Introduction to Combinatorics Lecture 14

Sum set in Additive Combinatorics

A classical problem dealing with the sum of primes is the well-known Goldbach Conjecture (1742). It was conjectured that for each even positive integer n, n is the sum of two primes. This is a revised version of the original conjecture that all integers larger than 5 can be written as a sum of three primes. Clearly, these two primes are odd integers. In the sense of sum set, we let $\mathcal{P}$ be the set of all primes in $\mathbb{N}$. The goal is to verify that $\mathcal{P}+\mathcal{P} \supseteq 2\mathbb{N}$ where $2\mathbb{N}$ denotes the set of even positive integers larger than 3. So far, this seeming easy conjecture remains unsettled though there are people who claim that they have verified the trueness of the Goldbach conjecture.

It is worth of noting that for (sufficiently) large odd integer n, n can be written as the sum of three primes. This result was obtained not too long ago. (2013, Harald Helfgott) That is, the weak version of Goldbach's conjecture has been proved. Another idea comes from using product sets. It was proved by 陳景潤 that every integer can be written as the sum "$p_1 + p_2 \cdot p_3$" where p_1, p_2 and p_3 are odd primes. (This is known as the 1+2 version of Goldbach conjecture.)

The most fundamental problem in additive combinatorics is to find the size of sum set $A+B := \{a+b | a \in A \text{ and } b \in B\}$ where A and B are two given sets in an abelian group with operation " + ". For example, $\mathbb{R}, \mathbb{Z}$ or $\mathbb{Z}_n$.

Definition 14.1 (Sum set, Difference set)**.**

- $A + B = \{a + b | a \in A \text{ and } b \in B\}$ is called the sum set of A and B. Similarly, $A - B = \{a - b | a \in A \text{ and } b \in B\}$ is called the difference set of A and B.
- The group considered for sum sets or difference sets is called an ambient group.

Remark.

- We shall consider the sum set in what follows, the idea on difference set $A - B$ can be dealt similarly.
- $A + A = 2A$ and $A + A + \cdots + A$ (n copies of A) $= nA$.

Facts.

1. Let Z be an ambient group (may be finite!). $\forall x \in Z$ and $A \subseteq Z$, $|A + x| = |A|$.
2. $\max\{|A|, |B|\} \leq |A + B| \leq |A| \cdot |B|$.
3. Let $A, B \subseteq \mathbb{Z}$ (the set of integers). Then, $|A + B| \geq |A| + |B| - 1$.

   Proof. Order the elements of A and B respectively as follows: $A = \{a_1, a_2, ..., a_n\}$, $a_1 < a_2 < \cdots < a_n$, $B = \{b_1, b_2, ..., b_m\}$, and $b_1 < b_2 < \cdots < b_n$. Then, in $A + B$, $a_1 + b_1, a_2 + b_1, ..., a_n + b_1, a_n + b_2, ..., a_n + b_m$ are $n + m - 1$ distinct integers in $\mathbb{Z}$. Hence, $|A + B| \geq |A| + |B| - 1$. □

   Note that if A, B are subsets of a finite set (finite additive group), then the size of $A + B$ may be smaller, but still we have $|A + B| \geq \max\{|A|, |B|\}$. (Fact 2)

4. $|A| \leq |A + A| \leq |A| \cdot (|A| + 1)/2 = \binom{|A| + 1}{2}$.
5. $|nA| \leq \binom{|A| + n - 1}{n}$.

   Proof. By induction on n. Clearly, it is true when $n = 1$. (The equality holds when the sum of any two elements are different.) On the other hand, if $|A| = 1$, the equality holds. (Both are "1".) So, consider $|A| > 1$ and $n \geq 2$, moreover the assertion is true for $n - 1$.

   Let $A = B \cup \{x\}$. $|B| = |A| - 1$. Then, $nA = \cup_{j=0}^{n}(jB + (n - j)x)$. ($j$ terms in B and $n - j$ x's) Hence,

$$|nA| = |\cup_{j=0}^{n} jB| \leq \sum_{j=0}^{n} |jB| \leq \sum_{j=0}^{n} \binom{|B| + j - 1}{j}$$
$$= \sum_{j=0}^{n} \binom{|A| + j - 2}{j}$$

$$
\begin{aligned}
&= \binom{|A|-2}{2} + \binom{|A|-1}{1} + \binom{|A|}{2} + \binom{|A|+1}{3} + \cdots + \binom{|A|+n-2}{n} \\
&= \binom{|A|}{1} + \binom{|A|}{2} + \binom{|A|+1}{3} + \cdots + \binom{|A|+n-2}{n} \\
&= \binom{|A|+1}{2} + \binom{|A|+1}{3} + \cdots + \binom{|A|+n-2}{n} \\
&= \binom{|A|+2}{3} + \cdots + \binom{|A|+n-2}{n} \\
&= \binom{|A|+n-1}{n}.
\end{aligned}
$$

□

Remark.

- If $A, B \subseteq \mathbb{Z}$, then $|A+B| \geq |A| + |B| - 1$, furthermore if $|A+B| = |A| + |B| - 1$, then either
  1. $|A| = 1,\ |B| = 1$ or
  2. both A, B are arithmetic progression integers with a common difference.
- In a finite group, the above conclusion may be wrong. (In $\mathbb{Z}_n$, $|A| + |B| - 1$ may be larger than n, but $A, B \subseteq \mathbb{Z}_n$ and thus $|A+B| \leq n$.)

Theorem 14.1 (Cauchy-Davenport, 1935)**.** *If p is a prime, $A, B \subseteq \mathbb{Z}_p$ are non-empty, then*

$$|A+B| \geq \min\{p, |A| + |B| - 1\}.$$

Proof. By induction on $|A| \geq 2$. ($|A| = 1$ is trivially true.) Moreover, assume that $|B| \neq p$. Note that $g + A$ is called a shift of A and $g + A = \{g\} + A$. "Shift" does not change the size of a set. By shifting the elements of A to obtain $\{0, g\} \subseteq A$, for some $g \neq 0 \in \mathbb{Z}_p$. Note that $\langle g \rangle \cong \mathbb{Z}_p$. Now, since $|B| \neq q$, we may shift the elements of B to obtain that $0 \in B$ and $g \notin B$. Hence, $A \cap B \neq \emptyset$, and $A \cap B \neq A$.

Let $x \in A \cap B,\ y \in A \cup B$. If $y \in A \setminus B$, then let $x \in B$. If $y \in B$, then let $x \in A$. Thus, $A \cap B + (A \cup B) \subseteq A + B$.

$$\begin{aligned} |A+B| &\geq |(A\cap B)+(A\cup B)| && (|A\cap B|<|A| \text{ by induction on } |A|)\\ &\geq \min\{p, |A\cap B|+|A\cup B|-1\}\\ &= \min\{p, |A|+|B|-1\}. \end{aligned}$$

□

Question. Is this theorem also true for $\mathbb{Z}_n$ where n is not a prime?

Problem. Given a sequence of $2n-1$ integers. Can we find an n-term sub-sequence such that the sum of these n terms is a multiple of n?
We can use Cauchy-Davenport's theorem to answer the problem when n is a prime. Clearly, the problem is equivalent to the following.

Problem. Given a set S of $2n-1$ elements in $\mathbb{Z}_n$. Can we find an n-subset of S such that its sum is 0 in $\mathbb{Z}_n$?
This is also known as a zero-sum problem.

Proposition 14.2. *If p is a prime and $a_1, a_2, ..., a_{2p-1} \in \mathbb{Z}_p$, then there is a subsequence with p terms such that the sum is 0.*

Proof. Based on $0, 1, 2, ..., p-1$, order the terms as follows: $a_1 \leq a_2 \leq \cdots \leq a_{2p-1}$. Now, consider $\{a_1, a_p\},\ \{a_2, a_{p+1}\}, \cdots, \{a_{p-1}, a_{2p-2}\},\ \{a_{2p-1}\}$.

1. If $a_i \neq a_{p+i}$ for each $i = 1, 2, ..., p-1$, then by Theorem 14.1, we conclude that the number of p-term sum in $\{a_1, a_p\} + \{a_2, a_{p+1}\} + \cdots + \{a_{p-1}, a_{2p-2}\} + \{a_{2p-1}\}$ has size p. That is, all elements in $\mathbb{Z}_p$ occur in the sum set. 0 is one of them.

2. If $\exists j,\ a_j = a_{j+p-1}$, then $a_j = a_{j+1} = \cdots = a_{j+p-1}$. Hence, $p \cdot a_j = a_j + a_{j+1} + \cdots + a_{j+p-1} = 0$ (in $\mathbb{Z}_p$).

□

For general n, we need to use the idea of Algebra to show that this proposition can be extended to the case where p is not a prime.

Lemma 14.3. *If G is an abelian group, then there exists a subgroup H such that G/H is of order p where $p|n$.*

Theorem 14.4 (Erdös-Ginzburg-Ziv)**.** *For each positive integer n let $S = \{a_1, a_2, ..., a_{2n-1}\}$ be a subset of $\mathbb{Z}_n$. Then, there exists an n-subset of S, S', such that $\sum_{x \in S'} x = 0$.*

Proof. By induction on n. If n is a prime, then the proof can be obtained by Proposition 14.2. So, assume that $n \geq 2$ and n is a composite integer. Let $p|n$ and p be a prime. For convenience, let $n = p \cdot h$ where $h > 1$. Now, let H be a subgroup of $\mathbb{Z}_n$ and $|H| = h$.

Since $\langle \mathbb{Z}_n, + \rangle$ is an abelian group, $\mathbb{Z}_n/H$ is a cyclic group of order p. Further, $\mathcal{H} = \{a_1 + H, a_2 + H, \cdots, A_{2n-1} + H\}$ is a set of $2n-1$ cosets. By the fact that $\mathbb{Z}_n/H \cong \mathbb{Z}_p$, and $2n - 1 > 2p - 1$, there exist p cosets in $\mathcal{H}$ whose sum is H. We can select such p cosets at a time to find a collection of $2h-1$ sets $\mathcal{H}_1, \mathcal{H}_2, ..., \mathcal{H}_{2h-1}$ satisfying $\sum_{y \in \mathcal{H}_i} y = H,\ i = 1, 2, ..., 2h-1$. $(2n - 1 = 2ph - 1 = p(2h-1) + p - 1.)$ For convenience, let $\mathcal{H}_i = \{a_{(i-1)p+1}, a_{(i-1)p+2}, ..., a_{ip}\} + \mathrm{H}$ and $\sum_{j=1}^{p} a_{(i-1)p+j} = b_i$. Since $\sum_{y \in \mathcal{H}_i} y = H,\ b_i \in H$. By the fact that $|H| < n$ and $\mathbb{B} = \{b_1, b_2, ..., b_{2h-1}\}$ is a set of $2h-1$ elements in H, we conclude that there are h elements in $\mathbb{B}$, $b_{i_1}, b_{i_2}, ..., b_{i_h}$, with 0 sum. Clearly, 0 in H is also 0 in $\mathbb{Z}_n$. The proof follows by letting the $ph = n$ elements be obtained from corresponding b_{i_j}'s, i.e., the union $\bigcup_{k=1}^{h} \{a_{(i_k-1)p+j} \mid j = 1, 2, ..., p\}$.

□

Definition 14.2 (Doubling constant)**.** The doubling constant of an additive set A (in Z) is $\sigma[A] =_{def} \frac{|2A|}{|A|} = \frac{|A+A|}{|A|}$.

Facts.

6. $1 \leq \sigma[A] \leq \frac{|A|+1}{2}$. (By Fact 4.)

Remark.

- The upper bound can be attained by letting $Z = \mathbb{Z}$ (set of integers), $A = \{1, 2, 2^2, ..., 2^{n-1}\}$. Now, the sum of any two elements (may be the same) from A are distinct and thus $\sigma[A] = \frac{|A|+1}{2}$. This kind of sets are known as Sidon sets.
- The lower bound of $\sigma[A]$ can be smaller. Let $A = \{0, d, 2d, ..., (n-1)d\}$ where $d \neq 0$ and $d \in \mathbb{Z}$. Then, $\sigma[A] \leq 2 - \frac{1}{n}$. This is by the fact that $\forall x, y \in A$, $x+y \in \{0, d, 2d, ..., (n-1)d, ..., 2(n-1)d\}$. Hence, $|A+A| \leq 2(n-1)+1 = 2n-1$, thus $\sigma[A] \leq \frac{2n-1}{n} = 2 - \frac{1}{n}$.

  As a matter of fact, since $kd \neq 0$ (in $\mathbb{Z}$) for any $k \in \mathbb{N}$, we also conclude that $\sigma[A] = 2 - \frac{1}{n}$ for the above set of arithmetic progression.
- If Z is a finite group, then the above equality holds if the order of d in Z is larger than $2(n-1)$.

Definition 14.3 (Difference constant)**.** Similarly, we can define the difference constant as $\delta[A] = \frac{|A-A|}{|A|}$.

Remark. Again, we have $\delta[A] \geq 1$ and $\delta[A] \leq |A| - 1 + \frac{1}{|A|}$.

Definition 14.4 (Product set)**.** Besides sum set, we can also consider the product set $A \cdot A$ where $A \cdot A := \{ab | a, b \in A\}$.

Remark. Clearly, $|A \cdot A|$ can be as large as the order of $|A|^2$. But, $|A \cdot A|$ can also be small if we take A as the set $\{1, r, r^2, ..., r^{n-1}\}$, then $|A \cdot A| = 2n-1$ (provided the multiplication order of r is larger than $2n-2$ in the multiplication group).

One of the important problems in Additive Combinatorics is to simultaneously estimate $|A+A|$ and $|A\cdot A|$ if both addition and multiplication are operations defined on A. Especially, is that possible to find a set A for which both $|A+A|$ and $|A\cdot A|$ are small.

Conjecture 14.1 (Erdös). *For every $\epsilon > 0$, every sufficiently large set $A \subseteq \mathbb{R}$ (or $\mathbb{Z}$) satisfies* $\max\{|A+A|, |A\cdot A|\} \geq |A|^{2-\epsilon}$.

The following theorem provides a good estimation of $\max\{|A+A|, |A\cdot A|\}$.

Theorem 14.5 (Elekes). *For any $A \subseteq \mathbb{R}$,*

$$|A+A|\cdot|A\cdot A| \geq \frac{1}{64}|A|^{5/2}.$$

To prove Theorem 14.5, we need the following theorems.

- If $cr(G) = 0$ (crossing number of G), then $||G|| \leq 3|G| - 6$.
- $cr(G) \geq ||G|| - 3|G| + 6 \leq ||G|| - 3|G|$.

  Proof. Let $cr(G) = c$. Convert G into a planar graph G' by letting the crossings be vertices, Hence, $|G'| = |G| + c$, $||G'|| = ||G|| + 2c$. By using $||G'|| \leq 3|G'| - 6$, we have $c \geq ||G|| - 3|G| + 6$.

  □

Theorem 14.6 (Crossing Lemma). *If $||G|| \geq 4|G|$, then $cr(G) \geq \dfrac{||G||^3}{64|G|^2}$.*

Proof. (Probabilistic method) Set $|G| = v, ||G|| = e$ and $cr(G) = c$. Consider a drawing of G with crossing number $cr(G) = c$. Let $p = \dfrac{4v}{e} \leq 1$. Choose $V' \subseteq V$ by selecting each vertex independently with probability p. Let $G' = \langle V'\rangle_G$. G' has a drawing from G and let $cr(G') = c'$. Then, $0 \leq c' - e' + 3v'$. Take expectation of them, we have

$$\begin{aligned} 0 &\leq E[c'] - E[e'] + 3E[v'] = p^4\cdot c - p^2\cdot e + 3p\cdot v \\ &= p^4(c - \frac{e}{p^2} + \frac{3v}{p^3}) = p^4(c - \frac{e^3}{64v^2}). \end{aligned}$$

□

Theorem 14.7 (Szemeredi-Trotter)**.** *Let V be a set of points and L be a set of lines in $\mathbb{R}^2$. let $m =_{def} \#\{(p,l) \in V \times L | p \sim l\}$. Then,*

$$m \leq 4(|V|^{2/3} \cdot |L|^{2/3} + |V| + |L|).$$

Proof. We may assume that every line contains a point. For each $l \in L$, add an edge between two consecutive pair of points on l, this gives $k-1$ edges for each line with k points. (Consider a graph $G = (V, E)$.) $|E| = \sum_{l \in L} (\#(\text{points on } l) - 1)$. $|L|^2 \geq \#\text{crossings}$.

Note that if $m - |L| = |E| < 4|V|$, then $m < 4|V| + |L|$. Done.

On the other hand, $|E| \geq 4|V|$, $|L|^2 \geq \#\text{crossings} \geq \frac{(m-|L|)^3}{64|V|^2}$ (by Crossing Lemma). Thus,

$$\begin{aligned}(m-|L|)^3 &\leq |L|^2 \cdot 64 \cdot |V|^2 \\ m - |L| &\leq 4|L|^{2/3} \cdot |V|^{2/3} \\ m &\leq 4(|V|^{2/3} \cdot |L|^{2/3} + |V| + |L|).\end{aligned}$$

□

Now, we can prove the theorem of Elekes.

Proof of Theorem 14.5. Let $A \subseteq \mathbb{R}$, define $V = (A+A) \times (A \cdot A)$. For $a, b \in A$, let $l_{a,b}$ be the line given by the equation $y = a(x-b)$. $L = \{l_{a,b} | a, b \in A\}$. For every $c \in A$, the point $(c+b, ac) \in V$. So, every $l_{a,b}$ hits at least $|A|$ points.

Since $|L| = |A|^2$,

$$\begin{aligned}|A|^3 = |A| \cdot |L| &\leq \#\text{point-line incident between } V \text{ and } L \\ &\leq 4(|V|^{2/3} \cdot |L|^{2/3} + |V| + |L|) = 4(|V|^{2/3} \cdot |A|^{4/3} + |V| + |A|^2) \\ &\leq 16|V|^{2/3}|A|^{4/3} \text{ (estimation)}\end{aligned}$$

Hence,

$$\begin{aligned}|V|^{2/3} &\geq \frac{1}{16}|A|^{5/3} \\ |V| &\geq (\frac{1}{16}|A|^{5/3})^{3/2} = \frac{1}{64}|A|^{5/2}.\end{aligned}$$

□