Additive Combinatorics

Jacob Denson

February 8, 2021

Table Of Contents

Chapter 1

Additive Combinatorics

Additive combinatorics studies the structure of *additive sets*, i.e. subsets of an additive group, often the set of integers. Unlike the methods of number theory, which look at particular number theoretic sets with a strong amount of structure (e.g. the set of primes, the set of squares) we look at more general sets with less structure.

Since we will almost always deal with subsets of the integers rather than subsets of the reals, for notational convenience we let the intervals $[a, b]$, $[a, b)$, (a, b), and $(a, b]$ denote the family of *integers* in these intervals rather than the family of all reals in these intervals, unless otherwise specified.

1.1 The Inverse Set Problem

If A and B are subsets of an additive group G, then

$$\max(\#(A), \#(B)) \leqslant \#(A + B) \leqslant \#(A)\#(B).$$

Accounting for symmetry in expressions when $A = B$, we have

$$\#(A) \leqslant \#(A + A) \leqslant \frac{\#(A)(\#(A) + 1)}{2}.$$

More generally, a simple stars and bars argument shows that

$$|A| \leqslant |A^{\oplus n}| \leqslant \binom{|A| + n - 1}{n},$$

where $A^{\oplus n}$ denotes the n-fold sumset. These bounds are all sharp, e.g. if $\#(A) = 1$. It is a general heuristic that for the 'average' pair of sets A and B, the quantity $\#(A + B)$ is much more likely to be close to $\#(A)\#(B)$, than to $\max(\#(A), \#(B))$.

Example. *Let A and B to be sets of size N and M choosen uniformly at random from $[0, 1]$. Then*
$$\mathbf{P}(\#(A + B) = \#(A)\#(B)) = 1.$$
If we let $A = \{X_1, \ldots, X_N\}$ and $B = \{Y_1, \ldots, Y_M\}$, then the variables $\{X_i, Y_j\}$ are independant and uniformly distributed on $[0, 1]$, and for any distinct indices (i_1, j_1) and (i_2, j_2),
$$\mathbf{P}(X_{i_1} + Y_{j_1} = X_{i_2} + Y_{j_2}) = 0.$$
Thus for any distinct pairs (i_1, j_1) and (i_2, j_2), $\mathbf{P}(X_{i_1} + Y_{j_1} = X_{i_2} + Y_{j_2}) = 0$. Taking a union bound shows that with probability one, the elements $\{X_i + Y_j\}$ are distinct for all i and j, and so $\mathbf{P}(\#(A + B) = NM) = 1$.

Example. *Let A and B be subsets of $\mathbf{Z}_N$ with each element in A and B uniformly chosen with probability K/N. Then for each $n \in \mathbf{Z}/N\mathbf{Z}$,*

$$\begin{aligned}
\mathbf{P}(n \in A + B) &= \sum_{i=0}^{N-1} \mathbf{P}(i \in A)\,\mathbf{P}(n - i \in B) \\
&\quad - \sum_{i \neq j} \mathbf{P}(i \in A)\,\mathbf{P}(n - i \in B)\,\mathbf{P}(j \in A)\,\mathbf{P}(n - j \in B) \\
&\geqslant K^2/N - K^4/N^2.
\end{aligned}$$

Thus $\mathbf{E}(\#(A + B)) \geqslant K^2 - K^4/N = \mathbf{E}(\#(A)\#(B)) - K^4/N$. Markov's inequality implies that with high probability, $\#(A + B) \approx \#(A)\#(B)$ if $K \ll N^{1/2}$.

Thus if $\#(A + B)$ is relatively small, it is natural to expect there to be some structure to the sets A and B. A fundamental problem in additive combinatorics is to characterize this structure. If $\#(A+B) = \max(\#(A), \#(B))$, then A and B have maximal additive structure, and in this case we have a strong characterization of this structure.

Theorem 1.1. *Suppose A and B are finite additive sets. Then the following are equivalent:*

1. $\#(A + B) = \#(A)$.

2. $\#(A - B) = \#(A)$.

3. $\#(A + B^{\oplus n} - B^{\oplus m}) = \#(A)$ *for some* $(n, m) \neq 0$.

4. $\#(A + B^{\oplus n} - B^{\oplus m}) = \#(A)$ *for all* n, m.

5. *There exists a finite subgroup H of G such that A is a union of cosets in G/H, and B is contained in a single coset of G/H.*

Proof. Assume $0 \in B$ without loss of generality. Then 1. implies $A + B = A$, so for each $a \in A$, $b \in B$, $a + b \in A$. But this means that B is a subset of the *symmetry group* of A, i.e. $\mathrm{Sym}(A) = \{x \in G : A + x = A\}$. If $v, w \in G$, $w \in A$, and $v - w \in \mathrm{Sym}(A)$, then $v = w + (v - w) \in A + (v - w) = A$. Thus A is a union of cosets. Thus we have shown 1. implies 5. A similar argument shows 2. and 3. imply 5., and the remaining implications are simple to verify. $\square$

Corollary 1.2. *If $A = A + A$, then A is a subgroup of G.*

If we have an exact equality $\#(A + B) = \#(A)\#(B)$, we do not have quite as much structure but we note the condition is equivalent to the following properties:

- $\#(A + B) = \#(A)\#(B)$.

- $\#(A - B) = \#(A)\#(B)$.

- $\#\{(a_1, a_2, b_1, b_2) \in A^2 \times B^2 : a_1 + b_1 = a_2 + b_2\} = \#(A)\#(B)$.

- $\#\{(a_1, a_2, b_1, b_2) \in A^2 \times B^2 : a_1 - b_1 = a_2 - b_2\} = \#(A)\#(B)$.

- $(A - A) \cap (B - B) = \{0\}$.

- For each $c \in A + B$, $\#((A - c) \cap B) = 1$.

- For each $c \in B - A$, $\#((A + c) \cap B) = 1$.

Remark. It is not always true that $\#(A + B) = \#(A - B)$. For instance, if $A = \{0, 1, 3\}$ in $\mathbf{Z}$ then $A + A = \{0, 1, 2, 3, 4, 6\}$ whereas $A - A = \{-1, 0, 1, 2, 3\}$. Thus in this case $\#(A + A) = 6$ but $\#(A - A) = 5$.

1.2 Doubling Constant

Another way to measure the arithmetic structure of sets is via the *doubling constant*

$$\sigma[A] = \frac{|A+A|}{|A|}.$$

If K is small, we have lots of arithmetic structure, whereas if K is large, we have little structure. Similarily we define the *difference constant*

$$\delta[A] = \frac{|A-A|}{|A|}.$$

We have already calculated that that

$$1 \leqslant \sigma[A] \leqslant (|A|+1)/2$$

and it is also easy to see that

$$1 \leqslant \delta[A] \leqslant (|A|-1)/2 + \frac{1}{|A|}$$

which follows from the fact that $|A-A| \leqslant |A|^2 - |A| + 1$. A set which achieve this upper bound (for either the doubling or difference constant, since they are equivalent) is called a *Sidon set*.

Theorem 1.3. *Let A be an additive set. Then any subset of A can have doubling constant at most $\sqrt{\sigma(A)|A|/2}$, and any Sidon set in A has cardinality at most $\sqrt{2\sigma(A)|A|}$.*

Proof. If $|B| \geqslant c|A|$, then $|B+B| \leqslant |A|\sigma(A) \leqslant |B|/c\sigma(A)$ so $\sigma(B) \leqslant \sigma(A)/c$. $\sqrt{2} \cdot \sigma(A)^{1/2}|A|^{1/2}/|A|^{1/2}$ □

The doubling constant behaves poorly under taking unions, except if one of the sets is very small with respect to the other set.

Example. *Find sets A and B each of cardinality N, with $\sigma(A) \leqslant 2$, $\sigma(B) \leqslant 2$, but $\sigma(A \cup B) = N/2$. TODO*

On the other hand, if $\delta(A) = 1$ or equivalently, $\sigma(A) = 1$, then A is a single coset of G/H, where H is a finite subgroup of G. If $A + A = A$, then we actually have $A = H$. An important heuristic is that if $\sigma(A)$ is close to one, then A is 'close' to being a subgroup of G; we think of them as *approximate groups*. On the other hand, we have yet to obtain tools to analyze the structure of sets A with $\sigma(A)$ close to $(|A|+1)/2$, i.e. 'approximately Sidon sets'.

1.3 Ruzsa Distance

For two additive sets A and B we let

$$d(A,B) = \log\left(\frac{|A-B|}{|A|^{1/2}|B|^{1/2}}\right).$$

This is known as the *Ruzsa distance* between the two sets. Indeed, the distance is a non-negative, symmetric function and obeys the triangle inequality

$$d(A,C) \leqslant d(A,B) + d(B,C)$$

which is equivalent to the fact that

$$|A-C| \leqslant \frac{|A-B||B-C|}{|B|}$$

which follows from the identity that $a - c = (a - b) + (b - c)$ for any $b \in B$. The only problem is that we do not have $d(A,A) = 0$ for all sets A. In fact, we know that $d(A,A) = 0$ precisely when A is a coset with respect to some finite subgroup H of G.

The *additive energy* $E(A,B)$ between two sets A and B is the number of solutions to the equation $a_1 + b_1 = a_2 + b_2$. Then

$$|A||B| \leqslant E(A,B) \leqslant |A||B|\min(|A|,|B|).$$

The lower bound follows because there are always at least this many trivial solutions, and the upper bound follows because given any three values in the equation, the fourth is always uniquely determined.

Lemma 1.4. *Let A and B be additive sets. Then*

$$|A||B| = \sum_{x \in A+B} |A \cap (x - B)| = \sum_{y \in A-B} |A \cap (B + y)|$$

and

$$\begin{aligned}
E(A,B) &= \sum_{x \in A+B} |A \cap (x - B)|^2 \\
&= \sum_{y \in A-B} |A \cap (B + y)|^2 \\
&= \sum_{z \in (A-A) \cap (B-B)} |A \cap (z + A)||B \cap (z + B)|.
\end{aligned}$$

6

1.4 The *e*-transform

In general, the bound $\#(A + B) = \max(\#(A), \#(B))$ holds only if A is a coset and B is a union of cosets of a finite group H of the ambient group G. In particular, in a group containing no nontrivial finite subgroups, like $\mathbf{Z}$ or $\mathbf{Z}_p$, this can only occur A or B is a singleton. A natural question to ask is what the minimum size of $\#(A + B)$ is if A is *not* a singleton. A useful construction in such a setting is the *e-transform*.

Given two sets A and B, and given $e \in A - B$, we set

$$A_e = A \cup (B + e)$$

and

$$B_e = B \cap (A - e).$$

If $C = B \setminus (A - e)$, then A_e is the disjoint union of A and $C + e$. In particular, this means $\#(A_e) + \#(B_e) = \#(A) + \#(B)$; we are, in some sense, rearranging the elements of $A \cup B$. Moreover,

$$
\begin{aligned}
A_e + B_e &= (A \cup (B + e)) + (B \cap (A - e)) \\
&= (A + (B \cap (A - e))) \cup ((B + e) + (B \cap (A - e))) \\
&\subset (A + B) \cup ((B + e) + (A - e)) = A + e.
\end{aligned}
$$

Now $\#(A_e) \geqslant \#(A)$ and $\#(B_e) \leqslant \#(B)$, with either of these equality occuring only if $A \subset B + e$.

Lemma 1.5. *For any set E,*

$$
\begin{aligned}
|A_e \cap E| + |B_e \cap E| = |A \cap E| + |B \cap E| \\
+ |((B + e) \setminus A) \cap (E \setminus (E + e))| \\
- |((B + e) \setminus A) \cap ((E + e) \setminus E)|.
\end{aligned}
$$

In particular, if $e \in Sym(E)$, then $\#(A_e \cap E) + \#(B_e \cap E) = \#(A \cap E) + \#(B \cap E)$.

Proof. For any set $E \subset G$,

$$
\begin{aligned}
|A_e \cap E| + |B_e \cap E| &= |A_e \cap E| + |(B_e + e) \cap (E + e)| \\
&= |(A_e \cap E) \cup ((B_e + e) \cap (E + e))| \\
&\quad + |A_e \cap E \cap (B_e + e) \cap (E + e)| \\
&= |A \cap E| + |B \cap E| \\
&\quad + |((B + e) \setminus A) \cap (E \setminus (E + e))| \\
&\quad - |((B + e) \setminus A) \cap ((E + e) \setminus E)|. \qquad \square
\end{aligned}
$$

Let us use the e-transform to improve the lower bound for sum sets over the integers, using the fact that the integers are ordered.

Lemma 1.6. *If $A, B \subset \mathbf{Z}$, then*

$$\#(A+B) \geqslant \#(A) + \#(B) - 1.$$

Proof. Let $e = \max(A) - \min(B)$. Then

$$A_e = A \cup (B+e)$$

is the disjoint union of $A \smallsetminus \{\max(A)\}$, $\{\max(A)\}$, and $(B+e) \smallsetminus \{\max(A)\}$, and

$$B_e = B \cap (A-e) = \{\min(B)\}.$$

Thus

$$\begin{aligned}
\#(A+B) &\geqslant \#(A_e + B_e) \\
&= \#(A_e) \\
&= (\#(A) - 1) + (\#(B) - 1) + 1 \\
&= \#(A) + \#(B) - 1.
\end{aligned}$$

$\square$

Now we exploit the fact that $\mathbf{Z}_p$ has no proper subgroups.

Lemma 1.7 (Cauchy-Davenport Inequality)**.** *If p is prime, and A and B are additive sets in $\mathbf{Z}_p$, then*

$$\#(A+B) \geqslant \min(\#(A) + \#(B) - 1, p).$$

Proof. Assume $\#(B) \leqslant \#(A)$. We induct on $\#(B)$. If $\#(B) = 1$ the theorem is trivial. In general, if we can find $e \in A - B$ such that $\#(B_e) < \#(B)$, then induction implies

$$\#(A+B) \geqslant \#(A_e + B_e) \geqslant \min(\#(A_e) + \#(B_e) - 1, p) = \min(\#(A) + \#(B) - 1, p).$$

On the other hand, if $\#(B_e) = \#(B)$ for all $e \in A - B$, then $B + e \subset A$. This means that $(A+B) - B \subset A$. In particular, $\#(A+B) \leqslant \#(A)$, which actually implies that $\#(A+B) = \#(A)$, that A is a union of cosets of a subgroup of $\mathbf{Z}_p$, and that B is a subset of a coset of this subgroup. Since $\mathbf{Z}_p$ has no nontrivial subgroups, this implies either B is a singleton or $A = \mathbf{Z}_p$; in both of these cases the theorem is trivial. $\square$

More generally, we have the following inequality.

Theorem 1.8 (Kneser's Theorem)**.** *For any additive sets A and B,*

$$\#(A+B) \geqslant \#(A+Sym(A+B)) + \#(B+Sym(A+B)) - \#(Sym(A+B))$$
$$\geqslant \#(A) + \#(B) - Sym(A+B).$$

Proof. Assume $\#(A) \geqslant \#(B)$. We perform a *triple* induction, first on $\#(A+B)$, then $\#(A) + \#(B)$, then $\#(B)$. We induct upwards on $\#(A+B)$, downwards on $\#(A) + \#(B)$, and upwards on $\#(B)$. The downward induction is not contradictory because we have the bound $\#(A) + \#(B) \leqslant 2\#(A+B)$. The base cases where $\#(A+B) = 1$, or $\#(A) + \#(B) = 2 \cdot \#(A+B)$, or where $\#(B) = 1$ are obvious and so it suffices to prove the inductive step.

Let $H = Sym(A+B)$. If H is nontrivial, we let $\pi : G \to G/H$ be the standard projection. Then

$$\#(A+B) = \#(H)\#(\pi(A+B)) = \#(H)\#(\pi(A) + \pi(B)).$$

Since $\#(H) > 1$, $\#(\pi(A) + \pi(B)) < \#(A+B)$, we can apply induction to give the required equation. Thus we may assume in our argument, without loss of generality, that $Sym(A+B) = \{e\}$. Our goal is then to show that $\#(A+B) = \#(A) + \#(B) - 1$.

Suppose $B_e = B$ for all $e \in A - B$. Then, as we saw in the Cauchy-Davenport theorem, we have $(A+B) - B \subset A$ and so B is contained in a coset of some finite subgroup H of G, and A is a union of such cosets. But then H is contained in $Sym(A+B)$, which implies that H is trivial, and so B is a singleton, so the theorem is trivial.

Thus we are reduced to the case where there exists some $e \in A - B$ such that $\#(B_e) < \#(B)$. Pick e which *maximizes* $\#(B_e)$ subject to the constraint that $\#(B_e) < \#(B)$. Without loss of generality, translating B if necessary, we may assume that $e = 0$. Thus $A_e = A \cup B$ and $B_e = A \cap B$.

Now we know that $\#(A \cup B) + \#(A \cap B) = \#(A) + \#(B)$ and that $\#(A \cap B) < \#(B)$. Thus we may apply induction to conclude that if $X = (A \cup B) + (A \cap B)$, then

$$\#(X) \geqslant \#((A \cup B) + H) + \#((A \cap B) + H) - \#(H),$$

where $H = Sym(X)$.

Let $C = (A \cap B) + H$. Then

$$(A \cup C) + (B \cup C) = A + B$$

Thus we can replace A and B with $A \cup C$ and $B \cup C$ without modifying $A+B$. If C is not contained in $A \cap B$, then $\#(A \cup C) + \#(B \cup C) > \#(A) + \#(B)$, and so we can apply induction to conclude that

$$
\begin{aligned}
\#(A + B) &\geqslant \#((A \cup C) + \mathrm{Sym}(A + B)) \\
&\quad + \#((B \cup C) + \mathrm{Sym}(A + B)) \\
&\quad - \#(\mathrm{Sym}(A + B)) \\
&\geqslant \#(A + \mathrm{Sym}(A + B)) \\
&\quad + \#((B \cup C) + \mathrm{Sym}(A + B)) \\
&\quad - \#(\mathrm{Sym}(A + B)).
\end{aligned}
$$

Thus, without loss of generality, we may assume that $C = A \cap B$. But this means that $(A \cap B) + H \subset (A \cap B)$, which implies that $H \subset \mathrm{Sym}(A \cap B)$, and so $C = A \cap B$ is a union of cosets of G/H.

We know that X is a subset of $A+B$. If $X = A+B$, then $H = \mathrm{Sym}(A+B) = \{e\}$, and so by the inductive hypothesis

$$
\#(A + B) = \#(X) \geqslant \#(A \cup B) + \#(A \cap B) - 1 = \#(A) + \#(B) - 1.
$$

Thus we may assume that X is a proper subset of $A + B$.

Let A' be the set of $a \in A$ such that $a + B$ is not a subset of X. Then A' is nonempty since X is a proper subset of $A + B$. If $a + b \notin X$, then $a + b + H$ is disjoint from X. Since $b \in A \cup B$, and $X = (A \cap B) + (A \cup B)$ this implies that $a + H$ is disjoint from $A \cap B$. Moreover, $((a + H) \cap A) + b$ is disjoint from X, but contained in $A + B$. Thus

$$
\#(A + B) \geqslant \#(((a + H) \cap A) + b) + \#(X).
$$

Using the fact that $a + H$ is disjoint from $A \cap B$, we conclude that

$$
\begin{aligned}
\#((A \cup B) + H) &\geqslant \#(A \cup B) + \#((a + H) \setminus (A \cup B)) \\
&= \#(A \cup B) + \#(H) - \#((a + H) \cap A) - \#((a + H) \cap B)
\end{aligned}
$$

But this means that

$$
\begin{aligned}
\#(A + B) &\geqslant \#(((a + H) \cap A) + b) + \#(X) \\
&\geqslant \#(((a + H) \cap A) + b) + \#((A \cup B) + H) + \#((A \cap B) + H) - \#(H) \\
&= \#(((a + H) \cap A) + b) + \#((A \cup B) + H) + \#(A \cap B) - \#(H) \\
&\geqslant \#(A \cup B) + \#(A \cap B) - \#((a + H) \cap B) \\
&= \#(A) + \#(B) - \#((a + H) \cap B).
\end{aligned}
$$

Our proof would be complete if there exists $a \in A'$ such that $\#((a+H)\cap B) \leqslant 1$. Thus we may assume that $\#((a + H) \cap B) > 1$ for all $a \in A'$.

For each $a \in A'$, let $A_a = (a + H) \cap A$ and let $B_a = (a + H) \cap B$. Note that for each such a, $A_a - B_a \subset H$. Suppose we can find $a_1, a_2 \in A'$ such that $A_{a_1} - B_{a_1} + B_{a_2}$ is not a subset of A_{a_2}. Then we can find $e \in A_{a_1} - B_{a_1}$ such that $e + B_{a_2}$ is not a subset of A_{a_2}. This means that B is not contained in $A - e$, and so B_e is strictly smaller than B. Since $e \in H$, we find that $B \cap (A - e)$ contains $B \cap ((A \cap B) - e) = B \cap (A \cap B) = A \cap B$. But this contradicts the maximality of the e-transform we chose at the beginning of the argument, i.e. which gave $B_0 = A \cap B$. Thus we must always have $A_{a_1} - B_{a_1} + B_{a_2} \subset A_{a_2}$ for all $a, a' \in A'$. Symmetry thus implies that $\#(A_{a_1}) = \#(A_{a_2})$ for all $a_1, a_2 \in A'$, and this means that there exists a finite subgroup $K_{a_1 a_2}$ of G such that A_{a_1} and A_{a_2} are a finite union of cosets of $G/K_{a_1 a_2}$, and that B_{a_1} and B_{a_2}, being subsets of $B_{a_1} - B_{a_2}$ and $B_{a_2} - B_{a_1}$ respectively, are contained in a single coset of $G/K_{a_1 a_2}$. Taking $K = \bigcap K_{a_1 a_2}$ shows that B_a is contained in a single coset of G/K for each a and that each A_a is a union of cosets of G/K for each a. Since $\#(B_a) > 1$ for all $a \in A'$, this means that $\#(K) \geqslant 2$. Now

$$A + B = X \cup \bigcup_{a \in A'} (A_a + B)$$

Since $A_a = (a + H) \cap A$ is contained in a coset of H, this means a coset of G/K is contained in a coset of G/H, and thus $K \subset H$. Thus X, which is a union of cosets in G/H, is a union of coset in G/K. On the other hand, $A_a + B$ is also a union of cosets in G/K, since A_a is. But this means $A + B$ is a union of cosets in G/K, hence $K \subset \mathrm{Sym}(A + B)$, contradicting the fact that $\mathrm{Sym}(A + B)$ is trivial. $\qquad\square$

Recall that we proved that $\sigma(A) = \delta(A) = 1$ if and only if A is a coset of a finite group. In this corollary to Kneser's theorem we show that if $\sigma(A)$ or $\delta(A)$ is small, then A is a large subset of a coset of a finite group.

Corollary 1.9. *The following are equivalent:*

1. *$\sigma(A) < 3/2$.*

2. *$\delta(A) < 3/2$.*

3. *$\#(A + B) < (3/2)\#(A)$ for some additive set B with $\#(A) \leqslant \#(B)$.*

4. *$\#(nA - mA) < (3/2)\#(A)$ for all n and m.*

11

5. *A is a subset of a coset of G/H for some subgroup H with $\#(H) < (3/2)\#(A)$.*

Proof. It is obvious that (1),(2), and (4) individually imply (3). If we assume (5), then $A + A$ is contained in a single coset of H, and so

$$\sigma(A) = \frac{\#(A + A)}{\#(A)} \leqslant \frac{\#(H)}{\#(A)} < (3/2).$$

Thus (1) holds. Thus it remains to show that (3) implies (5).

Suppose (3) is true. Applying Kneser's theorem, setting $H = \mathrm{Sym}(A + B)$, we conclude that

$$(3/2)\#(A) > \#(A) + \#(B) - \#(H) \geqslant 2\#(A) - \#(H).$$

Rearranging, we find that $\#(H) > \#(A)/2$. But since $\#(H)$ divides $\#(A + B) < 3\#(H)$, this implies that either $\#(A + B) = 2\#(H)$ or $\#(A + B) = \#(H)$. Suppose first that $A + B$ is a union of two cosets in H. Then $\#(H) = (1/2)\#(A + B) < (3/4)\#(A) \leqslant (3/4)\#(B)$, which implies that neither A nor B is contained in a single coset of H. Thus Kneser's theorem implies that

$$2\#(H) = \#(A + B) \geqslant \#(A + H) + \#(B + H) - \#(H) \geqslant 3\#(H),$$

which gives a clear contradiction. On the other hand, if $\#(A + B) = \#(H)$, then $A + B$ is a single coset of G/H, so A is contained in a single coset of G/H, which proves that (5) holds. $\square$

Corollary 1.10. *If $\#(A + B) \leqslant \#(A) + \#(B) - 1$, and $H = Sym(A + B)$, then*

$$\#(A + B) = \#(A + H) + \#(B + H) - \#(H).$$

Proof. By Kneser's theorem, it suffices to prove

$$\#(A + B) \leqslant \#(A + H) + \#(B + H) - \#(H).$$

Now $\#(H)$ divides $\#(A + B)$, $\#(A + H)$, and $\#(B + H)$. We have the trivial inequality

$$\#(A + B) \leqslant \#(A) + \#(B) - 1 \leqslant \#(A + H) + \#(B + H) - 1.$$

But since everything here is divisible by $\#(H)$, this inequality immediately improves to give the result that

$$\#(A + B) \leqslant \#(A + H) + \#(B + H) - \#(H). \qquad \square$$

Theorem 1.11. *Let A and B be additive sets. Prove that*

$$\#(A+B) \geqslant \#(A) + \#(B) - \min_{c \in A+B} \#\{(a,b) \in A \times B : a + b = c\}.$$

Proof. It suffices to prove that for each $c \in A + B$,

$$\#\{(a,b) \in A \times B : a + b = c\} \geqslant \#(A) + \#(B) - \#(A+B).$$

If $\#(A+B) \geqslant \#(A) + \#(B)$, this inequality is trivial. Thus we may assume $\#(A+B) \leqslant \#(A) + \#(B) - 1$, in which case it follows that

$$\#(A+B) = \#(A+H) + \#(B+H) - \#(H),$$

where $H = \mathrm{Sym}(A+B)$. We therefore have to prove that

$$\#\{(a,b) \in A \times B : a + b = c\} \geqslant \#(H) - \#((A+H) \smallsetminus A) - \#((B+H) \smallsetminus B).$$

Fix $(a_0, b_0) \in A \times B$ with $a_0 + b_0 = c$. Then for each $h \in H$,

$$(a_0 + h) + (b_0 - h) = c.$$

If $(a_0 + h, b_0 - h) \notin A \times B$, either $a_0 + h \in (A+H) \smallsetminus A$ or $b_0 - h \in (B+H) \smallsetminus B$. Thus

$$\begin{aligned}
\#\{(a,b) &\in A \times B : a + b = c\} \\
&\geqslant \#\{h \in H : (a_0 + h, b_0 - h) \in A \times B\} \\
&\geqslant \#(H) - \#\{h \in H : a_0 + h \in \#((A+H) \smallsetminus A)\} \\
&\qquad - \#\{h \in H : b_0 - h \in \#((B+H) \smallsetminus B)\} \\
&\geqslant \#(H) - \#((A+H) \smallsetminus A) - \#((B+H) \smallsetminus B). \qquad \square
\end{aligned}$$

Theorem 1.12 (Mann's Theorem). *Let $N \geqslant 1$, let $0 < \alpha < 1$, and let A and B be additive subsets of $\mathbf{Z}$, with $0 \in A \cap B$, and for all $1 \leqslant n \leqslant N$,*

$$\#(A \cap [1,n]) + \#(B \cap [1,n]) \geqslant \alpha n.$$

Then $\#((A+B) \cap [1,n]) \geqslant \alpha n$ for all $1 \leqslant n \leqslant N$.

Proof. We prove by induction, first on N and secondly on $\#(B)$. The case where $N = 1$ is trivial. By the inductive hypothesis, for $1 \leqslant n < N$,

$$\#((A+B) \cap [1,n]) \geqslant \alpha n,$$

13

and it suffices to show that $\#((A + B) \cap [1, N]) \geqslant \alpha n$. Without loss of generality we may assume that $A, B \subset [0, N]$ since adding other elements to A and B only helps us. It will suffice to find $e \in A \subset A - B$ such that $\#(B_e) < \#(B)$ and such that for all $1 \leqslant n \leqslant n$,

$$\#(A_e \cap [1, n]) + \#(B_e \cap [1, n]) \geqslant \alpha n,$$

for then we may apply the inductive hypothesis together with the fact that $A_e + B_e \subset A + B$. The fact that $e \in A$ implies immediately that $0 \in A_e \cap B_e$. If B was not strictly contained in A, then we could pick $e = 0$, since then $\#(B_e) < \#(B)$, and

$$\#(A_e \cap [1, n]) + \#(B_e \cap [1, n]) = \#(A \cap [1, n]) + \#(B \cap [1, n]) \geqslant \alpha n.$$

Thus we may assume $B \subset A$. Let e be the minimum element of A such that $e + B$ is not a subset of A. The largest element of A certainly belongs to this set so e is well defined. By construction, $(A \cap [0, e)) + B$ is contained in A. Then $\#(B_e) < \#(B)$. We calculate that

$$\#(A_e \cap [1, n]) + \#(B_e \cap [1, n]) = \#(A \cap [1, n]) + \#(B \cap [1, n])$$
$$+ \#(((B + e) \smallsetminus A) \cap [1, e]))$$

TODO $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

1.5 Freiman Homomorphisms

If A and B are additive sets in two groups G and H, a *Freiman homomorphism* of rank k is a map $f : A \to B$ such that for any $a_1, \ldots, a_k, a'_1, \ldots, a'_k \in A$, if $a_1 + \cdots + a_k = a'_1 + \cdots + a'_k$, then $\phi(a_1) + \cdots + \phi(a_k) = \phi(a'_1) + \cdots + \phi(a'_k)$. If ϕ is a Freiman homomorphism of rank k, it is also a Freiman homomorphism of rank k' for any $k' \leqslant k$. Many problems in additive combinatorics are invariant under Freiman isomorphisms of a suitably large rank.

Remark. One could have defined a 'rank k' additive set A intrinsically by considering an equivalence relation on A^k, i.e. where we imagine that $(a_1, \ldots, a_k) \sim (a'_1, \ldots, a'_k)$ if $a_1 + \cdots + a_k = a'_1 + \cdots + a'_k$. One can develop the theory of additive sets intrinsically, but there does not seem to be much advantage in working intrinsically vs extrinsically.

Obviously, all homomorphisms of groups are Freiman homomorphisms of arbitrarily large rank. Similarily, 'affine maps', i.e. homomorphisms composed by translations, are also Freiman homomorphisms. Let us consider some less trivial examples which cannot be induced by homomorphisms of ambient groups.

Example. *The quotient map $\phi : \mathbf{Z} \to \mathbf{Z}_n$, which is a Freiman homomorphism of any order. On the other hand, it is not a Freiman isomorphism, though it restricts to a rank k isomorphism from $[1,m]$ onto its image if $n > k(m-1)$.*

Example. *Fix $a,r \in G$, and let $P = \{a, a+r, \ldots, a+(n-1)r\}$ be the corresponding length n arithmetic progression in G. We have a natural Freiman homomorphism $\phi : [0,n) \to P$ which has arbitrarily large rank, because if $m_1 + \cdots + m_k = m_1' + \cdots + m_k'$, then*

$$\begin{aligned}
\phi(m_1) + \cdots + \phi(m_k) &= ka + (m_1 + \cdots + m_k)r \\
&= ka + (m_1' + \cdots + m_k')r \\
&= \phi(m_1') + \cdots + \phi(m_k').
\end{aligned}$$

On the other hand, this is a Freiman isomorphism of rank k if and only if $\mathrm{ord}(r) > k(n-1)$.

Example. *The sets $\{0,1,10,11\}$ and $\{0,1,100,101\}$ are Freiman isomorphic of order k for any $k < 10$, but not for any $k \geqslant 10$.*

Lemma 1.13. *Let $f : A \to B$ be a Freiman homomorphism of rank k. Then if $|\varepsilon_1| + \cdots + |\varepsilon_n| \leqslant k$, and $A_1, \ldots, A_n \subset A$, then*

$$\#(\varepsilon_1 \phi(A_1) + \cdots + \varepsilon_n \phi(A_n)) \leqslant \#(\varepsilon_1 A_1 + \cdots + \varepsilon_n A_n).$$

This inequality becomes an equality if f is an isomorphism.

Proof. Without loss of generality assume that $\varepsilon_i \in \{\pm 1\}$ for each i, and that $n = k$. Consider the two maps

$$\psi_1 : A_1 \times \cdots \times A_k \to B$$

and

$$\psi_2 : A_1 \times \cdots \times A_k \to A$$

where

$$\psi_1(a_1, \ldots, a_k) = \varepsilon_1 \phi(a_1) + \cdots + \varepsilon_k \phi(a_k),$$

15

and
$$\psi_2(a_1, \ldots, a_k) = \varepsilon_1 a_1 + \cdots + \varepsilon_k a_k.$$

Then $\psi_1(A_1 \times \cdots \times A_k) = \varepsilon_1 \phi(A_1) + \cdots + \varepsilon_n \phi(A_n)$ and $\psi_2(A_1 \times \cdots \times A_k) = \varepsilon_1 A_1 + \cdots + \varepsilon_n A_n$. If $\psi_2(a_1, \ldots, a_k) = \psi_2(a'_1, \ldots, a'_k)$, then

$$\sum_{\varepsilon_i = 1} a_i + \sum_{\varepsilon_i = -1} a'_i = \sum_{\varepsilon_i = 1} a'_i + \sum_{\varepsilon_i = -1} a_i$$

and so

$$\sum_{\varepsilon_i = 1} \phi(a_i) + \sum_{\varepsilon_i = -1} \phi(a'_i) = \sum_{\varepsilon_i = 1} \phi(a'_i) + \sum_{\varepsilon_i = -1} \phi(a_i)$$

which implies $\psi_1(a_1, \ldots, a_k) = \psi_2(a'_1, \ldots, a'_k)$. Thus the theorem follows. $\qquad \square$

If one wants to understand the sum sets of a *fixed* family of sets $A_1, \ldots, A_n$, one needs less structure than a Freiman homomorphism from $A_1 \cup \cdots \cup A_n$ to B. Indeed, individual translations of each set preserve all sum sets of the form $\#(k_1 A_1 + \cdots + k_n A_n)$ and so we would hope these are included in the family of transformations we consider. One trick to doing this is to consider the disjoint union

$$A = A_1 \times \{e_1\} \cup \cdots \cup A_n \times \{e_n\}$$

in $G \times \mathbf{Z}^n$. If $\phi : A \to B$ is a rank k Freiman homomorphism, if $|k_1| + \cdots + |k_n| = k$, and if we define $\phi_i(a) = \phi(a, e_i)$ for $i \in [1, n]$, then

$$\#(k_1 \phi_1(A_1) + \cdots + k_n \phi_n(A_n)) \leqslant \#(k_1 A_1 + \cdots + k_n A_n).$$

We note that the map $\phi : A \to G$ given by setting $\phi(a, e_i) = (a + s_i, e_i)$ is a Freiman isomorphism of arbitrary order.

Theorem 1.14. *Suppose $\phi : A \to B$ is a Freiman homomorphism of order $k \geqslant 2$. If P is a rank k arithmetic progression in A TODO*

An interesting corollary to these definitions is that for any additive set A in a torsion free group, and any $k > 0$, there is an order k Freiman isomorphism from A to a subset of $\mathbf{Z}$, or even to a subset of $\mathbf{Z}_n$ for some large n. We also have a 'compression lemma' in this setting.

Theorem 1.15. *Let A be an additive set in a group G which is torsion, or cyclic of prime order. Fix integers $n, m \geqslant 1$ such that*

$$2n\#(nA - nA) < m < \#(Z).$$

Then there is $A' \subset A$ with $\#(A') \geqslant (1/n)\#(A)$ and an order n Freiman isomorphism from A' to a subset of $\mathbf{Z}_m$.

Proof. Without loss of generality, in light of the preceding paragraph, we may assume A is a subset of $\mathbf{Z}_p$ for some prime p. Consider the three maps $\pi_p : \mathbf{Z} \to \mathbf{Z}_p$, $\pi_m : \mathbf{Z} \to \mathbf{Z}_m$, and $i : \mathbf{Z}_p \to \{0, \dots, p-1\}$, defined in the obvious manner, and set $\pi = \pi_m \circ i$. The function π is not a homomorphism of groups, but it is a Freiman homomorphism of order n when restricted to any set of the form $\pi_p(X)$, where X are a set of fewer than p/n contiguous elements of $\mathbf{Z}$; if $a_1, \dots, a_n$ and $a'_1, \dots, a'_n$ are elements of X, then

$$-p < (a_1 + \cdots + a_n) - (a'_1 + \cdots + a'_n) < p.$$

Thus if $\pi_p(a_1) + \cdots + \pi_p(a_n) = \pi_p(a'_1) + \cdots + \pi_p(a'_n)$, then $a_1 + \cdots + a_n = a'_1 + \cdots + a'_n$, and so $\pi_m(a_1) + \cdots + \pi_m(a_n) = \pi_m(a'_1) + \cdots + \pi_m(a'_n)$.

Consider the disjoint family of sets

$$\{Z_k = \pi_p((k(p/n), (k+1)p/n]) : 0 \leqslant k \leqslant n - 1\}.$$

whose union is $\mathbf{Z}_p$. Applying the pidgeonhole principle, for each λ, there exists k_λ such that $\#(Z_{k_\lambda} \cap (\lambda \cdot A)) \geqslant \#(A)/n$. If we set $A_\lambda = Z_{k_\lambda} \cap (\lambda \cdot A)$, then $\pi : A_\lambda \to \mathbf{Z}_n$ is an injective Freiman homomorphism of rank n. All that remains is to show that for some choice of $\lambda \in \mathbf{Z}_p^\times$, π is a Freiman isomorphism of rank n when restricted to the domain A_λ.

It suffices to show that there is some $\lambda \in \mathbf{Z}_p^\times$ such that for any family of elements $a_1, \dots, a_n$ and $a'_1, \dots, a'_n$ in $i(A_\lambda)$ with

$$\pi_m(a_1) + \cdots + \pi_m(a_n) = \pi_m(a'_1) + \cdots + \pi_m(a'_n),$$

then $a_1 + \cdots + a_n = a'_1 + \cdots + a'_n$. We call such an occurence a *collision*. This can only happen if $ni(A_\lambda) - ni(A_\lambda)$ contains a nonzero multiple of m. Since $-n(p-1) \leqslant ni(A_\lambda) - ni(A_\lambda) \leqslant n(p-1)$, if $ni(A_\lambda) - ni(A_\lambda)$ contains

km, then $|k| \leqslant n(p-1)/m$. Now for each such integer k,

$$
\begin{aligned}
\mathbf{P}\left(km \in ni(\lambda \cdot A) - ni(\lambda \cdot A)\right) &= \mathbf{P}\left(km + p\mathbf{Z} \in \lambda \cdot (nA - nA)\right) \\
&= \sum_{x \in nA - nA} \mathbf{P}(\lambda x = km + p\mathbf{Z}) \\
&= \sum_{x \in nA - nA} \mathbf{P}(\lambda^{-1} = x/(km + p\mathbf{Z})) \\
&= \sum_{x \in (nA - nA) \smallsetminus \{0\}} \frac{1}{p-1} \\
&\leqslant \frac{\#(nA - nA)}{p-1}.
\end{aligned}
$$

A union bound thus implies that a collision happens with probability at most $(2n/m)\#(nA - nA)$. Since $m > 2n\#(nA - nA)$ this probability is less than one, so there is some $\lambda \in \mathbf{Z}_p^\times$ for which no collision occurs. $\qquad\square$

18

Chapter 2

Miscellania

2.1 Sum Free Sets

Given a subset A of an abelian group, we say A is **sum free** if $A + A$ is disjoint from A.

Theorem 2.1. *If A is an arbitrary finite subset of positive natural numbers, then A contains a sum-free subset of size greater than $|A|/3$.*

Proof. The idea of this proof rests on two observations. If $B \subset [1, N]$, and $p > 2N$, then $B + p\mathbf{Z}$ is sumfree in $\mathbf{Z}_p$ if and only if B is sumfree. Thus we can turn out problem into a problem modulo p. Next, we notice that if f is an automorphism, then a subset B of an abelian group is sumfree if and only if $f(B)$ is sumfree. The presense of many automorphisms of $\mathbf{Z}_p$ (one for each natural number between 1 and $p - 1$) enables us to exploit randomness to construct a sumfree subset in A. If $X \subset \mathbf{Z}_p$ is sumfree, and does *not* contain zero, we consider the sets $X, 2X, \ldots, (p - 1)X$, which are all sumfree. For every $a \in X$, and nonzero $b \in \mathbf{Z}_p$, there is a unique $c \in [1, p)$ such that $ca = b$. Thus every nonzero $b \in \mathbf{Z}_p$ occurs in $|X|/(p - 1)$ of the sets $X, \ldots, (p - 1)X$. Thus means if we choose a nonzero $x \in \mathbf{Z}_p$ uniformly at random, then

$$\mathbf{E}\big|(A + \mathbf{Z}_p) \cap xX\big| = \sum_{a \in A + \mathbf{Z}_p} \mathbf{P}(a \in xX) = \frac{|A||X|}{p - 1}$$

Since xX is sumfree, so too is $(A + \mathbf{Z}_p) \cap xX$, and so lower bounding the expectation gives rise to a large sumfree sert. In $\mathbf{Z}_p$, a good candidate for

19

a sumfree set should be an interval, since an arithmetic progression has a small sumset, and all arithmetic progressions are mapped to an interval by an automorphism. Thus, taking $X = [k, 2k)$, where $4k - 2 < p + k$, we get a squarefree set. Thus taking p congruent to two modulo 3, and setting $3k = p + 1$, we find a sumfree set of size

$$\frac{k}{p-1}|A| = \frac{p+1}{3(p-1)}|A| > |A|/3$$

which completes the proof. $\qquad\square$

A fundamental problem in additive combinatorics is the *inverse sumset* problem. If $A + B$ or $A - B$ is small, what can one say about A and B? More specifically, if $A + A$ is small, what can one say about A? We have $|A| \leqslant |A + A| \leqslant [|A|^2 + |A|]/2$, and so we refer to the value $\sigma(A) = |A + A|/|A|$ as the **doubling constant** of the set A. We have $1 \leqslant |A| \leqslant (|A| + 1)/2$.

Example. *Geometric progressions have the largest doubling constant possible. If*

$$A = \{1, a, a^2, \ldots, a^{N-1}\}$$

then the sum of any two elements of A is distinct, so $|A + A| = (N^2 + N)/2$, and so $\sigma(A) = (N + 1)/2$.

A set A with $\sigma(A)$ maximal among sets of size N is known as a **Sidon set**. This means that all pairwise sums of any two $a_0, a_1 \in A$ are distinct, modulo the trivial equalities $a_0 + a_1 = a_1 + a_0$. This is a 'generic' behaviour: If A is a subset of N points chosen uniformly at random frmo $[0, 1]$, then A is Sidon with probability one. It is more interesting to characterize when $\sigma(A)$ is small.

Example. *In the other extreme, the main example of sets with small doubling constant is an arithmetic progression. If $A = b_0 + [0, N - 1]a$, then $A + A = 2b_0 + [0, 2N - 2]a$, which consists of $2N - 1$ points, so $\sigma(A) = 2 - 1/N$.*

Example. *If $A \subset B$, and $|A| = \alpha|B|$, then $|A + A| \leqslant |B + B|$, so*

$$\sigma(A) \leqslant \frac{|B + B|}{K|B|} = \sigma(B)/\alpha$$

Thus if $\sigma(B)$ is small, and A contains a large percentage of B, then $\sigma(A)$ is also small. In the other direction, if $|B| = \beta|A|$, then

$$|B+B| \leqslant |A+A|+|A+(B-A)|+|(B-A)+(B-A)| \leqslant \sigma(A)|A|+(\beta-1)|A|^2+\beta^2|A|^2$$

so
$$\sigma(B) \leqslant \sigma(A)/\beta + (\beta + 1 - 1/\beta)|B|$$

Thus if $\sigma(A)$ is small, and B doesn't contain many more points than A, then $\sigma(B)$ is also small.

Example. *If we consider N and M, and a resultant 'rank 2' arithmetic progression $A = c + [0,N]a + [0,M]b$, then $\sigma(A) \leqslant 4$. These sets can look very different from the original arithmetic progressions we were considering.*

If A and B are additive sets, and we form the graph G

2.2   Graph Theoretic Techniques

Theorem 2.2 (Turán). *Let G be a graph of n vertices. Then G contains an independant set of size at least*

$$\sum_{v \in G} \frac{1}{\deg(v) + 1}$$

In particular, if the vertices have degree bounded by d, then there is an independant set of size $|G|(d+1)^{-1}$.

Proof. Let $\pi : V \to [1,n]$ be a uniformly randomly chosen bijection. Let S be the set of all vertices v in V such that for any neighbour w of v, $\pi(v)$ is larger than $\pi(w)$. Then S is an independant set, and it suffices to show S is large in expectation. We find by the hockey stick identity that

$$\begin{aligned}
\mathbf{P}(v \in S) &= \frac{1}{n!} \sum_{m=1}^{n} \binom{m-1}{\deg(v)} \deg(v)!(n-1-\deg(v))! \\
&= \frac{\deg(v)!(n-1-\deg(v))!}{n!} \binom{n}{\deg(v)+1} \\
&= \frac{1}{\deg(v)+1}
\end{aligned}$$

and so

$$\mathbf{E}|S| = \sum_{v \in G} \mathbf{P}(v \in S) = \sum_{v \in G} \frac{1}{\deg(v)+1}$$

and this gives the required set. □

Given $B \subset A$, we say B is sumfree with respect to A if no element of A is the sum of two distinct elements of B. Given A, we let $\phi(A)$ denote the largest sumfree subset with respect to A. We let $\phi(n)$ be the smallest value of $\phi(A)$ among all sets $A \subset \mathbf{R}$ of size n.

Theorem 2.3 (Choi). *If A is any set of n real numbers, there is a set $B \subset A$ of cardinality $\log n - O(1)$ sumfree with respect to A. Thus $\phi(n) \geqslant \log n - O(1)$.*

Proof. Assume first that A is a subset of positive reals. Order $A = \{a_1 > a_2 > \cdots > a_n > 0\}$. Consider the graph G with vertices A, and edges (a_n, a_m) if $a_n + a_m \in A$. By Turán's theorem, since $\deg(a_i) \leqslant n - i$, we find an independant set S with

$$|S| \geqslant \sum_{i=1}^{n} \frac{1}{n-i+1} = \sum_{i=1}^{n} \frac{1}{i} = \log n - O(1)$$

In general, any set A of n real numbers either contains $n/2 - O(1)$ positive real numbers or $n/2 - O(1)$ negative real numbers, and the theorem then follows in this case. $\square$

The $n/(d+1)$ bound for graphs of bounded degree d cannot be improved for general graphs G. However, it is surprising that one can improve the bound by a $\log d$ factor, provided that the resultant graph has no three cycles.

Theorem 2.4. *If G has no three cycles with maximal degree d, then G contains an independant set of size $\Omega(n \log d / d)$.*

Proof. Choose a set I uniformly from the set of all independant sets in G. For each $v \in V$, define the random variable

$$X_v = d|I \cap \{v\}| + |N(v) \cap I| = \begin{cases} d & v \in I \\ |N(v) \cap I| & v \notin I \end{cases}$$

Any vertex can be in the neighbourhood of at most d other vertices, so

$$\sum_v X_v = d|I| + \sum_{v \notin I} |N(v) \cap I| \leqslant 2d|I|$$

Taking expectations gives that

$$\mathbf{E}|I| \geqslant \frac{1}{2d} \sum_v \mathbf{E}(X_v)$$

22

Thus it suffices to show that $\mathbf{E}(X_v)$ is large for each v. TODO: FINISH LATER. $\qquad\square$

The Balog-Szemerédi theorem says that if $E(A,B) \geqslant K_0 n^2$ and $|A +_G B| \leqslant K_1 n$, then one can find $A_0 \subset A$ and $B_0 \subset B$ such that $|A_0|, |B_0|$, and $|A_0 + B_0|$ are $\Theta_{K_0,K_1}(n)$. Gower's recently strengthened the theorem to showing the constants in the bound are polynomial in $1/K_0$ and K_1. We shall find that this result can be converted into a graph problem.

If $E(A,B) \gtrsim |A|^{3/2}|B|^{3/2}$, then there is $A_0 \subset A$ and $B_0 \subset B$ with $|A_0| \sim |A|$, $|B_0| \sim |B|$, and $|A_0 + B_0| \lesssim |A_0|^{1/2}|B_0|^{1/2}$. In particular, if A and B have n elements, and $E(A,B) \gtrsim n^3$, then there is $A_0 \subset A$ and $B_0 \subset B$ with $|A_0|, |B_0| \sim n$, and $|A_0 + B_0| \lesssim n$. Can we generalize this theorem to more general operations than addition, i.e. linear transformations of the coordinates?

Lemma 2.5. *If G is a bipartite graph with $|E| \geqslant |A||B|/K$ for some $K \geqslant 1$, then for any $0 < \varepsilon < 1$, there is $A_0 \subset A$ such that $|A_0| \geqslant |A|/K\sqrt{2}$, and such that $1 - \varepsilon$ of the pairs of vertices in A_0 are connected by $\varepsilon|B|/2K^2$ paths of length 2 in G.*

Proof. By decreasing K, we may assume that $|E| = |A||B|/K$. Now

$$\frac{\mathbf{E}_b|N(b)|}{|A|} = \frac{\mathbf{E}_a|N(a)|}{|B|} = \frac{|E|}{|A||B|} = \frac{1}{K}$$

and

$$\frac{\mathbf{E}_b|N(b)|^2}{|A|^2} = \mathbf{E}_{a,a'}\frac{|N(a) \cap N(a')|}{|B|}$$

$\qquad\square$

Let $A_1,\ldots,A_k$ be additive sets with cardinality n, and consider a k uniform k-partite hypergraph H on $A_1,\ldots,A_k$. If H has $\Omega(n^k)$ edges and $|\bigoplus^H A_i| = O(n)$, then we can find $A_i' \subset A_i$ with $|A_i'| = \Omega(n)$ and $|A_1' + \cdots + A_k'| = \Omega(n)$. If we let H be

23