only to difference sets and it is often necessary to care about general sumsets. This can be mitigated by clever applications, for example in the case $A = B$ taking $C = -A$ gives

$$|A - A| \leq \frac{|A - (-A)||(-A) - A|}{|A|} = \frac{|A + A|^2}{|A|}$$

but this is not always terribly efficient. Moreover the Ruzsa triangle inequality does not help very much when analysing iterated sumsets of the form

$$kA = \underbrace{A + \ldots + A}_k.$$

Such situations call for Plünnecke’s theorem, which was promulgated in this context by Ruzsa [38]. A much shorter proof was recently given by Petridis [36], for which we follow an exposition given by Gowers [17].

**Lemma 3.2** (Plünnecke). For sets $A, B$ there exists $A' \subseteq A$ such that for any natural number $k$,

$$|A' + kB| \leq \frac{|A'||A + B|^k}{|A|^k}. \quad (3.3)$$

In particular,

$$|kB| \leq \frac{|A + B|^k}{|A|^{k-1}}. \quad (3.4)$$

**Proof.** Pick $A' \subseteq A$ such that

$$\frac{|A' + B|}{|A'|} = K$$

is minimal across all subsets of $A$. It suffices to prove that

$$|A' + kB| \leq |A'|K^k \quad (3.5)$$

since by minimality in the choice of $K$,

$$K \leq \frac{|A + B|}{|A|}.$$

To establish (3.5) it suffices in turn to show that

$$|A' + B + C| \leq K|A' + C| \quad (3.6)$$
3.2. The Plünnecke-Ruzsa inequalities

for any set $C$. Indeed, once (3.6) is established, inequality (3.5) follows by induction on $k$, since taking $C = (k - 1)B$ gives

$$|A' + kB| \leq K|A' + (k - 1)B|$$

and by inductive hypothesis

$$|A' + (k - 1)B| \leq K^{k-1}|A'|.$$

Now to prove (3.6). Induction is used here as well, this time on $|C|$. When $|C| = 1$ it is immediate that $|D + C| = |D|$ for any set $D$. In particular,

$$|A' + B + C| = |A' + B|$$

and

$$K|A' + C| = K|A'|$$

so (3.6) holds with equality by definition of $K$. Now suppose that $|C| \geq 2$ and write

$$C = C' \cup \{x\}$$

so that by inductive hypothesis,

$$|A' + B + C'| \leq K|A' + C'|.$$

By inclusion-exclusion,

$$|A' + B + C| = |A' + B + (C' \cup \{x\})|$$

$$= |A' + B + C'| + |(A' + B + \{x\}) \setminus (A' + B + C')|$$

$$= |A' + B + C'| + |A' + B + \{x\}|$$

$$- |(A' + B + \{x\}) \cap (A' + B + C')|$$

$$\leq |A' + B + C'| + |A' + B| - |(A' + \{x\}) \cap (A' + C') + B|$$

$$= |A' + B + C'| + |A' + B| - |A' \cap (A' + C' - \{x\}) + B|$$

$$\leq K(|A' + C'| + |A'| - |A' \cap (A' + C' - \{x\})|)$$

(3.7)

On the other hand,

$$|A' + C| = |A' + C'| + |(A' + \{x\}) \setminus (A' + C')|$$

$$= |A' + C'| + |A' + \{x\}| - |(A' + \{x\}) \cap (A' + C')|$$

$$= |A' + C'| + |A'| - |A' \cap (A' + C' - \{x\})|$$

(3.8)
and so substituting (3.8) into (3.7) gives

$$|A' + B + C| \leq K|A' + C|$$

as required. \(\square\)

Lemma 3.2 is typically applied in the form (3.4) since this omits mention of the subset \(A' \subseteq A\). However Katz and Shen [29] observed that with a little work it is possible to show that \(|A'| \approx |A|\), which can in some circumstances make (3.3) a better bet.

**Corollary 3.3 (Katz, Shen).** For sets \(A, B\), there exists \(A' \subseteq A\) with \(|A'| \geq |A|^{\frac{3}{2}}\) such that

$$|A' + kB| \ll_k \frac{|A + B|^k}{|A|^{k-1}}.$$  

In situations where nothing is lost by passing to a constant-proportion subset of \(A\), Corollary 3.3 effectively implies that

$$|kA| \ll_k \frac{|A + A|^{k-1}}{|A|^{k-2}},$$

which is better than Lemma 3.2 by a factor of \(\frac{|A + A|}{|A|}\).

**Proof of Corollary 3.3.** Let \(A_*\) be any subset of \(A\) with \(|A_*| \geq |A|^{\frac{3}{2}}\). By Lemma 3.2 there exists \(A'_* \subseteq A_*\) such that

$$|A'_* + kB| \leq \frac{|A'_*||A_* + B|^k}{|A_*|^k} \ll_k \frac{|A'_*||A + B|^k}{|A|^k}.$$  

Apply the above observation recursively. Begin by taking \(A_* = A\) to find \(A_1 \subseteq A\) such that

$$|A_1 + kB| \ll_k \frac{|A_1||A + B|^k}{|A|^k}.$$  

Now take \(A_* = A \setminus A_1\) to find \(A_2 \subseteq A\) disjoint from \(A_1\) such that

$$|A_2 + kB| \ll_k \frac{|A_2||A + B|^k}{|A|^k}.$$  

Repeat this process until reaching \(n\) such that \(\bigcup_{i=1}^n A_i\) is of cardinality at least \(|A|^{\frac{3}{2}}\). Let

$$A' = \bigcup_{i=1}^n A_i.$$
3.3. Partial sumsets

Since all of the $A_i$ are disjoint it follows that $|A'| \geq \frac{|A|}{2}$ and

$$|A' + kB| \leq \sum_{i=1}^{n} |A_i + kB|$$

$$\ll_k \frac{|A + B|^k}{|A|^k} \sum_{i=1}^{n} |A_i|$$

$$\approx \frac{|A + B|^k}{|A|^{k-1}}$$

as required. \qed

3.3 Partial sumsets

For sets $A, B$ and $G \subseteq A \times B$, call the set

$$A^G + B = \{a + b : (a, b) \in G\}$$

a partial sumset. In the case where $G = A \times B$ this collapses into the sumset $A + B$, referred to as the complete sumset where there is a possibility for confusion. The motivation for using the letter $G$ is that $G \subseteq A \times B$ corresponds to (the edges of) a bipartite graph connecting $A$ and $B$.

Analysing partial sumsets is critical to the research in Chapters 4 and 5, but the Plünnecke-Ruzsa inequalities developed above apply only to complete sumsets and so are not immediately useful. Fortunately, it turns out that if $G$ is a reasonably large part of $A \times B$ and the partial sumset $A^G + B$ is small, then there are large subsets of $A$ and $B$ whose complete sumset is small as well. Results of this kind are called Balog-Szemerédi-Gowers type or BSG-type after the standard theorem in this area, due to Balog and Szemerédi and strengthened by Gowers.

Despite their usefulness, Balog-Szemerédi-Gowers type results are disproportionately costly. This is in the sense that they yield upper bounds on complete sumsets that are out of proportion to the density of $G$ in $A \times B$, and in particular they pick up more than a constant factor when $|G| \approx |A||B|$. Sections 3.3

---

1See Theorem 2.29 of [48] for a formulation and proof of the Balog-Szemerédi-Gowers theorem. Note however that the factor $K^4$ in equation (2.20) of that formulation should be replaced with $K^5$ due to an error in the text.
and 3.5 describe methods which do not see this distinction and so can in some circumstances be more efficient.

This thesis uses two BSG-type results. The first result is more efficient but is applicable only when $G$ is especially dense in $A \times B$. It is a consequence of Exercise 2.5.4 in [48].

**Lemma 3.4 (BSG-type for dense sets).** Let $0 < \epsilon < 1/4$ and let $G \subseteq A \times B$ with $|G| \geq (1 - \epsilon)|A||B|$. Then there exists $A' \subseteq A$ with $|A'| \geq (1 - \sqrt{\epsilon})|A|$ such that

$$|A' - A'| \ll \frac{|A - B|^2}{|A|}.$$ 

Lemma 3.4 does not seem to be widely used in the literature. But, where it is applicable, it is remarkably efficient for a BSG-type result.

The second BSG-type result, due to Bourgain and Garaev [5], is less efficient but is on the other hand applicable in a wider variety of situations.

**Lemma 3.5 (BSG-type for less-dense sets).** For sets $A, B$ and $G \subseteq A \times B$ there exists $A' \subseteq A$ with $|A'| \gg \frac{|G|}{|B|}$ such that

$$|A' - A'| \ll \frac{|A - B|^3}{|A\ G - B|^4}.$$ 

There is considerable overlap between the proofs of Lemma 3.4 and Lemma 3.5. For both it is convenient to adopt the following definitions.

For $a \in A$, write $B_G(a)$ for the set of $b \in B$ such that $(a, b) \in G$. Call the cardinality of $B_G(a)$ the $G$-degree of $a$, and the cardinality of the intersection $B_G(a_1) \cap B_G(a_2)$ the joint $G$-degree of $a_1$ and $a_2$.

For both results it suffices to find a large subset of $A$ with the property that the joint $G$-degree of any two distinct elements is large. This is demonstrated by the following Ruzsa-type lemma.

**Lemma 3.6.** Let $G \subseteq A \times B$. Suppose that $A' \subseteq A$ and $H \subseteq A' \times A'$ are such that every the joint $G$-degree of any $(a_1, a_2) \in H$ is at least $K$. Then

$$|A' - A'| \leq \frac{|A - B|^2}{K}.$$ 

52
3.3. Partial sumsets

In particular, if the joint $G$-degree of any pair of elements from $A'$ is at least $K$ then

$$|A' - A'| \leq \frac{|A - B|^2}{K}.$$  

Proof. For each $x \in A^H - A'$ pick $a_1^x, a_2^x \in A'$ such that

$$a_1^x - a_2^x = x.$$  

Let $Y \subseteq (A^H - A') \times B$ be given by

$$Y = \{(x, b) : b \in B_G(a_1^x) \cap B_G(a_2^x)\}.$$  

By hypothesis,

$$|Y| \geq |A^H - A'| K.$$  

On the other hand, the injection

$$(x, b) \mapsto (a_1^x - b, a_2^x - b)$$  

from $Y$ into $(A^G - B) \times (A^G - B)$ shows that

$$|Y| \leq |A^G - B|^2.$$  

Comparing the upper and lower bounds on $|Y|$ gives the result.  

The ease of finding sets satisfying the conditions of Lemma 3.6 depends on the density of $G$, corresponding to the two different BSG-type results. Over the next few pages, Section 3.3.1 builds the proof of Lemma 3.4 and Section 3.3.2 builds the proof of Lemma 3.5.  

3.3.1 High-density partial sumsets  

This section proves Lemma 3.4, the BSG-type result for high-density partial sumsets. In view of Lemma 3.6 this is accomplished by the following.

Lemma 3.7. Let $\epsilon > 0$ and $G \subseteq A \times B$ with $|G| \geq (1 - \epsilon)|A||B|$. There exists $A' \subseteq A$ with $|A'| \geq (1 - \sqrt{\epsilon})|A|$ such that the $G$-degree of every $a \in A'$ is at least $(1 - \sqrt{\epsilon})|B|$.  

In particular, any two elements of $A'$ have joint $G$-degree at least $(1 - 2\sqrt{\epsilon})|B|$ whenever $\epsilon < 1/4$.  

53
Applying Lemma 3.6 to the conclusion of Lemma 3.7 with \( H = A' \times A' \) and \( K = (1 - 2\sqrt{\epsilon})|B| \) immediately yields the statement of Lemma 3.4.

**Proof of Lemma 3.7.** Let \( A' \subseteq A \) be the set of \( a \in A \) with \( G \)-degree at least \((1 - \sqrt{\epsilon})|B|\). It suffices to show that \(|A'| \geq (1 - \sqrt{\epsilon})|A|\). To this end, observe that

\[
(1 - \epsilon)|A||B| \leq |G| = \sum_{a \in A} |B_G(a)| \\
\leq \sum_{a \in A'} |B| + \sum_{a \in A \setminus A'} (1 - \sqrt{\epsilon})|B| \\
= |A'||B| + (|A| - |A'|)(1 - \sqrt{\epsilon})|B|.
\]

It follows that

\[
(1 - \epsilon)|A| \leq |A'| + (|A| - |A'|)(1 - \sqrt{\epsilon})
\]

and so \(|A'| \geq (1 - \sqrt{\epsilon})\) as required. \(\square\)

### 3.3.2 Less-dense partial sumsets

This section proves Lemma 3.5, the BSG-type result for less-dense partial sumsets. The proof is not quite as direct as that for Lemma 3.4, and the main additional ingredient is the following lemma.

**Lemma 3.8.** Let \( G \subseteq A \times B \). Then for any \( \epsilon > 0 \) there exists \( A' \subseteq A \) with \(|A'| \gg \frac{|G|}{|B|}\), and \( H \subseteq A' \times A' \) with \(|H| \geq (1 - \epsilon)|A'|^2\), such that the joint \( G \)-degree of any \((a_1, a_2) \in H\) is at least \(\frac{\epsilon|G|^2}{2|A|^2|B|}\).

**Proof.** Let \( H \) be the set of \((a_1, a_2) \in A \times A\) with joint \( G \)-degree at least \(\frac{\epsilon|G|^2}{2|A|^2|B|}\). To prove the lemma it suffices to find \( A' \subseteq A \) with \(|A'| \gg \frac{|G|}{|B|}\) such that

\[
|(A' \times A') \cap H| \geq (1 - \epsilon)|A'|^2.
\]

It is clear that

\[
\sum_{(a_1, a_2) \notin H} |B_G(a_1) \cap B_G(a_2)| < \frac{\epsilon|G|^2}{2|B|}
\]

and hence

\[
\sum_{a_1, a_2 \in A} |B_G(a_1) \cap B_G(a_2)| \frac{1_{((a_1, a_2) \notin H)}}{\epsilon} < \frac{|G|^2}{2|B|}. \tag{3.9}
\]
3.3. Partial sumsets

Also, \( \sum_{b \in B} |B_G(b)| = |G| \) and so by Cauchy-Schwarz

\[
\sum_{a_1, a_2 \in A} |B_G(a_1) \cap B_G(a_2)| \geq \frac{|G|^2}{|B|}.
\]

Comparing (3.9) and (3.10) shows that

\[
\sum_{a_1, a_2 \in A} |B_G(a_1) \cap B_G(a_2)| \left( 1 - \frac{1}{\epsilon} \left( (a_1, a_2) \notin H \right) \right) \geq \frac{|G|^2}{2|B|}.
\]

Rearranging gives

\[
\sum_{b \in B} \sum_{a_1, a_2 \in A} \left( 1 - \frac{1}{\epsilon} \left( (a_1, a_2) \notin H \right) \right) \geq \frac{|G|^2}{2|B|^2}.
\]

So there exists \( b \in B \) such that

\[
\sum_{a_1, a_2 \in A} \left( 1 - \frac{1}{\epsilon} \left( (a_1, a_2) \notin H \right) \right) \geq \frac{|G|^2}{2|B|^2}.
\]

Evaluating the left hand side shows

\[
|A_G(b)|^2 - \frac{|(A_G(b) \times A_G(b)) \setminus H|}{\epsilon} \geq \frac{|G|^2}{2|B|^2}.
\]

Since the left hand side is at most \( |A_G(b)|^2 \) it follows that \( |A_G(b)| \gg \frac{|G|}{|B|} \). Take \( A' = A_G(b) \) so that

\[
|(A' \times A') \setminus H| \leq \epsilon \left( |A'|^2 - \frac{|G|^2}{2|B|^2} \right) \leq \epsilon |A'|^2.
\]

This implies \( |(A' \times A') \cap H| \geq (1 - \epsilon)|A'|^2 \) as required. \( \square \)

With this established, it is now possible to prove Lemma 3.5.

Proof of Lemma 3.5: Let \( \epsilon > 0 \) be sufficiently small, and fixed. By Lemma 3.8 there exists \( A' \subseteq A \) and \( H \subseteq A' \times A' \) with \( |A'| \gg \frac{|G|}{|B|} \) and \( |H| \geq (1 - \epsilon)|A'|^2 \), such that any \( (a_1, a_2) \in H \) have joint \( G \)-degree at least \( \Omega \left( \frac{|G|^2}{|A' |^2 |B|} \right) \). Lemma 3.6 implies

\[
|A' \setminus A' \setminus H| \ll \frac{|A' \setminus A \setminus H|^2 |A'|^2 |B|}{|G|^2}.
\]
By Lemma 3.7 there is a subset \( A'' \subseteq A' \) with \( |A''| \gg |A'| \) such that any two \( a_1, a_2 \in A'' \) have joint \( H \)-degree at least \( (1 - 2\sqrt{\epsilon})|A'| \). A final application of Lemma 3.6 shows that

\[
|A'' - A''| \ll \frac{|A' - A'|^2}{|A'|} \ll \frac{|A|^4 |B|^3 |A - B|^4}{|G|^5}
\]

as required.

### 3.4 Additive energy

Define the **additive energy** \( E_+(A, B) \) of set \( A \) and \( B \) to be the number of solutions to the equation

\[
a + b = a' + b'
\]

with \( a, a' \in A \) and \( b, b' \in B \). Note that

\[
E_+(A, B) = E_+(A, -B)
\]

since (3.11) holds if and only if \( a - b' = a' - b \). Define the additive energy of a single set \( A \) to be

\[
E_+(A) = E_+(A, A).
\]

When working with the multiplicative group of a field, we refer instead to **multiplicative energy**, denoted by \( E_\times(A, B) \), i.e. the number of solutions to \( ab = a'b' \) with \( a, a' \in A \) and \( b, b' \in B \).

Energy and sumsets are closely related, since the only way in which \( |A + B| \) can be small is if there are many pairs of elements giving the same sums. This would necessitate many solutions to (3.11) and thus a large energy. This intuition will shortly be made rigorous, but in order to do so it is first necessary to generalise and slightly reformulate energy.

Just like generalising sumsets to partial sumsets, given \( G \subseteq A \times B \) it is natural to consider the number of solutions to (3.11) with \( (a, b), (a', b') \in G \). This restriction is called the additive energy of \( G \) and denoted by \( E_+(G) \). As with partial sumsets, this collapses to the definition of \( E_+(A, B) \) when \( G = A \times B \).

The following lemma shows that additive energy can be helpfully formulated as an \( L^2 \) norm.
3.4. Additive energy

Lemma 3.9 (Reformulation of additive energy). For $G \subseteq A \times B$, let $\mu_G(x)$ denote the number of $(a,b) \in G$ with $a + b = x$. Then

$$E_+(G) = \sum_{x \in A + B} \mu_G(x)^2$$

(3.12)

$$= \sum_{(a,b) \in G} \mu_G(a + b).$$

(3.13)

Proof. It is clear that

$$E_+(G) = \sum_{x \in A + B} \# \{(a,b), (a',b') \in G : a + b = a' + b' = x\}$$

$$= \sum_{x \in A + B} \# \{(a,b) \in G : a + b = x\}^2$$

$$= \sum_{x \in A + B} \mu_G(x)^2$$

which establishes (3.12). Similarly,

$$E_+(G) = \sum_{(a,b) \in G} \# \{(a',b') : a + b = a' + b'\}$$

$$= \sum_{(a,b) \in G} \mu_G(a + b)$$

which gives (3.13).

When $G = A \times B$ we have $\mu_G(x) = |A \cap (x - B)|$ and the following corollary.

Corollary 3.10. For sets $A, B$ we have

$$E_+(A, B) = \sum_{x \in A + B} |A \cap (x - B)|^2$$

(3.14)

$$= \sum_{a \in A, b \in B} |A \cap (a + b - B)|$$

(3.15)

$$= \sum_{a,a' \in A} |(B + a) \cap (B + a')|$$

(3.16)

Proof. Equations (3.14) and (3.15) follow immediately from Lemma 3.9 and the fact that $\mu_G(x) = |A \cap (x - B)|$ when $G = A \times B$. Equation (3.16) follows by rearranging orders of summation from (3.15).

\[\square\]
Chapter 3. A handbook on sumset calculus

Now let’s make rigorous the earlier discussion of the relationship between energy and sumsets. The following estimates are trivial.

\[ |A||B| \leq E_+(A, B) \leq |A|^2|B|. \]

Energy close to \(|A||B|\) is ‘low’ and that close to \(|A|^2|B|\) is ‘high’. The following result shows that, as expected, small sumset implies high energy. This can be useful to know when analysing the structure of sets with small sumset.

**Lemma 3.11.** If \(G \subseteq A \times B\) then

\[ E_+(A, B) \geq E_+(G) \geq \frac{|G|^2}{|A + B|}. \]

Note that in this lemma, partial and complete sumsets are on the same footing: unlike BSG-type results there is no additional premium to pay for working with partial sumsets.

**Proof of Lemma 3.11.** It is clear that

\[ \sum_{x \in A^G + B} \mu_G(x) = |G|. \]

So by Cauchy-Schwarz,

\[ |G|^2 \leq |A + B| \sum_{x \in A^G + B} \mu_G(x)^2. \]

By Lemma 3.9 the right-hand side is equal to \(|A + B|E_+(G)|\) and so

\[ E_+(A, B) \geq E_+(G) \geq \frac{|G|^2}{|A + B|}, \]

which completes the proof.

The converse to Lemma 3.11 is false: high energy does not automatically imply small sumset. For example if \(A\) is the union of an arithmetic progression \(A_1\) and a geometric progression \(A_2\), each of cardinality \(|A|/2\), then \(A\) is high-energy since

\[ E_+(A) \gg E_+(A_1) \approx |A|^2. \]
However the sumset $|A + A|$ is also large since

$$|A + A| \geq |A_2 + A_2| \approx |A|^2.$$ 

What does hold, however, is the following partial converse. If a set has high energy then there must exist a fairly large $G \subseteq A \times A$ for which the partial sumset $A + A$ is small. This fact will not be required in the thesis and so precise details are omitted.

### 3.5 Covering methods

A covering lemma is a result of the following form. If $A$ and $B$ have small sumset then a large part of $A$ can be ‘covered by’ (contained in the union of) a small number of translates of $B$ or some modification of $B$. The canonical example is due to Ruzsa, for which a proof can be found in [48]:

**Lemma 3.12** (Ruzsa). For any sets $A, B$, the set $A$ is contained in the union of $\frac{|A - B|}{|B|}$ translates of $B - B$.

Another more recent example is due to Shen [40]. In a similar spirit to the improvement of Lemma 3.3 over Lemma 3.2 this enables a covering with translates of $B$ instead of $B - B$, at the expense of leaving a small part of $A$ uncovered.

**Lemma 3.13** (Shen). For sets $A, B$ and $\epsilon > 0$ there exists $A' \subseteq A$ with $|A'| \geq (1 - \epsilon)|A|$ such that $A'$ is contained in the union of $O_{\epsilon} \left(\frac{|A - B|}{|B|}\right)$ translates of $B$.

The proof of Shen’s covering result is based on additive energy, and so with some tweaks applies to the situation where only a dense partial sumset, rather than a complete sumset, is small. In certain situations, such as in Chapter 5, this can be a more-efficient substitute for Balog-Szemerédi-Gowers type methods, since there is no additional cost for dealing with partial rather than complete sumsets.

Two original variations on this theme are used in Chapter 5. The first, below, essentially restates Shen’s result with $A - B$ replaced by $A^G - B$.

**Lemma 3.14** (Shen variation 1). Let $G \subseteq A \times B$ and $0 < \epsilon < 1/4$. If $|G| \geq (1 - \epsilon)|A||B|$ then there exists $A' \subseteq A$ with $|A'| \geq (1 - 2\sqrt{\epsilon})|A|$ such that $A'$ is contained in the union of $O_{\epsilon} \left(\frac{|A - B|}{|B|}\right)$ translates of $B$. 

59
Chapter 3. A handbook on sumset calculus

Similarly, there is a subset \( A'' \subseteq A \) with \( |A''| \geq (1 - 2\sqrt{\epsilon})|A| \) such that \( A'' \) in contained in the union of \( O_\epsilon \left( \frac{|A \cap B|}{|B|} \right) \) translates of \(-B\).

Proof. We shall prove the case for covering with translates of \( B \), and remark on the slight alteration needed to cover with translates of \(-B\).

Since \( |G| \geq (1 - \epsilon)|A||B| \) there is by Lemma 3.7 a subset \( A_1 \subseteq A \) with \( |A_1| \geq (1 - \sqrt{\epsilon})|A| \) such that every element of \( A_1 \) has \( G \)-degree at least \((1 - \sqrt{\epsilon})|B|\).

Now for any subset \( A_* \subseteq A_1 \) let \( G_* = G \cap (A_1 \times B) \) so that

\[
|G_*| \geq (1 - \sqrt{\epsilon})|A_*||B|.
\]

By Lemma 3.11 it follows that

\[
E_+(A_*, B) \geq \frac{|G_*|^2}{|A \cap B|} \geq \frac{|A_*|^2(1 - \sqrt{\epsilon})^2|B|^2}{|A \cap B|}.
\]

By Corollary 3.10

\[
E_+(A_*, B) = E_+(A_*, -B) = \sum_{a \in A_*, b \in B} |A_* \cap (a - b) + B| \tag{3.17}
\]

and so for any \( A_* \subseteq A_1 \) there exist \( a \in A_*, b \in B \) such that

\[
|A_* \cap (a - b + B)| \geq \frac{|A_*|(1 - \sqrt{\epsilon})^2|B|}{|A \cap B|}.
\]

Apply the above discussion to a sequence of subsets of \( A_1 \). Begin by taking \( A_* = A_1 \) to find \( a_1 \in A_1, b_1 \in B \) such that

\[
|A_1 \cap (a_1 - b_1 + B)| \geq \frac{|A_1|(1 - \sqrt{\epsilon})^2|B|}{|A \cap B|}.
\]

The translate \((a_1 - b_1) + B\) covers \( \frac{|A_1|(1 - \sqrt{\epsilon})^2|B|}{|A \cap B|} \) elements of \( A_1 \). Discard \( A_1 \cap (a_1 - b_1) + B \) from \( A_1 \) and let \( A_2 \) be the set of elements remaining, now taking \( A_* = A_2 \) and repeating the process.

Iterate \( O_\epsilon \left( \frac{|A \cap B|}{|B|} \right) \) times until the set remaining is of cardinality no more than \( \sqrt{\epsilon}|A_1| \). Then take \( A' \) to be the set of elements discarded across all iterations, so that \( |A'| \geq (1 - \sqrt{\epsilon})|A_1| \). Since \( |A_1| \geq (1 - \sqrt{\epsilon})|A| \) we get \( |A'| \geq (1 - 2\sqrt{\epsilon})|A| \) as required, which completes the proof for covering with translates of \( B \).
3.5. Covering methods

The proof for covering with translates of $-B$ is identical, except that in place of (3.17), the identity

$$E_+(A_*, B) = \sum_{a \in A_*, b \in B} |A_* \cap (a + b - B)|$$

is used instead.

The second original variation on Shen’s result has a slightly different formulation. Instead of covering a large part of $A$ with translates of $B$, it yields a large part of $G$ whose corresponding partial difference set is covered by few translates of $B$.

**Lemma 3.15** (Shen variation 2). Let $G \subseteq A \times B$ and $0 < \epsilon < 1$. Then there exists $G' \subseteq G$ with $|G'| \geq (1 - \epsilon)|G|$ such that $A - B$ is contained in the union of $O_{\epsilon} \left( \frac{|A - B||A|}{|G|} \right)$ translates of $B$.

**Proof.** Let $G_*$ be any subset of $G$. By Lemma 3.11 there are at least $\frac{|G_*|^2}{|A - B||A|}$ solutions to

$$a - b = a' - b'$$

with $(a, b), (a', b') \in G_*$. So there exists $a_* \in A$ for which there are at least $\frac{|G_*|^2}{|A - B||A|}$ pairs $(a, b) \in G_*$ with

$$a - b \in a_* - B.$$

In other words, the translate $a_* - B$ accounts for the differences of $\frac{|G_*|^2}{|A - B||A|}$ pairs from $G_*$. Apply the above discussion to a sequence of subsets of $G$. Begin by taking $G_* = G$ to find a translate of $B$ accounting for the differences of $\frac{(1-\epsilon)^2|G|^2}{|A - B||A|}$ pairs from $G$. Discard these pairs from $G_1$ and let $G_2$ be the set of pairs remaining, now taking $G_* = G_2$ and repeating the process.

Iterate $O_{\epsilon} \left( \frac{|A - B||A|}{|G|} \right)$ times until the subset of $G$ remaining is of cardinality no more than $\epsilon|G|$. Then take $G'$ to be the set of discarded pairs, so that $|G'| \geq (1 - \epsilon)|G|$, and $A - B$ is contained in no more than $O_{\epsilon} \left( \frac{|A - B||A|}{|G|} \right)$ translates of $B$, as required. \qed
Chapter 1 introduced incidence geometry but deferred discussion of the finite field case. This chapter now picks up that baton.

The state of the art for incidences is weaker over finite fields than $\mathbb{R}$ and $\mathbb{C}$, and the discrepancy is much greater than the analogous gap for growth results. This chapter narrows that divide with two new theorems: an incidence bound and a line-counting result.

An earlier version of this work has been submitted to the European Journal of Combinatorics. A preprint [23], and a more recent update [24] also used here, are available on the arXiv.

4.1 Results

This section describes the two results proved in this chapter.

4.1.1 Incidence bounds

Chapter 1 showed that in the ‘critical’ case $|P|, |L| = N$, non-trivial incidence bounds are of the form

$$I(P, L) \ll N^{3/2-\epsilon}$$

for $\epsilon > 0$. The Szemerédi-Trotter theorem (Theorem 1.3) gives the sharp result $\epsilon = \frac{1}{6}$ in $\mathbb{R}^2$ and generalises to the complex setting $\mathbb{C}^2$ (Theorem 1.14).

Chapter 2 noted that any nontrivial results on growth in finite fields must be predicated on keeping away from subfields. The same is true for incidence
bounds, but this time one must keep away from subplanes. Just like growth, to keep these considerations as straightforward as possible it is often convenient to work with the field $\mathbb{F}_p$ of prime order $p$.

To see why avoiding subplanes is necessary, consider the example where $P$ is the whole plane $\mathbb{F}_p^2$ and $L$ is the set of all lines in $\mathbb{F}_p^2$. Taking $N = p^2$ gives $|P|, |L| \approx N$, but since every line in $L$ is incident to $p$ points in $P$ it follows that $I(P, L) \approx N^{3/2}$ and so in this case a non-trivial estimate is impossible.

Working over $\mathbb{F}_p^2$, Bourgain, Katz and Tao [6] proved the existence of a non-trivial $\epsilon > 0$, dependent on $\gamma > 0$, whenever $N < p^{2-\gamma}$. This has been made explicit in two cases:

- In the ‘small-set’ regime $N < p$, Helfgott and Rudnev [20] obtained a bound of $\epsilon \geq \frac{1}{10,678}$.
- In the ‘large-set’ regime $p^{1+\gamma} \leq N \leq p^{2-\gamma}$, Vinh [50] obtained $\epsilon \geq \frac{3}{4}$.

For the purposes of this thesis, where as explained in the prologue the emphasis is on small rather than large sets, the Helfgott-Rudnev bound is the one to beat.

It is instructive to compare with the situation for growth and sum-products. With sum-products there is a growth exponent of $\delta \geq \frac{1}{3} - o(1)$ over $\mathbb{R}$ and $\mathbb{C}$, and $\delta \geq \frac{1}{\pi} - o(1)$ over finite fields, which is a bit weaker but not excessively so. By contrast the finite field incidence bound of $\epsilon \geq \frac{1}{10,678}$ is an awfully long way from the real and complex bound of $\epsilon \geq \frac{1}{6}$ implied by Szemerédi-Trotter.

The first new result in this chapter goes some way to redressing this disparity by obtaining a much stronger finite field incidence bound of $\epsilon \geq \frac{1}{662} - o(1)$.

**Theorem 4.1.** Let $N < p$. If $P$ and $L$ are a set of points and lines over $\mathbb{F}_p$ with $|P|, |L| \leq N$ then

$$I(P, L) \lesssim N^{\frac{3}{2} - \frac{1}{662}}.$$ 

### 4.1.2 Line counting

Beck’s theorem from Chapter 1 shows that any set of points in $\mathbb{R}^2$ satisfies at least one of two extremes. Either there are at least $\Omega(|P|)$ collinear points, or the set $L(P)$ of lines determined by pairs of points in $P$ is of cardinality at least $\Omega(|P|^2)$. 

64
4.1. Results

As with incidence bounds, a nondegeneracy condition is required for Beck-type theorems over finite fields. For example if \( P = \mathbb{F}_p^2 \) then no more than \( |P|^{1/2} \) points are collinear, and so the first possible conclusion of Beck’s theorem cannot hold. But the second cannot hold either since \( |L(P)| \approx |P| \).

As with incidence bounds, there have been explicit finite field versions of Beck’s theorem in two instances:

- In the particular case of the ‘small-set’ regime \( |P| < p \) for which \( P = A \times A \) with \( A \subseteq \mathbb{F}_p \), Helfgott and Rudnev [20] showed that \( |L(P)| \gg |P|^{1 + \frac{1}{267}} \).
- In the ‘large-set’ regime, Iosevich, Rudnev and Zhai [22] recently showed that \( |L(P)| \approx p^2 \) whenever \( |P| > p \log p \).

The ‘small-set’ Helfgott-Rudnev result is again, for the purposes of this thesis, the one to beat. The second new result in this chapter does so in two respects. First, there is a stronger exponent of \( \frac{1}{133} - o(1) \) in place of \( \frac{1}{267} \). Second, the result holds for general \( P \subseteq \mathbb{F}_p^2 \) with \( |P| < p \) rather than simply those of the form \( P = A \times A \).

**Theorem 4.2.** If \( P \subseteq \mathbb{F}_p^2 \) and \( |P| < p \) then at least one of the following must occur:

1. At least \( \tilde{\Omega}(|P|) \) points from \( P \) are contained in a single line.
2. \( |L(P)| \gtrsim |P|^{1 + \frac{1}{133}} \).

4.1.3 Structure

The body of this chapter is concerned with the proofs of Theorems 4.1 and 4.2. Both follow from new observations that efficiently relate incidences to sum-product estimates.

Unlike the incidence material in Chapter 1, the analysis here makes crucial use of the properties of the projective plane. **Section 4.2** uses this to show that the existence of a certain configuration of points and lines would imply the existence of sets \( A, B \) and fairly dense \( G \subseteq A \times B \) such that the difference set \( A^G - B \) and the ratio set \( A^G / B \) are both small. Such a configuration is called a ‘sum-product configuration’.
Section 4.3 goes on to show that if there are too many points incident to too many lines, which would arise if either Theorem 4.1 or Theorem 4.2 were to fail, then there must also exist a large sum-product configuration.

The situation is then almost ripe for applying a finite field sum-product theorem from Chapter 2 to show that a sum-product configuration cannot arise, and so there cannot be too many incidences. The final hurdle is to relate the partial difference and ratio sets arising from a sum-product configuration to complete sets. This is accomplished in Section 4.4 using a concoction of Balog-Szemerédi-Gowers type results from Chapter 3.

Section 4.5 then uses the analysis from the previous sections to prove Theorems 4.1 and 4.2.

Section 4.6 gives pointers for further work.

4.2 Foci and configurations

This section uses the theory of projective transformations to interpret particular types of point-line configurations in terms of sum-product (or strictly speaking, difference-ratio) estimates. The key observation is that difference and ratio sets can be interpreted in terms of gradients, and that a projective transformation enables these to be viewed as point-line incidences.

There are several pictures of points and lines in this section, apparently treating them as if in $\mathbb{R}^2$. These are for illustration only; they do not correspond rigorously to the plane $\mathbb{F}_p^2$, although the definitions they illustrate do.

Let $F$ be any field. Recall, or quickly check Appendix B that the projective plane $\mathbb{P}F^2$ is given by equivalence classes of $F^3 \setminus \{(0,0,0)\}$ modulo dilation. More concretely, it can be viewed as the union of $F^2$ with a ‘line at infinity’ $l_\infty$. Points of $l_\infty$ correspond to gradients in $\mathbb{P}F^1 = F \cup \{\infty\}$, so that for each $\lambda \in \mathbb{P}F^1$ there is a point $p_\lambda \in l_\infty$ that is incident to all lines of gradient $\lambda$.

Recall also some theory on projective transformations of $\mathbb{P}F^2$. These are line-preserving permutations of $\mathbb{P}F^2$. A frame in $\mathbb{P}F^2$ is a set of four distinct points, no three of which are collinear. For any two frames there is a unique projective transformation that maps one to the other.

Now for some new definitions. If $P$ is a set of points in $\mathbb{P}F^2$ and $p \notin P$ is an individual point then say that $p$ is a $K$-focus for $P$ if $P$ is supported over at
4.2. Foci and configurations

most $K$ lines through $p$. This is illustrated in Figure 4.1 where the red point is a 3-focus for the blue points. Note that if $p$ is a $K$-focus for $P$ then it is also a $K$-focus for any subset of $P$.

![Figure 4.1. A point set and focus](image)

We also give a highly suggestive name to a particular configuration of points and foci. Let $P$ be a point set and $p_1, p_2, p_3, p_4 \notin P$ be distinct points. Say that $P$ and the $p_i$ form a $K$-sum-product configuration if

1. Each $p_i$ is a $K$-focus for $P$.

2. There is a line, which we call the base line, incident to $p_2, p_3$ and $p_4$ but not $p_1$.

3. No point in $P$ is incident to the base line.

Figure 4.2 illustrates sum-product configurations with (a) the base line in general position and (b) the base line at infinity. The red, black, green and yellow points are the foci $p_i$ and the blue points are the elements of $P$.

The following lemma justifies the choice of definition by showing that a sum-product configuration does indeed correspond to a (partial) sumset and product set.
Chapter 4. Incidence theorems over finite fields

Figure 4.2. Sum-product configurations

Lemma 4.3. Let $F$ be a field. Suppose $P$ is a set of points in $F^2$, and that it forms a $K$-sum-product configuration with points $p_1, p_2, p_3, p_4$. Then there exist sets $A, B \subseteq F$ and $G \subseteq A \times B$ with

$$|G| = |P|$$

and

$$|A|, |B|, |A \cap B|, |A/B| \leq K.$$
4.2. Foci and configurations

Proof. There is a projective transformation mapping any frame (four points, no three of which are collinear) to any other frame, and so certainly there is one mapping any non-collinear triple to any other. Since \( p_1, p_3 \) and \( p_4 \) are not collinear, and nor are \([1, 0, 0], [0, 1, 0], [0, 0, 1]\), it is possible to pick a projective transformation \( \tau \) such that

- \( p_1 \) is sent to the origin \([0, 0, 1]\) of \( F^2 \).
- \( p_3 \) is sent to \([0, 1, 0] \in l_\infty \).
- \( p_4 \) is sent to \([1, 0, 0] \in l_\infty \).

Since \( \tau \) preserves linearity and sends \( p_3 \) and \( p_4 \), which are incident to the base line, to points on \( l_\infty \), it in fact sends the whole base line to \( l_\infty \).

Let \( G = \tau(P) \). This is contained in \( F^2 \) since none of the points in \( P \) are incident to the base line and so none are mapped into \( l_\infty \).

The set \( G \) is supported over \( K \) ‘vertical’ lines \([1]\) and \( K \) ‘horizontal’ lines \([2]\) since lines incident to \([0, 1, 0] \in l_\infty \) are all vertical and those incident to \([1, 0, 0] \in l_\infty \) are all horizontal. Let \( A \) be the set of \( x \)-intercepts of the vertical lines and \( B \) be the set of \( y \)-intercepts of the horizontal lines, so that

\[
G \subseteq A \times B
\]

and

\[
|A|, |B| \leq K.
\]

These correspond to the yellow and black lines respectively in Figure 4.2.

Furthermore, \( G \subseteq A \times B \) is supported over \( K \) lines through the origin. These are identified by their gradient, and a point \((a, b) \in G \) is incident to the line with gradient \( \xi \) if and only if \[ a \frac{b}{a} = \xi. \] Thus each element of \( A/B \) corresponds to a different line through the origin, i.e. a red line in Figure 4.2 and so

\[
|A/B| \leq K.
\]

Finally, \( G \) is supported over \( K \) lines through \( \tau(p_2) \). Since \( \tau \) preserves linearity and sends the base line, which is incident to \( p_2 \), to \( l_\infty \) we know that \( \tau(p_2) \in l_\infty \). So

\[1\] that is, lines of the form \( x = c \) for some \( c \in F \)

\[2\] that is, lines of the form \( y = c \) for some \( c \in F \)
all lines incident to $\tau(p_2)$ have the same gradient, say $\lambda \in F$. These are identified
by the intercept with the $y$-axis, and a point $(a, b) \in G$ is incident to the line with
intercept $\rho$ if and only if $a + \lambda b = \rho$. Thus each element of $A + \lambda B$ corresponds
to a different line of gradient $\lambda$, i.e. a green line in Figure 4.2. Therefore
$$|A + \lambda B| \leq K.$$

Now let $B' = -\lambda B$ and $G' = \{(a, \lambda b) : (a, b) \in G\}$ to obtain
$$|G'| = |P|$$
and
$$|A|, |B'|, |A - B'|, |A / B'| \leq K.$$

4.3 How to find foci

Section 4.2 shows that a sum-product configuration of points and foci corresponds
to an upper bound on partial sum-products. This section shows that such a
configuration must arise whenever there are many point-line incidences. The
main result is the following lemma.

**Lemma 4.4** (Finding a sum-product configuration). Let $P$ and $L$ be a set of
points and lines respectively in a plane such that every point in $P$ is incident to
$\Theta(K)$ lines in $L$. Suppose that $K \gg \frac{|L|^{3/5}}{|P|^{1/5}}$, and that each line in $L$ is incident to
at most
$$O\left(\min\left\{\frac{|P|K^8}{|L|^4}, \frac{|P|K^4}{|L|^2}\right\}^{1-o(1)}\right)$$
points in $P$.

Then there exists a subset of $P$ of cardinality $\Theta\left(\frac{|P|K^8}{|L|}\right)$, and points $p_1, p_2, p_3, p_4$
with which it forms a $O(K)$-sum-product configuration.

It will be helpful when proving Lemma 4.4 to adopt some additional notation.
For points $p, q$ in the plane, let $l_{pq}$ be the line determined by $p$ and $q$. Given a
set $P$ of points and a set $L$ of lines, both satisfying the conditions of Lemma 4.4
and a single point $p \in P$, define
$$P_{pL} = \{q \in P : l_{pq} \in L\}$$
4.3. How to find foci

so that \( P_{pL} \) is the set of points in \( P \) incident to lines in \( L \) going through \( p \). Thus \( P_{pL} \) has \( p \) as an \( O(K) \)-focus, since every point in \( P \) is incident to \( \Theta(K) \) lines in \( L \).

As a first step, it is convenient to record the following standard result.

**Lemma 4.5** (Rich points and lines). Let \( P_1 \) be the set of points in \( P \) incident to at least \( \frac{I(P,L)}{2|P|} \) lines in \( L \). Then

\[
I(P_1, L) \approx I(P, L).
\]

Similarly, if \( L_1 \) is the set of lines in \( L \) incident to at least \( \frac{I(P,L)}{2|L|} \) points in \( P \) then

\[
I(P, L_1) \approx I(P, L).
\]

**Proof.** We prove the result for points, leaving that for lines as an exercise. Let \( P_2 \) be the set of points in \( P \) incident to at most \( \frac{I(P,L)}{2|P|} \) lines in \( L \). Then

\[
I(P_2, L) = \sum_{p \in P_2} \# \{ l \in L \text{ incident to } p \} \leq |P_2| \frac{I(P,L)}{2|P|} \leq \frac{I(P,L)}{2}.
\]

Since

\[
I(P, L) = I(P_1, L) + I(P_2, L)
\]

it follows that

\[
I(P_1, L) \geq \frac{I(P,L)}{2}
\]

as required. \( \square \)

Armed with Lemma 4.5 we now build incrementally towards Lemma 4.4. We shall show how to find individual foci, and simultaneous pairs of foci, before finding four arranged in a sum-product configuration. The initial results on finding singleton or paired foci are adjustments of similar methods used in the proofs of Bourgain-Katz-Tao [6] and Helfgott-Rudnev [20]; they are applied in the proof of the more developed Lemma 4.4.

The following result enables us find individual foci.

**Lemma 4.6** (Finding individual foci). Let \( P \) be a set of points and \( L \) a set of lines, such that every point is incident to \( \Theta(K) \) lines in \( L \). Then there exists \( P_1 \subseteq P \) with \( |P_1| \approx |P| \) such that

\[
|P_{pL}| \gg \frac{K^2|P|}{|L|}
\]

for each \( p \in P_1 \).
Proof. Since every point in \( P \) is incident to \( \Theta(K) \) lines in \( L \) it is immediate that

\[
I(P, L) \approx K|P|.
\]

Let \( L_1 \) be the set of lines in \( L \) incident to \( \Omega \left( \frac{I(P, L)}{|P|} \right) = \Omega \left( \frac{K|P|}{|L|} \right) \) points in \( P \) so that by Lemma 4.5 it follows that

\[
I(P, L) \approx I(P, L_1) \approx K|P|.
\]

Now let \( P_1 \) be the set of points in \( P \) incident to \( \Omega \left( \frac{I(P, L_1)}{|P_1|} \right) = \Omega \left( K \frac{|P|}{|L_1|} \right) \) points in \( P \) so that by Lemma 4.5 again,

\[
I(P_1, L_1) \approx I(P, L_1) \approx K|P|.
\]

Since \( P_1 \subseteq P \) and \( L_1 \subseteq L \), each point in \( P_1 \) is incident to at most \( O(K) \) lines in \( L_1 \) and so

\[
I(P_1, L_1) \ll K|P_1|.
\]

Comparing (4.1) and (4.2) shows that that \( |P_1| \gg |P| \), implying that \( |P_1| \approx |P| \) since \( P_1 \) is a subset of \( P \).

Now, each \( p \in P_1 \) is incident to \( \Omega (K) \) lines in \( L_1 \), and each of these lines is incident to \( \Omega \left( \frac{K|P|}{|L_1|} \right) \) points in \( P \). So

\[
|P_1 L| \geq |P_{pL}| \gg \frac{K^2|P|}{|L|}
\]

for each \( p \in P_1 \), as required.

The following result now takes the ability to find singleton foci, and uses it to find them in simultaneous pairs.

Lemma 4.7 (Finding paired foci). Let \( P \) and \( L \) be a set of points and lines respectively in a plane such that every point in \( P \) is incident to \( \Theta(K) \) lines in \( L \). Then there exist \( p_1, p_2 \in P \) such that

\[
|P_{p_1L} \cap P_{p_2L}| \gg \frac{|P|K^4}{|L|^2}.
\]

Proof. By Lemma 4.6 there exists \( P_1 \subseteq P \) with \( |P_1| \approx |P| \) such that \( |P_{pL}| \gg \frac{K^2|P|}{|L|} \) for each \( p \in P_1 \). In particular there is a point \( p_1 \in P \) such that

\[
|P_{p_1L}| \gg \frac{K^2|P|}{|L|}.
\]
4.3. How to find foci

Applying Lemma 4.6 again, this time to \( P_{p_1L} \) and \( L \), there is a point \( p_2 \in P_{p_1L} \) such that

\[
|P_{p_1L} \cap P_{p_2L}| \geq \frac{K^2|P_{p_1L}|}{|L|} \gg \frac{K^4|P|}{|L|^2}
\]

as required.

Now with the ability to find foci both one and two at a time, it is time for a proof of Lemma 4.4

**Proof of Lemma 4.4**. It suffices to find \( p_1, p_2, p_3, p_4 \in P \) with

\[
|P_{p_1L} \cap P_{p_2L} \cap P_{p_3L} \cap P_{p_4L}| \gg \frac{|P|K^8}{|L|^4}
\]

such that \( p_2, p_3 \) and \( p_4 \) are collinear along a line in \( L \) that is not incident to \( p_1 \). There will then exist an \( O(K) \)-sum-product configuration of appropriate size since:

- Each \( p \in P \) is incident to \( \Theta(K) \) lines in \( L \), and so is an \( O(K) \)-focus for \( P_{pL} \). This means that the first condition in the definition of an \( O(K) \)-sum-product configuration is satisfied for any subset of \( P_{p_1L} \cap P_{p_2L} \cap P_{p_3L} \cap P_{p_4L} \).

- The second condition of a sum-product configuration is clearly satisfied by virtue of the arrangement of \( p_1, p_2, p_3, p_4 \).

- No line in \( L \) is incident to more than \( O \left( \left( \frac{|P|K^4}{|L|^4} \right)^{1-o(1)} \right) \) points in \( P \), and so \( P_{p_1L} \cap P_{p_2L} \cap P_{p_3L} \cap P_{p_4L} \) contains a subset of cardinality \( \Theta \left( \frac{|P|K^8}{|L|^4} \right) \) for which the third condition of an \( O(K) \)-sum-product configuration is satisfied.

Observe first that by Lemma 4.7 there exist points \( p_1, p_2 \in P \) such that

\[
|P_{p_1L} \cap P_{p_2L}| \gg \frac{|P|K^4}{|L|^2}. \tag{4.3}
\]

For convenience, define

\[
Q = P_{p_1L} \cap P_{p_2L}. \tag{4.4}
\]

The proof will be complete if we can find distinct \( p_3, p_4 \in Q \) collinear with \( p_2 \) along a line in \( L \) that is not incident to \( p_1 \) such that

\[
|Q_{p_3L} \cap Q_{p_4L}| \gg \frac{|P|K^8}{|L|^4}. \tag{4.5}
\]
Now since \( Q \subseteq P \), every point in \( Q \) is incident to \( \Theta(K) \) lines in \( L \). So by Lemma 4.6 there exists \( Q_1 \subseteq Q \) with \(|Q_1| \approx |Q|\) such that

\[
|Q_{p|L}| \gg \frac{K^2|Q|}{|L|}
\]

(4.6)

for each \( p \in Q_1 \).

Since \( p_2 \) is an \( O(K) \)-focus for \( P_{p_2|L} \), and \( Q_1 \subseteq Q \subseteq P_{p_2|L} \), it follows that \( p_2 \) is an \( O(K) \)-focus for \( Q_1 \) as well. Let \( J \subseteq L \) be the set of \( O(K) \) lines through \( p_2 \) supporting \( Q_1 \), so that

\[
I(Q_1, J) = |Q_1|
\]

and

\[
|J| \ll K.
\]

Let \( J_1 \) be the set of \( l \in J \) incident to at least \( \Omega \left( \frac{I(Q,J)}{|J|} \right) = \Omega \left( \frac{|Q|}{K} \right) \) points in \( Q_1 \). Observe that \( |J_1| \geq 2 \). Indeed, Lemma 4.5 implies

\[
I(Q_1, J_1) \approx I(Q_1, J) \approx |Q|.
\]

(4.7)

But since \( Q_1 \subseteq P \) and \( J_1 \subseteq L \), and it is a hypothesis that each line in \( L \) is incident to at most

\[
O \left( \left( \frac{|P|K^4}{|L|^2} \right)^{1-o(1)} \right) = O(|Q|^{1-o(1)})
\]

points in \( P \), it follows that each line in \( J_1 \) is incident to at most \( O(|Q|^{1-o(1)}) \) points in \( Q \). Hence

\[
I(Q_1, J_1) \ll |J_1||Q|^{1-o(1)}.
\]

(4.8)

Comparing (4.7) and (4.8) gives \( |J_1| \gg |Q|^{o(1)} \). So by appropriate choice of constants in the statement of the theorem \( |J_1| \geq 2 \) as claimed.

Since there are at least two lines in \( J_1 \), and they are all incident to \( p_2 \), at least one of them is not incident to \( p_1 \). Fix this line \( l^* \in J_1 \), which will be the base line of the sum-product configuration. Since \( l^* \) is incident to \( p_2 \) but not \( p_1 \), it suffices to establish that (4.5) holds for some distinct \( p_3, p_4 \in Q \cap l^* \).

Because \( l^* \in J_1 \) we have

\[
|Q_1 \cap l^*| \gg \frac{|Q|}{K}
\]

(4.9)
4.4. Bounding partial sum-products

From (4.6) it follows that
\[
\frac{K^2|Q|}{|L|}|Q_1 \cap l^*| \ll \sum_{p \in Q \cap l^*} |Q_{pL}|.
\] (4.10)

On the other hand, by Cauchy-Schwarz,
\[
\sum_{p \in Q \cap l^*} |Q_{pL}| \leq |Q|^{1/2} \left( \sum_{p_3,p_4 \in Q \cap l^*} |Q_{p_3L} \cap Q_{p_4L}| \right)^{1/2}
\]
\[
= |Q|^{1/2} \left( \sum_{p \in Q \cap l^*} |Q_{pL}| + \sum_{p_3,p_4 \in Q \cap l^* \atop p_3 \neq p_4} |Q_{p_3L} \cap Q_{p_4L}| \right)^{1/2}.
\]

If the first summation on the right were to dominate then it would mean
\[
\sum_{p \in Q \cap l^*} |Q_{pL}| \ll |Q|.
\]
Comparing with (4.10) and applying (4.9) would then yield \(|Q|K \ll |L|\). By (4.3) and (4.4), this would mean \(K \ll |L|^{3/5}|P|^{1/5}\), and so by an appropriate choice of constant contradict the hypothesis \(K \gg |L|^{3/5}|P|^{1/5}\).

Thus the second summation on the right dominates instead and so by (4.10),
\[
\frac{K^4|Q||Q \cap l^*|^2}{|L|^2} \ll \sum_{p_3,p_4 \in Q \cap l^* \atop p_3 \neq p_4} |Q_{p_3L} \cap Q_{p_4L}|.
\]

Hence there exist distinct \(p_3,p_4 \in Q \cap l^* \subseteq Q \cap l^*\) such that
\[
|Q_{p_3L} \cap Q_{p_4L}| \gg \frac{K^4|Q|}{|L|^2} \gg \frac{|P|K^8}{|L|^4}
\]
as required. \(\square\)

4.4 Bounding partial sum-products

Between them, Section 4.2 and Section 4.3 show that the existence of too many incidences implies the existence of \(A,B \subseteq \mathbb{F}_p\) and a large \(G \subseteq A \times B\) for which the partial difference and ratio sets \(|A^G - B^G|\) and \(|A^G / B^G|\) are both small relative to \(|A|\) and \(|B|\).
This offends our sum-product sensibilities, since Chapter 2 showed that one or other of a product set or sumset must always be large. However those results were concerned with complete sumsets and product sets, whereas here only partial sum-products are controlled. Fortunately, Chapter 3 has tools designed for this kind of situation, in the form of Balog-Szemerédi-Gowers type results that allow passage from partial to complete sets.

The main results in this section are the following ‘partial sum-product’ results for $\mathbb{F}_p$. They will be proved using a combination of BSG-type results and Rudnev’s finite field sum-product estimate Theorem 2.2.

Lemma 4.8 (Partial sum-products in $\mathbb{F}_p$, v1). Let $A, B \subseteq \mathbb{F}_p$ and $G \subseteq A \times B$. If $|A| \ll p^{1/2}$ then

$$|G|^5 \lesssim |A|^3 |B|^3 |A - B|^{28} |A/B|^8.$$  

Lemma 4.9 (Partial sum-products in $\mathbb{F}_p$, v2). Let $A, B \subseteq \mathbb{F}_p$ and $G \subseteq A \times B$. If $|G| \ll p^{1/2}|B|$ then

$$|G|^6 \lesssim |A|^4 |B|^4 |A - B|^{28} |A/B|^{16}.$$  

Because sumsets and product sets are being analysed simultaneously, care is required when using Balog-Szemerédi-Gowers type results to prove these lemmata. Results from Chapter 3 yield the following preliminary result, which captures all the additive and multiplicative properties that need to be controlled.

Lemma 4.10 (BSG-type for sum-products). If $A, B \subseteq \mathbb{F}_p$ and $G \subseteq A \times B$ then there exists $A' \subseteq A$ with $|A'| \gg |G|/|B|$ such that

1. $|A' - A| \ll \frac{|A - B|^4 |A|^4 |B|^3}{|G|^3}$

2. $|A'/A| \ll \frac{|A/B|^4 |A|^4 |B|^3}{|G|^3}$

3. $E_x(A') \gg \frac{|G|^2 |A|^4}{|A/B|^2 |A|^2 |B|}$

Proof. Let $\epsilon > 0$ be sufficiently small. By Lemma 3.8 there exist $A_1 \subseteq A$ and $H \subseteq A_1 \times A_1$ with

$$|A_1| \gg \frac{|G|}{|B|}.$$
4.4. Bounding partial sum-products

and

\[ |H| \geq (1 - \epsilon)|A_1|^2 \]

such that every pair \((a_1, a_2) \in H\) has joint \(G\)-degree at least \(\frac{\epsilon|G|^2}{2|A||B|}\). By Lemma 3.6 applied once additively and once multiplicatively,

\[ |A_1^H - A_1| \ll \frac{|A - B| |A|^2 |B|}{|G|^2} \]
\[ |A_1^H / A_1| \ll \frac{|A/B|^2 |A|^2 |B|}{|G|^2}. \]

Apply Lemma 3.4 once additively and once multiplicatively to obtain \(A_2, A_3 \subseteq A_1\) with

\[ |A_2|, |A_3| \geq (1 - \sqrt{\epsilon})|A_1| \]

such that

\[ |A_2 - A_2| \ll \frac{|A_1^H - A_1|^2}{|A_1|} \ll \frac{|A - B| |A|^4 |B|^3}{|G|^5} \]
\[ |A_3 / A_3| \ll \frac{|A_1^H / A_1|^2}{|A_1|} \ll \frac{|A/B|^4 |A|^4 |B|^3}{|G|^5}. \]

Let \(A' = A_2 \cap A_3\). By the cardinalities of \(A_2\) and \(A_3\),

\[ |A'| \geq (1 - 2\sqrt{\epsilon})|A_1| \gg \frac{|G|}{|B|}. \]

That \(A'\) satisfies properties 1 and 2 is immediate from the above difference and ratio set estimates. To show that property 3 is also satisfied, let

\[ H' = H \cap (A' \times A'). \]

Since both \(H\) and \(A' \times A'\) are of cardinality at least \((1 - 2\sqrt{\epsilon})^2 |A_1|^2\) we have

\[ |H'| \geq (1 - 8\sqrt{\epsilon})|A_1|^2 \]

and so

\[ E_x(A') \geq \frac{|H'|^2_{H'}}{|A' / A'|} \gg \frac{|A_1|^4_{H}}{|A_1 / A_1|} \gg \frac{|G|^2 |A_1|^4}{|A - B|^2 |A|^2 |B|} \]

which completes the proof. \(\square\)
Chapter 4. Incidence theorems over finite fields

We can now prove Lemma 4.8 and Lemma 4.9.

Proof of Lemma 4.8. Apply Lemma 4.10 to obtain $A' \subseteq A$ with $|A'| \gg \frac{|G|}{|B|}$ such that

$$|A' - A| \ll \frac{|A - B|^4 |A|^4 |B|^3}{|G|^5}.$$  

$$E_\times(A') \gg \frac{|G|^2 |A'|^4}{|B|^2 |A|^2 |A/B|^2}.$$  

Since $|A| \ll p^{1/2}$, the set $A' \subseteq A$ is of cardinality at most $O\left(\frac{p}{B}\right)$ as well, and the sum-product estimate Theorem 2.2 can be applied to obtain

$$\frac{|G|^8 |A'|^{16}}{|B|^4 |A|^8 |A/B|^8} \ll E_\times(A')^4 \approx |A' - A'|^7 |A'|^4 \approx \frac{|A - B|^{28} |A^{28}| |B|^{21} |A'|^4}{|G|^{35}}.$$  

Rearranging gives

$$|G|^{43} |A'|^{12} \ll |A - B|^{28} |A/B|^8 |A|^36 |B|^{25}$$  

and so

$$|G|^{55} \approx \frac{A}{B}^{37} |A - B|^{28} |A/B|^8$$  

as required. 

Proof of Lemma 4.9. Apply Lemma 4.10 to obtain $A' \subseteq A$ with $|A'| \gg \frac{|G|}{|B|}$ such that

$$|A' - A| \ll \frac{|A - B|^4 |A|^4 |B|^3}{|G|^5}.$$  

$$|A'/A'| \ll \frac{|A/B|^4 |A|^4 |B|^3}{|G|^5}.$$  

Unlike in the proof of Lemma 4.8 where multiplicative energy was involved, these properties are preserved when passing to subsets of $A'$ and so we may assume that $|A'| \approx \frac{|G|}{|B|}$. Since $|G| \ll p^{1/2} |B|$ the set $A'$ is then of cardinality at most $O\left(p^{1/2}/B\right)$ and so Theorem 2.2 can be applied to obtain

$$\frac{|G|^{36}}{|B|^{28} |A|^{16} |A/B|^{16}} \ll \left(\frac{|A'|^4}{|A'/A'|}\right)^4 \ll E_\times(A')^4 \approx |A'|^4 |A' - A'|^7$$  

$$\ll \frac{|A - B|^{28} |A|^{28} |B|^{17}}{|G|^{31}}.$$  

78
Rearranging gives
\[ |G|^{67} \lesssim |A|^{44}|B|^{45}|A - B|^{28}|A/B|^{16} \]
as required.

4.5 Proving Theorems 4.1 and 4.2

This section uses the analysis from Sections 4.2, 4.3 and 4.4 to give proofs of
Theorems 4.1 and 4.2.

Proof of Theorem 4.1. Suppose that \( I(P, L) \gg N^{3/2-\epsilon} \). The aim is to show that
\( \epsilon \geq \frac{1}{662} - o(1) \).

First, a standard argument enables overly rich lines to be discarded. Let \( L_1 \)
be the set of lines in \( L \) incident to at most \( O(N^{1/2+\epsilon}/2) \) points in \( P \). We have
\[ I(P, L_1) \approx N^{3/2-\epsilon} \]
since if \( L_+ \) is the set of lines in \( L \) incident to at least \( CN^{1/2+\epsilon} \) points in \( P \) then
\[ I(P, L_+) = \sum_{l \in L_+} \sum_{p \in P} \delta_{pl} \leq \frac{1}{CN^{1/2+\epsilon}} \sum_{l \in L_+} \sum_{p_1, p_2} \delta_{p_1 l} \delta_{p_2 l} \leq \frac{1}{CN^{1/2+\epsilon}} \left( I(P, L) + |P|^2 \right) \leq \frac{2N^{3/2-\epsilon}}{C} \]
and so by an appropriate choice of constant \( I(P, L_+) \leq \frac{I(P, L)}{2} \).

By a dyadic pigeonholing there is a subset \( P_1 \subseteq P \) and an integer \( K \) with
\[ |P_1| K \gtrsim N^{3/2-\epsilon} \] (4.11)
such that every point in \( P_1 \) is incident to \( \Theta(K) \) lines in \( L_1 \). Note moreover that
\[ K \gtrsim N^{1/2-\epsilon} \] (4.12)
since \( |P_1| \leq N \). Applying Lemma 4.4 to \( P_1 \) and \( L_1 \), at least one of the following is true:
1. $K \ll \frac{|L_1|^{3/5}}{|P_1|^{1/10}}$.

2. There is a line in $L_1$ incident to at least $\Omega \left( \left( \frac{|P_1| K^4}{|L_1|^4} \right)^{1-o(1)} \right)$ points in $P_1$.

3. There is a line in $L_1$ incident to at least $\Omega \left( \left( \frac{|P_1| K^8}{|L_1|^4} \right)^{1-o(1)} \right)$ points in $P_1$.

4. There exists $P_2 \subseteq P_1$ with $|P_2| \approx \frac{|P_1| K^8}{|L_1|^4}$ and points $p_1, p_2, p_3, p_4$ in an $O(K)$-sum-product configuration.

The first three cases are quickly dispensed with. For the first, applying (4.11) and then (4.12) yields $\epsilon \geq \frac{1}{10} - o(1)$, which is far better than required. For the second, since every line in $L_1$ is incident to at most $N^{1/2+\epsilon}$ points in $P$, it follows that $\epsilon \geq \frac{1}{10} - o(1)$ as well. The third case is like the second, but this time yields $\epsilon \geq \frac{1}{18} - o(1)$.

The fourth case remains, and is the core of the argument. Apply Lemma 4.3 to $P_2$ to obtain $A, B \subseteq \mathbb{F}_p$ with

$$|A|, |B| \ll K$$

and $G \subseteq A \times B$ with

$$|G| \approx \frac{|P_1| K^8}{|L_1|^4}$$

such that

$$|A^G - B|, |A^G / B| \ll K. \quad (4.13)$$

Note that either $K < p^{1/2}$ or $K > \frac{|G|}{p^{1/2}}$ since if

$$p^{1/2} \leq K \leq \frac{|G|}{p^{1/2}}$$

then $|G| \geq p$ which contradicts the fact that $|G| \leq N < p$.

If $K < p^{1/2}$ then $|A| \ll p^{1/2}$ and an application of Lemma 4.8 gives

$$|G|^{55} \lesssim K^{109}$$

which implies by the cardinality of $G$ that

$$|P_1|^{55} K^{331} \ll |L|^{220}.$$
4.5. Proving Theorems 4.1 and 4.2

Applying (4.11), (4.12) and the fact that $|L| \leq N$ gives after rearranging

$$N^{1/2} \lesssim N^{331\epsilon}$$

which implies $\epsilon \geq \frac{1}{662} - o(1)$ as required.

On the other hand, suppose that $K > \frac{|G|}{p^{1/2}}$. Assume also that $K \geq p^{1/2}$ since otherwise we are done by the previous paragraph. Note that with a fixed $G$, the bounds in (4.13) are not affected by passing to supersets of $A$ or $B$. So assume $|B| \approx K$ and thus

$$\frac{|G|}{|B|} \ll p^{1/2}.$$ 

Hence Lemma 4.9 is applicable, giving

$$|G|^{67} \lesssim K^{133}$$

and therefore

$$|P_1|^{67} K^{403} \lesssim |L|^{268}.$$ 

Since $K \geq p^{1/2} > N^{1/2}$ and $|L| \leq N$ this means

$$|P_1|^{67} K^{67} \lesssim N^{100}.$$ 

By (4.11) it falls out that $\epsilon \geq \frac{1}{134} - o(1)$, which is far better than required. \(\square\)

Proof of Theorem 4.2. For $l \in L(P)$, write $\mu(l)$ for the number of points in $P$ incident to $l$. It is clear, as per the proof of Beck’s theorem in Chapter 1 that

$$|P|^2 \approx \sum_{l \in L(P)} \mu(l)^2.$$ 

By a dyadic pigeonholing there exists $L_1 \subseteq L(P)$ and an integer $k$ such that $\mu(l) \approx k$ for all $l \in L_1$ and

$$|L_1|^2 \gtrsim |P|^2.$$ (4.14)

To prove the theorem it suffices to show that either

$$k \gtrsim |P|,$$

in which case there are $\tilde{\Omega}(|P|)$ collinear points, or

$$k^{133} \lesssim |P|^{66}.$$
Chapter 4. Incidence theorems over finite fields

in which case $k^2 \lesssim |P|^{\frac{6}{13}}$ and so $|L(P)| \geq |L_1| \gtrsim |P|^{1+\frac{1}{13}}$ as required.

Note that $I(P, L_1) \approx |L_1|k$. So by another dyadic pigeonholing there exists $P_1 \subseteq P$ and an integer $K$ such that every point in $P_1$ is incident to $\Theta(K)$ lines in $L_1$ and

$$|P_1|K \gtrsim |L_1|k \gtrsim \frac{|P|^2}{k}. \quad (4.15)$$

Since $|P_1| \leq |P|$ we also have

$$K \gtrsim \frac{|P|}{k}. \quad (4.16)$$

Applying Lemma 4.4 to $P_1$ and $L_1$, at least one of the following is true:

1. $K \ll \frac{|L_1|^{3/5}}{|P_1|^{1/5}}$.

2. There is a line in $L_1$ incident to at least $\Omega\left(\left(\frac{|P_1|K^4}{|L_1|^2}\right)^{1-o(1)}\right)$ points in $P_1$.

3. There is a line in $L_1$ incident to at least $\Omega\left(\left(\frac{|P_1|K^8}{|L_1|^4}\right)^{1-o(1)}\right)$ points in $P_1$.

4. There exists $P_2 \subseteq P_1$ with $|P_2| \approx \frac{|P_1|K^8}{|L_1|^4}$ and points $p_1, p_2, p_3, p_4$ in a $K$-sum-product configuration.

As with Theorem 4.1, the first three cases fall easily. In the first case $(4.15)$ and $(4.16)$ give

$$\frac{|P|^6}{k^5} \lesssim |L_1|^3$$

By $(4.14)$ this in turn implies that $k \lesssim 1$ which is far better than required. In the second case, since all lines in $L_1$ are incident to $\Theta(k)$ points in $P$ and hence $O(k)$ points in $P_1$, it follows that

$$k \gg \frac{|P_1|K^4}{|L_1|^2}.$$

By $(4.15)$ and $(4.16)$ this gives $|L_1|^2k^5 \gtrsim |P|^5$ and so $(4.14)$ implies $k \gtrsim |P|$. The third case similarly yields $k \gtrsim |P|$.

The fourth case is left. Apply Lemma 4.3 to $P_2$ to obtain $A, B \subseteq \mathbb{F}_p$ with $|A|, |B| \ll K$ and $G \subseteq A \times B$ with

$$|G| \approx \frac{|P_1|K^8}{|L_1|^4}.$$
4.6. Further work

such that

\[ |A \overset{G}{\ominus} B|, |A \overset{G}{/} B| \ll K. \]

As with the proof of Theorem 4.1, either \( K < p^{1/2} \) or \( K > \frac{|G|}{p^{1/2}}. \) If \( K < p \) then as per the last proof, Lemma 4.8 gives

\[ |P_1|^{55} K^{331} \ll |L_1|^{220}. \]

By (4.15) and (4.16) this gives

\[ |P|^{386} \approx |L_1|^{220} K^{331}. \]

By (4.14) we then get \( k^{109} \ll |P|^{54} \), which is better than the sufficient \( k^{133} \ll |P|^{66} \).

On the other hand, if \( K > \frac{|G|}{p^{1/2}} \) and \( K > p^{1/2} \) then by Lemma 4.9

\[ |P_1|^{67} K^{403} \ll |L_1|^{268}. \]

By (4.15) and (4.16) this gives

\[ |P|^{470} \ll |L_1|^{268} K^{403}. \]

By (4.14) it follows that

\[ k^{133} \ll |P|^{66} \]

as required. \( \Box \)

4.6 Further work

An intrinsic proof in the plane? The work here, and the previous work of Bourgain-Katz-Tao and Helfgott-Rudnev uses finite field sum-product estimates as an animating force. The progress in this chapter comes down to finding a particularly efficient way of relating this to incidences by showing that a ‘sum-product’ configuration of points and lines must arise if there are too many incidences, but that such a configuration cannot exist since it would contradict known finite field growth results.

Is it instead possible to construct a proof that lives entirely in the plane? Such an approach, if it worked, could yield better incidence results. Since the Elekes result (Theorem 2.71) in Chapter 2 shows that incidence bounds yield growth results, this could also drive forward the study of growth in finite fields.
There may be some cause for optimism. The setup of points and lines that we called a ‘sum-product’ configuration is not only interesting because of its relationship to sum-products. It is also of intrinsic interest on the plane, where it invites comparison with Desargues’ theorem. This theorem, actually a defining property of the projective plane, says that two triangles of points are perspective to a point if and only if they are perspective to a line, the meaning of which is illustrated in Figure 4.3 below.

Figure 4.3. An illustration of Desargues’ theorem. The two shaded triangles are perspective to the red point, and are also perspective to the line carrying the black, green and yellow points.

A sum-product configuration should imply the existence of many pairs of triangles perspective to the same point and perspective to the same line, at the same three places. Could pursuing the geometric implications of such an arrangement be a fruitful line of inquiry?
Chapter 2 introduced expander functions as a type of growth result. Recall that for a field $F$, an $n$-variable expander is a function $f : F^n \rightarrow F$ such that for any subset $A$ of $F$, maybe satisfying some nondegeneracy conditions, the set $f(A) = \{f(a_1, \ldots, a_n)\}$ is of cardinality at least $|A|^{1+\delta}$ with $\delta > 0$. This chapter sets new records in two variables over finite fields, and in three and four variables over real and complex numbers.

An earlier version of the two-variable finite field work formed one half of a joint paper \[26\] with Oliver Roche-Newton has been accepted for publication in the Journal of Combinatorial Theory Series A.

An earlier version of the three and four-variable real and complex results has been submitted to Discrete and Computational Geometry, and a preprint [25] is available on the arXiv.

### 5.1 Results

This section describes the new theorems proved in the chapter.

#### 5.1.1 Two variables

As mentioned in Chapter 2, the strongest known two variable expanders is the function $f(a, b) = a + ab$, which was first studied by Garaev and Shen [16]. They

\[\text{The paper is the union of two theorems proved independently by the two authors.}\]
proved three results about the size of the set \( f(A) \), depending on the ambient field and the density of \( A \) within it:

\[
\begin{align*}
\text{If } A \subseteq \mathbb{F}_p \text{ with } |A| < p^{1/2} \text{ then } |f(A)| &\gtrsim |A|^{1+\frac{1}{105}} \\
\text{If } A \subseteq \mathbb{F}_p \text{ with } |A| \geq p^{2/3} \text{ then } |f(A)| &\gg |A|^{1/2}p^{1/2} \\
\text{If } A \subseteq \mathbb{R} \text{ is finite then } |f(A)| &\gg |A|^{1+\frac{1}{4}}.
\end{align*}
\]

Result (5.2) is sharp but (5.1) and (5.3) are not. The first new result of this chapter is the following improvement on (5.1).

**Theorem 5.1.** If \( A \subseteq \mathbb{F}_p \) and \( |A| < p^{1/2} \) then

\[ |f(A)| \gtrsim |A|^{1+\frac{1}{53}}. \]

Result (5.3), has also been recently improved, in the same paper [26] as that containing an earlier version of Theorem 5.1, to

\[ |f(A)| \gtrsim |A|^{1+\frac{5}{19}}. \]

5.1.2 Three and four variables

As mentioned in Chapter 2, the breakthrough result of Guth and Katz [18] on the distinct distances problem yields the function

\[ (a - b)^2 + (c - d)^2 \]

as a four-variable expander over \( \mathbb{R} \) with \( \delta \geq 1 - o(1) \). Iosevich, Roche-Newton and Rudnev [21] used the same methodology to show that

\[ ad - bc \]

is likewise a four-variable expander with \( \delta \geq 1 - o(1) \). Both this and the Guth-Katz result are sharp, as shown by the case where \( A \) is an arithmetic progression.

The next two new results in this chapter are examples of functions with stronger growth properties. The first is a function in only three real variables rather than four which is nevertheless an expander with \( \delta \geq 1 - o(1) \).

**Theorem 5.2.** Let \( g(a, b, c) = \frac{a-b}{a-c} \). For any finite \( A \subseteq \mathbb{C} \) we have

\[ |g(A)| \gtrsim |A|^2. \]
5.2. Two-variable finite field expanders

The second is a four-variable expander with \( \delta \geq 1 \) instead of \( 1 - o(1) \).

**Theorem 5.3.** Let \( h(a, b, c, d) = \frac{(a - b)(c - d)}{(b - c)(a - d)} \). For any finite \( A \subseteq \mathbb{R} \) we have

\[
|h(A)| \gg |A|^2.
\]

As with the results of Guth and Katz and Iosevich, Roche-Newton and Rudnev, the example of an arithmetic progression shows that Theorem 5.2 is sharp up to logarithmic factors. However it is not clear that the same is true of Theorem 5.3: in the case of an arithmetic progression one has \( |h(A)| \gtrsim |A|^3 \) and so there may well be scope for improvement.

5.1.3 Structure

In the rest of this chapter, Section 5.2 gives the proof of Theorem 5.1 and Section 5.3 gives the proofs of Theorems 5.2 and 5.3. Section 5.4 suggests possible further work.

5.2 Two-variable finite field expanders

This section is concerned with proving Theorem 5.1. The overall strategy is similar to the finite field case in the work of Garaev and Shen [16], exploiting the fact that

\[
f(A) = \{a + ab : a, b \in A\}
\]
can be written as simply the product set \( A(A + 1) \). A sum-product philosophy suggests that at least one of \( A(A + 1) \) and the difference set \( A - A \) must be large. If it is the former then of course there is nothing to prove, so the idea is to deal with the latter case. That is, we want to show that if the difference set is large then so is \( A(A + 1) \).

The innovation here is to take a more efficient route than [16], via the simple Ruzsa-type observation that if \( ab = cd \) then

\[
(a + ab) - (c + cd) = a - c.
\]

Section 5.2.1 establishes the new key result and shows three different ways of implementing it. Section 5.2.2 then deploys these in a modified sum-product type proof to establish Theorem 5.1.
5.2.1 Bounding partial sumsets

The following result is the key to our progress.

**Lemma 5.4.** Let $A, B \subseteq \mathbb{F}_p$, and let $\epsilon > 0$. There exists $G \subseteq A \times B$ with $|G| \geq (1 - \epsilon)|A||B|$ such that

$$|A^G - B| \ll_{\epsilon} \frac{|A(B + 1)||B(A + 1)||A/B|}{|A||B|}.$$

**Proof.** Without loss of generality assume that $0 \notin A, B$. Note that

$$\sum_{x \in A/B} |A \cap xB| = |A||B|.$$

Let $X$ be the set of $x \in A/B$ for which $|A \cap xB| \geq \epsilon|A||B|$. Then

$$|A||B| = \sum_{x \in X} |A \cap xB| + \sum_{x \notin X} |A \cap xB| \leq \sum_{x \in X} |A \cap xB| + \epsilon|A||B|$$

and so

$$\sum_{x \in X} |A \cap xB| \geq (1 - \epsilon)|A||B|.$$

Let $G \subseteq A \times B$ be given by

$$G = \{(a, b) \in A \times B : \frac{a}{b} \in X\}$$

so that

$$|G| = \sum_{x \in X} |A \cap xB| \geq (1 - \epsilon)|A||B|.$$

For each $\xi \in (A - B)$ pick $a_\xi \in A$ and $b_\xi \in B$ such that

$$a_\xi - b_\xi = \xi.$$

Let $S \subseteq (A - B) \times A \times B$ be given by

$$S = \left\{ (\xi, c, d) : \frac{c}{d} = \frac{a_\xi}{b_\xi} \right\}.$$
5.2. Two-variable finite field expanders

Note that

$$|S| \gg_\epsilon \frac{|A||B||A^G - B|}{|A/B|}$$

since there are $|A^G - B|$ choices of $\xi$, each of which has at least $\frac{|A||B|}{|A/B|}$ associated pairs $(c, d)$. We now show that

$$|S| \leq |A(B + 1)||B(A + 1)|.$$ This will follow after showing that the map

$$\psi : S \to A(B + 1) \times B(A + 1)$$

$$\psi(\xi, c, d) = (a\xi + a\xi d, b\xi + b\xi c)$$
is an injection, i.e. that for given $(t_1, t_2)$ in $\psi(S)$ there is only one choice of $(\xi, c, d) \in S$ for which $\psi(\xi, c, d) = (t_1, t_2)$. Indeed for given $(t_1, t_2)$ in $\psi(S)$ it is immediate that

$$\xi = a\xi - b\xi = (a\xi + a\xi d) - (b\xi + b\xi c) = t_1 - t_2$$

so we know $\xi$ and therefore $a\xi$ and $b\xi$. We therefore also know $(c, d)$ since

$$t_1 = a\xi + a\xi d$$
$$t_2 = b\xi + b\xi c.$$ So $\psi$ is indeed an injection and the upper bound on $|S|$ holds.

Comparing the upper and lower bounds on $|S|$ shows that

$$\frac{|A||B||A^G - B|}{|A/B|} \ll_\epsilon |S| \leq |A(B + 1)||B(A + 1)|$$

and so

$$|A^G - B| \ll_\epsilon \frac{|A(B + 1)||B(A + 1)||A/B|}{|A||B|}$$
as required.

Let’s now apply this, using sumset calculus from Chapter 3 to give upper bounds on $|A - A|$ in terms of $|A(A + 1)|$. First, a Balog-Szemerédi-Gowers approach yields a quite-efficient relationship:
Corollary 5.5. For any set $A \subseteq \mathbb{F}_p$ there exists $A' \subseteq A$ with $|A'| \approx |A|$ such that $|A' - A'| \ll \frac{|A(A+1)|^8}{|A|^7}$.

Proof. Apply Lemma 5.4 with $A = B$ and some fixed sufficiently small $\epsilon > 0$ to find $G \subseteq A \times A$ with $|G| \geq (1 - \epsilon)|A|^2$ such that

$$|A^G - A| \ll \frac{|A(A + 1)|^2|A/A|}{|A|^2}.$$ 

By Lemma 3.4 there exists $A' \subseteq A$ with $|A'| \approx |A|$ such that

$$|A' - A'| \ll \frac{|A^G - A|^2}{|A|} \ll \frac{|A(A + 1)|^4|A/A|^2}{|A|^5}.$$ 

By Lemma 3.1 applied multiplicatively, $|A/A| \leq \frac{|A(A + 1)|^2}{|A|}$ and so the result follows.

Corollary 5.5 is useful by itself, but we can do better in some places. Covering results are often used in sum-product proofs, and applying Lemma 3.14 to Lemma 5.4 yields one that will be helpful here.

Corollary 5.6. Let $A, B, C \subseteq \mathbb{F}_p$, and $A, B \subseteq xC + y$ for some $x \in \mathbb{F}_p, y \in \mathbb{F}_p$. Let $0 < \epsilon < \frac{1}{16}$. Then $(1 - \epsilon)|A|$ elements of $A$ can be covered by

$$O\left(\frac{|C(C + 1)|^2|C/C|}{|A||B|^2}\right)$$

translates of $B$. Similarly, $(1 - \epsilon)|A|$ elements of $A$ can be covered by this many translates of $-B$.

Proof. Applying Lemma 5.4 to the sets $A_{xy} = \frac{A - y}{x} \subseteq C$ and $B_{xy} = \frac{B - y}{x} \subseteq C$ there exists $G_{xy} \subseteq A_{xy} \times B_{xy}$ of cardinality at least $\left(1 - \frac{\epsilon^2}{4}\right)|A||B|$ such that

$$|A_{xy}^{G_{xy}} - B_{xy}| \ll \frac{|A_{xy}(B_{xy} + 1)||B_{xy}(A_{xy} + 1)||A_{xy}/B_{xy}|}{|A||B|} \leq \frac{|C(C + 1)|^2|C/C|}{|A||B|}.$$ 

Then let

$$G = \left\{(a, b) : \left(\frac{a - y}{x}, \frac{b - y}{x}\right) \in G_{xy}\right\}$$

to obtain

$$|A^G - B| = |A_{xy}^{G_{xy}} - B_{xy}| \ll \frac{|C(C + 1)|^2|C/C|}{|A||B|}.$$ 

The result follows by applying Lemma 3.14 to $G$, with $\epsilon$ replaced by $\epsilon^2/4$.\qed
5.2. Two-variable finite field expanders

Applying Lemma 3.15 yields another helpful covering result.

**Corollary 5.7.** Let $0 < \epsilon < \frac{1}{2}$. Then there exists $G \subseteq A \times A$ with $|G| \geq (1-\epsilon)|A|^2$ such that $A - A$ is covered by $O_\epsilon \left( \frac{|A(A+1)|^2}{|A|^3} \right)$ translates of $A$.

**Proof.** By Lemma 5.4, with $\epsilon$ replaced by $\frac{\epsilon}{2}$, there exists $G' \subseteq A \times A$ with $|G'| \geq (1-\frac{\epsilon}{2})|A|^2$ such that

$$|A - A| \ll \epsilon \frac{|A(A+1)|^2}{|A|^2}.$$ 

Then by Lemma 3.15 there exists $G \subseteq G'$ with

$$|G| \geq \left(1 - \frac{\epsilon}{2}\right)|G'| \geq \left(1 - \frac{\epsilon}{2}\right)^2 |A|^2 \geq (1-\epsilon)|A|^2$$

such that $A - A$ is covered by $O_\epsilon \left( \frac{|A(A+1)|^2}{|A|^3} \right)$ translates of $A$, as required.  

### 5.2.2 Proving Theorem 5.1

This section uses Corollaries 5.5, 5.6, and 5.7 to prove Theorem 5.1.

By Corollary 5.5 and passing to a subset of $A$ if necessary we may assume that

$$|A - A| \ll \frac{|A(A+1)|^8}{|A|^7}. \quad (5.4)$$

By Corollary 3.3 and again passing to a subset if necessary we may assume that

$$|A - A - A - A| \ll \frac{|A - A|^3}{|A|^2}. \quad (5.5)$$

Now, by Corollary 3.10 and Lemma 3.11

$$\sum_{a,b \in A} |a(A+1) \cap b(A+1)| = E_\times(A, A+1) \geq \frac{|A|^4}{|A(A+1)|}.$$ 

So there exists $b_0 \in A$ such that

$$\sum_{a \in A} |a(A+1) \cap b_0(A+1)| \geq \frac{|A|^3}{|A(A+1)|}.$$ 

By dyadic pigeonholing there exists $A_1 \subseteq A$ and $N \in \mathbb{N}$ such that

$$|a(A+1) \cap b_0(A+1)| \approx N.$$
for all \( a \in A_1 \) and
\[
N|A_1| \gtrsim \frac{|A|^3}{|A(A + 1)|}. 
\] (5.6)

Since \( |A_1| \leq |A| \) this also implies
\[
N \gtrsim \frac{|A|^2}{|A(A + 1)|}. 
\] (5.7)

Now consider the set
\[
R(A_1) = \left\{ \frac{\alpha - \beta}{\gamma - \delta} : \alpha, \beta, \gamma, \delta \in A_1, \alpha \neq \beta, \gamma \neq \delta \right\}
\]
and break into two cases according to whether or not \( R(A_1) = \mathbb{F}_p \).

5.2.2.1 \( R(A_1) \neq \mathbb{F}_p \)

The important thing about \( R(A_1) \) is that if \( \xi / \in R(A_1) \) then
\[
|A_1 + \xi A_1| = |G| 
\] (5.8)
for any \( G \subseteq A_1 \times A_1 \). Note that (5.8) employs a slight abuse of notation: by \( A_1 + \xi A_1 \) it means the set of \( a + \xi b \) for which \( (a, b) \in G \). It holds because there can be no repetition in \( A_1 + \xi A_1 \), since if
\[
a + \xi b = c + \xi d
\]
with \( (a, b) \neq (c, d) \) then \( \xi = \frac{a - c}{b - d} \in R(A_1) \) which is a contradiction.

Now since \( R(A_1) \neq \mathbb{F}_p \) there must exist \( \xi = \frac{\alpha - \beta}{\gamma - \delta} \in R(A_1) \) such that \( \xi - 1 \notin R(A_1) \). For any \( G \subseteq A_1 \times A_1 \) it follows that
\[
|G| = |A_1 + (\xi - 1)A_1| 
\]
\[
= \left| A_1 + \frac{\alpha - \beta - \gamma + \delta}{\gamma - \delta} A_1 \right| 
\]
\[
= |(\gamma - \delta)A_1 + (\alpha - \beta - \gamma + \delta)A_1| 
\] (5.9)
where there is another abuse of notation in (5.9).

Now proceed to fix a particular choice of \( G \). Let \( \epsilon > 0 \) be sufficiently small and for convenience write
\[
\lambda = \frac{|A(A + 1)|^2|A/A|}{N^2|A_1|}.
\]
5.2. Two-variable finite field expanders

Applying Corollary 5.6 to the sets \( \alpha(A_1 + 1) \) and \( b_0(A_1 + 1) \cap \alpha(A_1 + 1) \) shows that there is a set \( A_\alpha \subseteq A_1 \) with
\[
|A_\alpha| \geq (1 - \epsilon)|A_1|
\]
such that \( \alpha A_\alpha \) is contained in the union of \( O_\epsilon(\lambda) \) translates of \( b_0A \). Similarly, there are sets \( A_\beta, A_\gamma, A_\delta \subseteq A_1 \) with
\[
|A_\beta|, |A_\gamma|, |A_\delta| \geq (1 - \epsilon)|A_1|
\]
such that \( \beta A_\beta \) and \( \gamma A_\gamma \) are contained in the union of \( O(\lambda) \) translates of \( b_0A \), and \( \delta A_\delta \) is contained in the union of \( O(\lambda) \) translates of \( -b_0A \). Let
\[
A_2 = A_\alpha \cap A_\beta \cap A_\gamma \cap A_\delta
\]
so that
\[
|A_2| \geq (1 - 4\epsilon)|A_1|.
\]
By Corollary 5.7 there exists \( G \subseteq A_2 \times A_2 \) with
\[
|G| \approx |A_2|^2 \approx |A_1|^2
\]
such that \( A_2^G - A_2 \) is covered by \( O\left(\frac{|A(A+1)|^2|A/A|}{|A_1|^3}\right) \) translates of \( A_2 \). Fix this choice of \( G \) so that from (5.9),
\[
|A_1|^2 \ll \left| (\gamma - \delta)A_2^G - (\gamma - \delta)A_2 + (\alpha - \beta)A_2 \right| \leq \frac{|(\gamma - \delta)A_2 - (\alpha - \beta)A_2||A(A+1)|^2|A/A|}{|A_1|^3}
\]
\[
\leq \frac{|\alpha A_2 - \beta A_2 - \gamma A_2 + \delta A_2||A(A+1)|^2|A/A|}{|A_1|^3}.
\]
Since
\[
A_2 \subseteq A_\alpha, A_\beta, A_\gamma, A_\delta,
\]
and \( \alpha A_\alpha, \beta A_\beta, \gamma A_\gamma \) are each contained in the union of \( O(\lambda) \) translates of \( b_0A \), and \( \delta A_\delta \) is contained in the union of \( O(\lambda) \) translates of \( -b_0A \), it follows that
\[
|A_1|^2 \ll \frac{|\alpha A_\alpha - \beta A_\beta - \gamma A_\gamma + \delta A_\delta||A(A+1)|^2|A/A|}{|A_1|^3}
\]
\[
\leq \frac{\lambda^4|b_0A - b_0A - b_0A - b_0A||A(A+1)|^2|A/A|}{|A_1|^3}
\]
\[
= \frac{|A - A - A - A||A(A+1)|^2|A/A|}{N^8|A_1|^7}.
\]
93
Chapter 5. Expander functions

By (5.5), this gives
\[ |A_1|^2 \leq \frac{|A - A|^3 |A(A + 1)|^8 |A/A|^5}{N^8 |A_1|^7 |A|^2}. \]  
(5.11)

Now by the Ruzsa triangle inequality (Lemma 3.1) applied multiplicatively,
\[ |A/A| \leq \frac{|A(A + 1)|^2}{|A|}. \]  
(5.12)

Applying (5.4) and (5.12) to (5.11) yields
\[ |A_1|^2 \leq \frac{|A(A + 1)|^{44}}{N^8 |A_1|^7 |A|^{28}}. \]

Rearranging and applying (5.6) and (5.7) gives
\[ |A(A + 1)|^{44} \gg |A_1|^9 N^8 |A|^28 \gg \frac{|A_1|^52 |A|^54}{|A(A + 1)|^8} \gg \frac{|A|^54}{|A(A + 1)|^9} \]
and so
\[ |A(A + 1)| \gtrsim |A|^{54/53} \]
as required.

5.2.2.2 $R(A_1) = \mathbb{F}_p$

Let $E$ be the number of solutions to
\[ a + \xi b = c + \xi d \]  
(5.13)
with $a, b, c, d \in A_1$ and $\xi \in R(A_1)$. Moreover, for each $\xi \in R(A_1)$ recall that $E_+(A_1, \xi A_1)$ is the additive energy of $A_1$ and $\xi A_1$, i.e. the number of solutions to (5.13) with $\xi$ fixed, so that
\[ E = \sum_{\xi \in R(A_1)} E_+(A_1, \xi A_1). \]

There are no more than $|R(A_1)||A_1|^2 = p|A_1|^2$ solutions to (5.13) for which $(a, b) = (c, d)$. And there are no more than $|A_1|^4$ solutions with $(a, b) \neq (c, d)$. So in total
\[ \sum_{\xi \in R(A_1)} E_+(A_1, \xi A_1) = E \leq |A_1|^4 + p|A_1|^2. \]
5.3. Three and four-variable real and complex expanders

Since \(|A_1| \leq |A| < p^{1/2}\) this gives

\[
\sum_{\xi \in R(A_1)} E_+(A_1, \xi A_1) \ll p|A_1|^2.
\]

So there exists \(\xi = \frac{\alpha - \beta}{\gamma - \delta} \in R(A_1)\) such that

\[
E_+(A_1, \xi A_1) \ll |A_1|^2.
\]

Moreover, for any \(A_2 \subseteq A_1\) we have also

\[
E_+(A_2, \xi A_2) \ll |A_1|^2.
\]

Now by Lemma 3.11,

\[
E_+(A_2, \xi A_2) \geq \frac{|A_2|^4}{|A_2 - \xi A_2|}
\]

and so if \(|A_2| \approx |A_1|\) then

\[
|\alpha A_2 - \beta A_2 - \gamma A_2 + \delta A_2| \geq \left| A_2 - \frac{\alpha - \beta}{\gamma - \delta} A_2 \right|
\]

\[
= |A_2 - \xi A_2|
\]

\[
\gg |A_1|^2.
\]

Let \(A_\alpha, A_\beta, A_\gamma, A_\delta\) be as before and fix

\[
A_2 = A_\alpha \cap A_\beta \cap A_\gamma \cap A_\delta.
\]

This yields the same situation as considered at (5.10) in the \(R(A_1) \neq F_p\) case, but with one less factor of \(\frac{|A(A+1)|^2[A/A]}{|A_1|^3}\) to deal with. So we obtain (and in fact exceed) the required bound.

\[
\square
\]

5.3 Three and four-variable real and complex expanders

This section is concerned with the proofs of Theorems 5.2 and 5.3.

Let’s first place these in the context of the Guth-Katz proof on distinct distances, which led to the example of the function \((a-b)^2 + (c-d)^2\) as a four-variable expander. This was based on a framework of Elekes and Sharir \[11\]. The idea,
when counting the number of distinct objects determined by a set, is to analyse functions under which that object is invariant.

In the Guth-Katz proof, where the objective is to count distances, the approach is to analyse the group $SE_2$ of orientation-preserving isometries. This analysis can be parameterised as an incidence problem of points and lines in $\mathbb{R}^3$. The usual Szemerédi-Trotter theorem is too weak to be of help directly, but Guth and Katz were able to amplify it to Theorem 1.9, which they developed using a novel ‘polynomial partitioning’ technique specifically for this purpose.

The approach of Iosevich, Roche-Newton and Rudnev for showing that $ad - bc$ is a four-variable expander follows the same approach, but uses the fact that this function can be viewed as a determinant. Thus the approach is to analyse determinant-preserving maps, i.e. elements of $SL_2$. Like the Guth-Katz result, this too required an application of Theorem 1.9.

The approach for proving Theorems 5.2 and 5.3 is likewise founded on the Elekes-Sharir paradigm. Whereas Guth and Katz counted distances by analysing isometries from the group $SE_2$, and Iosevich, Roche-Newton and Rudnev counted determinants by analysing $SL_2$, the functions $g$ and $h$ considered here are instances of cross ratios, which are preserved by the group $PSL_2$ of projective transformations of the line.

There are several advantages to working with cross ratios:

- As shown in Section 5.1, we are able to prove quantitatively stronger expander results.

- A smaller arsenal is required. Neither Theorem 1.9 nor any other application of the Guth-Katz polynomial partitioning technique is required. Instead only the Szemerédi-Trotter theorem on points and lines is used when proving Theorem 5.2 and only the Edelsbrunner-Guibas-Sharir theorem on points and planes is used when proving Theorem 5.3.

- Parameterising as an incidence problem is more straightforward. The Guth and Katz and Iosevich, Roche-Newton and Rudnev proofs go through a certain amount of hassle in order to make the parameterisation and verify that appropriate nondegeneracy conditions are satisfied. But with cross ratios and projective transformations everything falls out naturally.
In what follows, Section 5.3.1 describes the standard theory of cross ratios, establishing that they are invariants of projective transformations. Section 5.3.2 then identifies projective transformations in a natural way with points in three-dimensional projective space, and establishes how the transformations’ behaviour corresponds to line and plane structures of points. Section 5.3.3 then uses the results of the preceding two sections to prove Theorems 5.2 and 5.3.

5.3.1 Cross ratios

This section records some standard theory on cross ratios.

Recall, or consult Appendix B, that for a field $F$ the projective line $\mathbb{P}F^1$ is the set of equivalence classes $[x, y]$ of $F^2 \setminus \{(0, 0)\}$ modulo dilation, and that it can more concretely be viewed as the extended line $\overline{F} = F \cup \{\infty\}$ by identifying $[1, 0]$ with $\infty$ and $[x, 1]$ with $x \in F$. Under this identification, the cross ratio $X(a, b, c, d) \in \overline{F}$ of four elements $a, b, c, d \in \overline{F}$ is given by

$$X(a, b, c, d) = \frac{(a - b)(c - d)}{(b - c)(a - d)}$$

interpreted in the sense of limits where necessary. Note that the functions $g$ and $h$ from Theorems 5.2 and 5.3 are given by

$$g(a, b, c) = X(\infty, a, b, c)$$
$$h(a, b, c, d) = X(a, b, c, d)$$

So to prove the theorems it suffices respectively to show that

$$\# \{X(\infty, a, b, c) : a, b, c \in A\} \gg |A|^{2-o(1)} \text{ for any finite } A \subseteq \mathbb{C}$$
$$\# \{X(a, b, c, d) : a, b, c, d \in A\} \gg |A|^2 \text{ for any finite } A \subseteq \mathbb{R}.$$

The importance of the cross ratio is that it is a projective invariant of quadruples, in the sense of the following result which can be found in [39], for example.

**Lemma 5.8.** Let $a_i \in \overline{F}$ be distinct for $i = 1, 2, 3, 4$ and the same for $b_i \in \overline{F}$. Then $X(a_1, a_2, a_3, a_4) = X(b_1, b_2, b_3, b_4)$ if and only if there is a projective transformation in $\text{PSL}_2(F)$ that sends each $a_i$ to $b_i$.

**Proof.** By Lemma B.1, there is a unique projective transformation $\mu \in \text{PSL}_2(F)$ that sends $a_i$ to $b_i$ for $i = 1, 2, 3$. We shall show that $X(a_1, a_2, a_3, a_4) = X(b_1, b_2, b_3, b_4)$ if and only if $\mu$ also sends $a_4$ to $b_4$. 

97
First note that \( X(a, b, c, d) = \tau_{abc}(d) \) where \( \tau_{abc} \in \text{PSL}_2(F) \) is the unique projective transformation that sends \((a, b, c)\) to \((\infty, 1, 0)\). To see this it suffices simply to check that
\[
\tau_{abc} = \begin{bmatrix}
 b - a & (a - b) c \\
 c - b & (b - c) a
\end{bmatrix}
\]
and then that \( \tau_{abc}(d) = X(a, b, c, d) \).

It follows that \( X(a_1, a_2, a_3, a_4) = X(b_1, b_2, b_3, b_4) \) if and only if
\[
\tau_{a_1a_2a_3}(a_4) = \tau_{b_1b_2b_3}(b_4).
\]
Now by definition of \( \mu \) we have
\[
\tau_{a_1a_2a_3} = \tau_{b_1b_2b_3} \circ \mu
\]
and so by injectivity of \( \tau_{b_1b_2b_3} \), equation (5.14) holds precisely when \( \mu(a_4) = b_4 \).
This completes the proof.

5.3.2 Points, planes and transformations

This section contains two results that apply to any field \( F \). The first, a ‘points lemma’, identifies projective transformations from \( \text{PSL}_2(F) \) with points in \( \mathbb{P}F^3 \). The second, a ‘planes lemma’, establishes that the behaviour of transformations corresponds to line and plane structures of their associated points.

**Lemma 5.9** (Points lemma). Define \( \psi : \text{PSL}_2(F) \to \mathbb{P}F^3 \) by
\[
\psi \begin{bmatrix} p & q \\ r & s \end{bmatrix} = [p, q, r, s].
\]
The map \( \psi \) is well-defined and injective, and its image is \( \mathbb{P}F^3 \setminus Q \) where \( Q \) is the quadratic surface given by \( ps = qr \).

**Proof.** That \( \psi \) is well-defined and injective follows from checking that if \( t_1, t_2 \in \text{SL}_2(F) \) then \( \psi[t_1] = \psi[t_2] \) if and only if \( t_1 = \pm t_2 \). That the image is \( \mathbb{P}F^3 \setminus Q \) follows from the definition \( \text{PSL}_2(F) = \text{SL}_2(F)/\pm I \).

**Lemma 5.10** (Planes lemma). Let \( \psi \) be as in the points lemma. For each \((a, b) \in F \times F\) there is a plane \( \pi_{ab} \subseteq \mathbb{P}F^3 \) with the following properties.
5.3. Three and four-variable real and complex expanders

1. If \( \tau \in PSL_2(F) \) then \( \tau(a) = b \) if and only if \( \psi(\tau) \in \pi_{ab} \).

2. No three planes are collinear.

3. Different pairs \((a, b) \in F \times F\) determine different planes \(\pi_{ab}\).

4. Different pairs of planes \(\{\pi_{ab}, \pi_{cd}\}\) intersect in different lines \(\pi_{ab} \cap \pi_{cd}\).

5. For any \(A \subseteq F\), a point \(p \in \mathbb{P}F^3 \setminus Q\) is incident to at most \(|A|\) of the planes from \(\{\pi_{ab} : a, b \in A\}\).

Proof. From the theory in Appendix B, a projective transformation \(\tau = \begin{pmatrix} p & q \\ r & s \end{pmatrix}\) sends \(a\) to \(b\) if and only if

\[
\frac{ap + q}{ar + s} = b,
\]

which is the same as

\[
ap + q - bar - bs = 0.
\]

For fixed \(a, b\) this is an homogeneous linear constraint on \(\psi(\tau) = [p, q, r, s] \in \mathbb{P}F^3\) and so describes a plane in \(\mathbb{P}F^3\), which we define to be \(\pi_{ab}\). Property 1 is satisfied by construction, and it is now straightforward to establish properties 2 to 5 in turn.

2. It suffices to show that three planes intersect in a point. Let \((a, b, c)\) and \((d, e, f)\) be two triples of distinct elements of \(F\). By Lemma B.1 there is a unique \(\tau \in PSL_2(F)\) that sends \((a, b, c)\) to \((d, e, f)\). So

\[
\pi_{ad} \cap \pi_{be} \cap \pi_{ef} = \psi(\tau)
\]

which is a single point in \(\mathbb{P}F^3\).

3. If \(\pi_{ab} = \pi_{cd}\) for some \((a, b) \neq (c, d)\) then \(\pi_{ab} \cap \pi_{cd} \cap \pi_{ef}\) is either a line or a plane for any third pair \((e, f)\), which contradicts property 2.

4. Suppose that

\[
\pi_{ab} \cap \pi_{cd} = \pi_{a'd'b'} \cap \pi_{c'd'}.
\]

Then

\[
\pi_{ab} \cap \pi_{cd} \cap \pi_{a'd'b'} = \pi_{a'b'} \cap \pi_{c'd'}.
\]

But by property 2 the set on the left hand side is a point, whereas that on the right is a line, unless \(\pi_{ab} \in \{\pi_{a'b'}, \pi_{c'd'}\}\). Similarly a contradiction follows unless \(\pi_{cd} \in \{\pi_{a'b'}, \pi_{c'd'}\}\).
Chapter 5. Expander functions

5. Let $p$ be a point in $\mathbb{P}F^3 \setminus Q$, so that $p = \psi(\tau)$ for some $\tau \in \text{PSL}_2(F)$. For each $a \in A$ there is at most one $b \in A$ for which $p$ is incident to $\pi_{ab}$, as otherwise $\tau(a)$ would take two different values. Counting over all $a \in A$ shows that $p$ is incident to at most $|A|$ planes.

5.3.3 Proving Theorems 5.2 and 5.3

This section uses the results from Sections 5.3.1 and 5.3.2 to prove Theorems 5.2 and 5.3. We first give the proof of Theorem 5.2, which uses the Szemerédi-Trotter theorem.

Proof of Theorem 5.2. Since

$$|g(A)| = \# \{X(\infty, a, b, c) : a, b, c \in A\},$$

we want to show that

$$\# \{X(\infty, a, b, c) : a, b, c \in A\} \gg |A|^{2-o(1)}$$

for any finite $A \subseteq \mathbb{C}$. To this end write $E(A)$ for the number of solutions to the equation

$$X(\infty, a_1, a_2, a_3) = X(\infty, b_1, b_2, b_3) \quad (5.15)$$

with each of the $a_i$ and $b_i$ in $A$. Write $\mu(x)$ for the number of $a_1, a_2, a_3 \in A$ with $X(\infty, a_1, a_2, a_3) = x$. Then

$$\sum_{x \in g(A)} \mu(x) \approx |A|^3$$

and Cauchy-Schwarz implies that

$$|A|^6 \approx \left( \sum_{x \in g(A)} \mu(x) \right)^2 \leq |g(A)|E(A).$$

So it suffices to show

$$E(A) \ll |A|^{4+o(1)}.$$
By Lemma 5.8, equation (5.15) is satisfied precisely when there exists \( \tau \in PSL_2(\mathbb{C}) \) that fixes \( \infty \) and sends each \( a_i \) to \( b_i \). So if we define
\[
T = \bigcup_{a,b \in A} \{ \tau : \tau(\infty) = \infty, \tau(a) = b \}
\]
and write \( N(\tau) \) for the number of \( (a, b) \in A^2 \) for which \( \tau(a) = b \), then
\[
E(A) = \sum_{a_1, a_2, a_3 \in A} \sum_{b_1, b_2, b_3 \in A} \mathbf{1}(X(\infty, a_1, a_2, a_3) = X(\infty, b_1, b_2, b_3))
\]
\[
\leq \sum_{a_1, a_2, a_3 \in A} \sum_{b_1, b_2, b_3 \in A} \sum_{\tau \in T} \mathbf{1}(\tau(a_i) = b_i \text{ for each } i)
\]
\[
\leq \sum_{\tau \in T} N(\tau)^3. 
\tag{5.16}
\]

Let \( \psi \) be as in the points lemma. Define a set \( P \) of points by
\[
P = \psi(T)
\]
and a set of \( L \) lines by
\[
L = \{ \pi_{ab} \cap \pi_{\infty\infty} : a, b \in A \}
\]
so that \( |L| \approx |A|^2 \). The points and lines all lie in the plane \( \pi_{\infty\infty} \). Moreover, if we write \( m(p) \) for the number of lines from \( L \) incident to a point \( p \in P \), then
\[
N(\tau) = \# \{ (a, b) : \tau(a) = b \}
\]
\[
= \# \{ (a, b) : \psi(\tau) \in \pi_{ab} \cap \pi_{\infty\infty} \}
\]
\[
= m(\psi(\tau)). 
\tag{5.17}
\]
Combining (5.16) and (5.17) gives
\[
E(A) \leq \sum_{p \in P} m(p)^3. 
\tag{5.18}
\]

For each \( j \in \mathbb{N} \) write \( P_j \) for the set of \( p \in P \) with \( m(p) \in [2^j, 2^{j-1}) \). Applying the complex Szemerédi-Trotter theorem in the form of Corollary 1.15 gives
\[
|P_j| \ll \frac{|L|^2}{2^{3j}} + \frac{|L|}{2^j}
\]
Chapter 5. Expander functions

and so

\[ E(A) \ll \sum_{j=0}^{\log |A|} |P_j|2^{3j} \]

\[ \ll \sum_{j=0}^{\log |A|} \left( \frac{|L|^2}{2^{3j}} + \frac{|L|}{2^j} \right) 2^{3j} \]

\[ \approx |A|^4 \log |A| \]

as required. \( \square \)

We now give the proof of Theorem 5.3, which uses the Edelsbrunner-Guibas-Sharir theorem.

Proof of Theorem 5.3. This time we want to show that

\[ |h(A)| = \# \{ X(a_1, a_2, a_3, a_4) : a_i \in A \} \gg |A|^2 \]

for any finite \( A \subseteq \mathbb{R} \). To this end, we this time write \( E(A) \) for the number of solutions to the equation

\[ X(a_1, a_2, a_3, a_4) = X(b_1, b_2, b_3, b_4) \] \hspace{1cm} (5.19)

with the \( a_i, b_i \in A \). Using Cauchy-Schwarz as in Theorem 5.2 shows that

\[ |h(A)| \gg \frac{|A|^8}{E(A)} \]

so it suffices to show that

\[ E(A) \ll |A|^6. \]

Equation (5.19) is satisfied precisely when there exists \( \tau \in PSL_2(\mathbb{R}) \) that sends \( a_i \) to \( b_i \) for each \( i \). Define

\[ T = \bigcup_{a,b \in A} \{ \tau : \tau(a) = b \} \]

and write \( N(\tau) \) for the number of \( (a, b) \in A^2 \) for which \( \tau(a) = b \). Then

\[ E(A) \ll \sum_{\tau \in T} N(\tau)^4. \]
5.4. Further work

Let $\psi$ be as in the points lemma. Define a set $P \subseteq \mathbb{R}^3$ of points by

$$P = \psi(T)$$

and a set $\Pi$ of planes by

$$\Pi = \{\pi_{ab} : a, b \in A\}$$

so that $|\Pi| \approx |A|^2$. If we write $m(p)$ for the number of planes from $\Pi$ incident to
a point $p$, then

$$N(\tau) = m(\psi(\tau)).$$

So, following the same argument as for (5.18) in the preceding proof,

$$E(A) \leq \sum_{p \in P} m(p)^4.$$

For each $j \in \mathbb{N}$ write $P_j$ for the set of $p \in P$ with $m(p) \in [2^j, 2^{j+1})$. Then
applying the Edelsbrunner-Guiber-Sharir theorem on points and planes in the
form of Corollary 1.11 gives

$$E(A) \ll \sum_{j=0}^{\log |A|} |P_j| 2^{4j} \ll \sum_{j=0}^{\log |A|} \left(\frac{|\Pi|^3}{2^{2j}} + \frac{|\Pi|}{2^j}\right) 2^{4j} \ll |A|^6 \sum_{j=0}^{\infty} \frac{1}{2^j} \approx |A|^6$$

as required.

5.4 Further work

• Where else could Ruzsa-type observations lead? The Ruzsa-type observation that

$$(a + ab) - (c + cd) = a - c$$

whenever $ab = cd$ leads to a doubling of the growth exponent for the function $a + ab$ over finite fields. Could similar observations enable the construction of other expander functions, or give improved bounds on existing ones like $a + \frac{1}{b}$, $a + b^2$, or $a^2 + ab$?
Chapter 5. Expander functions

- **What more can be proved using cross ratios?** Cross ratios turned out to be very useful when constructing expanders in Theorems 5.2 and 5.3. For example, Theorem 5.2 is a sharp three-variable result whose bound of $|g(A)| \gtrsim |A|^2$ is as strong as the best previously known four-variable results. What more can be said?

As a start, it seems reasonable to conjecture that the four-variable Theorem 5.3 can be strengthened to $|h(A)| \gtrsim |A|^3$. This would bring the theorem into line with the case where $A$ is an arithmetic progression, which is a sharp example for Theorem 5.2.

In terms of applications to other growth results, might the strength of cross-ratio estimates enable new things to be said about two-variable expanders, or about sum-product estimates?
A sum-product theorem in function fields

This chapter proves a sum-product theorem in the function field \( \mathbb{F}_q(t) \), showing that if \( A \) is a finite subset of \( \mathbb{F}_q(t) \) then

\[
\max \{ |A + A|, |AA| \} \gg_q |A|^{1 + \frac{1}{5} - o(1)}.
\]

The exponent of \( \frac{1}{5} - o(1) \) lies between the \( \frac{1}{3} - o(1) \) known in the complex setting and the \( \frac{1}{11} - o(1) \) known in the finite field setting. This reflects the fact that unlike finite fields, \( \mathbb{F}_q(t) \) has an associated norm and topology, but that this norm is very different from that on \( \mathbb{C} \), with an unusually rigid ‘non-archimedean’ geometry.

The material in this chapter is joint work with Thomas Bloom.

6.1 Results

Recall that \( \mathbb{F}_q \) denotes the finite field of order \( q \), where \( q = p^\alpha \) is a prime power. The prime \( p \) is the characteristic of \( \mathbb{F}_q \), i.e. the least \( n \) such that

\[
x + \ldots + x = 0
\]

for all \( x \in \mathbb{F}_q \). Throughout this chapter the letter \( p \) will be reserved for this characteristic. The function field \( \mathbb{F}_q(t) \) is the field of rational functions of a transcendental element \( t \) over \( \mathbb{F}_q \). Elements are therefore of the form

\[
x = \sum_{j=-\infty}^{N} x_j t^j
\]
where the $x_j$ are elements of $\mathbb{F}_q$. Note that this means that although $\mathbb{F}_q$ is finite, the function field $\mathbb{F}_q(t)$ is not.

As with finite fields, it is necessary to rule out the possibility of finite subfields to be able to say anything non-trivial about growth in $\mathbb{F}_q(t)$. But unlike finite fields, this is accomplished by a *minimum* rather than a *maximum* condition on the cardinality of a set $A$. The reason is that the only finite subfields of $\mathbb{F}_q(t)$ are $\mathbb{F}_q$ and its subfields: to rule these out it suffices to insist that $A$ is a bit bigger than $q$. Think therefore of $q$ as being small and fixed, as opposed to finite fields where it is taken to be very large indeed. A convenient way of capturing this necessity is to introduce an implicit dependency on $q$ when formulating sum-product estimates, so that results are of the form

$$\max \{ |A + A|, |AA| \} \gg_q |A|^{1+\delta}$$

for an absolute $\delta > 0$ and any finite $A \subseteq \mathbb{F}_q(t)$.

As mentioned in Chapter 2, Li and Roche-Newton [33] obtained a sum-product estimate $\delta \geq \frac{1}{11} - o(1)$ for a finite field $\mathbb{F}_q$ whose order is not necessarily prime, extending the applicability of Rudnev’s result for $\mathbb{F}_p$. Because of its combinatorial generality this proof should go through in the function field setting without additional complication.

It is possible to do better by developing techniques specific to function fields. This chapter proves the following theorem

**Theorem 6.1.** If $A \subseteq \mathbb{F}_q(t)$ is finite then

$$|A + A|^3 |AA|^2 \gg_q |A|^{6-\omega(1)}.$$

A sum-product result for function fields with $\delta \geq \frac{1}{3} - o(1)$ follows immediately.

**Corollary 6.2.** If $A \subseteq \mathbb{F}_q(t)$ is finite then

$$\max \{ |A + A|, |AA| \} \gg_q |A|^{1+\frac{1}{3} - o(1)}.$$

The next section provides more background on function fields, and explains the structure of the proof and the rest of the chapter.
6.2 Function fields

This section has two parts. The first part, Section 6.2.1, gives some standard background on the geometry of function fields. The second part, Section 6.2.2, explains how this will be used in the proof of Theorem 6.1 and describes how the rest of the chapter is organised.

6.2.1 Background

The field $\mathbb{F}_q(t)$ has a norm or valuation $| \cdot |$ given by

$$|x| = \begin{cases} q^{\deg(x)} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

where $\deg(x)$ is the degree of $x$, i.e. the maximal $j$ for which $x_j$ is non-zero. This valuation has the non-archimedean property that

$$|x + y| \leq \max\{|x|, |y|\}$$

which is stronger than the usual triangle inequality. As a consequence $\mathbb{F}_q(t)$ has an unusually rigid geometry, which will be exploited when proving sum-product estimates. A particular concern will be the behaviour of balls, which are as usual sets of the form

$$B(x, r) = \{y \in \mathbb{F}_q(t) : |x - y| \leq r\}.$$

In $\mathbb{F}_q(t)$, the non-archimedean property implies the following fact, which is considered to be standard.

**Lemma 6.3.** If $B_1$ and $B_2$ are balls in $\mathbb{F}_q(t)$ then either they are disjoint, or $B_1 \subseteq B_2$, or $B_2 \subseteq B_1$. If in addition $B_1$ and $B_2$ have the same radius then either they are disjoint or $B_1 = B_2$.

**Proof.** Let $B_1 = B(x, r)$ and $B_2 = B(y, s)$. If there exists $a \in B(x, r) \cap B(y, s)$ then

$$|x - y| \leq \max\{|a - x|, |a - y|\} \leq \max\{r, s\}.$$ 

If $r \leq s$ then this implies $B(x, r) \subseteq B(y, s)$ since if $b \in B(x, r)$ then

$$|y - b| \leq \max\{|y - x|, |b - x|\} \leq \max\{r, s\} = s.$$ 

Conversely if $s \leq r$ then $B(y, s) \subseteq B(x, r)$. Hence if $r = s$ then $B(x, r) = B(y, s)$. 

$$\square$$
6.2.2 Sum-product considerations

The $\mathbb{F}_q(t)$ sum-product proof builds upon an approach of Solymosi [41] for sum-products in $\mathbb{C}$. When adapting this method, the non-archimedean geometry of $\mathbb{F}_q(t)$ turns out to be a mixed blessing.

First, the bad news. Solymosi’s argument fails at a critical point in the $\mathbb{F}_q(t)$ setting, for the following reason. For each $a \in A$, let $a' \in A \setminus \{a\}$ be such that $|a - a'|$ is minimal, and let $B_a$ be the ball of radius $|a - a'|$ centred on $a$. Solymosi’s method uses the crucial fact that a single complex number can be contained in at most $O(1)$ of the $B_a$. This fails spectacularly in $\mathbb{F}_q(t)$: in this setting an element could be contained in as many as $|A|$ of the $B_a$, as demonstrated by the example

$$A = \{t^j : 0 \leq j \leq n\}$$

where

$$B_{t^j} = \{x \in \mathbb{F}_q(t) : |x| \leq q^j\}$$

for $j \geq 1$ and $B_1 = B_t$, meaning that every one of the $|A|$ balls contains 0 as an element.

But all is not lost. In the example above, the astute reader will notice that $|A + A| \approx |A|^2$, and so a strong-sum product estimate holds despite the failure of Solymosi’s argument. In fact we will be able to show that something like this is possible whenever the Solymosi argument fails, by defining and considering separable sets.

Say that a set $A$ is **separable** if its elements can be indexed as

$$A = \{a_1, \ldots, a_{|A|}\}$$

in such a way that for any $1 \leq j \leq |A|$ there is a ball $B_j$ with

$$A \cap B_j = \{a_1, \ldots, a_j\}.$$  

Separability is fairly unexciting in the complex setting, but in the non-archimedean regime of $\mathbb{F}_q(t)$ it is a stronger notion. The rigid geometry makes it harder to find separable sets, but where they do exist it will in fact imply the existence of large sumsets. The idea, therefore, is to show that a large separable sets must exist whenever the Solymosi approach fails. Combining this with an analysis of separable sets as having large sumsets will lead to a proof of Theorem 6.1.
6.3. Separable sets

In what follows, Section 6.3 analyses separable sets and develops the necessary results about their sumsets. Section 6.4 then adapts Solymosi’s proof from [41] to establish that if $|A + A|$ and $|AA|$ are both small then $A$ must contain a large separable set. Section 6.5 uses these results to prove Theorem 6.1. Section 6.6 considers some directions for further work.

6.3 Separable sets

This section analyses sumsets of separable sets. Recall that a set $A \subseteq \mathbb{F}_q(t)$ is separable if its elements can be indexed as

$$A = \{a_1, \ldots, a_{|A|}\}$$

in such a way that for each $1 \leq j \leq |A|$ there is a ball $B_j$ with

$$A \cap B_j = \{a_1, \ldots, a_j\}.$$

Say that the balls $B_j$ separate $A$. It is an immediate consequence of the definition that a subset of a separable set is itself separable.

As in Chapter 3 write $kA$ for the $k$-fold sumset of $A$. The following lemma shows that if $A$ is separable then $kA$ has essentially maximum-possible cardinality.

Lemma 6.4. If $A$ is separable then

$$|kA| \gg_k |A|^k$$

for any natural number $k$.

Proof. Let $E_k(A)$ denote the $k$-fold additive energy of $A$, i.e. the number of solutions to

$$a_1 + \ldots + a_k = b_1 + \ldots + b_k \quad (6.1)$$

with the $a_i, b_i \in A$. For $x \in kA$ write $\mu(x)$ for the number of solutions to $x = a_1 + \ldots + a_k$. By Cauchy-Schwarz as in Lemma 3.11

$$|A|^{2k} \approx \left( \sum_{x \in kA} \mu(x) \right)^2 \leq |kA|E_k(A)$$

and so it suffices to show that $E_k(A) \ll_k |A|^k$, i.e. that there are at most $O_k(|A|^k)$ solutions to (6.1).
Say that a solution to (6.1) is **trivial** if at least $2^k - 1$ of the $2^k$ terms occur with multiplicity at least 2. By elementary counting there are at most $O_k(|A|^k)$ trivial solutions, so it suffices to show that there are no non-trivial solutions.

Suppose for a contradiction that a non-trivial solution to (6.1) exists. Gathering terms gives an expression of the form

$$n_1c_1 + \ldots + n_tc_t = 0$$

(6.2)

where the $c_i$ are distinct elements of $A$ and, since we are in characteristic $p$, the $n_i$ are integers non-strictly between 1 and $p - 1$. The assumption of non-triviality implies that $t \geq 2$, since at least two of the terms have $n_i \in \{1, p - 1\}$. Additionally, note that

$$n_1 + \ldots + n_t \equiv 0 \pmod{p}.$$  

(6.3)

Indeed after gathering terms on the left the different multiplicities $n_i$ must sum to zero, since there are the same number of terms on the left of (6.1) as on the right. Any $c_i$ for which $n_i \equiv 0 \pmod{p}$ is discarded, meaning that the sum of the remaining multiplicities is 0 (mod $p$) as well.

Now since $A$ is separable and the $c_i$ are in $A$ we may relabel them if necessary to assume the existence of a ball $B(x, r)$ such that $c_1 \notin B$ but $c_2, \ldots, c_t \in B$. By (6.2),

$$|c_1 - x| = |n_1c_1 - n_1x|$$

$$= |n_2c_2 + \ldots + n_tc_t + n_1x|.$$  

Then by (6.3) and the non-archimedean property it follows that

$$|c_1 - x| = |n_2(c_2 - x) + \ldots + n_t(c_t - x)|$$

$$\leq \max \{|c_2 - x|, \ldots, |c_t - x|\}$$

$$\leq r$$

and hence $c_1 \in B(x, r)$ which is a contradiction. Thus there are no non-trivial solutions and the proof is complete.

**6.4 Finding many separable sets**

The goal in this section is to show that if the sumset and product set of a set $A$ are both small then it must contain a large separable set. For this it adapts the
6.4. Finding many separable sets

argument of Solymosi [41] for complex sum-products discussed in Section 6.2.2. Note that all of the analysis remains in the $\mathbb{F}_q(t)$ setting; indeed some of the facts of non-archimedean geometry deployed here are manifestly false in $\mathbb{C}$.

A couple of new definitions are required. Define

$$r_A(a) = \min_{a' \in A \setminus \{a\}} |a - a'|$$

$$B_A(a) = B(a, r_A(a)).$$

Additionally, say that $C \subseteq A$ is an $A$-chain if its elements can be indexed as $C = \{c_1, \ldots, c_n\}$ in such a way that

$$B_A(c_1) \subseteq \ldots \subseteq B_A(c_n).$$

The following argument, a strengthened form of that found in [41], finds a large chain in $A$ as long as the sumset and product set are both small. In the event that this condition were to fail there would of course be nothing to prove.

**Lemma 6.5.** Any set $A$ contains an $A$-chain of cardinality

$$\Omega \left( \frac{|A|^5}{|A + A|^2|AA|^2 \log^3 |A|} \right).$$

**Proof.** For each $a \in A$ write $N(a)$ for the maximal cardinality $N$ of an $A$-chain $C = \{c_1, \ldots, c_N\}$ for which $c_N = a$. Note for future reference that

$$N(a) \leq |B_A(a) \cap A|$$

since if $C$ is such an $A$-chain then $C \subseteq A$ by definition and for each $c \in C$ we have $c \in B_A(c) \subseteq B_A(a)$.

It suffices to find $a \in A$ such that

$$N(a) \gg \frac{|A|^5}{|A + A|^2|AA|^2 \log^3 |A|}.$$ 

Begin with a dyadic pigeonholing. For each $0 \leq j \leq \log_2 |A|$ define $A_j$ to be the set of $a \in A$ for which $2^j \leq N(a) < 2^{j+1}$. The $A_j$ partition $A$ and so

$$\sum_{j=0}^{\log_2 |A|} |A_j| = |A|.$$
Hence there exists $j$ for which $|A_j| \gg |A|/ \log |A|$. We shall show that

$$2^j \gg \frac{|A|^5}{|A + A|^2 |AA|^2 \log^4 |A|}.$$  

To this end, say that a pair $(a,c) \in A \times A$ is **additively good** if

$$|(A + A) \cap (B_A(a) + c)| \leq \frac{2^{j+3}|A + A|}{|A_j|}$$

and that $(a,d) \in A \times A$ is **multiplicatively good** if

$$|(AA) \cap (d \cdot B_A(a))| \leq \frac{2^{j+3}|AA|}{|A_j|}.$$

Say that a quadruple $(a,b,c,d) \in A^4$ is **good** if

1. $a \in A_j$.
2. $b \in B_A(a) \cap A$.
3. $(a,c)$ is additively good.
4. $(a,d)$ is multiplicatively good.

Write $Q$ for the number of good quadruples. We shall bound $Q$ from below to obtain

$$Q \gg 2^j |A_j||A|^2$$  \hspace{1cm} (6.4)

and bound it from above to obtain

$$Q \ll \frac{2^{2j}|A + A|^2 |AA|^2}{|A_j|^2}.$$  \hspace{1cm} (6.5)

Comparing (6.4) and (6.5) will then give the required bound on $2^j$ since $|A_j| \gg |A|/ \log |A|$. Let’s first establish (6.4). For fixed $c \in A$ we have

$$\sum_{a \in A_j} |(A + A) \cap (B_A(a) + c)| = \sum_{a \in A_j} \sum_{u \in A + A} 1 \ (u \in B_A(a) + c) = \sum_{v \in A + A - c} \sum_{a \in A_j} 1 \ (v \in B_A(a)) = \sum_{v \in A + A - c} |C_j(v)|$$

where $C_j(v)$ is the set of $a \in A_j$ with $v \in B_A(a)$. 
6.4. Finding many separable sets

Note that $C_j(v)$ is an $A$-chain. This follows from Lemma 6.3 since for any two $a, b \in C_j(v)$ we have $v \in B_A(a) \cap B_A(b)$ and so either $B_A(a) \subseteq B_A(b)$ or $B_A(b) \subseteq B_A(a)$.

Now since $C_j(v) \subseteq A_j$ and $C_j(v)$ is an $A$-chain, there is an $a \in A_j$ for which

$$|C_j(v)| \leq N(a) \leq 2^{j+1}.$$

We therefore have

$$\sum_{a \in A_j} |(A + A) \cap (B_A(a) + c)| \leq 2^{j+1}|A + A|$$

and hence

$$|(A + A) \cap (B_A(a) + c)| \leq \frac{2^{j+3}|A + A|}{|A_j|}$$

holds for at least $3|A_j|/4$ elements $a \in A_j$. So for fixed $c \in A$ there are at least $3|A_j|/4$ elements $a \in A_j$ for which $(a, c)$ is additively good.

By the same argument we may show that for fixed $d \in A \setminus \{0\}$ there are at least $3|A_j|/4$ elements $a \in A_j$ for which $(a, d)$ is multiplicatively good.

Thus for any $c \in A$ and $d \in A \setminus \{0\}$ there are at least $|A_j|/2$ elements $a \in A_j$ for which $(a, c)$ is additively good and $(a, d)$ is multiplicatively good, i.e. for which conditions 3 and 4 hold. Furthermore for each such $a \in A_j$ there are at least $2^j$ elements $b \in A$ for which condition 2 holds, since

$$2^j \leq N(a) \leq |B_A(a) \cap A|.$$

In total therefore,

$$Q \gg |A|^2|A_j|2^j$$

which concludes the proof of (6.4).

We now prove (6.5). Note that the map

$$(a, b, c, d) \mapsto (a + c, b + c, ad, bd)$$

is injective and so it suffices to bound the number of possibilities for this latter expression, subject to the constraint that $(a, b, c, d)$ is good. There are certainly at most $|A + A|$ possibilities for $a + c$ and at most $|AA|$ for $ad$, so it suffices to show that if these are fixed then there are at most $O(2^j|A + A|/|A_j|)$ possibilities for $b + c$ and at most $O(2^j|AA|/|A_j|)$ for $bd$. 

113
First establish the bound on the number of $b + c$. Note that if
\[ a + c = a' + c' \]
then either
\[ B_A(a) + c \subseteq B_A(a') + c' \]
or
\[ B_A(a') + c' \subseteq B_A(a) + c \]
since both sets are balls with the same centre $a + c$.

As a consequence, if $G \subseteq A \times A$ is the set of additively good pairs $(a, c)$, then for any $x \in A + A$ there is a fixed additively good pair $(a_x, c_x)$ such that
\[ B_A(a) + c \subseteq B_A(a_x) + c_x \]
whenever $a + c = x$ and $(a, c)$ is additively good. Thus if $a + c = x$ is the fixed first co-ordinate and $b + c$ is a possible second co-ordinate then since $b \in B_A(a) \cap A$ and $c \in A$ we have
\[ b + c \in (A + A) \cap (B_A(a) + c) \]
\[ \subseteq (A + A) \cap (B_A(a_x) + c_x). \]
Since $(a_x, c_x)$ is additively good, there are, as required, at most $O(2^{j+4}|A + A|/|A_j|)$ possibilities for $b + c$. The argument that there are at most at most $O(2^j AA|/|A_j|)$ for $bd$ is similar.

In total therefore
\[ Q \leq 2^{2j+4}|A + A|^2|AA|^2 \]
which concludes the proof of (6.5) and thus of the lemma.

The following result shows that any chain contains a large separable subset, allowing Lemma 6.4 to be applied to the chain found in Lemma 6.5.

**Lemma 6.6.** If $C$ is an $A$-chain then $C$ contains a separable set of cardinality at least $|C|/q$.

**Proof.** Observe that any subset $\{c_1, \ldots, c_n\} \subseteq C$ with
\[ B_A(c_1) \subset \ldots \subset B_A(c_n). \]
6.5. Proof of Theorem 6.1

is separable. Indeed, such a set is separated by the balls $B_A(c_i)$ because if $c_{i+1}$ were an element of $B_A(c_i)$ it would follow that $r_A(c_{i+1}) = r_A(c_i)$ and so by Lemma 6.3 we would have the contradiction $B_A(c_i) = B_A(c_{i+1})$.

Define an equivalence relation on elements of $A$ by $a ∼ b$ if and only if $B_A(a) = B_A(b)$. To prove the lemma it suffices to show that each equivalence class contains at most $q$ elements of $A$.

Note first that if $a ∼ b$ then $|a - b| = r_A(a) = r_A(b)$.

Indeed, since $B_A(a) = B_A(b)$ it follows that $b ∈ B_A(a)$ and so $|a - b| ≤ r_A(a)$. However by minimality, $|a - b| ≥ r_A(a)$ and so $|a - b| = r_A(a)$. Similarly $|a - b| = r_A(b)$.

Suppose for a contradiction that there is an equivalence class containing elements $a_1, \ldots, a_{q+1}$. Consider differences $a_1 - a_i$ for $2 ≤ i ≤ q + 1$. By the last paragraph we have $|a_1 - a_i| = r_A(a_1) = r_A(a_i)$.

Now look at the leading terms of the $a_1 - a_i$. Since the leading term must be non-zero, there are only $q - 1$ possibilities and so by the pigeonhole principle there must exist $i ≠ j$ such that $a_1 - a_i$ and $a_1 - a_j$ have the same leading term. Since $|a_1 - a_i| = |a_1 - a_j| = r_A(a_1)$ it follows that $a_1 - a_i$ and $a_1 - a_j$ have the same degree, and that this is strictly greater than the degree of $a_i - a_j = (a_1 - a_j) - (a_1 - a_i)$.

This yields the contradiction $r_A(a_i) = |a_i - a_j| < |a_1 - a_i| = r_A(a_i)$ and so concludes the proof.

6.5 Proof of Theorem 6.1

Theorem 6.1 now follows by combining Lemma 6.4 from Section 6.3 with Lemma 6.5 and Lemma 6.6 from Section 6.4.
Proof of Theorem 6.1. By Lemma 6.5 the set $A$ contains an $A$-chain of cardinality 
\[ \Omega \left( \frac{|A|^5}{|A + A|^2|AA|^2 \log^3 |A|} \right). \]
By Lemma 6.6 it therefore contains a separable set $S$ of cardinality 
\[ \Omega \left( \frac{|A|^5}{q|A + A|^2|AA|^2 \log^3 |A|} \right) \]
and so Lemma 6.4 implies 
\[ |kA| \gg |kS| \gg \left( \frac{|A|^5}{q|A + A|^2|AA|^2 \log^3 |A|} \right)^k. \]
Plünnecke’s inequality (Lemma 3.2) shows that $|kA| \ll \frac{|A + A|^k}{|A|^k}$ for any $k \in \mathbb{N}$ and so combining upper and lower bounds on $|kA|$ gives 
\[ |A + A|^k \gg_k \left( \frac{|A|^5}{q|A + A|^2|AA|^2 \log^3 |A|} \right)^k |A|^{k-1}. \]
Taking $k$-th roots, we get 
\[ |A + A| \gg_k \frac{|A|^{6-\frac{1}{k}}}{q|A + A|^2|AA|^2 \log^3 |A|}. \]
Letting $k$ tend to infinity and rearranging then yields 
\[ |A + A|^3|AA|^2 \gg_q |A|^{6-o(1)} \]
as required.

\[
\]

6.6 Further work

• Incidences and expanders. Now that we have a sum-product estimate in $\mathbb{F}_q(t)$, it is possible to obtain expander results like Theorem 5.1 and incidence results like Theorem 4.1 and Theorem 4.2 for function fields without much fuss. These will be stronger than the finite field case, but the gap versus the $\mathbb{F}_q(t)$ sum-product estimate will be quite large.

Is it possible to do better by working directly with the non-archimedean geometry of $\mathbb{F}_q(t)$? For example, can we obtain an incidence bound in $\mathbb{F}_q(t)^2$ that is almost as strong as the Szemerédi-Trotter theorem in $\mathbb{R}^2$?
6.6. Further work

- **Computer science applications.** At the start of the thesis, we said that we would not worry about applications. However it is worth mentioning that some existing applications of finite field growth results to theoretical computer science may be improved by considering $\mathbb{F}_q(t)$ instead.

- **Other arithmetic combinatorics problems.** There are many more topics in arithmetic combinatorics than considered in this chapter, or indeed in this thesis. Thomas Bloom, with whom the work in this chapter is joint, is investigating a number of such problems in $\mathbb{F}_q(t)$. His upcoming thesis is likely to be worth a read.
Appendix A

Pigeonholing

This appendix summarises some standard pigeonholing results used throughout the thesis.

A.1 Averaging

Averaging results enable us to take information about the average behaviour of a set and deduce the existence of elements with particular behaviour. The most basic result of this kind shows that at least one element must be at least average, and at least one element must be at most average. It follows by elementary pigeonholing and so we record it without proof.

**Lemma A.1.** Let $A$ be a finite set of real numbers. Then at least one element of $A$ must be greater than or equal to $\frac{1}{|A|} \sum_{a \in A} a$ and at least one must be less than or equal to $\frac{1}{|A|} \sum_{a \in A} a$.

Lemma A.1 is so standard that it is used throughout the thesis without reference. More developed versions of this approach enable one to show that not just one element but many must exhibit behaviour not very much different from the average. These are constructed in the main body of the thesis to deal with particular situations.
Appendix A. Pigeonholing

A.2 Dyadic pigeonholing

The phrase ‘dyadic pigeonholing’ refers to the following result, which at the price of a logarithmic factor allows us to assume that a variable is essentially constant.

**Lemma A.2** (Dyadic pigeonholing). Let $A$ be a finite set of real numbers strictly greater than one and less than or equal to $\alpha$. Then there exists an integer $k$ and a subset $A'$ of $A$ such that every element of $A'$ lies in the interval $(k, 2k]$ and

$$|A'|k \gg \sum_{a \in A} \frac{a}{\log \alpha}.$$ 

*Proof.* For each integer $0 \leq j \leq \lceil \log_2 \alpha \rceil$, let

$$A_j = A \cap (2^j, 2^{j+1}].$$

The sets $A_j$ partition $A$ and so

$$\sum_{a \in A} a \approx \sum_{j=1}^{\lceil \log_2 \alpha \rceil} |A_j|2^j.$$

Hence there is a $j$ for which

$$|A_j|2^j \gg \sum_{a \in A} \frac{a}{\log \alpha}.$$

Setting $k = 2^j$ and $A' = A_{2j}$, the proof is complete. \qed

A.3 The Cauchy-Schwarz inequality

The phrase ‘by Cauchy-Schwarz’ typically has two meanings in the literature, and this is reflected in the thesis. It should be clear from the context which meaning is implied. The first meaning is the standard Cauchy-Schwarz inequality.

**Lemma A.3** (Cauchy-Schwarz). Let $A, B$ be finite sets of real numbers. Then

$$\sum_{a \in A, b \in B} ab \leq \left( \sum_{a \in A} a^2 \right)^{1/2} \left( \sum_{b \in B} b^2 \right)^{1/2}.$$ 

The second meaning is a particular application of Cauchy-Schwarz to pairwise intersection of sets, as follows.
A.3. The Cauchy-Schwarz inequality

**Corollary A.4.** Let $A$ be a finite set, and suppose we have a collection of subsets $A_i \subseteq A$, indexed by a finite set $I$. Then

$$\sum_{i \in I} |A_i| \leq |A|^{1/2} \left( \sum_{i,j \in I} |A_i \cap A_j| \right)^{1/2}.$$

**Proof.** We have

$$\sum_{i \in I} |A_i| = \sum_{a \in A} \left( \sum_{i \in I} 1 \ (a \in A_i) \right).$$

Hence by the Cauchy-Schwarz inequality,

$$\sum_{i \in I} |A_i| \leq |A|^{1/2} \left( \sum_{a \in A} \sum_{i,j \in I} 1 \ (a \in A_i) \ 1 \ (a \in A_j) \right)^{1/2}$$

$$= |A|^{1/2} \left( \sum_{i,j \in I} |A_i \cap A_j| \right)^{1/2}.$$

□
Appendix B

Projective geometry

This appendix gives the background in projective geometry necessary for Chapters 4 and 5.

First it defines projective space $\mathbb{P}F^n$ over a field $F$. Then it shows how $\mathbb{P}F^n$ can be considered as the union of affine space $F^n$ and a hyperplane ‘at infinity’. Lastly, it defines projective transformations and establishes some useful facts about their transitivity.

The approach is based on that in [39].

B.1 Projective space

Let $F$ be a field. Define projective $n$-space $\mathbb{P}F^n$ to be

$$\mathbb{P}F^n = (F^{n+1} \setminus \{0\}) / \sim$$

where $\sim$ is the equivalence relation given by dilation, i.e if $x, y \in F^{n+1} \setminus \{0\}$ then $x \sim y$ if and only if $\lambda x = y$ for some $\lambda \in F \setminus \{0\}$. Elements of $\mathbb{P}F^n$ are therefore equivalence classes, and we write $[x]$ for the equivalence class containing $x \in F^{n+1} \setminus \{0\}$.

We will be concerned with linear subspaces of $\mathbb{P}F^n$. In affine space $F^n$ an $(n - 1)$-dimensional hyperplane is the locus of zeroes $x = (x_1, \ldots, x_n)$ of a linear, possibly inhomogeneous, equation in $n$ variables

$$a_1x_1 + \ldots + a_nx_n + a_{n+1} = 0$$

(B.1)
where the $a_i$ are fixed elements of $F$.

But in projective space, an $(n-1)$-dimensional hyperplane is the set of $[x] \in \mathbb{P}F^n$ for which $x \in F^{n+1}$ satisfies a linear homogeneous equation in $n+1$ variables

$$a_1x_1 + \ldots + a_nx_n + a_{n+1}x_{n+1} = 0. \quad (B.2)$$

The homogeneity of (B.2) ensures that this is well-defined. Note that $[x]$ and $[y]$ lie in the same $(n-1)$-dimensional projective hyperplane of $\mathbb{P}F^n$ if and only if $x$ and $y$ lie in the same $n$-dimensional affine hyperplane of $F^{n+1}$.

### B.2 The hyperplane at infinity

It is often helpful to think of $\mathbb{P}F^n$ as the union of $F^n$ with an $(n-1)$-dimensional hyperplane ‘at infinity’. The idea is that two parallel $(n-1)$-spaces are disjoint in $F^n$, but in $\mathbb{P}F^n$ they intersect in an $(n-2)$-space on the hyperplane at infinity. Moreover, all $(n-1)$-spaces of the same gradient will intersect in the same $(n-2)$-space at infinity.

For example, the projective line $\mathbb{P}F^1$ can be viewed as the extended line $F \cup \{\infty\}$. And the projective plane $\mathbb{P}F^2$ can be viewed as $F^2 \cup l_\infty$ where $l_\infty$ is the projective line at infinity. In this latter case, two parallel lines in $F^2$ intersect in a point on $l_\infty$, and all lines of the same gradient intersect at the same such point.

To justify this interpretation, view $F^{n+1}$ as $F^n \times F$ and identify $x \in F^n$ with $[x, 1] \in \mathbb{P}F^n$. This preserves hyperplanes, since $x$ lies in the affine hyperplane given by (B.1) if and only if $[x, 1]$ lies in the projective hyperplane given by (B.2). This accounts for all elements of $\mathbb{P}F^n$ apart from those of the form $[x, 0]$ with $x \in F^n$. These form the projective $(n-1)$-space given by $x_{n+1} = 0$, which we call the ‘hyperplane at infinity’. The verification of the claim that all $(n-1)$-spaces of the same gradient intersect in the same $(n-2)$-space is left as an exercise.

### B.3 Projective transformations

The group $PSL_{n+1}(F)$ of projective transformations of $\mathbb{P}F^n$ is defined by

$$PSL_{n+1}(F) = SL_{n+1}(F)/\pm I$$
B.3. Projective transformations

where $I$ is the identity. Elements are therefore equivalence classes $[T]$ of linear transformations $T \in SL_{n+1}(F)$. The group has an action on $PF^n$ given by

$$[T][x] = [T(x)].$$

It is easy to check that this action is well-defined, that elements of $PSL_{n+1}(F)$ are permutations of $PF^n$, and that they preserve linear subspaces.

The action also has an important transitivity property. Say that $(n+2)$ distinct points $p_i \in PF^n$ are a frame if no $n+1$ of them lie in the same $(n-1)$-dimensional projective hyperplane. For example, three distinct elements of $PF^1$ are a frame, and four distinct elements of $PF^2$ are a frame if no three of them are collinear. The following result shows that the action of $PSL_n(F)$ is sharply transitive on frames.

**Lemma B.1** (Sharp transitivity on frames). Let $(p_1, \ldots, p_{n+2})$ and $(q_1, \ldots, q_{n+2})$ be two frames of points in $PF^n$. There is a unique projective transformation $\tau \in PSL_{n+1}(F)$ such that

$$(q_1, \ldots, q_{n+2}) = (\tau(p_1), \ldots, \tau(p_{n+2})).$$

**Proof.** Let $e_i$ with $1 \leq i \leq n+1$ be elements of the canonical basis of $F^{n+1}$, let $f_i = [e_i] \in PF^n$ and define

$$e_* = \sum_{i=1}^{n+1} e_i$$

and $f_* = [e_*]$.

It suffices to show that for any frame $(p_1, \ldots, p_{n+2})$ of points in $PF^n$ there exists a unique $\tau \in PSL_{n+1}(F)$ that sends $p_i$ to $f_i$ for each $1 \leq i \leq n+1$ and sends $p_{n+2}$ to $f_*$. Indeed, if this is established then given frames $(p_1, \ldots, p_{n+2})$ and $(q_1, \ldots, q_{n+2})$ there are unique $\tau_1, \tau_2 \in PSL_{n+1}(F)$ that send both to $(f_1, \ldots, f_{n+1}, f_*)$. Then $\mu = \tau_2^{-1}\tau_1$ is the unique map that sends $(p_1, \ldots, p_{n+2})$ to $(q_1, \ldots, q_{n+2})$.

We first prove the existence of an appropriate $\tau$. Say that $p_i \in PF^n$ is given by $p_i = [t_i]$ with $t_i \in F^{n+1}$. Since the first $n+1$ points $p_i$ are not coplanar in $PF^n$, the corresponding $t_i$ are not coplanar in $F^{n+1}$ and so form a basis. Hence we can write

$$t_{n+2} = \sum_{i=1}^{n+1} \lambda_i t_i$$

(B.3)
Appendix B. Projective geometry

with the $\lambda_i$ all elements of $F$. Let $T \in SL_{n+1}(F)$ be a linear transformation that sends the $F^{n+1}$-basis $(\lambda_1 t_1, \ldots, \lambda_{n+1} t_{n+1})$ to a scalar multiple of the canonical basis $(e_1, \ldots, e_{n+1})$. Let $\tau = [T]$. Then

$$\tau(p_i) = [T][t_i] = [T(t_i)] = [e_i] = f_i$$

for each $1 \leq i \leq n+1$. Additionally (B.3) and the choice of $T$ imply that

$$\tau(p_{n+2}) = [T(t_{n+2})] = \left[ \sum_{i=1}^{n} \tau(\lambda_i t_i) \right] = f_*$$

and so we have established existence.

We now prove uniqueness, for which it suffices to show that our choice of $[T] \in PSL_{n+1}(F)$ is the only one that sends $p_i$ to $f_i$ for $1 \leq i \leq n+1$ and $p_{n+2}$ to $f_*$. So suppose that $[T]$ has this property. Then $[T][t_i] = [e_i]$ and so there exists $\mu_i \in F$ such that

$$T(\lambda_i t_i) = \mu_i e_i.$$ 

But since $[T][t_{n+2}] = [e_*]$ there exists $\mu_*$ such that

$$T\left( \sum_{i=1}^{n+1} \lambda_i t_i \right) = \mu_* \sum_{i=1}^{n+1} e_i.$$ 

Combining gives

$$\sum_{i=1}^{n+1} \mu_i e_i = \mu_* \sum_{i=1}^{n+1} e_i$$

and so by linear independence $\mu_i = \mu_*$ for all $i$. In other words, $T$ sends the basis $(\lambda_1 t_1, \ldots, \lambda_{n+1} t_{n+1})$ to a scalar multiple of the canonical basis $(e_1, \ldots, e_{n+1})$, as required. \qed
References

[1] P. Agarwal and B. Aronov. Counting facets and incidences. *Discrete Comput. Geom.*, 7(4):359–369, 1992.

[2] R. Apfelbaum and M. Sharir. Large complete bipartite subgraphs in incidence graphs of points and hyperplanes. *SIAM J. Discrete Math.*, 21(3):707–725, 2007.

[3] J. Beck. On the lattice property of the plane and some problems of Dirac, Motzkin and Erdős in combinatorial geometry. *Combinatorica*, 3(3-4):281–297, 1983.

[4] J. Bourgain. More on the sum-product phenomenon in prime fields and its applications. *Int. J. Number Theory*, 1(1):1–32, 2005.

[5] J. Bourgain and M. Garaev. On a variant of sum-product estimates and explicit exponential sum bounds in prime fields. *Math. Proc. Cambridge Philos. Soc.*, 146(1):1–21, 2009.

[6] J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields and applications. *Geom. Funct. Anal.*, 14(1):27–57, 2004.

[7] P. Brass and C. Knauer. On counting point-hyperplane incidences. *Comput. Geom.*, 25(1-2):13–20, 2003.

[8] H. Edelsbrunner, L. Guibas, and M. Sharir. The complexity of many cells in arrangements of planes and related problems. *Discrete Comput. Geom.*, 5(2):197–216, 1990.
References

[9] G. Elekes. On the number of sums and products. *Acta. Arith.*, 81(4):365–367, 1997.

[10] G. Elekes, M. Nathanson, and I. Ruzsa. Convexity and sumsets. *Journal of Number Theory*, 83(2):194–201, 1999.

[11] G. Elekes and M. Sharir. Incidences in three dimensions and distinct distances in the plane. *Computational geometry (SCG’10)*, pages 413–422, 2010.

[12] G. Elekes and C. Tóth. Incidences of not-too-degenerate hyperplanes. *Computational geometry (SCG’05)*, pages 16–21, 2005.

[13] P. Erdős and E. Szemerédi. On sums and products of integers. *Stud. Pure Math.*, pages 213–218, 1983.

[14] K. Ford. Sums and products from a finite set of real numbers. *Ramanujan J.*, 2(1-2):59–66, 1998.

[15] M. Garaev. The sum-product estimate for large subsets of prime fields. *Proc. Amer. Math. Soc.*, 136(8):2735–2739, 2008.

[16] M. Garaev and C. Shen. On the size of the set $A(A + 1)$. *Math. Z.*, 265(1):125–132, 2010.

[17] W. Gowers. A new way of proving sumset estimates. (http://gowers.wordpress.com/2011/02/10/a-new-way-of-proving-sumset-estimates), 2011.

[18] L. Guth and N. Katz. On the Erdős distinct distance problem in the plane. Preprint arXiv:1011.4105v3, 2010.

[19] D. Hart, L. Li, and C. Shen. Fourier analysis and expanding phenomena in finite fields. *Proc. Amer. Math. Soc.* To appear, posted online June 19 2012.

[20] H. Helfgott and M. Rudnev. An explicit incidence theorem in $\mathbb{F}_p$. *Mathematika*, 57(1):135–145, 2011.

[21] A. Iosevich, O. Roche-Newton, and M. Rudnev. On an application of Guth-Katz theorem. *Math. Res. Lett.*, 18(4):691–697, 2011.
References

[22] A. Iosevich, M. Rudnev, and Y. Zhai. Areas of triangles and Beck's theorem in planes over finite fields. Preprint arXiv:1205.0107, 2012.

[23] T. Jones. An improved incidence bound over fields of prime order. Preprint arXiv:1110.4752v2, 2011.

[24] T. Jones. Further improvements to incidence and Beck-type bounds over prime finite fields. Preprint arXiv:1206.4517, 2012.

[25] T. Jones. New results on the growth of sets of real numbers. Preprint arXiv:1202.4972, 2012.

[26] T. Jones and O. Roche-Newton. Improved bounds on the set $A(A+1)$. *Journal of Combinatorial Theory, Series A*, 120:515–526, 2013.

[27] H. Kaplan, J. Matoušek, and M. Sharir. Simple proofs of classical theorems in discrete geometry using the Guth-Katz polynomial partitioning technique. *Discrete Comput. Geom.*, 4(3):499–517, 2012.

[28] N. Katz and C. Shen. Garaev’s inequality in finite fields not of prime order. *Online J. Anal. Comb.*, (3), 2008.

[29] N. Katz and C. Shen. A slight improvement to Garaev’s sum product estimate. *Proc. Amer. Math. Soc.*, 136(7):2499–2504, 2008.

[30] S. Konyagin and M. Rudnev. On new sum-product type estimates. Preprint arXiv:1207.6785, 2012.

[31] L. Li. Slightly improved sum-product estimates in fields of prime order. *Acta Arith.*, 147(2):153–160, 2011.

[32] L. Li and O. Roche-Newton. Convexity and a sum-product type estimate. Preprint arXiv:1111.5159, 2011.

[33] L. Li and O. Roche-Newton. An improved sum-product estimate for general finite fields. *SIAM J. Discrete Math.*, 25(3):1285–1296, 2011.

[34] M. Nathanson. On sums and products of integers. *Proc. Amer. Math. Soc.*, 125(1):9–16, 1997.

[35] J. Pach and M. Sharir. On the number of incidences between points and curves. *Combin. Probab. Comput.*, 7(1):121–127, 1998.
References

[36] G. Petridis. Plünnecke’s inequality. *Combin. Probab. Comput.*, 20(6):921–938, 2011.

[37] M. Rudnev. An improved sum-product inequality in fields of prime order. *Int. Math. Res. Notices*, 16:3693–3704, 2012.

[38] I. Ruzsa. An application of graph theory to additive number theory. *Sci. Ser. A Math. Sci. (N.S.)*, 3:97–109, 1989.

[39] P. Samuel. *Projective Geometry*. Springer-Verlag, 1988. ISBN 0-387-96752-4.

[40] C. Shen. Quantitative sum product estimates on different sets. *Electron. J. Combin.*, 15(1), 2008.

[41] J. Solymosi. On sum-sets and product-sets of complex numbers. *J. Théor. Nombres Bordeaux*, 17(3):921–924, 2005.

[42] J. Solymosi. On the number of sums and products. *Bull. London Math. Soc.*, 37(4):491–494, 2005.

[43] J. Solymosi. Bounding multiplicative energy by the sumset. *Adv. Math.*, 222(2):402–408, 2009.

[44] J. Solymosi and T. Tao. An incidence theorem in higher dimensions. *Discrete Comput. Geom.*, 48(2):255–280, 2012.

[45] L. Székely. Crossing numbers and hard Erdős problems in discrete geometry. *Combin. Probab. Comput.*, 6(3):353–358, 1997.

[46] E. Szemerédi and W. T. Trotter. Extremal problems in discrete geometry. *Combinatorica*, 3(3-4):381–392, 1983.

[47] T. Tao. The sum-product phenomenon in arbitrary rings. *Contrib. Discrete Math.*, 4(2):59–82, 2009.

[48] T. Tao and V. Vu. *Additive Combinatorics*. Cambridge University Press, 2006. ISBN 0-521-85386-9.

[49] C. Tóth. The Szemerédi-Trotter theorem in the complex plane. Preprint arXiv:0305283v4, 2011.

[50] L. Vinh. The szemerédi-trotter type theorem and the sum-product estimate in finite fields. *European J. Combin.*, 32(8):1177–1181, 2011.
[51] J. Zahl. A Szemerédi-Trotter type theorem in $\mathbb{R}^4$. Preprint arXiv:1203.4600, 2012.