On Randomness of Subsets of $\mathbb{Z}_N$, as Described by Uniformity of Fourier Coefficients

Bryan Gillespie, Pennsylvania State University

August 9, 2010

Randomness is a notion which everyone is familiar with in a practical, day-to-day manner—we learn how to identify it and cope with it as a matter of course, just through the need to exist in a world full of uncertainty and unpredictability. It's not something that most people think too hard about—we just know it when we see it.

However, when we try to rigorously define what it means for something to be "random", we find that it is actually rather difficult to pin down. What properties should hold in a random system? What structure would one expect to find? What structure would one *not* expect to find? What should you be able to predict about such a system? It turns out that these questions are not easy to answer definitively—the inquisitive among us have come up with a surprising number of ways to define randomness, each with its own advantages and disadvantages, strengths and limitations.

In the following discussion, we will consider one such notion of randomness, that of "Fourier Pseudorandomness", which provides a measure of randomness for subsets of $\mathbb{Z}_N$. We will

- Define and prove the existence (in fact abundance) of Fourier Pseudorandom sets;
- Explore some of the nice properties of this type of randomness; and
- Discuss how Fourier Pseudorandomness is limited as a definition of randomness.

We will begin by recalling some tools and vocabulary from probability theory which will be useful in the technical details to follow.

1 Tools from Probability Theory

The notation and proofs of this section are drawn to some extent from Terence Tao and Van H. Vu's book *Additive Combinatorics*, sections 1.1 to 1.3 [1].

We assume some basic familiarity with the definitions and properties of probability theoretic notions such as probability spaces, events and random variables, independence and joint independence, and expectation and variance.

For an event E, we let $I(E)$ denote the indicator function of E, and for a set $U \subseteq \mathbb{Z}_N$, we let $U(x)$ denote the indicator function of the event $(x \in U)$. The probability of an event E will be denoted by $\Pr(E)$, and the expectation and variance of a random variable X will be denoted by $\mathbb{E}(X)$ and $\operatorname{Var}(X)$ respectively. We recall without proof some basic identities related to expectation and variance which will be particularly useful.

Proposition 1. *Let X, X_i, $i = 1, \dots, n$ be random variables, let c_i, $i = 1, \dots, n$ be constants, and let E, E_i, $i = 1, \dots n$ be events. Then*

- $\Pr(E_1 \cup \cdots \cup E_n) \le \Pr(E_1) + \cdots + \Pr(E_n)$
- $\mathbb{E}(I(E)) = \Pr(E)$, *and* $\operatorname{Var}(I(E)) = \Pr(E) - \Pr(E)^2$
- $\mathbb{E}(c_1 X_1 + \cdots + c_n X_n) = c_1 \mathbb{E}(X_1) + \cdots + c_n \mathbb{E}(X_n)$
- $\operatorname{Var}(X) := \mathbb{E}\Big(|X - \mathbb{E}(X)|^2\Big) = \mathbb{E}\Big(|X|^2\Big) - |\mathbb{E}(X)|^2$
- *If the X_i are pairwise independent, then*
  $\operatorname{Var}(X_1 + \cdots + X_n) = \operatorname{Var}(X_1) + \cdots + \operatorname{Var}(X_n)$.

The following inequality, called Markov's inequality, is quite important as a basic bound which controls the probability that a non-negative real-valued random variable is large, in terms of its expected value.

Theorem 2 (Markov's Inequality)**.** *Let X be a non-negative random variable. Then for any positive real $\lambda > 0$, we have*

$$\Pr(X \ge \lambda) \le \frac{\mathbb{E}(X)}{\lambda}. \tag{1}$$

Proof. For all a in the sample space, the trivial inequality $X(a) \ge \lambda \cdot I(X \ge \lambda)(a)$ holds. Taking the expectation of both sides and applying linearity of expectation yields the result. □

Markov's inequality is particularly valuable because it can be applied easily in rather diverse situations—it does not restrict the random variable X, and it only requires computation of the expectation, which is often a more manageable feat than higher moments. However, its usefulness is also limited by the fact that it fails to provide any bound on the probability that a random variable is small. A more useful inequality in this regard is Chebyshev's inequality, which bounds the probability that a random variable is far from its expected value, in terms of its variance.

Theorem 3 (Chebyshev's Inequality)**.** *Let X be a random variable. Then for any real $\lambda > 0$, we have*

$$\Pr(|X - \mathbb{E}(X)| \ge \lambda\sigma) \le \frac{1}{\lambda^2}, \tag{2}$$

where $\sigma := \operatorname{Var}(X)^{1/2}$.

Proof. In the case that $\operatorname{Var}(X) = 0$, we have that $X = \mathbb{E}(X)$ with full probability, so the inequality holds trivially. If $\operatorname{Var}(X) > 0$, we may apply Markov's inequality to the random variable $|X - \mathbb{E}(X)|^2$ to find

$$\Pr\Big(|X - \mathbb{E}(X)|^2 \ge \lambda^2 \operatorname{Var}(X)\Big) \le \frac{\mathbb{E}\Big(|X - \mathbb{E}(X)|^2\Big)}{\lambda^2 \operatorname{Var}(X)} = \frac{1}{\lambda^2},$$

and the result follows directly. □

So Chebyshev's inequality gives an improvement over Markov's in terms of lower-end behavior and in terms of decay with respect to the choice of constant λ. However, in certain more restricted cases, it is possible to do better than inverse-square decay. The following inequality, called Chernoff's inequality, in fact offers exponential decay as lambda varies, at the cost of a strong restriction on the form of the random variable.

Theorem 4 (Chernoff's Inequality)**.** *Let $X := X_1 + \cdots + X_n$ be a random variable, where $X_1, \ldots, X_n$ are simple, real-valued random variables which are jointly independent and have $|X_i - \mathbb{E}(X_i)| \leq 1$ for all i. Then for any real $\lambda > 0$, we have*

$$\Pr(|X - \mathbb{E}(X)| \geq \lambda\sigma) \leq 2\max\left(e^{-\lambda^2/4}, e^{-\lambda\sigma/2}\right), \tag{3}$$

where $\sigma := \mathrm{Var}(X)^{1/2}$

In order to approach this result, we need a lemma which bounds the so-called "exponential moment" $\mathbb{E}\big(e^{tX_i}\big)$ of the variables X_i. To this end we have

Lemma 4.1. *Let X be a simple real-valued random variable with $|X| \leq 1$ and $\mathbb{E}(X) = 0$. Then for any $-1 \leq t \leq 1$ we have $\mathbb{E}\big(e^{tX}\big) \leq \exp(t^2\,\mathrm{Var}(X))$.*

Proof. Since $|tX| \leq 1$, we may compare the exponential e^{tX} to its Taylor series expansion by

$$e^{tX} = \sum_{k=0}^{\infty} \frac{(tX)^k}{k!} = 1 + tX + \frac{t^2X^2}{2} + \frac{t^2X^2}{2}\sum_{k=3}^{\infty}\frac{(tX)^{k-2}}{k!/2} \leq 1 + tX + t^2X^2.$$

Taking the expectation of both sides and using linearity of expectation and the fact that $\mathbb{E}(X) = 0$, we find

$$\mathbb{E}\big(e^{tX}\big) \leq 1 + t^2\mathbb{E}\big(X^2\big) = 1 + t^2\mathbb{E}\Big(|X - \mathbb{E}(X)|^2\Big) \leq \exp(t^2\,\mathrm{Var}(X)).$$

□

Using this bound on the exponential moment, we may proceed with a proof of Chernoff's inequality.

Proof of Chernoff's inequality. Notice that $X - \mathbb{E}(X)$ is invariant under addition of a constant to the X_i. Therefore, assume without loss of generality that the variables X_i have $\mathbb{E}(X_i) = 0$ by working with the functions $\tilde{X}_i := X_i - \mathbb{E}(X_i)$. In particular this implies that $\mathbb{E}(X) = 0$.

Notice next that $\Pr(|X| \geq \lambda\sigma) = \Pr(X \geq \lambda\sigma) + \Pr(X \leq -\lambda\sigma)$. In applying a symmetric argument for $-X$ to the latter case, it thus suffices to show that

$$\Pr(X \geq \lambda\sigma) \leq e^{-t\lambda\sigma/2}$$

where $t := \min(\lambda/2\sigma, 1)$. We can apply Markov's inequality to obtain

$$\Pr(X \geq \lambda\sigma) = \Pr\Big(e^{tX} \geq e^{t\lambda\sigma}\Big) \leq e^{-t\lambda\sigma}\mathbb{E}\big(e^{tX_1}\cdots e^{tX_n}\big).$$

Since the variables X_i are jointly independent, so are the e^{tX_i}, and we may split the latter expectation. Using this and applying the lemma, we find that

$$\mathbb{E}\big(e^{tX_1}\cdots e^{tX_n}\big) = \mathbb{E}\big(e^{tX_1}\big)\cdots\mathbb{E}\big(e^{tX_n}\big) \leq \exp(t^2\,\mathrm{Var}(X_1))\cdots\exp(t^2\,\mathrm{Var}(X_n)).$$

Again by joint independence of the X_i, we know that $\mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = \mathrm{Var}(X) = \sigma^2$. Putting together the inequalities and using the fact that $t \leq \lambda/2\sigma$, we see that

$$\Pr(X \geq \lambda\sigma) \leq e^{-t\lambda\sigma}e^{t^2\sigma^2} \leq e^{-t\lambda\sigma}e^{t\sigma^2\cdot\lambda/2\sigma} = e^{-t\lambda\sigma/2},$$

and this completes the proof. □

2 The Fourier Transform on $\mathbb{Z}_N$

We now proceed to define the main player in our discussion of Fourier pseudorandomness—the Fourier transform on $\mathbb{Z}_N$.

Definition 5. *Let $f : \mathbb{Z}_N \to \mathbb{C}$. The Fourier transform $\widehat{f} : \mathbb{Z}_N \to \mathbb{C}$ of f is defined by*

$$\widehat{f}(\xi) := \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f(x) \exp\left(-x\xi \cdot \frac{2\pi i}{N}\right). \tag{4}$$

The Fourier transform has a number of interesting properties which are useful in applications. We note some of them here as a proposition, but leave the simple proofs as an exercise for the reader.

Proposition 6. *Let $f, g : \mathbb{Z}_N \to \mathbb{C}$. Then*

- *Fourier inversion formula:*

$$f(x) = \sum_{\xi \in \mathbb{Z}_N} \widehat{f}(\xi) \exp\left(x\xi \cdot \frac{2\pi i}{N}\right) \tag{5}$$

- *Parseval's identity:*

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} f(x)\overline{g(x)} = \sum_{\xi \in \mathbb{Z}_N} \widehat{f}(\xi)\overline{\widehat{g}(\xi)} \tag{6}$$

- *Plancherel's identity:*

$$\frac{1}{N} \sum_{x \in \mathbb{Z}_N} |f(x)|^2 = \sum_{\xi \in \mathbb{Z}_N} \left|\widehat{f}(\xi)\right|^2 \tag{7}$$

- *Uniform boundedness:*

$$\max_{\xi \in \mathbb{Z}_N} \left|\widehat{f}(\xi)\right| \leq \frac{1}{N} \sum_{x \in \mathbb{Z}_N} |f(x)| \tag{8}$$

A further definition which is useful in dealing with the Fourier transform is that of the convolution.

Definition 7. *Let $f, g : \mathbb{Z}_n \to \mathbb{C}$. The convolution $f * g : \mathbb{Z}_N \to \mathbb{C}$ of f with g is defined by*

$$f * g(x) := \frac{1}{N} \sum_{y \in \mathbb{Z}_N} f(y)g(x-y). \tag{9}$$

This function serves as an average of sorts between the functions f and g. Convolution has some nice properties which are summarized in the following proposition.

Proposition 8. *Let $f, g, h : \mathbb{Z}_N \to \mathbb{C}$. Then*

- *Convolution is commutative: $f * g = g * f$*
- *Convolution is associative: $(f * g) * h = f * (g * h)$*
- *Fourier transformation distributes over convolution: $\widehat{f * g} = \widehat{f} \cdot \widehat{g}$*

Again we omit the proofs of these facts, but it is worthwhile for the interested reader to work out these identities in detail for a full understanding of the mechanisms at work.

We conclude the section by connecting the notions of Fourier transformation and convolution with the language of probability theory.

Proposition 9. *Let $f_1, \dots, f_n : \mathbb{Z}_N \to \mathbb{C}$, where we consider f_i as a random variable on the probability space $\mathbb{Z}_N$ with a uniform probability distribution. Then*

- *Expectation identity:*

$$\mathbb{E}(f_1 * \cdots * f_n) = \prod_{k=1}^{n} \widehat{f_k}(0)$$

- *Variance identity:*

$$\mathrm{Var}(f_1 * \cdots * f_n) = \sum_{\xi \neq 0} \prod_{k=1}^{n} \left| \widehat{f_k}(\xi) \right|^2$$

Proof. Because of the distributive property of the Fourier transform over convolutions, it is sufficient to prove each property for a single function $f : \mathbb{Z}_N \to \mathbb{C}$. For the expectation identity we have

$$\mathbb{E}(f) = \sum_{x \in \mathbb{Z}_N} \frac{1}{N} f(x) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f(x) \exp\left(0 \cdot \frac{2\pi i}{N}\right) = \widehat{f}(0).$$

For the variance identity we see that

$$\begin{aligned} \mathrm{Var}(f) &= \mathbb{E}\Big(|f - \mathbb{E}(f)|^2\Big) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \left| f(x) - \widehat{f}(0) \right|^2 \\ &= \sum_{\xi \in \mathbb{Z}_N} \left| \left(\widehat{f - \widehat{f}(0)} \right)(\xi) \right|^2 = \sum_{\xi \in \mathbb{Z}_N} \left| \widehat{f}(\xi) - \widehat{f}(0) \cdot I(\xi = 0) \right|^2 = \sum_{\xi \neq 0} \left| \widehat{f}(\xi) \right|^2. \end{aligned}$$

□

3 Fourier Pseudorandomness

And finally we can introduce the topic of interest. To motivate the definition of Fourier pseudorandomness, take another look at the Fourier inversion formula:

$$f(x) = \sum_{\xi \in \mathbb{Z}_N} \widehat{f}(\xi) \exp\left(x\xi \cdot \frac{2\pi i}{N}\right)$$

The exponential part of the sum may be considered more transparently as a periodic trigonometric function of x. Indeed, in light of Euler's formula that $e^{i\theta} = \cos(\theta) + i \sin(\theta)$, the Fourier inversion formula can be rewritten as

$$f(x) = \sum_{\xi \in \mathbb{Z}_N} \widehat{f}(\xi) \cos\left(2\pi \cdot \frac{x\xi}{N}\right) + i \sum_{\xi \in \mathbb{Z}_N} \widehat{f}(\xi) \sin\left(2\pi \cdot \frac{x\xi}{N}\right).$$

This illustrates that the Fourier coefficients $\widehat{f}(\xi)$ can be viewed as weights of a decomposition of f into periodic parts of different periods. Notice in particular that the zero-th Fourier coefficient $\widehat{f}(0)$ plays a special role, in that the periodic part corresponding to that coefficient is constant throughout the sum. With these thoughts in mind, we make a definition.

Definition 10. *For $f : \mathbb{Z}_N \to \mathbb{C}$, we define the Fourier bias $\|f\|_{\mathrm{u}}$ of f by*

$$\|f\|_{\mathrm{u}} := \max_{\xi \neq 0} \left| \widehat{f}(\xi) \right|. \tag{10}$$

A subset $A \subseteq \mathbb{Z}_N$ is called ϵ-uniform, or ϵ-pseudorandom if $\|A\|_{\mathrm{u}} \leq \epsilon$. The property of having low Fourier bias is called linear uniformity.

In terms of our motivation, this definition in some sense states that no periodic part of the Fourier decomposition dominates. We will explore a few of the formal properties of linearly uniform sets later.

For now, we note some basic properties of Fourier bias on subsets of $\mathbb{Z}_N$.

Proposition 11. *Let $A \subseteq \mathbb{Z}_N$ with $|A| = \delta N$. Then*

- *Symmetry properties:*
$$\|A\|_{\mathrm{u}} = \|\mathbb{Z}_N \setminus A\|_{\mathrm{u}} = \|-A\|_{\mathrm{u}} = \|A + x\|_{\mathrm{u}}$$
- *Triangle inequality [1, Ex. 4.3.3]: If $B \subseteq \mathbb{Z}_N$ with $A \cap B = \emptyset$, then*
$$|\|A\|_{\mathrm{u}} - \|B\|_{\mathrm{u}}| \leq \|A \cup B\|_{\mathrm{u}} \leq \|A\|_{\mathrm{u}} + \|B\|_{\mathrm{u}}$$
- *Uniform boundedness:*
$$\|A\|_{\mathrm{u}} \leq \min(\delta, 1-\delta) \leq \widehat{A}(0)$$
- *Variance bound:*
$$\|A\|_{\mathrm{u}}^2 \geq \frac{\delta(1-\delta)}{N-1}$$

Proof. The first three properties we leave as an exercise to the reader. For the last, notice that we have

$$\mathbb{E}(A) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} A(x) = \frac{|A|}{N} = \delta.$$

Using this, the variance identity, and the definition of Fourier bias, we see

$$\begin{aligned} \delta(1-\delta) &= \mathbb{E}(A) - \mathbb{E}(A)^2 = \mathbb{E}(|A|) - |\mathbb{E}(A)|^2 = \mathrm{Var}(A) \\ &= \sum_{\xi \neq 0} \left| \widehat{A}(\xi) \right|^2 \leq (N-1) \max_{\xi \neq 0} \left| \widehat{A}(\xi) \right|^2 = (N-1)\|A\|_{\mathrm{u}}^2, \end{aligned}$$

and this proves the inequality. □

4 Properties of Pseudorandom Sets

It turns out that Fourier pseudorandom sets have a number of nice properties that one might expect from "random" sets. The first such property which we discuss concerns the density of intersections of a set with translates of another set. Indeed, if one or both of the sets in question are pseudorandom, then this density does not differ too much from the expected density, on average:

Theorem 12. *Let $A, B \subseteq \mathbb{Z}_N$ with $|A| = \delta_A N$ and $|B| = \delta_B N$. Then*

$$\mathbb{E}\left(\left|\frac{|A \cap (B+x)|}{N} - \delta_A\delta_B\right|\right) \leq \sqrt{\|A\|_{\mathrm{u}}\|B\|_{\mathrm{u}} \cdot \sqrt{\delta_A\delta_B}} \leq \sqrt{\|A\|_{\mathrm{u}}\|B\|_{\mathrm{u}}},$$

and in particular,

$$\#\left\{x \in \mathbb{Z}_N : \left|\frac{|A \cap (B+x)|}{N} - \delta_A\delta_B\right| \geq \sqrt[4]{\|A\|_{\mathrm{u}}\|B\|_{\mathrm{u}}}\right\} \leq N \cdot \sqrt[4]{\|A\|_{\mathrm{u}}\|B\|_{\mathrm{u}}}.$$

Proof. We use without proof the simple identities

$$\frac{|A \cap B|}{N} = \sum_{\xi \in \mathbb{Z}_N} \widehat{A}(\xi)\overline{\widehat{B}(\xi)},$$

and

$$\widehat{A+x}(\xi) = \widehat{A}(\xi) \exp\left(-x\xi \cdot \frac{2\pi i}{N}\right).$$

We have

$$\left|\frac{|A \cap (B+x)|}{N} - \delta_A\delta_B\right| = \left|\sum_{\xi \in \mathbb{Z}_N} \widehat{A}(\xi)\overline{\widehat{B}(\xi)} \exp\left(x\xi \cdot \frac{2\pi i}{N}\right) - \delta_A\delta_B\right|$$
$$= \left|\sum_{\xi \neq 0} \widehat{A}(\xi)\overline{\widehat{B}(\xi)} \exp\left(x\xi \cdot \frac{2\pi i}{N}\right)\right| = \sigma(x) \cdot \sum_{\xi \neq 0} \widehat{A}(\xi)\overline{\widehat{B}(\xi)} \exp\left(x\xi \cdot \frac{2\pi i}{N}\right).$$

Here $\sigma(x)$ is a function with $|\sigma(x)| = 1$ for each x. From this we see that

$$\mathbb{E}\left(\left|\frac{|A \cap (B+x)|}{N} - \delta_A\delta_B\right|\right) = \frac{1}{N}\sum_{x \in \mathbb{Z}_N} \sigma(x) \sum_{\xi \neq 0} \widehat{A}(\xi)\overline{\widehat{B}(\xi)} \exp\left(x\xi \cdot \frac{2\pi i}{N}\right)$$
$$= \sum_{\xi \neq 0} \widehat{A}(\xi)\overline{\widehat{B}(\xi)} \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \sigma(x) \exp\left(-x(-\xi) \cdot \frac{2\pi i}{N}\right) = \sum_{\xi \neq 0} \widehat{A}(\xi)\overline{\widehat{B}(\xi)}\widehat{\sigma}(-\xi)$$
$$\leq \sum_{\xi \neq 0} \left|\widehat{A}(\xi)\right| \left|\widehat{B}(\xi)\right| |\widehat{\sigma}(-\xi)| \leq \left(\sum_{\xi \neq 0} \left|\widehat{A}(\xi)\right|^4\right)^{1/4} \cdot \left(\sum_{\xi \neq 0} \left|\widehat{B}(\xi)\right|^4\right)^{1/4} \cdot \left(\sum_{\xi \neq 0} |\widehat{\sigma}(\xi)|^2\right)^{1/2},$$

where the last inequality can be seen as a double application of the Cauchy-Schwarz inequality. We can estimate the three terms of this product individually to obtain the desired bound. Using Plancherel's identity, we have

$$\sum_{\xi \neq 0} \left|\widehat{A}(\xi)\right|^4 \leq \|A\|_{\mathrm{u}}^2 \cdot \sum_{\xi \in \mathbb{Z}_N} \left|\widehat{A}(\xi)\right|^2 = \|A\|_{\mathrm{u}}^2 \cdot \frac{1}{N} \sum_{x \in \mathbb{Z}_N} |A(x)|^2 = \|A\|_{\mathrm{u}}^2 \cdot \delta_A,$$

and an equivalent bound holds for the second term as well. Even more simply for the third term, we see

$$\sum_{\xi\neq 0} |\widehat{\sigma}(\xi)|^2 \leq \sum_{\xi\in\mathbb{Z}_N} |\widehat{\sigma}(\xi)|^2 = \frac{1}{N} \sum_{\xi\in\mathbb{Z}_N} |\sigma(\xi)|^2 = 1,$$

and the bound follows.

To see the second bound, simply apply Markov's inequality (1) with $\lambda = \|A\|_{\mathrm{u}}^{1/4} \|B\|_{\mathrm{u}}^{1/4}$.

□

The second property of pseudorandom sets which we consider concerns the propensity of a pseudorandom set to contain roots of a linear polynomial in several variables. If $P(x_1, \dots, x_n)$ is a linear polynomial in n variables, then one might expect to find $|A|^{n-1}\delta$ roots of P in a "random" set A of density δ. To see this, notice that given certain restrictions on N, we may find for each choice of $x_1, \dots, x_{n-1}$ from A a unique x_n such that $(x_1, \dots, x_n)$ is a root of P. Thus in a random set, we could expect a proportion of approximately δ of the $|A|^{n-1}$ values of x_n to lie within A, giving $|A|^{n-1}\delta$ roots. Indeed, we will see that a pseudorandom set contains approximately the expected number of roots of a given polynomial.

Theorem 13. *Let $n \geq 2$, and let $P(x_1, \dots, x_n) := c_1x_1 + \dots + c_nx_n$ be a fixed linear polynomial in n variables. For $A \subseteq \mathbb{Z}_N$ with $|A| = \delta N$, denote the number of solutions to $P(x) = 0$ contained in A by*

$$\mathcal{N}_P(A) := \#\left\{(x_1, \dots, x_n) \in A^n : P(x_1, \dots, x_n) = 0\right\}.$$

Then for N with greatest common divisor $(c_1c_2\cdots c_n, N)$ equal to 1,

$$\left|\frac{\mathcal{N}_P(A)}{N^{n-1}} - \delta^n\right| \leq \|A\|_{\mathrm{u}}^{n-2} \cdot \delta.$$

Proof. Recall that an orthogonality condition exists for sums over roots of unity. Indeed, for $x \in \mathbb{Z}$,

$$\frac{1}{N} \sum_{\xi\in\mathbb{Z}_N} \exp\left(x\xi \cdot \frac{2\pi i}{N}\right) = I(x \equiv 0 \mod N).$$

Making use of this, we have

$$\begin{aligned}\mathcal{N}_P(A) &= \sum_{x_1,\dots,x_n\in A} I(P(x_1, \dots, x_n) \equiv 0 \mod N)\\ &= \sum_{x_1,\dots,x_n\in\mathbb{Z}_N} \prod_{k=1}^{n} A(x_k) \left(\frac{1}{N} \sum_{\xi\in\mathbb{Z}_N} \exp\left(-(c_1x_1 + \dots + c_nx_n)\xi \cdot \frac{2\pi i}{N}\right)\right)\\ &= N^{n-1} \sum_{\xi\in\mathbb{Z}_N} \prod_{k=1}^{n} \frac{1}{N} \sum_{x_k\in\mathbb{Z}_N} A(x_k) \exp\left(-x_kc_k\xi \cdot \frac{2\pi i}{N}\right) = N^{n-1} \sum_{\xi\in\mathbb{Z}_N} \prod_{k=1}^{n} \widehat{A}(c_k\xi),\end{aligned}$$

and further,

$$\begin{aligned}\left|\frac{\mathcal{N}_P(A)}{N^{n-1}} - \delta^n\right| &= \left|\sum_{\xi\in\mathbb{Z}_N} \prod_{k=1}^{n} \widehat{A}(c_k\xi) - \delta^n\right| = \left|\sum_{\xi\neq 0} \prod_{k=1}^{n} \widehat{A}(c_k\xi)\right|\\ &\leq \sum_{\xi\neq 0} \prod_{k=1}^{n} \left|\widehat{A}(c_k\xi)\right| \leq \prod_{k=1}^{n} \left(\sum_{\xi\neq 0} \left|\widehat{A}(c_k\xi)\right|^n\right)^{1/n}.\end{aligned}$$

This last inequality is an application of the generalized Hölder's inequality. Notice that since $(c_1 c_2 \cdots c_n, N) = 1$, we have that $(c_k, N) = 1$ for each $k = 1, \ldots, N$. In particular, this means that as ξ traverses a complete set of non-zero congruence classes, $c_k \xi$ does as well. Thus we have

$$\begin{aligned}\left|\frac{\mathcal{N}_P(A)}{N^{n-1}} - \delta^n\right| &\leq \prod_{k=1}^{n}\left(\sum_{\xi\neq 0}\left|\widehat{A}(c_k\xi)\right|^n\right)^{1/n} = \prod_{k=1}^{n}\left(\sum_{\xi\neq 0}\left|\widehat{A}(\xi)\right|^n\right)^{1/n} \\ &\leq \prod_{k=1}^{n}\left(\|A\|_{\mathrm{u}}^{n-2}\cdot\sum_{\xi\neq 0}\left|\widehat{A}(\xi)\right|^2\right)^{1/n} \leq \prod_{k=1}^{n}\left(\|A\|_{\mathrm{u}}^{n-2}\cdot\sum_{\xi\in\mathbb{Z}_N}\left|\widehat{A}(\xi)\right|^2\right)^{1/n} \\ &= \prod_{k=1}^{n}\left(\|A\|_{\mathrm{u}}^{n-2}\cdot\frac{1}{N}\sum_{x\in\mathbb{Z}_N}|A(x)|^2\right)^{1/n} = \prod_{k=1}^{n}\left(\|A\|_{\mathrm{u}}^{n-2}\cdot\frac{|A|}{N}\right)^{1/n} = \|A\|_{\mathrm{u}}^{n-2}\cdot\delta.\end{aligned}$$

□

5 Existence of Pseudorandom Sets

Now that we have proven fairly non-trivial properties about linearly uniform sets of numbers, it is natural to ask whether such sets actually exist. Certainly it is not unthinkable that such sets are a rarity or an impossibility, and that the preceding exposition is, while valid, vacuous. Fortunately, it is possible to give explicit constructions of pseudorandom sets, and in fact, one can formally show that pseudorandom sets are in a sense the norm rather than the exception.

First we note some sets which are trivially pseudorandom. Clearly if $A = \emptyset, \mathbb{Z}_N$ then $\|A\|_{\mathrm{u}} = 0$, and it fairly easy to show that these choices of A are the only ones with this property. Further, we may use the uniform boundedness property of Proposition 11 to see that any set with density near 0 or 1 necessarily has small linear bias.

So a more interesting question to consider is whether we can find pseudorandom sets of non-trivial density. We first discuss a construction which produces a family of pseudorandom sets of density approximately 1/2 which has members with arbitrarily small Fourier bias. The construction of these sets is straightforward, but the proof that they are pseudorandom relies on some facts from elementary number theory, most of which are proven in a first course on the subject.

Proposition 14. *Let p be an odd prime, and let $A_p \subseteq \mathbb{Z}_p$ be the set of all (non-zero) quadratic residues of $\mathbb{Z}_p$, defined by*

$$A_p := \left\{x^2 : x \in \mathbb{Z}_p^*\right\}.$$

Then A_p has density $(p-1)/2p$, and

$$\|A_p\|_{\mathrm{u}} = \begin{cases} \frac{\sqrt{p}+1}{2p}, & p \equiv 1 \pmod 4 \\ \frac{\sqrt{p+1}}{2p}, & p \equiv 3 \pmod 4 \end{cases}.$$

In particular, $\lim_{p\to\infty}\|A_p\|_{\mathrm{u}} = 0$.

Proof. There are as many quadratic residues as non-residues in $\mathbb{Z}_p^*$, so A_p has exactly $\left|\mathbb{Z}_p^*\right|/2 = (p-1)/2$ elements, which gives a density of $(p-1)/2p$ in $\mathbb{Z}_p$.

Now notice that we can formulate the indicator function of A_p in a convenient manner in terms of the Legendre symbol, mainly,

$$A_p(x) = \frac{\left(\frac{x}{p}\right) + 1}{2} - \frac{I(x \equiv 0)}{2}.$$

We calculate for non-zero ξ:

$$\begin{aligned}\widehat{A_p}(\xi) &= \frac{1}{p}\sum_{x\in\mathbb{Z}_p} A_p(x)\exp\left(-x\xi\cdot\frac{2\pi i}{p}\right) = \frac{1}{2p}\sum_{x\in\mathbb{Z}_p}\left(\left(\frac{x}{p}\right)+1\right)\exp\left(-x\xi\cdot\frac{2\pi i}{p}\right) - \frac{1}{2p}\\ &= \frac{1}{2p}\sum_{x\in\mathbb{Z}_p}\left(\frac{x}{p}\right)\exp\left(-x\xi\cdot\frac{2\pi i}{p}\right) - \frac{1}{2p} = \left(\frac{-\xi^{-1}}{p}\right)\cdot\frac{1}{2p}\sum_{x\in\mathbb{Z}_p}\left(\frac{x}{p}\right)\exp\left(x\cdot\frac{2\pi i}{p}\right) - \frac{1}{2p}.\end{aligned}$$

The sum of Legendre symbols is what is known as a quadratic Gauss sum, and takes on values $\sqrt{p}$ for $p \equiv 1 \pmod 4$ and $i\sqrt{p}$ for $p \equiv 3 \pmod 4$. So in the case of $p \equiv 1$ we have

$$\widehat{A_p}(\xi) = \frac{\pm\sqrt{p} - 1}{2p},$$

and when $p \equiv 3$,

$$\widehat{A_p}(\xi) = \frac{\pm i\sqrt{p} - 1}{2p},$$

where the signs depend on whether $-\xi^{-1}$ is a quadratic residue. The respective values of $\|A_p\|_{\mathrm{u}}$ follow by calculating the magnitudes of these quantities. □

It is worth noting that in the case of $p \equiv 3 \pmod 4$, A_p is actually as uniform as is possible—the variance bound from Proposition 11 gives us that

$$\|A_p\|_{\mathrm{u}}^2 \geq \frac{\delta(1-\delta)}{p-1} = \frac{\frac{p-1}{2p}\cdot\frac{p+1}{2p}}{p-1} = \frac{p+1}{4p^2},$$

and indeed we have shown that for $p \equiv 3$, the Fourier bias achieves this value.

So we see that we can construct uniform sets with density around 1/2. Next we describe a similar construction based on lecture notes by Stanford professor K. Soundararajan which gives uniform sets of arbitrary non-trivial density.

Proposition 15 (Soundararajan [2]). *Fix $0 < \delta < 1$ a desired density. Let p be prime with $p \equiv 3 \pmod 4$, and let $A_p \subseteq \mathbb{Z}_p$ be defined by*

$$A_p := \left\{x \in \mathbb{Z}_p : x^2 \equiv y \pmod p \text{ with } |y| \leq \delta p/2\right\}.$$

Then A_p has density $(2\lfloor\delta p/2\rfloor + 1)/p$, and

$$\|A_p\|_{\mathrm{u}} = O_\delta\left(\frac{\log p}{\sqrt{p}}\right).$$

Proof. Since $p \equiv 3 \pmod 4$, -1 is not a quadratic residue, and we know that x is a quadratic residue if and only if $-x$ is not. Thus if there are k quadratic residues in $[1, \lfloor\delta p/2\rfloor]$, there are exactly $\lfloor\delta p/2\rfloor - k$ residues in $[-\lfloor\delta p/2\rfloor, -1]$, and this implies that $|A_p|/p = (2\lfloor\delta p/2\rfloor + 1)/p$ because every non-zero quadratic residue has exactly two square-roots in $\mathbb{Z}_p$.

Now to prove the estimate on the uniformity, we calculate for $\xi \neq 0$,

$$\begin{aligned}\widehat{A_p}(\xi) &= \frac{1}{p}\sum_{x\in\mathbb{Z}_p} A_p(x)\exp\left(-x\xi\cdot\frac{2\pi i}{p}\right)\\ &= \frac{1}{p}\sum_{x\in\mathbb{Z}_p}\left(\frac{1}{p}\sum_{|b|\le\delta p/2}\sum_{r\in\mathbb{Z}_p}\exp\left(r(x^2-b)\cdot\frac{2\pi i}{p}\right)\right)\exp\left(-x\xi\cdot\frac{2\pi i}{p}\right)\\ &= \frac{1}{p^2}\sum_{r\in\mathbb{Z}_p}\left(\sum_{x\in\mathbb{Z}_p}\exp\left((rx^2-\xi x)\cdot\frac{2\pi i}{p}\right)\right)\left(\sum_{|b|\le\delta p/2}\exp\left(-br\cdot\frac{2\pi i}{p}\right)\right).\end{aligned}$$

The sum over x is called a generalized Gauss sum and has modulus bounded by $\sqrt{p}$. We further bound the modulus of the sum over b by $\min(\delta p+1, 1/(2\|r/p\|))$, where $\|r/p\|$ denotes the distance from r/p to the nearest integer. The bound of $\delta p+1$ follows from simple application of the triangle inequality. To get the second bound, we first need a short lemma.

Lemma 16. *For $\theta \in [-\pi,\pi]$,*

$$|1-\exp(i\theta)| \ge \frac{2}{\pi}\cdot|\theta|, \quad \theta\in[-\pi,\pi].$$

Proof. By symmetry, it is sufficient to show that this holds for $\theta\in[0,\pi]$. We calculate an equivalent statement by noting

$$|1-\exp(i\theta)| = \sqrt{(1-\cos(\theta))^2+\sin(\theta)^2} = \sqrt{2}\cdot\sqrt{1-\cos(\theta)},$$

and seeing that

$$|1-\exp(i\theta)| = \frac{2}{\pi}\cdot\theta \iff 1-\cos(\theta)-\frac{2}{\pi^2}\cdot\theta^2 = 0.$$

The inequality we wish to show is actually equality at the endpoints $\theta = 0,\pi$, and it also holds at $\theta=\pi/2$. Thus it suffices to show that the endpoints are the only points of equality in the domain $[0,\pi]$. This follows if we can show that $1-\cos(\theta)-2\theta^2/\pi^2$ has zero derivative at at most a single point in $(0,\pi)$. Indeed, we have

$$\begin{aligned}\frac{d}{d\theta}\left(1-\cos(\theta)-\frac{2}{\pi^2}\cdot\theta^2\right) &= \sin(\theta)-\frac{4}{\pi^2}\cdot\theta = 0\\ &\iff \sin(\theta) = \frac{4}{\pi^2}\cdot\theta.\end{aligned}$$

sin is concave in the domain $[0,\pi]$ and strictly concave in $(0,\pi)$, and so any linear function can intersect it at at most two points in $[0,\pi]$. In particular, $4\theta/\pi^2$ intersects $\sin(\theta)$ at $\theta=0$, so it cannot have more than a single further intersection in $(0,\pi)$. The claim follows. □

To prove the second bound, we make use of the geometric sum formula and the inequality above to find

$$\begin{aligned}&\left|\sum_{|b|\le\delta p/2}\exp\left(-br\cdot\frac{2\pi i}{p}\right)\right| = \left|\sum_{b=0}^{2\lfloor\delta p/2\rfloor+1}\exp\left(r\cdot\frac{2\pi i}{p}\right)^b\right| = \left|\frac{1-\exp\left(r\cdot\frac{2\pi i}{p}\right)^{2\lfloor\delta p/2\rfloor+2}}{1-\exp\left(r\cdot\frac{2\pi i}{p}\right)}\right|\\ &\le \frac{2}{\left|1-\exp\left(r\cdot\frac{2\pi i}{p}\right)\right|} = \frac{2}{|1-\exp(\|r/p\|\cdot 2\pi i)|} \le \frac{2}{2/\pi\cdot|2\pi\|r/p\||} = \frac{1}{2\|r/p\|}.\end{aligned}$$

So we have

$$\begin{aligned}\left|\widehat{A_p}(\xi)\right| &= \left|\frac{1}{p^2}\sum_{r\in\mathbb{Z}_p}\left(\sum_{x\in\mathbb{Z}_p}\exp\left((rx^2-\xi x)\cdot\frac{2\pi i}{p}\right)\right)\left(\sum_{|b|\leq\delta p/2}\exp\left(-br\cdot\frac{2\pi i}{p}\right)\right)\right| \\ &\leq \frac{\sqrt{p}}{p^2}\sum_{r\in\mathbb{Z}_p}\min\left(\delta p+1,\frac{1}{2\|r/p\|}\right) = \frac{\sqrt{p}}{p^2}\left(\delta p + 2\cdot\sum_{r=1}^{(p-1)/2}\min\left(\delta p+1,\frac{p}{2r}\right)\right).\end{aligned}$$

In particular, we notice that $\delta p + 1 \leq p/2r$ only when $r \leq \lfloor p/2(\delta p+1)\rfloor \leq k := \lfloor 1/2\delta\rfloor$, and so we have

$$\begin{aligned}\left|\widehat{A_p}(\xi)\right| &\leq \frac{\sqrt{p}}{p^2}\left((k+1)\delta p + 2\cdot\sum_{r=1}^{(p-1)/2}\frac{p}{2r}\right) = \frac{\sqrt{p}}{p^2}\left(p\log\left(\frac{p-1}{2}\right)+O_\delta(p)\right) \\ &= \frac{\log(p-1)}{\sqrt{p}} + O_\delta\left(\frac{1}{\sqrt{p}}\right) = O_\delta\left(\frac{\log p}{\sqrt{p}}\right),\end{aligned}$$

and this proves the estimate for $\|A_p\|_{\mathrm{u}}$. □

This proposition shows that the example of quadratic residues is not just an isolated instance of pseudorandomness—in fact we can generate pseudorandom sets of arbitrary density. Indeed, as a rule of thumb we find that sets whose primary structure is non-linear tend to be linearly uniform.

We now show that the linear structure indicated by high Fourier bias is probabilistically uncommon. To complete the proof, we make use of a Corollary of Chernoff's inequality (Theorem 4) which extends the use of the inequality to complex-valued random variables.

Corollary 17 (Chernoff's Inequality II). *Let $X := X_1 + \cdots + X_n$ be a random variable, where $X_1, \ldots, X_n$ are simple, complex-valued random variables which are jointly independent and have $|X_i - \mathbb{E}(X_i)| \leq 1$ for all i. If we denote by $\sigma := \mathrm{Var}(X)^{1/2}$ the standard deviation of X, then for any real $\lambda > 0$, we have*

$$\Pr(|X-\mathbb{E}(X)|\geq\lambda\sigma)\leq 4\max\left(e^{-\lambda^2/8}, e^{-\lambda\sigma/2\sqrt{2}}\right).$$

The proof, which uses the real-valued version of Chernoff's inequality by splitting the random variable into real and imaginary parts, we leave as an exercise. Using this variant, we now show that on average, a random subset of $\mathbb{Z}_N$ is relatively pseudorandom:

Proposition 18. *Let $0 < \tau \leq 1$, and let A be a random subset of $\mathbb{Z}_N$ defined by letting the events $a \in A$ be independent with probability τ. Then for any $\lambda > 0$ we have*

$$\Pr(\|A\|_{\mathrm{u}}\geq\lambda\sigma)\leq 4(N-1)\max\left(e^{-\lambda^2/8}, e^{-\lambda\sigma/2\sqrt{2}}\right),$$

where $\sigma^2 := \tau(1-\tau)/N$.

This result is based on Lemma 4.16 from Tao and Vu's book *Additive Combinatorics* [1].

Proof. We apply Corrolary 17 to the complex-valued random variable $\widehat{A}(\xi)$ for fixed non-zero ξ. Notice that because the $A(x)$ are jointly independent random variables (with range $\{0,1\}$), $\widehat{A}(\xi)$ is a random variable of the proper form to apply Chernoff's inequality II:

$$\widehat{A}(\xi) = \frac{1}{N}\sum_{x\in\mathbb{Z}_N}A(x)\exp\left(-x\xi\cdot\frac{2\pi i}{N}\right) = \sum_{x\in\mathbb{Z}_N}\frac{\exp\left(-x\xi\cdot\frac{2\pi i}{N}\right)}{N}\cdot A(x) =: \sum_{x\in\mathbb{Z}_N}A_x(\xi).$$

Clearly if the $A(x)$ are jointly independent, then for arbitrary constants c_x the variables $c_x A(x)$ are also jointly independent. Further, since the coefficients $c_x := \exp(-x\xi \cdot \frac{2\pi i}{N})/N$ have modulus at most 1, the range of $A_x(\xi)$ has diameter bounded by 1 for each x, and hence $|A_x(\xi) - \mathbb{E}(A_x(\xi))| \leq 1$. Thus $\widehat{A}(\xi) = A_1(\xi) + \cdots + A_N(\xi)$ satisfies the conditions necessary to apply Chernoff's inequality. We calculate

$$\begin{aligned}
\mathbb{E}\Big(\widehat{A}(\xi)\Big) &= \sum_{A \subseteq \mathbb{Z}_N} \tau^{|A|}(1-\tau)^{N-|A|} \frac{1}{N} \sum_{x \in \mathbb{Z}_N} A(x) \exp\left(-x\xi \cdot \frac{2\pi i}{N}\right) \\
&= \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \exp\left(-x\xi \cdot \frac{2\pi i}{N}\right) \sum_{A \subseteq \mathbb{Z}_N} A(x) \tau^{|A|}(1-\tau)^{N-|A|} \\
&= \frac{1}{N} \sum_{x \in \mathbb{Z}_N} \exp\left(-x\xi \cdot \frac{2\pi i}{N}\right) \tau \cdot \sum_{B \subseteq \mathbb{Z}_N \setminus \{x\}} \tau^{|B|}(1-\tau)^{(N-1)-|B|} \\
&= \frac{\tau}{N} \sum_{x \in \mathbb{Z}_N} \exp\left(-x\xi \cdot \frac{2\pi i}{N}\right) = 0,
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}\Big(\widehat{A}(\xi)\Big) &= \mathbb{E}\left(\Big|\widehat{A}(\xi)\Big|^2\right) = \sum_{A \subseteq \mathbb{Z}_N} \tau^{|A|}(1-\tau)^{N-|A|} \left| \frac{1}{N} \sum_{x \in \mathbb{Z}_N} A(x) \exp\left(-x\xi \cdot \frac{2\pi i}{N}\right) \right|^2 \\
&= \sum_{A \subseteq \mathbb{Z}_N} \tau^{|A|}(1-\tau)^{N-|A|} \frac{1}{N^2} \sum_{x \in \mathbb{Z}_N} A(x) \exp\left(-x\xi \cdot \frac{2\pi i}{N}\right) \sum_{y \in \mathbb{Z}_N} A(y) \exp\left(y\xi \cdot \frac{2\pi i}{N}\right) \\
&= \frac{1}{N^2} \sum_{x,y \in \mathbb{Z}_N} \exp\left((y-x)\xi \cdot \frac{2\pi i}{N}\right) \sum_{A \subseteq \mathbb{Z}_N} A(x)A(y)\tau^{|A|}(1-\tau)^{N-|A|}.
\end{aligned}$$

For $x = y$, we have

$$\begin{aligned}
&\sum_{A \subseteq \mathbb{Z}_N} A(x)A(y)\tau^{|A|}(1-\tau)^{N-|A|} \\
&\qquad = \tau \cdot \sum_{B \subseteq \mathbb{Z}_N \setminus \{x\}} \tau^{|B|}(1-\tau)^{(N-1)-|B|} = \tau,
\end{aligned}$$

and for $x \neq y$ we have

$$\begin{aligned}
&\sum_{A \subseteq \mathbb{Z}_N} A(x)A(y)\tau^{|A|}(1-\tau)^{N-|A|} \\
&\qquad = \tau^2 \cdot \sum_{B \subseteq \mathbb{Z}_N \setminus \{x,y\}} \tau^{|B|}(1-\tau)^{(N-2)-|B|} = \tau^2.
\end{aligned}$$

So we find

$$\begin{aligned}
\operatorname{Var}\left(\widehat{A}(\xi)\right) &= \frac{1}{N^2}\left(\tau\sum_{x=y\in\mathbb{Z}_N}\exp\left((y-x)\xi\cdot\frac{2\pi i}{N}\right)+\tau^2\sum_{x\neq y\in\mathbb{Z}_N}\exp\left((y-x)\xi\cdot\frac{2\pi i}{N}\right)\right)\\
&= \frac{1}{N^2}\left(N\tau+\tau^2\sum_{k\neq 0}\sum_{y-x\equiv k}\exp\left((y-x)\xi\cdot\frac{2\pi i}{N}\right)\right)\\
&= \frac{1}{N^2}\left(N\tau+N\tau^2\sum_{k\neq 0}\exp\left((y-x)\xi\cdot\frac{2\pi i}{N}\right)\right)=\frac{1}{N^2}\left(N\tau-N\tau^2\right)=\frac{\tau(1-\tau)}{N},
\end{aligned}$$

and by Chernoff's inequality II we have

$$\Pr\left(\left|\widehat{A}(\xi)\right|\geq\lambda\sigma\right)\leq 4\max\left(e^{-\lambda^2/8},e^{-\lambda\sigma/2\sqrt{2}}\right)$$

for any $\lambda>0$. In particular, if we apply this inequality over all non-zero ξ, we find

$$\begin{aligned}
\Pr(\|A\|_{\mathrm{u}}\geq\lambda\sigma) &= \Pr\left(\left|\widehat{A}(\xi)\right|\geq\lambda\sigma\text{ for some }\xi\neq 0\right)\\
&\leq\sum_{\xi\neq 0}\Pr\left(\left|\widehat{A}(\xi)\right|\geq\lambda\sigma\right)\leq 4(N-1)\max\left(e^{-\lambda^2/8},e^{-\lambda\sigma/2\sqrt{2}}\right).
\end{aligned}$$

□

6 Limitations of Linear Uniformity

Based on the properties we've explored, it is apparent that the notion of Fourier pseudorandomness mirrors the intuitive notion of "randomness" in a variety of ways. However, it is important to point out that Fourier pseudorandomness fails to capture the idea of randomness when looking at local structures which are not in some sense linear. As a demonstration of this, we return to the example of Soundararajan from Proposition 15.

Proposition 19 (Soundararajan [2])**.** *Fix $0<\delta<1$ and let A_p be defined as in Proposition 15. Then for large p, there exist at least $\delta^3p^2/(2\cdot 7^3)$ 4-term arithmetic progressions in A_p.*

Proof. Define B_p as A_p but with density $\delta/7$. As per Proposition 15, the Fourier bias of B_p goes to 0 as $\log p/\sqrt{p}$, and so by Theorem 13, the number of 3-term arithmetic progressions (the number of roots $\mathcal{N}_P$ of the polynomial $P(x_1,x_2,x_3)=x_1-2x_2+x_3$) approaches the expected number, mainly

$$\begin{aligned}
\left|\frac{\mathcal{N}_P(B_p)}{p^2}-(\delta/7)^3\right| &\leq\|A\|_{\mathrm{u}}\cdot\delta\leq C_\delta\cdot\frac{\log p}{\sqrt{p}}\\
\implies\left|\mathcal{N}_P(B_p)-\delta^3p^2/7^3\right| &\leq C_\delta\cdot p^{3/2}\log p.
\end{aligned}$$

In particular, for large enough p, the difference $C_\delta\cdot p^{3/2}\log p$ is small relative to $\delta^3p^2/7^3$, and this implies that B_p has at least $\delta^3p^2/(2\cdot 7^3)$ 3-term arithmetic progressions for all p which are sufficiently large.

Conveniently enough, it turns out that for each 3-term arithmetic progression $a, a+d, a+2d$ which lies in $B_p\subseteq A_p$, the fourth term $a+3d$ also lies in A_p. To see this, note the identity

$$a^2-3(a+d)^2+3(a+2d)^2-(a+3d)^2=0.$$

Since $a, a+d$ and $a+2d$ are assumed to be in B_p, their squares are congruent mod p to numbers with magnitude at most $\delta p/14$. Thus in particular, $(a+3d)^2$ is congruent mod p to some b with $|b| \leq 7(\delta p/14) = \delta p/2$, and hence is contained in A_p.

Thus each 3-term arithmetic progress in B_p represents a distinct 4-term arithmetic progression in A_p, and we see that there are at least $\delta^3 p^2/(2 \cdot 7^3)$ such progressions for large p. □

For a random subset in $\mathbb{Z}_p$, one would expect to find about $\delta^4 p^2$ arithmetic progressions of length 4, so this result implies that for small δ, we find significantly more 4-term arithmetic progressions in A_p than expected. So we see that Fourier pseudorandomness does not control the quantity of 4-term arithmetic progressions in a set in the same way that it does the quantity of 3-term arithmetic progressions or roots of arbitrary linear polynomials.

We conclude by noting that there exists an alternate formulation of pseudorandomness called Gowers uniformity which avoids the limitations noted here. This notion is similar in spirit to Fourier pseudorandomness, making use of the Gowers uniformity norms in place of Fourier bias, and it has the advantage of controlling the density of structures defined by higher degree polynomials. However, balancing these advantages is the fact that Gowers uniformity is significantly more technical and difficult to work with than Fourier pseudorandomness, so both notions have a useful place in the general theory. A notable application of Gowers uniformity is in a proof of Szemerédi's theorem on the existence of arbitrarily long arithmetic progressions in a subset of the integers with positive density.

Acknowledgement. The author would like to acknowledge the support and assistance of Neil Lyall, Mariah Hamel and Alex Rice throughout the 2010 University of Georgia REU in mathematics, which led to the preparation of this manuscript.

References

[1] Tao, Terence and Vu, Van H. *Additive Combinatorics*. Cambridge University Press, New York, 2007. Print.

[2] Soundararajan, K. *Additive Combinatorics: Winter 2007*. Web. 20 July 2010.