An Application of the Probabilistic Method to Sum-Free Sets

Bharath Jaladi

Lecture Given at PACT 2017: July 25^{th}, 2017 | Updated: November 12^{th}, 2017
Inspired by and expanded upon a lecture given by Mark Braverman at PACT 2014

Definition A set S is said to be sum-free when, for any two elements $a, b \in S$ (not necessarily distinct), $a + b \notin S$.

Theorem (Erdös) Every nonempty set $B = \{b_1, b_2, ..., b_n\}$ of n non-zero integers contains a sum-free subset A such that $|A| > \frac{n}{3}$.

Proof.

Lemma 1 There exist infinitely many primes p of the form $p = 3j + 2, j \in \mathbb{Z}^+$.

Proof. Assume for the sake of contradiction that there exist finitely many primes p of the form $p = 3j + 2, j \in \mathbb{Z}^+$. In particular, note here that because $j > 0$, any prime of this form must be greater than 2 and odd. Suppose there are h such primes, $p_1, p_2, ..., p_h$. Clearly, $p_1 p_2 \cdots p_h$ must be some odd positive integer, g. Consider $P = 3p_1 p_2 \cdots p_h + 2$. Note that P is odd and an integer greater than 1. Thus, it must be the case that P is composite, else there would be a contradiction, as it is of the form $3g + 2$. As such, P must be able to be expressed as the product of odd primes. It can be seen, however, that P is not divisible by 3 or any prime $p_1, p_2, ..., p_h$, as each is a divisor of $P - 2$ and greater than 2. Further, note that any number of the form $3j + 3$ for $j \in \mathbb{Z}^+$ is divisible by 3 and thus not prime. As a result, P must be able to be expressed as the product of odd primes of the form $3j + 1$ for $j \in \mathbb{Z}^+$. Suppose the prime factorization of P consists of f such primes (not necessarily distinct), $p_1^{''}, p_2^{''}, ..., p_f^{''}$. That is, P can be expressed as $(3p_1^{'} + 1)(3p_2^{'} + 1) \cdots (3p_f^{'} + 1)$. However, this product will be of the form $3e + 1$ for some $e \in \mathbb{Z}^+$ (that is, $P \equiv 1 \pmod 3$), which is a contradiction, as P was defined to be $3g + 2$ (that is, $P \equiv 2 \pmod 3$). □

Let $r = 3k + 2$ be a prime such that $r > 2 \max_i b_i$. Such a prime must exist by Lemma 1. Consider the set $C = \{k + 1, k + 2, ..., 2k + 1\}$. Note that the elements of C are a subset of the possible values of $a \bmod r$ for $a \in \mathbb{Z}$. Let c, d be arbitrary but particular elements of C (not necessarily distinct). It can be seen that $k + 1 \leq c, d$, and as such, $(k + 1) + (k + 1) = 2k + 2 \leq c + d$. As $2k + 2$ is greater than the largest element of C, $2k + 1$, it is clear that C is a sum-free set.

Definition Denote a set S as sum-free with respect to mod c when, for any two elements $a, b \in S$ (not necessarily distinct), $a + b \pmod c \notin S$.

C was already shown to be sum-free. With an additional observation, it can be seen that C is sum-free with respect to mod r. Once again consider arbitrary but particular $c, d \in C$ (not necessarily distinct). Clearly, $2k + 1 \geq c, d$. Thus, $c + d \leq (2k + 1) + (2k + 1) = 4k + 2 \equiv k$

(mod r). As such, C can be denoted as sum-free with respect to mod r.

Lemma 2 For any integers x, y, z and $u \in \mathbb{R}$, if ux mod w, uy mod w, and uz mod w are elements of a set D that is sum-free and sum-free with respect to mod w, then $x + y \neq z$.

Proof. Let ux mod $w = x'$, uy mod $w = y'$, and uz mod $w = z'$. As D is sum-free, it is immediately seen that $x' + y' \neq z'$. ux, uy, and uz can be expressed $ux = ra + x'$, $uy = rb + y'$, and $uz = rc + z'$ for some integers a, b, c. Assume for the sake of contradiction that $x + y = z$, that is, $\frac{ra+x'}{u} + \frac{rb+y'}{u} = \frac{rc+z'}{u}$. Thus, we have the following.

$$\frac{ra + x'}{u} + \frac{rb + y'}{u} = \frac{rc + z'}{u}$$

$$ra + x' + rb + y' = rc + z'$$

$$a + \frac{x'}{r} + b + \frac{y'}{r} = c + \frac{z'}{r}$$

$$(a + b) + \frac{x' + y'}{r} = c + \frac{z'}{r}$$

In the exact same manner as in the proof of Lemma 2, it can be seen that $0 \leq x', y', z' \leq r - 1$, thus, $0 \leq x' + y' \leq 2r - 2$. a, b, c are integers, thus, it must be the case that the fractional part of $\frac{x'+y'}{r}$ must equal $\frac{z'}{r}$. That is, $x' + y' \pmod r = z'$, which is a contradiction. □

Select a q uniformly at random from $[1..r - 1]$ and consider the set $A = \{b_i \mid qb_i \pmod r \in C\}$. By Lemma 2, A is sum-free. Note that for all i, qb_i is not divisible by r because $q, b_i < r$ and r is prime. Thus, there are $3k + 1$ possible values qb_i $(1, 2, ..., 3k + 1)$ for all i. Note that for a particular b_i, as q ranges over $[1..r - 1]$, $qb_i \pmod r$ ranges over all elements of C. As a result, for each i

$$Pr[qb_i \in C] = \frac{|C|}{3k + 1} = \frac{k + 1}{3k + 1} > \frac{1}{3}.$$

Using this result, it can be seen that

$$\mathbb{E}[|A|] = \sum_{i=1}^{n} Pr[qb_i \in C] > \sum_{i=1}^{n} \frac{1}{3} > \frac{n}{3}.$$

As $\mathbb{E}[|A|] > \frac{n}{3}$, there must exist some A such that $|A| > \frac{n}{3}$. Thus, there exists a sum-free subset A of B such that $|A| > \frac{n}{3}$. □