Sum-Free Subsets of Integers

Talk by Aaron Berger
Notes by Sanjana Das
February 17, 2023

§1 Introduction

The question we'll consider is the following.

Question 1.1. Given a set A of positive integers with $|A| = n$, how large of a sum-free subset $B \subseteq A$ can we find?

Definition 1.2. A set B is *sum-free* if there do not exist $x, y, z \in B$ with $x + y = z$.

Example 1.3

If $A = \{1, 2, 3\}$, we can't take B to be the entire subset $\{1, 2, 3\}$ as $1 + 2 = 3$, and we can't take B to be $\{1, 2\}$ as $1 + 1 = 2$. However, we *can* take B to be $\{1, 3\}$, which is sum-free.

The first-order result is due to Erdős.

Theorem 1.4 (Erdős)

We can always find a sum-free subset $B \subseteq A$ of size at least $\frac{n}{3}$.

Proof. Define a random variable $x \sim \text{Unif}[0, 1]$, and let

$$A_x = \left\{ a \in A \mid \{ax\} \in \left[\frac{1}{3}, \frac{2}{3}\right) \right\}.$$

Claim — The set A_x is always sum-free.

Proof. If $\{ax\} \in [\frac{1}{3}, \frac{2}{3})$ and $\{bx\} \in [\frac{1}{3}, \frac{2}{3})$, then

$$\begin{aligned} \{(a+b)x\} &= \{ax\} + \{bx\} \mod 1 \\ &\in \left[\frac{2}{3}, \frac{4}{3}\right) \mod 1, \end{aligned}$$

which consists of $[0, \frac{1}{3})$ and $[\frac{2}{3}, 1)$, so $\{(a+b)x\}$ cannot be in $[\frac{1}{3}, \frac{2}{3})$. This means if a and b are in A_x, then $a + b$ is not in A_x. □

So for any choice of x, the set A_x is a sum-free subset of A; our goal is to show that for some x, this produces a subset of size at least $\frac{n}{3}$. We'll do this by looking at the *expectation* of the size of this subset.

Claim — We have $\mathbb{E}[|A_x|] = \frac{n}{3}$.

Proof. We can write $|A_x| = \sum_{a \in A} \mathbf{1}_{a \in A_x}$ as a sum of indicator functions, so by linearity of expectation

$$\mathbb{E}[|A_x|] = \sum_{a \in A} \mathbb{P}[a \in A_x] = \frac{n}{3},$$

since for each $a \in A$ the probability that $\{ax\} \in [\frac{1}{3}, \frac{2}{3})$ is $\frac{1}{3}$. □

So for some choice of x, we must have $|A_x| \geq \frac{n}{3}$, as desired. □

It is possible to improve the theorem in the case where $3 \mid n$ — if we can get even a tiny improvement on this bound (e.g. if $n = 9$, instead of showing we can produce $|A_x| \geq 3$ we show we can produce $|A_x| \geq 3.01$), then we can improve the bound by a full integer in the $3 \mid n$ case, since subset sizes are integers.

To get this small improvement, let $\varepsilon = \frac{1}{3 \max(A)}$ (where $\max(A)$ denotes the largest element in A). Then for any $x < \varepsilon$, we have $\{ax\} = ax < \frac{1}{3}$ for all $a \in A$, so $|A_x| = 0$. This means $\mathbb{P}[|A_x| = 0] \geq \varepsilon$, so

$$\mathbb{E}[|A_x|] \leq \varepsilon \cdot 0 + (1 - \varepsilon) \cdot \max_x |A_x|$$

(where $\max_x |A_x|$ denotes the maximum size of $|A_x|$ over all $x \in [0, 1]$). Since $\mathbb{E}[|A_x|] = \frac{n}{3}$, this means

$$\max |A_x| \geq \frac{n}{3(1 - \varepsilon)} > \frac{n}{3}.$$

So taking B to be the set A_x of maximal size, we get that there exists sum-free $B \subseteq A$ with $|B| > \frac{n}{3}$, and since $|B|$ must be an integer, this means $|B| \geq \frac{n}{3} + \frac{1}{3}$.

The result we will talk about is the following.

Theorem 1.5 (Bourgain)
We can always find a sum-free subset $B \subseteq A$ of size at least $\frac{n}{3} + \frac{2}{3}$.

This may not seem like a big improvement — it only even gives a better bound in the case $n \equiv 2 \pmod 3$ — but it really means that instead of bounding $\max_x |A_x| - \frac{n}{3}$ by a tiny number as in the above proof (our bound of $\frac{n}{3(1-\varepsilon)}$ can be arbitrarily close to $\frac{n}{3}$ if ε is tiny), we need to bound it by some constant (greater than $\frac{1}{3}$). So in this result, we're really going from 'arbitrarily small' to 'constant.'

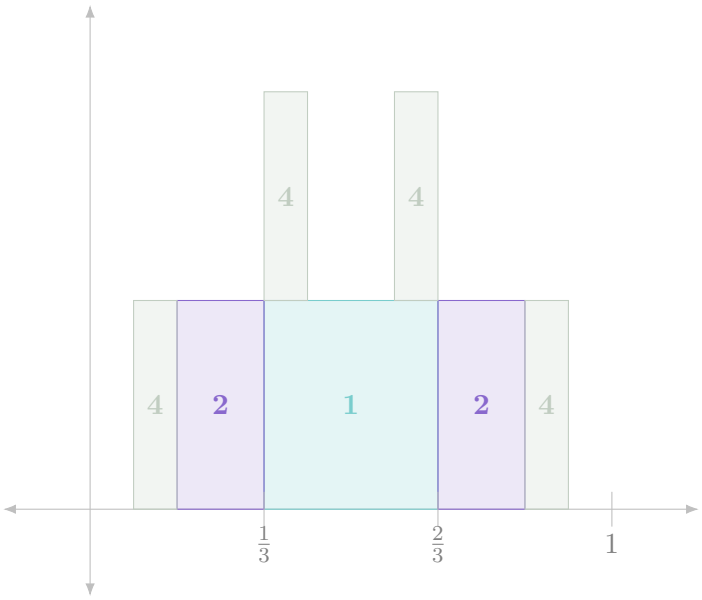
Remark 1.6. The constant of $\frac{1}{3}$ cannot be improved — there is a (difficult) construction of sets where there is no sum-free subset of size $(\frac{1}{3} + \varepsilon)n$.

Remark 1.7. In the case where $A = \{1, \ldots, n\}$, the answer is $\frac{n}{2}$ — we can get a set of this size by taking the top half, or all the odd numbers. Meanwhile, this is tight because if the maximum number we choose is m, then the numbers smaller than it pair up (we can't take both x and $m - x$), so we can take at most half of them.

We won't go through the entire proof, as it involves some messy casework, but we will see one illustrative case. We'll also see ingredients in the Bourgain paper that lead to a possible path towards a bound of $\frac{n}{3} + \omega(1)$ (where $\omega(1)$ denotes a term that goes to ∞ as n increases). Finally, we'll also discuss a bound on $\mathbb{P}[|A_x| = 0]$.

§2 Proof of Bourgain's Theorem

Imagine graphing $|A_x|$ as a function of x. Then each element $a \in A$ creates a periodic set of 'bumps' — for example, if $A = \{1, 2, 4\}$ then we get the following graph:



For example, $\{1x\}$ is in our interval when x is in $[\frac{1}{3}, \frac{2}{3})$, contributing one bump in that interval; $\{2x\}$ is in our interval when x is in $[\frac{1}{6}, \frac{1}{3})$ or $[\frac{2}{3}, \frac{5}{6})$, contributing two bumps there; and similarly $\{4x\}$ contributes bumps at $[\frac{1}{12}, \frac{2}{12})$, $[\frac{4}{12}, \frac{5}{12})$, $[\frac{7}{12}, \frac{8}{12})$, and $[\frac{10}{12}, \frac{11}{12})$.

In particular, as we can see here, if we keep on adding powers of 2 then we halve the space where $|A_x| = 0$ each time; this gives a construction of a set A for which $\mathbb{P}[|A_x| = 0] \leq c \cdot 2^{-n}$. Later we'll see a *lower* bound on $\mathbb{P}[|A_x| = 0]$.

Notation 2.1. We define $f(x) = \mathbf{1}_{x \in [\frac{1}{3}, \frac{2}{3})} - \frac{1}{3}$, and $F(x) = |A_x| - \frac{n}{3} = \sum_{a \in A} f(ax)$.

We refer to $F(x)$ as the *discrepancy* (it's the amount by which $|A_x|$ is above or below its mean). We've previously shown that for some x the discrepancy must be at least some tiny number $\varepsilon > 0$; now we want to show that for some x it must be at least a constant (greater than $\frac{1}{3}$ — then this produces a subset of size greater than $\frac{n}{3} + \frac{1}{3}$).

§2.1 Fourier Series

The first ingredient of the proof is to find the Fourier series of f. By definition, the coefficient corresponding to each integer m is

$$\hat{f}(m) = \int_0^1 f(x) e^{-2\pi i m x} \, dx.$$

For $m = 0$ this is equal to 0 (this is why we subtracted $\frac{1}{3}$ in the definition of f). For $m \neq 0$ we have $\int_0^1 \frac{1}{3} e^{-2\pi i m x} = 0$, so we can ignore it and only consider the part of f from the indicator function; this gives

$$\hat{f}(m) = \int_{1/3}^{2/3} e^{-2\pi i m x} \, dx = -\frac{1}{2\pi i m x} e^{-2\pi i m x} \Big|_{1/3}^{2/3} = -\frac{\chi(m)}{m} \cdot \frac{\sqrt{3}}{2\pi},$$

where $\chi(m)$ is the multiplicative character mod 3 given by

$$\chi(m) = \begin{cases} 0 & \text{if } m \equiv 0 \pmod 3 \\ 1 & \text{if } m \equiv 1 \pmod 3 \\ -1 & \text{if } m \equiv 2 \pmod 3. \end{cases}$$

So then we can write

$$f(x) = \sum_{\substack{m \in \mathbb{Z} \\ m \neq 0}} -\frac{\sqrt{3}}{2\pi} \cdot \frac{\chi(m)}{m} \cdot e(mx),$$

where $e(t)$ denotes $e^{2\pi i t}$. Now the Fourier series of $F(x) = \sum_{a \in A} f(ax)$ is

$$F(x) = -\frac{\sqrt{3}}{2\pi} \sum_{a \in A} \sum_{\substack{m \in \mathbb{Z} \\ m \neq 0}} \frac{\chi(m)}{m} e(amx).$$

In particular, the Fourier series has a contribution for every multiple of an element in A. We can also group terms with the same exponent by summing over $r = am$ instead; this gives

$$F(x) = -\frac{\sqrt{3}}{2\pi} \sum_{\substack{r \in \mathbb{Z} \\ r \neq 0}} \sum_{\substack{a \in A \\ a \mid r}} \frac{\chi(r/a)}{r/a} e(rx).$$

§2.2 The Main Idea

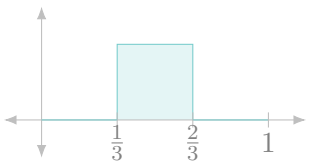
Our goal is to have $|A_x|$ a constant amount above its expectation. One idea for how we might do this is to condition on a certain element being in A_x. Suppose we fix some $u \in A$, and condition on u lying in A_x. (This is equivalent to a simple condition on x — for example, if $u = 1$ then the conditioning constrains x to lie in $[\frac{1}{3}, \frac{2}{3})$.) If whether each other element is present in A_x were independent of whether u is present, then we would have

$$\mathbb{E}[|A_x| \mid u \in A_x] = 1 + \frac{n-1}{3} = \frac{n}{3} + \frac{2}{3}.$$

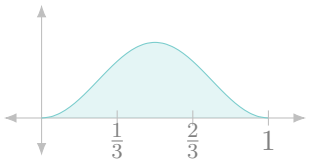
(It would suffice if on average, their presence were uncorrelated with u.) This would give us that $|A_x|$ is at least a constant for some x (i.e., $|A_x|$ contains at least a constant number of elements more than its expectation), which is what we're looking for.

This can't possibly work because the other elements are not uncorrelated with u — for example, if we require 1 to be in A_x, this forces 2 to *not* be in A_x; similarly if we require 2 to be in, this forces both 1 and 4 to be out. So this idea has obvious problems — but Bourgain somehow makes it work. The main idea is that instead of doing a *hard* condition, we do a *soft* condition that manages to avoid these issues.

Right now, we're trying to draw x from the set where $\{ux\} \in [\frac{1}{3}, \frac{2}{3})$ — for example, if $u = 1$, then we're drawing x from $[\frac{1}{3}, \frac{2}{3})$.



Instead, we'll draw x from a more smooth distribution — we'll draw x according to the distribution $\mu_u(x) = 1 - \cos(2\pi u x)$. (This is a valid distribution as $\mu_u(x)$ is always nonnegative and has integral 1 — the integral of the cosine is 0.) For example, if $u = 1$ then the distribution looks like the following.



This distribution makes it more *likely* that $u \in A_x$, but it's not guaranteed. The reason this distribution is useful is that the expansion of $\cos(2\pi ux)$ only has two terms — $e(ux)$ and $e(-ux)$ — which interact with our Fourier series in very predictable ways, so it's more predictable what this type of conditioning does to our function.

Our goal is to show that $\mathbb{E}_{x\sim\mu_u} F(x) > \frac{1}{3}$ — if we can show that $\mathbb{E}[F(x)]$ for x drawn from any one distribution is greater than $\frac{1}{3}$, then there must exist some x for which $F(x) > \frac{1}{3}$ and we're done. We have

$$\begin{aligned}\mathbb{E}_{x\sim\mu_u} F(x) &= \int_0^1 F(x)\mu_u(x)\,dx \\ &= \int_0^1 F(x)\left(1 - \frac{1}{2}e(ux) - \frac{1}{2}e(-ux)\right)\,dx.\end{aligned}$$

But $\int_0^1 F(x) = 0$ and integrating F against an exponential $e^{2\pi iux}$ gives the corresponding Fourier coefficient, so this is equal to

$$-\frac{1}{2}\hat{F}(-u) - \frac{1}{2}\hat{F}(u) = \frac{\sqrt{3}}{2\pi}\sum_{\substack{a\in A\\ a|u}} \frac{\chi(u/a)}{u/a}$$

(using our calculation of the Fourier series of F — the coefficients corresponding to u and $-u$ are the same). This expression is somewhat complicated, but it becomes a lot simpler if we choose $u \in A$ to have no proper divisors in A — for example, by choosing u to be the smallest element of A. Then the only contribution to the sum is from $a = u$, which gives

$$\mathbb{E}_{x\sim\mu_u} F(x) = \frac{\sqrt{3}}{2\pi}\cdot\frac{\chi(1)}{1} = \frac{\sqrt{3}}{2\pi}.$$

Unfortunately, this is around 0.28, which is slightly less than $\frac{1}{3}$ — so this method did give us some constant, but it wasn't good enough. But this gives us hope that the method might work. There's two ways to improve the analysis:

- Choose a distribution μ_u that looks closer to the original step function, for example by adding another cosine term. We're not going to do this, but it is used in the full proof (which involves more complicated casework).
- Choose another element v (preferably with no proper divisors as well) and use *both* u and v.

§2.3 Proof of Special Case

Proposition 2.2

Assume that $\gcd(A) = 1$ and $1 \notin A$. Then Theorem 1.5 holds (i.e., we can find B of size at least $\frac{n}{3} + \frac{2}{3}$).

We can assume $\gcd(A) = 1$ without loss of generality — multiplying or dividing every element of A by a constant doesn't change anything. The assumption that $1 \notin A$ is what makes this a special case.

Proof. Let u be the minimal element of A, and let v be the minimal element of $\{a \in A \mid u \nmid a\}$ (which must exist because $u > 1$ and $\gcd(A) = 1$, so u cannot divide all elements of A — this is why we need the assumption).

Claim — Both u and v have no proper divisors in A.

Proof. It's clear that u has no proper divisors in A. Meanwhile for v, since v itself is not a multiple of u, all divisors of v are not multiples of u either. So if v had a proper divisor in A, this divisor would be smaller than v but not a multiple of u, which is not possible. □

Now we'll draw x from the distribution

$$\mu_{u,v}(x) = (1 - \cos(2\pi ux))(1 - \cos(2\pi vx)).$$

To check that this is a valid distribution, we need to check that it's nonnegative (which is clear) and that its integral is 1; we can show this by rewriting the cosines to get

$$\mu_{u,v}(x) = \left(1 - \frac{1}{2}e(ux) - \frac{1}{2}e(-ux)\right)\left(1 - \frac{1}{2}e(vx) - \frac{1}{2}e(-vx)\right).$$

If we expand this out and integrate, the constant term gives 1, and all the other terms are exponentials and integrate to 0.

Now we have

$$\begin{aligned}\mathbb{E}_{x\sim\mu_{u,v}}F(x) &= \int_0^1 F(x)\mu_{u,v}(x)\,dx \\ &= \int_0^1 F(x)\,dx - \frac{1}{2}(\hat{F}(u) + \hat{F}(-u) + \hat{F}(v) + \hat{F}(-v)) \\ &\quad + \frac{1}{4}(\hat{F}(u+v) + \hat{F}(-u-v) + \hat{F}(u-v) + \hat{F}(v-u)).\end{aligned}$$

(The first term is from the constant in the above expansion, the next terms are from the linear terms, and the last terms are from the cross terms.)

The first term $\int_0^1 F(x)\,dx$ is 0. We've already computed that if a number $u \in A$ has no proper divisors in A, then $\hat{F}(u) = -\frac{\sqrt{3}}{2\pi}$, so the linear terms together contribute $2 \cdot \frac{\sqrt{3}}{2\pi}$. Now we'll look at the cross terms.

Claim — $v - u$ is not in A and has no divisors in A.

Proof. Since v is not a multiple of u and u is, then $v - u$ is not a multiple of u either, and neither is any of its divisors. Additionally, $v - u$ and all its divisors are smaller than v, so since v is the smallest non-multiple of u in A, none of them can be in A. □

This means $\hat{F}(v - u) = 0$ — since the only terms which can contribute to the coefficient are from $a \in A$ with $a \mid v - u$, of which there are none. This implies $\hat{F}(u - v) = 0$ as well.

Claim — No *proper* divisor of $v + u$ can be in A.

Proof. For the same reason as before, $v + u$ is not a multiple of u, so none of its proper divisors are either. But any proper divisor of $v + u$ is at most

$$\frac{v+u}{2} < \frac{v+v}{2} = v.$$

So $v + u$ may or may not be in A, but none of its proper divisors can be. □

This means we have

$$\hat{F}(u+v) = \hat{F}(-u-v) = -\frac{\sqrt{3}}{2\pi} \cdot \mathbf{1}_{u+v\in A}.$$

So plugging back into the original expression, we have

$$\int_0^1 F(x)\mu_{u,v}(x)\,dx = 0 + 2\cdot\frac{\sqrt{3}}{2\pi} - \frac{1}{2}\cdot\frac{\sqrt{3}}{2\pi} \geq \left(2-\frac{1}{2}\right)\frac{\sqrt{3}}{2\pi} = \frac{3\sqrt{3}}{4\pi}.$$

This is approximately 0.41, which is bigger than $\frac{1}{3}$, so we are done. □

Remark 2.3. For a heuristic reason why we should expect this to work out (i.e., that the constants obtained from the $\frac{\sqrt{3}}{2\pi}$ terms are of the right order of magnitude), the original heuristic was that forcing an element to be in the set would give us an improvement of $\frac{2}{3}$. So we'd expect smoothing out the distribution (i.e., drawing from a distribution that makes it more likely but not guaranteed that the element is in the set) should give a reasonable fraction of that improvement. This reasonable fraction turns out to be less than $\frac{1}{2}$, but it is greater than $\frac{1}{4}$ (i.e., $\frac{1}{6} < \frac{\sqrt{3}}{2\pi} < \frac{2}{3}$), so it's reasonable that doing it twice might work.

§3 Discrepancy and the Littlewood Conjecture

Next we'll talk about a result of Bourgain from trying to find a different way to show that the discrepancy can be large — by relating it to a quantity considered in the Littlewood conjecture.

To hear more about this, you can reference the talk by Thomas Bloom called *Sets with small L^1 Fourier norm* at IAS.

Question 3.1. Given $A \subseteq \mathbb{Z}$ with $|A| = n$, what can we say about the function $g_A(x) = \sum_{a\in A}\cos(2\pi a x)$ — in particular, can we lower bound its L^1 norm $\|g_A(x)\|_1$?

Example 3.2

Some examples for specific A (these are not obvious, and ignore constants):

- If A is an arithmetic progression, $\|g_A(x)\|_1 \sim \log(n)$.
- If A is a *generalized arithmetic progression*, $\|g_A(x)\|_1 \sim \log(n)^t$.
- If A is random, then $\|g_A(x)\|_1 \sim n^{1/2}$.

The Littlewood conjecture is the following; it is now a theorem. (There are multiple Littlewood conjectures; googling *Littlewood conjecture* may not give this one.)

Theorem 3.3

There exists a constant $c > 0$ such that $\|g_A(x)\|_1 \geq c\log(n)$.

The reason this is relevant is because of the following result Bourgain proved (in the same paper).

Proposition 3.4 (Bourgain)

There exists a constant $c > 0$ such that $\max F(x) \geq \frac{c}{\log n}\|g_A(x)\|_1$.

Using the above theorem, this tells us that the (maximum) discrepancy is at least some absolute constant. For our purposes — to prove a bound on $|B|$ better than $\frac{n}{3} + \frac{1}{3}$ — we would need this constant to be big enough. However, an idea for how we might be able to prove a bound of $\frac{n}{3} + \omega(1)$ comes from here — the point is that for most sets, $\|g_A(x)\|_1$ should be much bigger than $\log(n)$ (for example, $\log(n)^2$). This norm should only be small for sets which are very structured (e.g., sets which look like an arithmetic progression), and for such sets we might be able to find a sum-free subset by hand.

§3.1 Proof of Bourgain's Result

First we'll relate $F(x)$ to the function $g_A(x)$ — using our original formula and swapping the order of summation and combining the positive and negative terms, we have

$$\begin{aligned} F(x) &= -\frac{\sqrt{3}}{2\pi} \sum_{\substack{m \in \mathbb{Z} \\ m \neq 0}} \frac{\chi(m)}{m} \sum_{a \in A} e(amx) \\ &= -\frac{\sqrt{3}}{\pi} \sum_{m>0} \frac{\chi(m)}{m} \sum_{a \in A} \cos(2\pi amx) \\ &= -\frac{\sqrt{3}}{\pi} \sum_{m>0} \frac{\chi(m)}{m} g_A(mx). \end{aligned}$$

This writes F as a sum of an infinite number of copies of g_A. Ideally we'd like to show that the contribution from the $m = 1$ term is somewhat comparable to the entire summation, but this is difficult.

First, we will bound the size of this tail (i.e., the contribution from terms with m greater than some value).

Claim 3.5 — The tail of $f(x)$ is bounded above by $\left\| \sum_{m>\ell} \hat{f}(m) e(mx) \right\|_1 \lesssim \frac{1}{\sqrt{\ell}}$.

Note that here we're using f instead of F — recall that

$$f(x) = -\frac{\sqrt{3}}{2\pi} \sum_{\substack{m \in \mathbb{Z} \\ m \neq 0}} \frac{\chi(m)}{m} \cdot e(mx)$$

(and F is a sum of copies of f).

Proof. Let $f_{>\ell}$ denote this tail $\sum_{m>\ell} \hat{f}(m) e(mx)$. Instead of considering its L^1 norm, we'll consider its L^2 norm — we wil show that $\|f_{>}\ell\|_2^2 \leq \frac{1}{\ell}$. (The left-hand side denotes the square of the L^2 norm.)

By Plancherel's theorem, we have

$$\|f_{>\ell}\|_2^2 = \sum_{|m|>\ell} \hat{f}(m)^2.$$

Since $\hat{f}(m) = -\frac{\sqrt{3}}{2\pi} \cdot \frac{\chi(m)}{m}$ and $\chi(m)^2 \in \{0, 1\}$ for all m, this means

$$\|f_{>\ell}\|_2^2 \lesssim \sum_{|m|>\ell} \frac{1}{m^2} \lesssim c \int_{\ell}^{\infty} \frac{1}{m^2}\, dm = \frac{c}{\ell}.$$

Then we have $\|f_{>\ell}\|_1 \leq \|f_{>\ell}\|_2 \lesssim \frac{1}{\sqrt{\ell}}$. □

Next we'll relate the quantity $\max(F)$ we care about to the L^1 norm of F.

Claim 3.6 — We have $\max(F) \geq \frac{1}{2}\|F\|_1$.

Proof. We know F has integral 0, so if we write $F = F^+ + F^-$ (where F^+ corresponds to the parts of F which are positive, and F^- to the parts which are negative), we must have $\int F^+ = -\int F^- = \frac{1}{2}\|F\|_1$. Meanwhile we have $\max(F^+) \geq \int F^+$, so $\max(F) = \max(F^+) \geq \frac{1}{2}\|F\|_1$. □

Now we'd like to bound $\|F\|_1$ in terms of $\|g_A\|_1$. Right now we have an expression for F as a sum of copies of g_A. However, if we tried to apply the triangle inequality, it would point in the wrong direction — this would tell us that $\|F\|_1$ is *at most* something times $\|g_A\|_1$, while we want to show $\|F\|_1$ is *at least* such an expression. So we'd really like to take our expression for F as a sum of copies of g_A, and turn it into an expression for g_A as a sum of copies of F (then the triangle inequality would point in the right direction). This can be done using *Möbius inversion*.

Definition 3.7. The *Möbius function* $\mu(k)$ is defined as

$$\mu(k) = \begin{cases} (-1)^\ell & \text{if } k \text{ is a product of } \ell \text{ distinct primes} \\ 0 & \text{otherwise.} \end{cases}$$

Fact 3.8 — For all positive integers d,

$$\sum_{m|d} \mu(m) = \begin{cases} 1 & \text{if } d = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Now let p be a big prime, and consider the sum (over $k \geq 1$)

$$\sum_{k|p!} \frac{\mu(k)\chi(k)}{k} \cdot F(kx) = -\frac{\sqrt{3}}{\pi} \sum_{k|p!} \frac{\mu(k)\chi(k)}{k} \sum_{m\geq 1} \frac{\chi(m)}{m} g_A(kmx).$$

Since $\chi(k)\chi(m) = \chi(km)$, we can parametrize one of the summations by $r = km$ instead of m; then we can rewrite this sum as

$$-\frac{\sqrt{3}}{\pi} \sum_{r\geq 1} \frac{\chi(r)}{r} g_A(rx) \sum_{\substack{k|p! \\ k|r}} \mu(k).$$

But we can rewrite the condition $k \mid p!$ and $k \mid r$ as $k \mid \gcd(p!, r)$, so by the above fact this summation is 1 if $\gcd(p!, r) = 1$ and 0 otherwise; this means

$$\sum_{k|p!} \frac{\mu(k)\chi(k)}{k} \cdot F(kx) = -\frac{\sqrt{3}}{\pi} \sum_{\substack{r\geq 1 \\ \gcd(r,p!)=1}} \frac{\chi(r)}{r} g_A(rx).$$

On the right-hand side, we have a term of $g_A(x)$ from $r = 1$, and the remaining terms all come from $r \geq p$. These extra terms can be bounded in terms of $\|\sum_{a\in A} f_{>p}\|_1$ (we don't have all the terms, but the proof of the tail bound still works).

Now if we take $p \gg n^2$, the tail bound on $\|f_{>p}\|_1$ is $o(\frac{1}{n})$, and there are n terms (one for each a), so the triangle inequality gives

$$\|g_A(x)\|_1 \leq o(1) + \sum_{k|p!} \frac{|\mu(k)|}{k} \|F\|_1 .$$

(All the functions $F(kx)$ are the same as F but scaled down and squished.) Now the sum factors as

$$\sum_{k|p!} \frac{|\mu(k)|}{k} = \left(1+\frac{1}{2}\right)\left(1+\frac{1}{3}\right)\cdots\left(1+\frac{1}{p}\right)$$

(as $\mu(k)$ is 0 when k is not squarefree). This is approximately $\log p$, which is on the same order as $\log n$. So

$$\|g_A(x)\|_1 \le o(1) + 2\log n \cdot \|F\|_1,$$

which means

$$\max(F) \ge \frac{1}{2}\|F\|_1 \ge \frac{c\,\|g_A(x)\|_1}{\log n},$$

as desired.

§4 Probability of Size 0

Finally we'll talk about the probability that $|A_x| = 0$ (for x chosen uniformly at random). We've previously seen there exists a set A for which $\mathbb{P}[|A_x| = 0] = c2^{-n}$; we'll now see a lower bound.

Proposition 4.1
For any set A, we have $\mathbb{P}[|A_x| = 0] \ge 3^{-n}$.

One way to prove this is that the set $B = \{x \mid |A_x| = 0\}$ is a *Bohr set* of dimension n and radius $\frac{1}{3}$, which implies that $|B| \ge (\frac{1}{3})^n$. For what a Bohr set is, B is defined as the set of $x \in \mathbb{R}/\mathbb{Z}$ which satisfy a list of equations $\{a_1 x\} \in [-\frac{1}{3}, \frac{1}{3})$, $\{a_2 x\} \in [-\frac{1}{3}, \frac{1}{3})$, and so on. If you replace $\mathbb{R}/\mathbb{Z}$ with a general group, these equations with arbitrary homomorphisms, and $\frac{1}{3}$ with ρ, then you get the definition of a Bohr set.

We'll now see another proof of this result.

Claim 4.2 — We can cover $\mathbb{R}/\mathbb{Z}$ with 3^n translates of B.

This would prove that $|B| \ge 3^{-n}$.

In fact, we'll prove a more specific statement: consider the partition of $\mathbb{R}/\mathbb{Z}$ into at most 3^n parts, with the parts indexed by $(t_1, \dots, t_n)$ for $t_i \in \{0,1,2\}$ and defined as

$$\mathcal{P}_{t_1,\dots,t_n} = \{x \in \mathbb{R}/\mathbb{Z} \mid \{a_i x\} \in [\frac{t_i}{3}, \frac{t_i+1}{3}) \text{ for all } i \in [n]\}.$$

(Here $A = \{a_1, \dots, a_n\}$. In other words, we're splitting up all possible x by the list of which third of the unit interval each $\{a_i x\}$ lies in.)

Claim 4.3 — Each of the parts $\mathcal{P} = \mathcal{P}_{t_1,\dots,t_n}$ can be covered by a single translate of B.

Proof. If $\mathcal{P}$ is empty, this is clearly true. Otherwise fix x to be any element of $\mathcal{P}$. Then for any $x' \in \mathcal{P}$ we must have $x' - x \in B$ — for each i we have $\{a_i(x'-x)\} = \{a_i x'\} - \{a_i x\} \mod 1$, and since $\{a_i x'\}$ and $\{a_i x\}$ are both in $[\frac{t_1}{3}, \frac{t_1+1}{3})$, their difference must be (strictly) between $-\frac{1}{3}$ and $\frac{1}{3}$. This means $\mathcal{P}$ is covered by $x + B$. □