Continuous Ramsey Theory and Sidon Sets

Greg Martin (gerg@math.ubc.ca)
and
Kevin O'Bryant (kobryant@math.ucsd.edu)

September 28, 2002

Abstract

A symmetric subset of the reals is one that remains invariant under some reflection $x \mapsto c - x$. Given $0 < \varepsilon \le 1$, there exists a real number $\Delta(\varepsilon)$ with the following property: if $0 \le \delta < \Delta(\varepsilon)$, then every subset of $[0,1]$ with measure ε contains a symmetric subset with measure δ, while if $\delta > \Delta(\varepsilon)$, then there exists a subset of $[0,1]$ with measure ε that does not contain a symmetric subset with measure δ. In this paper we establish upper and lower bounds for $\Delta(\varepsilon)$ of the same order of magnitude: for example, we prove that $\Delta(\varepsilon) = 2\varepsilon - 1$ for $\frac{11}{16} \le \varepsilon \le 1$ and that $0.59\varepsilon^2 < \Delta(\varepsilon) < 0.8\varepsilon^2$ for $0 < \varepsilon \le \frac{11}{16}$.

This continuous problem is intimately connected with a corresponding discrete problem. A set S of integers is called a $B^*[g]$ set if for any given m there are at most g ordered pairs $(s_1, s_2) \in S \times S$ with $s_1 + s_2 = m$; in the case $g = 2$, these are better known as Sidon sets. We also establish upper and lower bounds of the same order of magnitude for the maximal possible size of a $B^*[g]$ set contained in $\{1, \dots, n\}$, which we denote by $R(g,n)$. For example, we prove that $R(g,n) < 1.31\sqrt{gn}$ for all $n \ge g \ge 2$, while $R(g,n) > 0.79\sqrt{gn}$ for sufficiently large integers g and n.

These two problems are so interconnected that both continuous and discrete tools can be applied to each problem with surprising effectiveness. The harmonic analysis methods and inequalities among various L^p norms we use to derive lower bounds for $\Delta(\varepsilon)$ also provide uniform upper bounds for $R(g,n)$, while the techniques from combinatorial and probabilistic number theory that we employ to obtain constructions of large $B^*[g]$ sets yield strong upper bounds for $\Delta(\varepsilon)$.

Contents

1 Introduction

A set $C \subseteq \mathbb{R}$ is *symmetric* if there exists a number c (the center of C) such that $c + x \in C$ if and only if $c - x \in C$. Given a set $A \subseteq [0, 1)$ of positive measure, is there necessarily a symmetric subset $C \subseteq A$ of positive measure? The answer turns out to be "yes", and the main topic of this paper is to determine how large, in terms of the Lebesgue measure of A, one may take the symmetric set C. In other words, for each $\varepsilon > 0$ we are interested in

$$\Delta(\varepsilon) := \sup \left\{ \delta \colon \begin{array}{c} \text{every measurable subset of } [0,1) \text{ with measure } \varepsilon \\ \text{contains a symmetric subset with measure } \delta \end{array} \right\}. \tag{1}$$

It is not immediately obvious that $\Delta(\varepsilon) > 0$.

We have dubbed this sort of question "continuous Ramsey theory", and we direct the reader to Section 2 for problems with a similar flavor; some of these have appeared in the literature and some are given here for the first time.

We determine a lower bound for $\Delta(\varepsilon)$ using tools and ideas from harmonic analysis, nonstandard analysis, and the theory of wavelets. We also construct sets without large symmetric subsets using results from the theory of finite fields and probabilistic number theory. These two lines of attack complement each other, and our bounds yield new results in additive number theory as well.

Consider first the analogous discrete problem: given a set $A \subseteq \{1, 2, \dots, n\}$, how large is the largest symmetric subset of A? There are $\sim \frac{1}{2}|A|^2$ pairs of distinct elements of A (where $|A|$ denotes the cardinality of A), and each pair (a, b) has a center $\frac{a+b}{2}$ which is among the $\sim 2n$ values $\{\frac{3}{2}, 2, \frac{5}{2}, 3, \dots, \frac{2n-1}{2}\}$. By the pigeonhole principle, there is some c that is the center of at least $\sim \frac{1}{2}|A|^2/(2n)$ pairs of elements of A. The union of those pairs is a symmetric set, i.e., there is a symmetric set $C \subseteq A$ with

$$\frac{|C|}{n} \gtrsim \frac{2}{n} \frac{\frac{1}{2}|A|^2}{2n} = \frac{1}{2} \left(\frac{|A|}{n} \right)^2.$$

In other words, the density of C is roughly at least one-half the square of the density of A.

A *Sidon set* is a set A of integers with the property that distinct pairs of elements have distinct sums: if $a, b, c, d \in A$ and $a + b = c + d$, then $\{a, b\} = \{c, d\}$. This is equivalent to asserting that A has no symmetric subsets with more than 2 elements. It is known [Sin38] that there is a Sidon set A contained in $\{1, 2, \dots, n\}$ with $|A| \sim \sqrt{n}$. Thus, if $C \subseteq A$ is a symmetric set, then

$$\frac{|C|}{n} \leq \frac{2}{n} \sim 2 \left(\frac{|A|}{n} \right)^2.$$

In other words, the density of C is roughly at most twice the square of the density of A.

For the discrete version of the problem, at least, we see that one can guarantee a "quadratically large" symmetric subset, and that one cannot do better in general.

1.1 Continuous Results

Let λ denote Lebesgue measure on $\mathbb{R}$. We are led by analogy with the discrete problem to guess that every subset $A \subseteq [0, 1)$ has a symmetric subset with measure $\frac{1}{2}\lambda(A)^2$, and that there are subsets $A \subseteq [0,1)$ that do not have symmetric subsets with measure larger than $2\lambda(A)^2$. That is, we are led to guess that $\frac{1}{2}\varepsilon^2 \le \Delta(\varepsilon) \le 2\varepsilon^2$.

We find the following equivalent definition of $\Delta(\varepsilon)$ easier to work with than the definition given in Eq. (1): if we define

$$D(A) := \sup\{\lambda(C) \colon \quad C \subseteq A,\, C \text{ is symmetric}\}, \tag{2}$$

then

$$\Delta(\varepsilon) := \inf\{D(A) \colon \quad A \subseteq [0,1),\, \lambda(A) = \varepsilon\}. \tag{3}$$

Write $A(x)$ for the indicator function of a set A, and define the sumset $A + A := \{a_1 + a_2 \colon a_i \in A\}$. We notice first that the maximal symmetric subset of A with center c is $A \cap (2c - A)$, where $2c - A = \{2c - a \colon a \in A\}$; this intersection has measure $\int A(x)A(2c-x)\,dx$, which can be written as the value of the convolution $A * A(2c)$. This means that $D(A)$ simply equals the supremum norm $\|A * A\|_\infty$. Since $\operatorname{supp}(A * A)$, the support of the function $A * A$, equals $A + A$ (up to a set of measure zero) and is thus contained in $[0,2)$, we see that

$$\|A * A\|_\infty \ge \frac{\|A * A\|_1}{\lambda(\operatorname{supp}(A * A))} = \frac{\lambda(A)^2}{\lambda(A + A)} \ge \frac{1}{2}\lambda(A)^2,$$

and hence $\Delta(\varepsilon) \ge \frac{1}{2}\varepsilon^2$ (as we had guessed from the analogy with the discrete case). This lower bound, which we shall call the trivial lower bound on $\Delta(\varepsilon)$, is not so far from the best we can derive! In fact, the bulk of this paper is devoted to improving the constant in this lower bound from $\frac{1}{2}$ to 0.591389. Moreover, we are able to establish a complementary upper bound for $\Delta(\varepsilon)$ in a manner that we shall discuss in the next section.

In addition to a heavy reliance on Fourier analysis, we make use of wavelets, albeit only in a rather naive manner. Although much of our argument (and indeed the initial problem of bounding $\Delta(\varepsilon)$) was initially motivated by nonstandard analysis, we do not make direct use of it here.

Figure 1 shows the precise upper and lower bounds we obtain for $\Delta(\varepsilon)/\varepsilon^2$ as functions of ε, and Theorem 1.1 gives the highlights:

Theorem 1.1. *We have:*

i. $\Delta(\varepsilon) = 2\varepsilon - 1$ *for* $\frac{11}{16} \le \varepsilon \le 1$, *and* $\Delta(\varepsilon) \ge 2\varepsilon - 1$ *for* $\frac{1}{2} \le \varepsilon \le \frac{11}{16}$;

ii. $\Delta(\varepsilon) \ge 0.591389\varepsilon^2$ *for all* $0 < \varepsilon \le 1$;

iii. $\Delta(\varepsilon) \ge 0.5546\varepsilon^2 + 0.088079\varepsilon^3$ *for all* $0 < \varepsilon \le 1$;

iv. $\Delta(\varepsilon) \le \frac{96}{121}\varepsilon^2 < 0.7934\varepsilon^2$ *for* $0 < \varepsilon \le \frac{11}{16}$;

v. $\Delta(\varepsilon) \le \frac{\pi\varepsilon^2}{(1+\sqrt{1-\varepsilon})^2} = \frac{\pi}{4}\varepsilon^2 + O(\varepsilon^3)$ *for all* $0 < \varepsilon \le 1$.

Note that $\frac{\pi}{4} < 0.7854$. The upper bound in part (v) of the theorem is superior to the one in part (iv) in the range $0 < \varepsilon < \frac{11}{96}(8\sqrt{6\pi} - 11\pi) \doteq 0.0201$. The five parts of the theorem are proved separately in Proposition 7.1, Proposition 4.12, Proposition 4.16, Corollary 7.4, and Proposition 7.5, respectively.

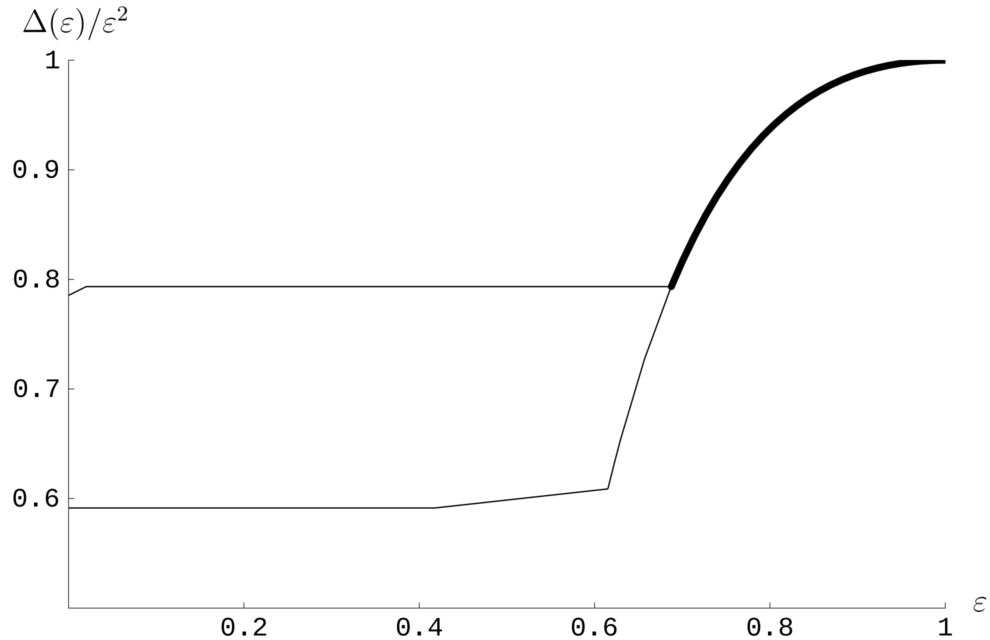

Figure 1: Upper and Lower Bounds for $\Delta(\varepsilon)/\varepsilon^2$

Thus we have established that ε^2 is the correct order of magnitude for the function Δ, and we have improved upon both the constants $\frac{1}{2}$ and 2 which appeared in our heuristic. In the following subsection, we discuss how we might improve the corresponding constants for the discrete problem as well.

1.2 Connecting the Continuous to the Discrete

It turns out that the upper bounds given in Theorem 1.1(iv)–(v) are derived from number-theoretic considerations. A set S of integers is called a $B_2[g]$ set if for any given m there are at most $2g$ ordered pairs $(s_1, s_2) \in S \times S$ with $s_1 + s_2 = m$ (equivalently, if the coefficients of $\left(\sum_{n \in S} z^n\right)^2$ are bounded by $2g$). Sidon used $B_2[1]$ sets, which are the Sidon sets mentioned earlier, as a tool in his study of Fourier series. It is perhaps fitting that we now use Fourier analysis as a tool in our study of $B_2[g]$ sets.

Many constructions of $B_2[g]$ sets have appeared in recent years. Our constructions of sets of reals without large symmetric subsets are based on known constructions of $B_2[g]$ sets. This geometric aspect of $B_2[g]$ sets has led us to generalize and optimize these constructions further and to give improved upper bounds on the density of $B_2[g]$ sets contained in $\{1, 2, \ldots, n\}$.

Given a $B_2[g]$ set $S \subseteq \{1, 2, \ldots, n\}$ and any integer m, the union of all pairs $(s_1, s_2) \in S \times S$ with $s_1 + s_2 = m$ is the largest symmetric subset of S with center $\frac{m}{2}$, and all symmetric subsets of S arise in this way. From the definition of a $B_2[g]$ set, we see that S contains no symmetric subset with more than $2g$ integers. In light of this, it is not surprising (though it certainly requires proof—see Proposition 5.1) that the largest symmetric subset of the set of real numbers

$$A(S) := \bigcup_{s\in S} \Big[\frac{s-1}{n}, \frac{s}{n}\Big) \subseteq [0,1) \tag{4}$$

has measure at most $\frac{2g}{n}$. In other words, $\Delta\big(\frac{|S|}{n}\big) \le D(A(S)) \le \frac{2g}{n}$.

For technical reasons, it is more convenient for us to speak of $B^*[g]$ sets rather than $B_2[g]$ sets. A set S of integers is called a $B^*[g]$ set if for any given m there are at most g ordered pairs $(s_1, s_2) \in S \times S$ with $s_1 + s_2 = m$. Note that the definitions of $B_2[g]$ sets and $B^*[2g]$ sets coincide, but we shall also consider $B^*[g]$ sets with g odd. We introduce the function

$$R(g,n) := \max\{|S|\colon \quad S \subseteq \{1,2,\ldots,n\},\, S \text{ is a } B^*[g] \text{ set}\}. \tag{5}$$

Since there are $|S|^2$ sums of pairs from S, while there are fewer than $2n$ possible sums each of which can be realized at most g times, we immediately deduce the upper bound $R(g,n) \le \sqrt{2gn}$, which we shall call the trivial upper bound for $R(g,n)$.

The construction of $A(S)$ indicated above thus gives the bound $\Delta\big(\frac{R(g,n)}{n}\big) \le \frac{g}{n}$, and we can use this inequality to give an *upper* bound on $R(g,n)$. The trivial lower bound on $\Delta(\varepsilon)$ gives $\frac{1}{2}\big(\frac{R(g,n)}{n}\big)^2 \le \Delta\big(\frac{R(g,n)}{n}\big) \le \frac{g}{n}$, whence $R(g,n) \le \sqrt{2gn}$. Thus, the trivial lower bound on $\Delta(\varepsilon)$ implies the trivial upper bound on $R(g,n)$, and any improvement we can make on the trivial lower bound for $\Delta(\varepsilon)$ will immediately provide stronger upper bounds for $R(g,n)$. In the same way, we shall use lower bounds on $R(g,n)$ to derive upper bounds on $\Delta(\varepsilon)$. All this is made rigorous in Sections 5 and 7.

In fact much more than $\Delta\big(\frac{R(g,n)}{n}\big) \le \frac{g}{n}$ is true. Proposition 5.3 below states that

$$\Delta(\varepsilon) = \inf\{\tfrac{g}{n}\colon n \ge g \ge 1,\ R(g,n) \ge n\varepsilon\}.$$

Thus a sufficient understanding of the dependence of $R(g,n)$ on g and n would completely solve our continuous problem. Unfortunately, this understanding is still somewhat lacking.

1.3 Discrete Results

The true size of Sidon sets is essentially known. Erdős and Turán [ET41] exploited the fact that the pairwise differences $s_1 - s_2$ from a Sidon set are distinct to establish the improved upper bound $R(2,n) \lesssim \sqrt{n}$. (By the notation $f(n) \lesssim g(n)$, we mean that $\limsup_{n\to\infty} \frac{f(n)}{g(n)} \le 1$.) Ruzsa [Ruz93] has observed that the Erdős/Turán argument can be modified to give also $R(3,n) \lesssim \sqrt{n}$. A construction of Singer [Sin38] (see Proposition 6.1(iii) below) yields the complementary lower bounds $R(3,n) \ge R(2,n) \gtrsim \sqrt{n}$.

Present knowledge of $R(g,n)$ for $g > 3$ is much less impressive. Constructions of large $B^*[g]$ sets have so far yielded only moderate lower bounds on $R(g,n)$. The trivial upper bound on $R(g,n)$ has been improved for general g, but only quite recently. In this paper, we present the strongest upper bound to date on $R(g,n)$ for all $g \ge 21$ as well as for all odd $g \ge 5$, and we also improve or match the best known lower bounds on $R(g,n)$.

The seminal paper of Cilleruelo, Ruzsa, and Trujillo [CRT] gave the first upper bound for $R(g,n)$ that was nontrivial for infinitely many g, namely $R(2g,n) \lesssim 1.3181\sqrt{2gn}$, and Green [Gre01] improved this to $R(2g,n) \lesssim 1.3038\sqrt{2gn}$. Green also proved

$$R(2g,n) \lesssim \sqrt{\tfrac{7}{4}(2g-1)n}, \tag{6}$$

which is stronger than our results for even integers $g \le 20$. (In both these papers, the results were stated in terms of the function $F_2(g,n)$, a notational difference only as $F_2(g,n) = R(2g,n)$ for all g and n.)

It seems likely that $\lim_{n\to\infty} \frac{R(g,n)}{\sqrt{gn}}$ exists for every g, but this has been proved only for $g = 2$ and $g = 3$. We define

$$\begin{aligned}
\underline{\rho}(g) &:= \liminf_{n\to\infty} \frac{R(g,n)}{\sqrt{gn}}, \\
\overline{\rho}(g) &:= \limsup_{n\to\infty} \frac{R(g,n)}{\sqrt{gn}}, \\
\rho(g) &:= \lim_{n\to\infty} \frac{R(g,n)}{\sqrt{gn}}.
\end{aligned}$$

Thus $\rho(2) = \frac{1}{\sqrt{2}}$, $\rho(3) = \frac{1}{\sqrt{3}}$, and $\rho(g)$ is not known to be well-defined for $g \ge 4$. Green's result [Gre01] is equivalent to $\overline{\rho}(g) \le \sqrt{\frac{7}{4}}\sqrt{1-\frac{1}{g}}$.

The following theorem gives the bounds on $R(g,n)$ that we prove in this work. Note that for even $g \le 20$, Green's bound (6) is superior.

Theorem 1.2. *$R(g,n) \le 1.30036\sqrt{gn}$ for all g and $R(g,n) \le 1.31925\sqrt{(g-1)n} + \frac{1}{3}$ if g is odd. Further, we have the following upper bounds for $\overline{\rho}(g)^2$:*

$$\overline{\rho}(g)^2 \le \begin{cases} 1.74043 - \frac{1.00483}{g}, & g \le 8 \text{ and even;} \\ 1.58337 - \frac{0.026335}{g} + \sqrt{0.011572 - \frac{0.083397}{g} + \frac{0.00069356}{g^2}}, & g \ge 10 \text{ and even;} \\ 1.74043 - \frac{4.75492}{g}, & g \le 23 \text{ and odd;} \\ 1.58337 - \frac{0.071949}{g} + \sqrt{0.011572 - \frac{0.22784}{g} + \frac{0.0051768}{g^2}}, & g \ge 25 \text{ and odd.} \end{cases}$$

We comment that our result is an improvement in two aspects: we improve the numerical constants in the bound on $\overline{\rho}(g)$ and we replace "$\lesssim$" with "$\le$" in the bounds on $R(g,n)$. These minor improvements aside, we believe that our geometrical approach is easier to follow and promises future improvements. Accordingly, throughout this paper we discuss the quality of the results, what could possibly be improved and what is best possible. In addition, to the best of our knowledge, Theorem 1.2 gives the first general upper bound for $R(2g-1,n)$ that improves upon the trivial inequality $R(2g-1,n) \le R(2g,n)$. Theorem 1.2 is proved in Corollary 5.2 and Proposition 5.4.

We turn now to lower bounds for $R(g,n)$. Constructions of "$B^*[2] \pmod{n}$" sets have been given by Singer [Sin38], Bose [Bos42], and Ruzsa [Ruz93]. These constructions were extended to $B^*[g]$ sets in [CRT]. In Propositions 6.1 and 6.4 below, we generalize the first three constructions and further optimize the extension given in [CRT].

Theorem 1.3 presents our improved lower bounds for $R(g,n)$, stated in terms of the function $\underline{\rho}(g)$.

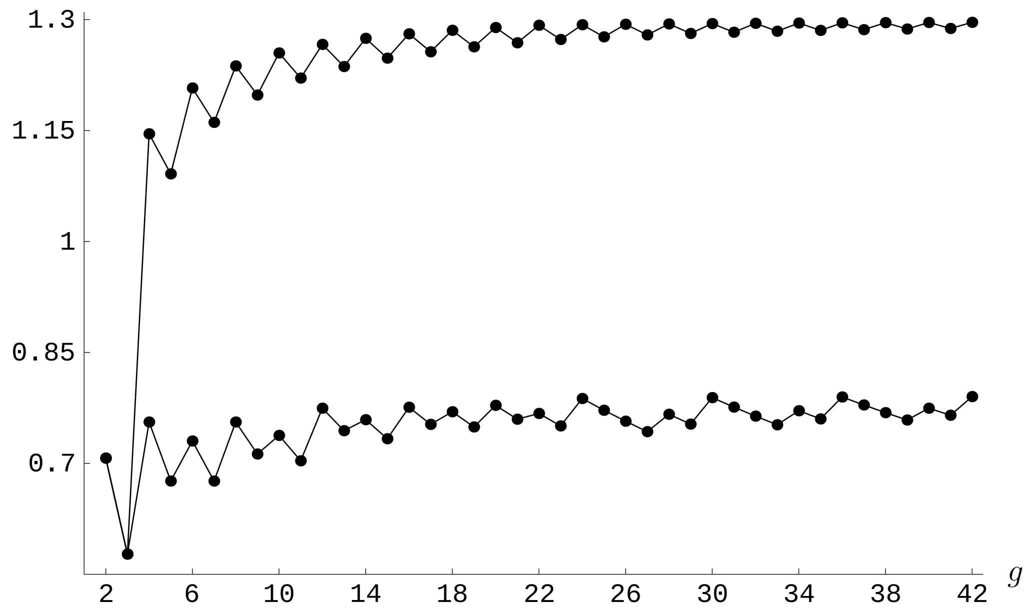

Figure 2: Lower bounds on $\underline{\rho}(g)$ and upper bounds on $\overline{\rho}(g)$

Theorem 1.3. *We have*

$$\begin{array}{rlrl}
\underline{\rho}(4) \geq \frac{2}{\sqrt{7}} & > 0.755, & \underline{\rho}(14) \geq \frac{11}{\sqrt{210}} & > 0.759, \\
\underline{\rho}(6) \geq \frac{2\sqrt{2}}{\sqrt{15}} & > 0.730, & \underline{\rho}(16) \geq \frac{17}{4\sqrt{30}} & > 0.775, \\
\underline{\rho}(8) \geq \frac{2}{\sqrt{7}} & > 0.755, & \underline{\rho}(18) \geq \frac{4}{3\sqrt{3}} & > 0.769, \\
\underline{\rho}(10) \geq \frac{7}{3\sqrt{10}} & > 0.737, & \underline{\rho}(20) \geq \frac{2\sqrt{5}}{\sqrt{33}} & > 0.778, \\
\underline{\rho}(12) \geq \frac{\sqrt{3}}{\sqrt{5}} & > 0.774, & \underline{\rho}(22) \geq \frac{18}{5\sqrt{22}} & > 0.767,
\end{array}$$

and for any $g \geq 12$,

$$\underline{\rho}(2g) \geq \frac{g + 2\lfloor g/3 \rfloor + \lfloor g/6 \rfloor}{\sqrt{6g^2 - 2g\lfloor g/3 \rfloor + 2g}}.$$

In particular, for any $\delta > 0$ we have $R(g,n) > (\frac{11}{8\sqrt{3}} - \delta)\sqrt{gn}$ if both g and $\frac{n}{g}$ are sufficiently large in terms of δ.

We note that $\frac{11}{8\sqrt{3}} > 0.7938$. The lower bound for $\underline{\rho}(4)$ reproduces a result of Habsieger and Plagne [HP], while the lower bounds for $\underline{\rho}(6)$ and $\underline{\rho}(10)$ reproduce results of [CRT]; for other even g, our lower bounds are new. These lower bounds on $\underline{\rho}$, together with the strongest known upper bounds on $\overline{\rho}$ including those derived from Theorem 1.2, are plotted for $2 \leq g \leq 42$ in Figure 2.

To the authors' knowledge, nothing more is known about lower bounds for $R(2g+1,n)$ for general n and g than the obvious inequality $R(2g,n) \leq R(2g+1,n)$. In particular,

$$\sqrt{\frac{2g}{2g+1}}\,\frac{R(2g,n)}{\sqrt{2gn}} \leq \frac{R(2g+1,n)}{\sqrt{(2g+1)n}},$$

which implies that

$$\underline{\rho}(2g+1) \geq \sqrt{\frac{2g}{2g+1}}\,\underline{\rho}(2g).$$

It seems likely that this inequality is actually an equality (i.e., when g is fixed, the quotient $R(2g+1,n)/R(2g,n)$ should tend to 1 as n tends to infinity), but this is known only for $g = 1$ [Ruz93].

Theorem 1.3 contains our best results for fixed g, but we can obtain a better constant in the lower bound if we allow g to grow slowly with n:

Theorem 1.4. *For any $\delta > 0$, we have $R(g,n) > \left(\frac{2}{\sqrt{\pi}} - \delta\right)\sqrt{gn}$ if both $\frac{g}{\log n}$ and $\frac{n}{g}$ are sufficiently large in terms of δ.*

We note that $2/\sqrt{\pi} > 1.128$. Theorems 1.3 and 1.4 are both important for obtaining the upper bounds on $\Delta(\varepsilon)$ given in Theorem 1.1. Theorem 1.3 is proved in Section 6.6 through explicit constructions, while Theorem 1.4 is proved in Section 6.5 by a probabilistic argument.

As mentioned above, in Section 2 we describe several other problems in continuous Ramsey theory, along with their discrete analogues. In Section 3 we list some easy-to-derive properties of the function $\Delta(\varepsilon)$, while in Section 4 we describe how we establish our strongest lower bounds for $\Delta(\varepsilon)$. Section 4 is the most involved part of this paper, employing many techniques from harmonic analysis, most notably Fourier series. The connection between the continuous and discrete problems measured by $\Delta(\varepsilon)$ and $R(g,n)$, respectively, is given in Section 5, and our upper bounds for $R(g,n)$ are derived therein. In that section, we also establish in Theorem 5.7 that large $B^*[g]$ sets must have many pairwise sums that repeat at least αg times for suitable α, and in Theorem 5.9 we show a stronger upper bound for $B^*[g]$ sets that are uniformly distributed in $\{1,\ldots,n\}$ (which we conjecture is typical for the largest possible $B^*[g]$ sets). Section 6 is devoted to lower bounds for $R(g,n)$, which rely upon improved constructions of $B^*[g]$ sets and their "mod n" counterparts. In Section 7 we apply the results of Section 6 to bound $\Delta(\varepsilon)$ from above. Finally, in Section 8 we collect together several open questions and problems relating to our methods.

2 Some Problems in Continuous Ramsey Theory

A "coloring Ramsey theorem" has the form:

> Given a sufficiently large set of mathematical objects colored with a finite number of colors, there is a highly structured monochromatic subset.

The prototypical example is Ramsey's Theorem itself: However one colors the edges of the complete graph K_n with r colors, there is a monochromatic complete subgraph on t vertices, provided that n is sufficiently large in terms of r and t. Another example is van der Waerden's Theorem: However one colors the integers $\{1, 2, \ldots, n\}$ with r colors, there is a monochromatic arithmetic progression with t terms, provided that n is sufficiently large in terms of r and t.

In many cases, the coloring aspect of a Ramsey-type theorem is a ruse and one may prove a stronger statement with the form:

> Given a sufficiently large set R of mathematical objects, any large subset of R contains a highly structured subset.

Such a result is called a "density Ramsey theorem." For example, van der Waerden's Theorem is a special case of the density Ramsey theorem of Szemerédi: Every subset of $\{1, 2, \ldots, n\}$ with cardinality at least $n\delta$ contains a t-term arithmetic progression, provided that n is sufficiently large in terms of δ and t.

Ramsey theory is the study of such theorems on different types of structures. By "continuous Ramsey theory" we refer to Ramsey-type problems on continuous measure spaces. In particular, this thesis is concerned with a density-Ramsey problem on the structure $[0, 1) \subseteq \mathbb{R}$ with Lebesgue measure. The type of substructure we focus on are symmetric subsets.

There is frequently a connection between problems concerning $[0, 1)$ with Lebesgue measure and problems concerning $\{1, 2, \ldots, n\}$ with a uniform probability measure. We attempt to make this connection more explicit in our descriptions of problems in the remainder of this section.

2.1 The Correlation Problem

Erdős asked for upper and lower bounds on $M_n(\frac{1}{2})$, where

$$M_n(\alpha) := \inf_{\substack{|S|=\lfloor \alpha n \rfloor \\ T=\{1,2,\ldots,n\}\setminus S}} \sup_{c\in\mathbb{Z}} \frac{|(S+c)\cap T|}{n}.$$

In essence, Erdős asked for a formalization of the observation: a large subset of $\{1, 2, \ldots, n\}$ can always be translated so as to have a large intersection with its complement.

S. Świerczkowski [Świ58] showed that $\lim_{n\to\infty} M_n(\alpha) = M(\alpha)$, where

$$M(\alpha) := \inf_{\substack{\lambda(S)=\alpha \\ T=[0,1)\setminus S}} \sup_{c\in\mathbb{R}} \lambda((S+c)\cap T).$$

Working in the continuous setting, Świerczkowski showed that

$$M(\alpha) > \frac{2-\sqrt{4-10\alpha(1-\alpha)}}{5},$$

which improved on the asymptotic bounds then known for $M_n(\alpha)$.

It is known (see [Guy94, problem C17]) that $0.178 < M(\frac{1}{2}) < 0.2$, but the precise value remains unknown.

2.2 The Convolution Problem

Banakh, Verbitsky, and Vorobets [BVV00] considered the coloring version of the problem considered in the present paper: given a measurable finite coloring of $[0,1)$, how large a monochromatic symmetric set is guaranteed? They claim that if 2 colors are used, then there is necessarily a monochromatic symmetric subset with measure $\frac{1}{4+\sqrt{6}} \approx 0.155$, and if $r > 2$ colors are used then there is a monochromatic subset with measure $\frac{1/2}{r^2}$. Unfortunately, the proofs of several crucial lemmas of this interesting paper appear only in earlier Russian-language articles, and we have been unable to verify their proofs. If r colors are used, then there is a monochromatic set with measure at least $\frac{1}{r}$. In Section 4.6, we show that *any* subset of $[0,1)$ with measure at least $\frac{1}{r}$ contains a symmetric subset with measure $\frac{0.591389}{r^2}$, and in Section 4.8, we show further that any subset with measure at least $\frac{1}{2}$ has a symmetric subset with measure 0.1496. This paper is concerned with the density version of their coloring problem.

We call this the convolution problem because, if f is the indicator function of $E \subseteq \mathbb{R}$, then $f * f(c) := \int_{\mathbb{R}} f(x)f(c-x)\,dx$ is the measure of the maximal symmetric subset of E with center $c/2$. Therefore the Banakh–Verbitsky–Vorobets problem is equivalent to finding the infimum of the possible values of

$$\max\{\|f*f\|_\infty, \|(1-f)*(1-f)\|_\infty\}$$

as f ranges over all indicator functions of measurable subsets of $[0,1)$. We exploit this connection to Fourier analysis to give lower bounds for the function $\Delta(\varepsilon)$ defined in Eq. (3).

We note that the discrete versions of these problems are closely related to the continuous versions. Specifically, define $MS(n,r)$ to be the maximum of integers M such that for any coloring of $\{1,2,\ldots,n\}$ with r colors, there is a monochromatic symmetric subset with cardinality M. Also, define $MS([0,1],r)$ to the supremum of real μ such that for any measurable coloring of $[0,1]$ with r colors, there is a monochromatic subset with measure μ. It is shown in [BVV00] that $MS([0,1],r) = \lim_{n\to\infty} MS(n,r)/n$. Proposition 5.3 below gives the analogous connection between $\Delta(\varepsilon)$ and the function $R(g,n)$ defined in Eq. (5).

2.3 The Linear Equation Problem

Following Chung, Erdős, and Graham [CEG], for any matrix A, we say that $S \subseteq \{1,2,\ldots,n\}$ is A-hitting if for every vector $\bar{x} = (x_1,x_2,\ldots,x_s)$ with $x_i \in \{1,2,\ldots,n\}$ and $\bar{x}A = \bar{0}$, either there is some $x_i \in S$ or $x_1 = x_2 = \cdots = x_s$ (this second condition is included to avoid certain trivialities). Set δ_n to be the minimum density of an A-hitting set, i.e.,

$$\delta_n := \min\left\{\tfrac{|S|}{n}: \quad S \subseteq \{1,2,\ldots,n\},\, S \text{ is } A\text{-hitting}\right\}.$$

If, for example, $A = (1,-2,1)$, then δ_n is the minimum density of a subset of $\{1,2,\ldots,n\}$ that intersects every 3-term arithmetic progression in $\{1,2,\ldots,n\}$.

We also say that $S \subseteq [0,1]$ is A-hitting if for every vector $\bar{x} = (x_1, x_2, \ldots, x_s)$ with $x_i \in [0,1]$ and $\bar{x}A = \bar{0}$, either $x_1 = x_2 = \cdots = x_s$ or there is some $x_i \in S$. Set δ to be the infimum of the measures of A-hitting sets, i.e.,

$$\delta := \inf\left\{\lambda(S): \quad S \subseteq [0,1],\, S \text{ is } A\text{-hitting}\right\}.$$

Given our experience with the above correlation and convolution problems, one might guess that $\liminf_n \delta_n = \delta$. In [CEG], the values of both δ and $\liminf_n \delta_n$ are derived for several matrices A, and indeed equality seems to be the typical case. Surprisingly, however, there are non-trivial examples where this is not the case. For

$$A = \begin{pmatrix} 2 & 3 \\ -1 & 0 \\ 0 & -1 \end{pmatrix},$$

for example, it is shown in [CEG] that $\delta = \frac{1}{5}$ whereas $0.199 < \liminf_n \delta_n < 0.1997$.

3 Easy Bounds for $\Delta(\varepsilon)$

We now turn our attention to the investigation of the function $\Delta(\varepsilon)$ defined in Eq. (3). In this section we establish several simple lemmas describing basic properties of Δ.

Lemma 3.1. $\Delta(\varepsilon) \geq 2\varepsilon - 1$ *for all* $0 \leq \varepsilon \leq 1$.

Proof. For $A \subseteq [0,1)$, the centrally symmetric set $A \cap (1-A)$ has measure equal to

$$\lambda(A) + \lambda(1-A) - \lambda(A \cup (1-A)) = 2\lambda(A) - \lambda(A \cup (1-A)) \geq 2\lambda(A) - 1.$$

Therefore $D(A) \geq 2\lambda(A) - 1$ from the definition (2) of the function D. Taking the infimum over all subsets A of $[0,1)$ with measure ε, this becomes $\Delta(\varepsilon) \geq 2\varepsilon - 1$ as claimed. ■

While this bound may seem obvious, it is in many situations the state of the art. As we show in Proposition 7.1 below, $\Delta(\varepsilon)$ actually equals $2\varepsilon - 1$ for $\frac{11}{16} \leq \varepsilon \leq 1$; and $\Delta(\varepsilon) \geq 2\varepsilon - 1$ is the best lower bound of which we are aware in the range $0.61522 \leq \varepsilon < \frac{11}{16} \doteq 0.6875$.

One is tempted to try to sharpen the bound $\Delta(\varepsilon) \geq 2\varepsilon - 1$ by considering the symmetric subsets with center $1/3$, $1/2$, or $2/3$, for example, instead of merely $1/2$. Unfortunately, it can be shown that given any $\varepsilon \geq \frac{1}{2}$ and any finite set $\{c_1, c_2, \ldots, c_n\}$, one can construct a sequence S_k of sets, each with measure ε, that satisfies

$$\lim_{k\to\infty} \left(\max_{1\leq i\leq n} \{\lambda\left(S_k \cap (2c_i - S_k)\right)\} \right) = 2\varepsilon - 1.$$

Thus, no improvement is possible with this sort of argument.

Lemma 3.2 (Trivial Lower Bound). $\Delta(\varepsilon) \geq \frac{1}{2}\varepsilon^2$ *for all* $0 \leq \varepsilon \leq 1$.

Proof. We repeat the argument given briefly in Section 1.1. Given a subset A of $[0,1)$ of measure ε, let $A(x)$ denote the indicator function of A, so that the integral of $A(x)$ over any interval containing $[0,1)$ equals ε. If we define $f(c) := \int_{-\infty}^{\infty} A(x)A(2c-x)\,dx$, then $f(c)$ is the measure of the largest symmetric subset of A with center c, and we seek to maximize $f(c)$. But f is clearly supported on $[0,1)$, and so

$$\begin{aligned} D(A) = \max_c f(c) \geq \int_0^1 f(c)\,dc = \int_{-\infty}^{\infty}\int_0^1 A(x)A(2c-x)\,dc\,dx \\ = \int_{-\infty}^{\infty} A(x)\left(\frac{1}{2}\int_{-x}^{2-x} A(w)\,dw\right)dx = \tfrac{1}{2}\varepsilon^2. \end{aligned}$$

Since A was an arbitrary subset of $[0,1)$ of measure ε, we have shown that $\Delta(\varepsilon) \geq \frac{1}{2}\varepsilon^2$. ■

Lemma 3.3. $\Delta(\varepsilon) \leq \frac{\Delta(x)}{x}\varepsilon$ *for all* $0 \leq \varepsilon \leq x \leq 1$. *In particular,* $\Delta(\varepsilon) \leq \varepsilon$.

Proof. If $tA := \{ta \colon a \in A\}$ is a scaled copy of a set A, then clearly $D(tA) = tD(A)$. Applying this with any set $A \subseteq [0,1)$ of measure x and with $t = \frac{\varepsilon}{x} \leq 1$, we see that $\frac{\varepsilon}{x}A$ is a subset of $[0,1)$ with measure ε, and so by the definition of Δ we have $\Delta(\varepsilon) \leq D(\frac{\varepsilon}{x}A) = \frac{\varepsilon}{x}D(A)$. Taking the infimum over all sets $A \subseteq [0,1)$ of measure x, we conclude that $\Delta(\varepsilon) \leq \frac{\varepsilon}{x}\Delta(x)$. The second assertion of the lemma is obvious, and indeed it follows from the first assertion in light of the trivial value $\Delta(1) = 1$. ■

It is obvious from the definition of Δ that $\Delta(\varepsilon)$ is an increasing function; Lemma 3.3 shows that $\frac{\Delta(\varepsilon)}{\varepsilon}$ is also an increasing function. Later in this paper (see Proposition 7.2), we will show that in fact even $\frac{\Delta(\varepsilon)}{\varepsilon^2}$ is an increasing function.

Let

$$S \diamond T := (S \setminus T) \cup (T \setminus S)$$

denote the symmetric difference of S and T. (While this operation is more commonly denoted with a triangle rather than with a diamond, we would rather avoid any potential confusion with the function Δ featured prominently in this paper.)

Lemma 3.4. *If S and T are two sets of real numbers, then $|D(S) - D(T)| \leq 2\lambda(S \diamond T)$.*

Proof. Let E be any symmetric subset of S, and let c be the center of E, so that $E = 2c - E$. Define $F = E \cap T \cap (2c - T)$, which is a symmetric subset of T with center $2c$. We can write $\lambda(F)$ using the inclusion-exclusion formula

$$\begin{aligned}\lambda(F) = \lambda(E) + \lambda(T) + \lambda(2c - T)& \\ - \lambda(E \cup T) - \lambda(E \cup (2c - T)) - \lambda(T \cup (2c - T)) + \lambda(E \cup T \cup (2c - T)).\end{aligned}$$

Rearranging terms, and noting that $T \cup (2c - T) \subseteq E \cup T \cup (2c - T)$, we see that

$$\lambda(E) - \lambda(F) \leq -\lambda(T) - \lambda(2c - T) + \lambda(E \cup T) + \lambda(E \cup (2c - T))$$

Because reflecting a set in the point c does not change its measure, this is the same as

$$\begin{aligned}\lambda(E) - \lambda(F) &\leq -\lambda(T) - \lambda(T) + \lambda(E \cup T) + \lambda(E \cup (2c - T)) \\ &= -\lambda(T) - \lambda(T) + \lambda(E \cup T) + \lambda((2c - E) \cup T) \\ &= 2\big(\lambda(E \cup T) - \lambda(T)\big) \\ &\leq 2\big(\lambda(S \cup T) - \lambda(T)\big) = 2\lambda(S \setminus T) \leq 2\lambda(S \diamond T),\end{aligned}$$

Therefore, since F is a symmetric subset of T,

$$\lambda(E) \leq \lambda(F) + 2\lambda(S \diamond T) \leq D(T) + 2\lambda(S \diamond T).$$

Taking the supremum over all symmetric subsets E of S, we conclude that $D(S) \leq D(T) + 2\lambda(S \diamond T)$. If we now exchange the roles of S and T, we see that the proof is complete. ∎

Lemma 3.5. *The function Δ satisfies the Lipschitz condition $|\Delta(x) - \Delta(y)| \leq 2|x - y|$ for all x and y in $[0, 1]$. In particular, Δ is continuous.*

Proof. Without loss of generality assume $y < x$. In light of the monotonicity $\Delta(y) \leq \Delta(x)$, it suffices to show that $\Delta(y) \geq \Delta(x) - 2(x - y)$. Let $S \subseteq [0, 1)$ have measure y. Choose any set $R \subseteq [0, 1) \setminus S$ with measure $x - y$, and set $T = S \cup R$. Then $S \diamond T = R$, and so by Lemma 3.4, $D(T) - D(S) \leq 2\lambda(R) = 2(x - y)$. Therefore $D(S) \geq D(T) - 2(x - y) \geq \Delta(x) - 2(x - y)$ by the definition of Δ. Taking the infimum over all sets $S \subseteq [0, 1)$ of measure y yields $\Delta(y) \geq \Delta(x) - 2(x - y)$ as desired. ∎

4 Lower Bounds for $\Delta(\varepsilon)$

Section 4.1 below makes explicit the connection between $\Delta(\varepsilon)$ and harmonic analysis. Section 4.2 gives a simple, but quite good, lower bound on $\Delta(\varepsilon)$. In Section 4.3, we give a more general form of the argument of Section 4.2. Using an analytic inequality established in Section 4.4, we investigate in Section 4.5 the connection between $\|f * f\|_\infty$ and the Fourier coefficients of f. In Section 4.6, we combine the results of Sections 4.2 and 4.5 to show $\Delta(\varepsilon) \geq 0.591389\varepsilon^2$. The bound given in Section 4.2 and improved in Section 4.6 depends on a kernel function with certain properties; in Section 4.7 we discuss how we chose our kernel. In Section 4.8, we use a different approach to derive a lower bound on $\Delta(\varepsilon)$ which is superior for $\frac{3}{8} < \varepsilon < \frac{5}{8}$.

4.1 Introduction and Notation

There are many ways to define the basic objects of Fourier analysis; we follow [Fol84]. Unless specifically noted otherwise, all integrals are over the circle group $\mathbb{T} := \mathbb{R}/\mathbb{Z}$; for example, L^1 denotes the class of functions f for which $\int_{\mathbb{T}} |f(x)|\,dx$ is finite. For each integer j, we define $\hat{f}(j) := \int f(x)e^{-2\pi ijx}\,dx$, so that for any function $f \in L^1$, we have $f(x) = \sum_{j=-\infty}^{\infty} \hat{f}(j)e^{2\pi ijx}$ almost everywhere. We define the convolution $f * g(c) = \int f(x)g(c-x)\,dx$, and we note that $\widehat{f * g}(j) = \hat{f}(j)\hat{g}(j)$ for every integer j; in particular, $\widehat{f * f}(j) = \hat{f}(j)^2$.

We define the usual L^p norms

$$\|f\|_p = \left(\int |f(x)|^p\,dx\right)^{1/p}$$

and

$$\|f\|_\infty = \lim_{p\to\infty} \|f\|_p = \sup\left\{y\colon \quad \lambda(\{x\colon |f(x)| > y\}) > 0\right\}.$$

With these definitions, Hölder's Inequality is valid: if p and q are conjugate exponents—that is, $\frac{1}{p} + \frac{1}{q} = 1$—then $\|fg\|_1 \leq \|f\|_p\|g\|_q$. We also note that $\|f * g\|_1 = \|f\|_1\|g\|_1$; in particular, $\|f * f\|_1 = \|f\|_1^2 = \hat{f}(0)^2$. We shall also employ the ℓ^p norms for bi-infinite sequences: if $a = \{a_j\}_{j\in\mathbb{Z}}$, then $\|a\|_p = \left(\sum_{j\in\mathbb{Z}} |a_j|^p\right)^{1/p}$ and $\|a\|_\infty = \lim_{p\to\infty} \|a\|_p = \sup_{j\in\mathbb{Z}} |a_j|$. Although we use the same notation for the L^p and ℓ^p norms, no confusion should arise, as the object inside the norm symbol will either be a function on $\mathbb{T}$ or its sequence of Fourier coefficients, respectively. With this notation, we recall Parseval's identity

$$\int f(x)g(x)\,dx = \sum \hat{f}(j)\hat{g}(-j)$$

(assuming the integral and sum both converge); in particular, if $f = g$ is real-valued (so that $\hat{f}(-j)$ is the conjugate of $\hat{f}(j)$ for all j), this becomes $\|f\|_2 = \|\hat{f}\|_2$. The Hausdorff-Young inequality, $\|\hat{f}\|_q \leq \|f\|_p$ whenever p and q are conjugate exponents with $1 \leq p \leq 2 \leq q \leq \infty$, can be thought of as a generalization of this latter version of Parseval's identity. We also require the definition

$$_m\|a\|_p = \left(\sum_{|j|\geq m} |a(j)|^p\right)^{1/p} \tag{7}$$

for any sequence $a = \{a_j\}_{j\in\mathbb{Z}}$, so that ${}_0\|a\|_p = \|a\|_p$, for example.

We recall that the Fourier coefficients of any function $f \in L^1$ satisfy the estimate $\hat{f}(j) = O\left(\frac{1}{j}\right)$; in particular, $\|\hat{f}\|_p$ is finite for all $p > 1$. Moreover, if $f \in L^1$ is continuous, then $\hat{f}(j) = O\left(\frac{1}{j^2}\right)$. We also note that for any fixed sequence $a = \{a_j\}_{j\in\mathbb{Z}}$, the ℓ^p-norm $\|a\|_p$ is a decreasing function of p. To see this, suppose that $1 \le p \le q < \infty$ and $a \in \ell^p$. Then $|a_j| \le \|a\|_p$ for all $j \in \mathbb{Z}$, whence $|a_j|^{q-p} \le \|a\|_p^{q-p}$ (since $q-p \ge 0$) and so $|a_j|^q \le \|a\|_p^{q-p}|a_j|^p$. Summing both sides over all $j \in \mathbb{Z}$ yields $\|a\|_q^q \le \|a\|_p^{q-p}\|a\|_p^p = \|a\|_p^q$, and taking qth roots gives the desired inequality $\|a\|_q \le \|a\|_p$.

Finally, we define a "pdf", short for "probability density function", to be a nonnegative function in L^2 whose L^1-norm (which is necessarily finite, since $\mathbb{T}$ is a finite measure space) equals 1. Also, we single out a special type of pdf called an "nif", short for "normalized indicator function", which is a pdf that only takes one nonzero value, that value necessarily being the reciprocal of the measure of the support of the function. (We exclude the possibility that an nif takes the value 0 almost everywhere.) Specifically, we define for each $E \subseteq \mathbb{T}$ the nif

$$f_E(x) := \begin{cases} \lambda(E)^{-1} & x \in E, \\ 0 & x \notin E. \end{cases}$$

Note that if f is a pdf, then $1 = \hat{f}(0) = \hat{f}(0)^2 = \|f\|_1^2 = \|f * f\|_1$.

We are now ready to reformulate the function $\Delta(\varepsilon)$ in terms of this notation.

Lemma 4.1. *We have*

$$\tfrac{1}{2}\varepsilon^2 \inf_g \|g * g\|_\infty \le \tfrac{1}{2}\varepsilon^2 \inf_f \|f * f\|_\infty = \Delta(\varepsilon),$$

the first infimum being taken over all pdfs g that are supported on $[-\frac{1}{4}, \frac{1}{4}]$, and the second infimum being taken over all nifs f whose support is a subset of $[-\frac{1}{4}, \frac{1}{4}]$ of measure $\frac{\varepsilon}{2}$.

Proof. The inequality is trivial, since every nif is a pdf; it remains to prove the equality.

For each measurable $A \subseteq [0, 1)$, define $E_A := \{\frac{1}{2}(a - \frac{1}{2}) \colon a \in A\} \subseteq [-\frac{1}{4}, \frac{1}{4}]$. The sets A and E_A differ only by translation and scaling, so that $\lambda(A) = 2\lambda(E_A)$ and $D(A) = 2D(E_A)$. Thus

$$\begin{aligned}
\Delta(\varepsilon) &:= \inf\{D(A) \colon A \subseteq [0,1),\, \lambda(A) = \varepsilon\} \\
&= \varepsilon^2 \inf\left\{\frac{D(A)}{\lambda(A)^2} \colon A \subseteq [0,1),\, \lambda(A) = \varepsilon\right\} \\
&= \varepsilon^2 \inf\left\{\frac{2D(E_A)}{(2\lambda(E_A))^2} \colon A \subseteq [0,1),\, \lambda(A) = \varepsilon\right\} \\
&= \frac{1}{2}\varepsilon^2 \inf\left\{\frac{D(E)}{\lambda(E)^2} \colon E \subseteq [-\tfrac{1}{4}, \tfrac{1}{4}],\, \lambda(E) = \tfrac{\varepsilon}{2}\right\}.
\end{aligned}$$

For each $E \subseteq [-\frac{1}{4}, \frac{1}{4}]$ with $\lambda(E) = \frac{\varepsilon}{2}$, the function $f_E(x)$ is an nif supported on a subset of $[-\frac{1}{4}, \frac{1}{4}]$ with measure $\frac{\varepsilon}{2}$, and it is clear that every such nif arises from some set E. Thus, it remains only to show that $\frac{D(E)}{\lambda(E)^2} = \|f_E * f_E\|_\infty$, i.e., that $D(E) = \lambda(E)^2\|f_E * f_E\|_\infty$.

Fix $E \subseteq [-\frac{1}{4}, \frac{1}{4}]$, and let $E(x)$ be the indicator function of E. Note that $f_E(x) = \lambda(E)^{-1}E(x)$. The maximal symmetric subset of E with center c is $E \cap (2c - E)$, and this has measure $\int E(x)E(2c-x)\,dx$. Thus

$$\begin{aligned}
D(E) &:= \sup\{\lambda(C) \colon \quad C \subseteq E,\, C \text{ is symmetric}\} \\
&= \sup_c \left(\int E(x)E(2c-x)\,dx \right) \\
&= \sup_c \left(\int \lambda(E) f_E(x) \lambda(E) f_E(2c-x)\,dx \right) \\
&= \lambda(E)^2 \sup_c \left(\int f_E(x) f_E(2c-x)\,dx \right) \\
&= \lambda(E)^2 \sup_c f_E * f_E(2c) \\
&= \lambda(E)^2 \left\| f_E * f_E \right\|_\infty ,
\end{aligned}$$

as desired. ∎

The convolution in Lemma 4.1 may be taken over $\mathbb{R}$ or over $\mathbb{T}$, the two settings being equivalent since $f * f$ is supported on an interval of length 1. In fact, the reason we scale f to be supported on an interval of length $1/2$ is so that we may replace convolution over $\mathbb{R}$, which is the natural place to study $\Delta(\varepsilon)$, with convolution over $\mathbb{T}$, which is the natural place to do harmonic analysis.

4.2 The Basic Argument

We begin the process of improving upon the trivial lower bound for $\Delta(\varepsilon)$ by stating a simple version of our method that illustrates the ideas and techniques involved.

Proposition 4.2. *Let K be any continuous function on $\mathbb{T}$ satisfying $K(x) \geq 1$ when $x \in [-\frac{1}{4}, \frac{1}{4}]$, and let f be a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$. Then*

$$\|f * f\|_\infty \geq \|f * f\|_2^2 \geq \|\hat{K}\|_{4/3}^{-4}.$$

Proof. We have

$$1 = \int f(x)\,dx \leq \int f(x)K(x)\,dx = \sum_j \hat{f}(j)\hat{K}(-j)$$

by Parseval's identity. Hölder's Inequality now gives $1 \leq \|\hat{f}\|_4 \|\hat{K}\|_{4/3}$, which we restate as the inequality $\|\hat{K}\|_{4/3}^{-4} \leq \|\hat{f}\|_4^4$.

Now $\|\hat{f}\|_4^4 = \sum_j |\hat{f}(j)|^4 = \sum_j |\widehat{f * f}(j)|^2 = \|f * f\|_2^2$ by another application of Parseval's identity. Since $(f*f)^2 \leq \|f*f\|_\infty (f*f)$, integration yields $\|f*f\|_2^2 \leq \|f*f\|_\infty \|f*f\|_1 = \|f*f\|_\infty$. Combining the last three sentences, we see that $\|\hat{K}\|_{4/3}^{-4} \leq \|\hat{f}\|_4^4 = \|f*f\|_2^2 \leq \|f*f\|_\infty$ as claimed. ∎

This reasonably simple theorem already allows us to give a nontrivial lower bound for $\Delta(\varepsilon)$.

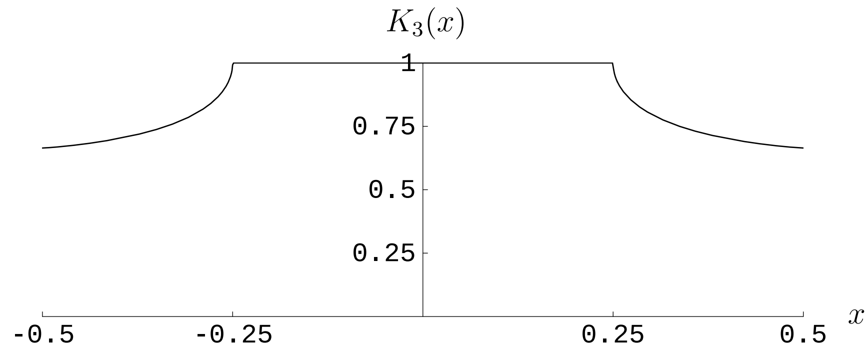

Figure 3: The function $K_3(x)$

The step function

$$K_1(x) := \begin{cases} 1 & 0 \le |x| \le \frac{1}{4} \\ 1 - \frac{2\pi^4}{\pi^4 + 24\zeta(\frac{4}{3})^3\left(5+2^{4/3}-2^{8/3}\right)} & \frac{1}{4} < |x| \le \frac{1}{2} \end{cases}$$

has $\|\hat{K_1}\|_{4/3}^{-4} = 1 + \frac{\pi^4}{8\left(2^{4/3}-1\right)^3\zeta(\frac{4}{3})^3} > 1.074$ (the elaborate constant used in the definition of K_1 was chosen to minimize $\|\hat{K_1}\|_{4/3}$); a careful reader may complain that K_1 is not continuous. The continuity condition is not essential, however, as we may approximate K_1 by a continuous function L with $\|L\|_{4/3}$ arbitrarily close to $\|K_1\|_{4/3}$. Green [Gre01] used a discretization of the kernel function

$$K_2(x) := \begin{cases} 1 & 0 \le |x| \le \frac{1}{4} \\ 1 - \alpha + \alpha\left(40(2x-1)^4 - \frac{3}{2}\right) & \frac{1}{4} < |x| \le \frac{1}{2} \end{cases}$$

with a suitably chosen α to get $\|\hat{K_2}\|_{4/3}^{-4} > \frac{8}{7} > 1.142$. We get a slightly larger value of $\|\hat{K}\|_{4/3}^{-4}$ in Corollary 4.3 with a much more complicated kernel. See Section 4.7 for a discussion of how we came to find our kernel.

Corollary 4.3. *If f is a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$, then*

$$\|f * f\|_2^2 \ge 1.14915.$$

Consequently, $\Delta(\varepsilon) \ge 0.574575\varepsilon^2$ for all $0 \le \varepsilon \le 1$.

Proof. Set

$$K_3(x) := \begin{cases} 1 & 0 \le |x| \le \frac{1}{4}, \\ 0.6644 + 0.3356\left(\frac{2}{\pi}\tan^{-1}\left(\frac{1-2x}{\sqrt{4x-1}}\right)\right)^{1.2015} & \frac{1}{4} \le |x| \le \frac{1}{2}. \end{cases} \tag{8}$$

$K_3(x)$ is pictured in Figure 3.

We do not know how to rigorously bound $\|\hat{K_3}\|_{4/3}$, but we can rigorously bound $\|\hat{K_4}\|_{4/3}$ where K_4 is a piecewise linear function 'close' to K_3. Specifically, let $K_4(x)$ the even piecewise linear function with corners at

$$(0,1), \left(\frac{1}{4},1\right), \left(\frac{1}{4}+\frac{t}{4\times 10^4}, K_3\Big(\frac{1}{4}+\frac{t}{4\times 10^4}\Big)\right) \quad (t=0,1,\ldots,10^4).$$

We calculate (using Proposition 4.14 below) that $\|\hat{K_4}\|_{4/3} < 0.9658413$. Therefore, by Proposition 4.2 we have

$$\|f*f\|_2^2 \geq (0.9658413)^{-4} > 1.14915.$$

Using Lemma 4.1, we now have $\Delta(\varepsilon) > \frac{1}{2}\varepsilon^2(1.14915) > 0.574575\varepsilon^2$. ∎

The constants in the definition (8) of $K_3(x)$ were numerically optimized to minimize $\|\hat{K_4}\|_{4/3}$ and otherwise have no special significance. The definition of $K_3(x)$ is certainly not obvious; there are much simpler kernels that do give nontrivial bounds. In Section 4.7 below, we indicate the experiments that led to our choice.

We note that the function

$$b(x) := \begin{cases} \frac{4/\pi}{\sqrt{1-16x^2}}, & -1/4 < x < 1/4, \\ 0, & \text{otherwise} \end{cases}$$

has $\int b = 1$ and $\|b*b\|_2^2 < 1.14939$. Although b is not a pdf (it is not in L^2), it gives strong testimony that the bound on $\|f*f\|_2^2$ given in Corollary 4.3 is not far from best possible.

Note that this corollary, along with Corollary 5.4, gives $R(g,n) \leq 1.31925\sqrt{gn}$, only slightly worse than the bound given in [CRT].

This bound on $\|f*f\|_2^2$ may be nearly correct, but the resulting bound on $\|f*f\|_\infty$ is not: we prove below that $\|f*f\|_\infty \geq 1.182778$, and believe that $\|f*f\|_\infty \geq \pi/2$. We have tried to improve the argument given in Proposition 4.2 in the following four ways:

1. Instead of considering the sum $\sum_j \hat{f}(j)\hat{K}(-j)$ as a whole, we single out the central terms (which dominate the sum) and the tails (which contribute essentially nothing), and deal with the three resulting sums separately. This generalized form of the above argument is expounded in the next section. The success of this generalization relies on certain inequalities restricting the possible values of these central coefficients; establishing these restrictions is the goal of Sections 4.4 and 4.5. The final lower bound derived from these methods is given in Section 4.6.

2. We can try to find more advantageous kernel functions $K(x)$ for which we can compute $\|\hat{K}\|_{4/3}$ in an accurate way. A detailed discussion of our search for the best kernel functions is in Section 4.7.

3. The application of Parseval's identity can be replaced with the Hausdorff-Young inequality, which leads to the conclusion $\|f*f\|_\infty \geq \|\hat{K}\|_p^{-q}$, where $p \leq \frac{4}{3}$ and $q \geq 4$ are conjugate exponents. Numerically, the values $(p,q) = (\frac{4}{3}, 4)$ appear to be optimal. However, Beckner's sharpening [Bec75] of the Hausdorff-Young inequality leads to the stronger conclusion $\|f*f\|_\infty \geq C(q)\|\hat{K}\|_p^{-q}$ where $C(q) = \frac{q}{2}(1-\frac{2}{q})^{q/2-1} = \frac{q}{2e} + O(1)$. We have not experimented to see whether a larger lower bound can be obtained from this stronger inequality by taking $q > 4$.

4. Notice that we used the inequality $\|g\|_2^2 \le \|g\|_\infty \|g\|_1$ with the function $g = f * f$. This inequality is sharp exactly when the function g takes only one nonzero value (i.e., when g is a nif), but the convolution $f*f$ never seems to behave that way. Perhaps for these autoconvolutions, an analogous inequality with a stronger constant than 1 could be established. Unfortunately, we have not been able to realize any success with this idea, although we believe Conjecture 4.4 below. If true, Conjecture 4.4 implies the bound $\Delta(\varepsilon) \ge 0.651\varepsilon^2$.

Conjecture 4.4. *If f is a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$, then*

$$\frac{\|f*f\|_\infty}{\|f*f\|_2^2} \ge \frac{\pi/2}{\log 4},$$

with equality only if either $f(x)$ or $f(-x)$ equals $\sqrt{\frac{2}{4x+1}}$ on the interval $|x| \le \frac{1}{4}$.

We remark that Proposition 4.2 can be extended from a twofold convolution in one dimension to an h-fold convolution in d dimensions.

Proposition 4.5. *Let K be any continuous function on $\mathbb{T}^d$ satisfying $K(\bar{x}) \ge 1$ when $\bar{x} \in [-\frac{1}{2h}, \frac{1}{2h}]^d$, and let f be a pdf supported on $[-\frac{1}{2h}, \frac{1}{2h}]^d$. Then*

$$\|f^{*h}\|_\infty \ge \|f^{*h}\|_2^2 \ge \|\hat{K}\|_{2h/(2h-1)}^{-2h}.$$

Every subset of $[0,1]^d$ with measure ε contains a symmetric subset with measure $(0.574575)^d\varepsilon^2$.

Proof. The proof proceeds as above, with the conjugate exponents $(\frac{2h}{2h-1}, 2h)$ in place of $(\frac{4}{3}, 4)$, and the kernel function $K(x_1, x_2, \ldots, x_d) = K(x_1)K(x_2)\cdots K(x_d)$ in place of the kernel function $K(x)$ defined in the proof of Corollary 4.3. The second assertion of the proposition follows on taking $h = 2$. ∎

4.3 The Main Bound

We now present a more subtle version of Proposition 4.2. Recall that the notation ${}_n\|a\|_p$ was defined in Eq. (7). We also use $\Re z$ to denote the real part of the complex number z.

Proposition 4.6. *Let $1 \le m < m' \le \infty$. Suppose that f is a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$ and that K is even, continuous, satisfies $K(x) = 1$ for $-\frac{1}{4} \le x < \frac{1}{4}$, and ${}_{m'}\|\hat{K}\|_p > 0$. Set $M := 1 - \hat{K}(0) - 2\sum_{j=1}^{m-1} \hat{K}(j)\Re\hat{f}(j)$. Then*

$$\|f*f\|_2^2 \ge \sum_{|j|\le m'} |\hat{f}(j)|^4 \ge 1 + \left(\frac{M}{{}_m\|\hat{K}\|_{4/3}}\right)^4 + 2\sum_{j=1}^{m-1} |\Re\hat{f}(j)|^4 - o(1) \tag{9}$$

as $m' \to \infty$.

Proof. The first inequality follows from Parseval's formula

$$\|f*f\|_2^2 = \sum_j |\widehat{f*f}(j)|^2 = \sum_j |\hat{f}(j)|^4 \ge \sum_{|j|\le m'} |\hat{f}(j)|^4.$$

As in the proof of Proposition 4.2, we have

$$1 = \int f(x)K(x)\,dx = \sum_j \hat{f}(j)\hat{K}(-j) = \sum_{|j|<m} \hat{f}(j)\hat{K}(-j) + \sum_{|j|\geq m} \hat{f}(j)\hat{K}(-j).$$

Since K is even, $\hat{K}(-j) = \hat{K}(j)$ is real, and since f is real valued, $\hat{f}(-j) = \overline{\hat{f}(j)}$. We have

$$1 = \hat{K}(0) + 2\sum_{j=1}^{m-1} \hat{K}(j)\Re\hat{f}(j) + \sum_{|j|\geq m} \hat{f}(j)\hat{K}(j),$$

which we can also write as $M = \sum_{|j|\geq m} \hat{f}(j)\hat{K}(j)$. Set $\eta := \sum_{|j|>m'} |\hat{K}(j)|$, and note that $\eta = o(1)$ as $m' \to \infty$ since K is continuous. Taking absolute values and applying Hölder's inequality, we have

$$\begin{aligned}
|M| &\leq \sum_{|j|\geq m} |\hat{f}(j)\hat{K}(j)| \\
&\leq \sum_{m'\geq|j|\geq m} |\hat{f}(j)\hat{K}(j)| + \sum_{|j|>m'} |\hat{K}(j)| \\
&\leq \left(\sum_{m'\geq|j|\geq m} |\hat{f}(j)|^4\right)^{1/4} \left(\sum_{m'\geq|j|\geq m} |\hat{K}(j)|^{4/3}\right)^{3/4} + \eta \\
&\leq \left(\sum_{m'\geq|j|\geq m} |\hat{f}(j)|^4\right)^{1/4} {}_m\|\hat{K}\|_{4/3} + o(1),
\end{aligned}$$

which we recast in the form

$$\sum_{m'\geq|j|\geq m} |\hat{f}(j)|^4 \geq \left(\frac{|M| - o(1)}{{}_m\|\hat{K}\|_{4/3}}\right)^4 = \left(\frac{M}{{}_m\|\hat{K}\|_{4/3}}\right)^4 - o(1).$$

We add $\sum_{|j|<m} |\hat{f}(j)|^4$ to both sides and observe that $\hat{f}(0) = 1$ and $|\hat{f}(j)| \geq |\Re\hat{f}(j)|$ to finish the proof of the second inequality. ∎

Corollary 4.7. *If f is a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$, then*

$$\sum_{|j|\leq m'} |\hat{f}(j)|^4 \geq 1.14915 - o(1)$$

as $m' \to \infty$

Proof. The kernel function used in the proof of Corollary 4.3 has $\hat{K}_4(0) \doteq 0.870250799$ and ${}_1\|\hat{K}_4\|_{4/3} \doteq 0.208784534$. Now apply Proposition 4.6 with $m = 1$. ∎

Corollary 4.8. *Let f be a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$, and set $x_1 := \Re\hat{f}(1)$. Then as $m' \to \infty$*

$$\|f * f\|_2^2 \geq \sum_{|j|\leq m'} |\hat{f}(j)|^4 \geq 1 + 2x_1^4 + \left(\frac{0.368067372 - 0.541553784x_1}{0.239175395}\right)^4 - o(1).$$

Proof. Set

$$K_5(x) = \begin{cases} 1 & |x| \le \frac{1}{4}, \\ 1-(1-(4(\frac{1}{2}-x))^{1.61707})^{0.546335} & \frac{1}{4} < |x| \le \frac{1}{2}. \end{cases}$$

Denote by $K_6(x)$ the even piecewise linear function with corners at

$$(0,1), \left(\frac{1}{4}, 1\right), \left(\frac{1}{4}+\frac{t}{4\times 10^4}, K_5(\frac{1}{4}+\frac{t}{4\times 10^4})\right) \text{ (where } t=0,1,\ldots,10^4).$$

We find (using Proposition 4.14) that $\hat{K_6}(0) \doteq 0.631932628$, $\hat{K_6}(1) \doteq 0.270776892$, and ${}_2\|\hat{K_6}\|_{4/3} \doteq 0.239175395$. Apply Proposition 4.6 with $m=2$ to finish the proof. ■

With $m' = \infty$ and $x_j = \Re \hat{f}(j)$, the bound of Proposition 4.6 becomes

$$1 + \left(\frac{1-\hat{K}(0) - 2\sum_{j=1}^{m-1}\hat{K}(j)x_j}{{}_m\|\hat{K}\|_{4/3}}\right)^4 + 2\sum_{j=1}^{m-1} x_j^4.$$

This is a quartic polynomial in the x_j, and consequently it is not difficult to minimize, giving an absolute lower bound on $\|f*f\|_2^2$. This minimum occurs at

$$x_j = \frac{(\hat{K}(j))^{1/3}\left(1-\hat{K}(0)-2\sum_{i=1}^{j-1}\hat{K}(i)x_i\right)}{{}_j\|\hat{K}\|_{4/3}^{4/3}},$$

where $(\hat{K}(j))^{1/3}$ is the real cube root of $\hat{K}(j)$. Consequently,

$$\inf_{x_j\in\mathbb{R}}\left\{1 + \left(\frac{1-\hat{K}(0) - 2\sum_{j=1}^{m-1}\hat{K}(j)x_j}{{}_m\|\hat{K}\|_{4/3}}\right)^4 + 2\sum_{j=1}^{m-1} x_j^4\right\} = 1 + \left(\frac{1-\hat{K}(0)}{{}_1\|\hat{K}\|_{4/3}}\right)^4,$$

which is nothing more than the bound that Proposition 4.6 gives with $m=1$. Moreover,

$$1 + \left(\frac{1-\hat{K}(0)}{{}_1\|\hat{K}\|_{4/3}}\right)^4 = \sup_{0\le\alpha\le 1} \|(\alpha+(1-\alpha)K)^\wedge\|_{4/3}^{-4},$$

(the details of this calculation are given in Section 4.7.1) so that Proposition 4.6 by itself does not give a different bound on $\|f*f\|_2^2$ than Proposition 4.2. However, we shall obtain additional information on $\hat{f}(j)$ in terms of $\|f*f\|_\infty$ in subsection 4.5 below, and this information can be combined with Proposition 4.6 to provide a stronger lower bound on $\|f*f\|_\infty$ than that given by Proposition 4.2.

4.4 Some Useful Inequalities

Hardy, Littlewood, and Pólya [HLP88] call a function $u(x)$ *symmetric decreasing* if $u(x) = u(-x)$ and $u(x) \ge u(y)$ for all $0 \le x \le y$, and they call

$$f^{\mathrm{sdr}}(x) := \inf\{y\colon \quad \lambda(\{t\colon f(t)\ge y\}) \le 2|x|\}$$

the *symmetric decreasing rearrangement* of f. For example, if f is the indicator function of a set with measure μ, then f^{sdr} is simply the indicator function of the interval $(-\frac{\mu}{2}, \frac{\mu}{2})$. Another example is any function f defined on an interval $[-a, a]$ and is periodic with period $\frac{2a}{n}$, where n is a positive integer, and that is symmetric decreasing on the subinterval $[-\frac{a}{n}, \frac{a}{n}]$; then $f^{\text{sdr}}(x) = f(\frac{x}{n})$ for all $x \in [-a, a]$. In particular, on the interval $[-\frac{1}{4}, \frac{1}{4}]$, we have $\cos^{\text{sdr}}(2\pi j x) = \cos(2\pi x)$ for any nonzero integer j. We shall need the following result.

Lemma 4.9.

$$\int f(x)u(x)\,dx \le \int f^{sdr}(x)u^{sdr}(x)\,dx.$$

Proof. This is Theorem 378 of [HLP88]. ■

We say that $\bar{f}$ *is more focused than* f (and f is less focused than $\bar{f}$) if for all $z \in [0, \frac{1}{2}]$ and all $r \in \mathbb{T}$ we have

$$\int_{r-z}^{r+z} f \le \int_{-z}^{z} \bar{f}.$$

For example, f^{sdr} is more focused than f. In fact, we introduce this terminology because it refines the notion of symmetric decreasing rearrangement in a way that is useful for us. To give another example, if f is a nonnegative function, set $\bar{f}$ to be $\|f\|_\infty$ times the indicator function of the interval $[-\frac{1}{2\|f\|_\infty}, \frac{1}{2\|f\|_\infty}]$; then $\bar{f}$ is more focused than f.

Lemma 4.10. *Let $u(x)$ be a symmetric decreasing function, and let $h, \bar{h}$ be pdfs with $\bar{h}$ more focused than h. Then for all $r \in \mathbb{T}$,*

$$\int h(x-r)u(x)\,dx \le \int \bar{h}(x)u(x)\,dx.$$

Proof. Without loss of generality we may assume that $r = 0$, since if $\bar{h}(x)$ is more focused than $h(x)$, then it is also more focused than $h(x-r)$. Also, without loss of generality we may assume that $h, \bar{h}$ are continuous and strictly positive on $\mathbb{T}$, since any nonnegative function in L^1 can be L^1-approximated by such.

Define $H(z) = \int_{-z}^{z} h(t)\,dt$ and $\bar{H}(z) = \int_{-z}^{z} \bar{h}(t)\,dt$, so that $H(\frac{1}{2}) = \bar{H}(\frac{1}{2}) = 1$, and note that the more-focused hypothesis implies that $H(z) \le \bar{H}(z)$ for all $z \in [0, \frac{1}{2}]$. Now h is continuous and strictly positive, which implies that H is differentiable and strictly increasing on $[0, \frac{1}{2}]$ since $H'(z) = h(z) + h(-z)$. Therefore H^{-1} exists as a function from $[0,1]$ to $[0, \frac{1}{2}]$. Similar comments hold for $\bar{H}^{-1}$.

Since $H \le \bar{H}$, we see that $\bar{H}^{-1}(s) \le H^{-1}(s)$ for all $s \in [0,1]$. Then, since $H^{-1}(s)$ and $H^{-1}(s)$ are positive and u is decreasing for positive arguments, we conclude that $u(H^{-1}(s)) \le u(\bar{H}^{-1}(s))$, and so

$$\int_0^1 u(H^{-1}(s))\,ds \le \int_0^1 u(\bar{H}^{-1}(s))\,ds. \tag{10}$$

On the other hand, making the change of variables $s = H(t)$, we see that

$$\int_0^1 u(H^{-1}(s))\,ds = \int_0^{H^{-1}(1)} u(t)H'(t)\,dt = \int_0^{1/2} u(t)(h(t)+h(-t))\,dt = \int_{\mathbb{T}} u(t)h(t)\,dt$$

since u is symmetric. Similarly $\int_0^1 u(\bar{H}^{-1}(s))\,ds = \int_{\mathbb{T}} u(t)\bar{h}(t)\,dt$, and so inequality (10) becomes $\int u(t)h(t)\,dt \le \int u(t)\bar{h}(t)\,dt$ as desired. ∎

4.5 Fourier Coefficients of Density Functions

To use Proposition 4.6 to bound $\Delta(\varepsilon)$, we need to develop a better understanding of the central Fourier coefficients $\hat{f}(j)$ for small j. In particular, we wish to apply Proposition 4.6 with $m = 2$, i.e., we need to develop the connections between $\|f * f\|_\infty$ and the real part of the Fourier coefficient $\hat{f}(1)$.

We turn now to bounding $|\hat{f}(j)|$ in terms of $\|f * f\|_\infty$. The guiding principle is that if $f * f$ is very concentrated then $\|f * f\|_\infty$ will be large, and if $f * f$ is not very concentrated then $|\hat{f}(j)|$ will be small. Green [Gre01, Lemma 26] proves the following lemma in a discrete setting, but since we need a continuous version we include a complete proof.

Lemma 4.11. *Let f be a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$. For $j \neq 0$,*

$$|\hat{f}(j)|^2 \le \frac{\|f * f\|_\infty}{\pi} \sin\left(\frac{\pi}{\|f * f\|_\infty}\right).$$

Proof. Let $f_1 : \mathbb{T} \to \mathbb{R}$ be defined by $f_1(x) := f(x - x_0)$, with x_0 chosen so that $\hat{f}_1(j)$ is real and positive (clearly $\hat{f}_1(j) = |\hat{f}(j)|$ and $\|f * f\|_\infty = \|f_1 * f_1\|_\infty$). Set $h(x)$ to be the symmetric decreasing rearrangement of $f_1 * f_1$, and $\overline{h}(x) := \|f * f\|_\infty I(x)$, where $I(x)$ is the indicator function of $[-\frac{1}{2\|f*f\|_\infty}, \frac{1}{2\|f*f\|_\infty}]$. We have

$$|\hat{f}(j)|^2 = \hat{f}_1(j)^2 = \widehat{f_1 * f_1}(j) = \int f_1 * f_1(x) \cos(2\pi j x)\,dx \le \int h(x) \cos(2\pi x)\,dx$$

by Lemma 4.9. We now apply Lemma 4.10 to find

$$\begin{aligned} |\hat{f}(j)|^2 &\le \int \overline{h}(x) \cos(2\pi x)\,dx \\ &= \int_{-1/(2\|f*f\|_\infty)}^{1/(2\|f*f\|_\infty)} \|f * f\|_\infty \cos(2\pi x)\,dx = \frac{\|f * f\|_\infty}{\pi} \sin\left(\frac{\pi}{\|f * f\|_\infty}\right). \end{aligned}$$

∎

4.6 The Full Bound

With the technical result of Section 4.5 in hand, we can finally establish the lower bound on $\Delta(\varepsilon)$ given in Theorem 1.1(ii).

Proposition 4.12. $\Delta(\varepsilon) \ge 0.591389\varepsilon^2$ *for all* $0 \le \varepsilon \le 1$.

Proof. Let f be a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$, and assume that

$$\|f * f\|_\infty < 1.182778. \tag{11}$$

Set $x_1 = \Re\hat{f}(1)$ and $x_2 = \Re\hat{f}(2)$. Since f is supported on $[-\frac{1}{4}, \frac{1}{4}]$, we see that $x_1 > 0$. By Lemma 4.11,

$$0 < x_1 < 0.4191447. \tag{12}$$

Corollary 4.8 with $m' = \infty$, gives

$$\|f * f\|_\infty \geq \|f * f\|_2 \geq 1 + 2x_1^4 + \left(\frac{0.368067372 - 0.541553784x_1}{0.239175395}\right)^4. \tag{13}$$

Routine calculus shows that there are no simultaneous solutions to Inequalities (11), (12), and (13). Therefore $\|f*f\|_\infty \geq 1.182778$, whence Lemma 4.1 implies that $\Delta(\varepsilon) \geq 0.591389\varepsilon^2$. ■

This gist of the proof of Proposition 4.12 is that if $\|f * f\|_\infty$ is small, then $\Re\hat{f}(1)$ is small by Lemma 4.11, and so $\|f * f\|_2^2$ is not very small by Corollary 4.8, whence $\|f * f\|_\infty$ is not small. If $\|f * f\|_\infty \leq 1.182778$, then we get a contradiction. We can actually prove a meaningful result about $\|f * f\|_2^2$ under the condition that $\|f * f\|_\infty$ is not much larger than 1.182778. The following result will be useful in Section 5.

Lemma 4.13. *Let f be a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$. If $1.182778 \leq \|f*f\|_\infty \leq 1.229837$, then*

$$\|f * f\|_2^2 \geq \sum_{|j| \leq m'} |\hat{f}(j)|^4 \geq 21.922911 - 33.711941\|f * f\|_\infty + 13.676987\|f * f\|_\infty^2 - o(1)$$

as $m' \to \infty$.

Proof. The first inequality has already been shown in Proposition 4.6. Set

$$B(x_1) := 1 + 2x_1^4 + \left(\frac{0.368067372 - 0.541553784x_1}{0.239175395}\right)^4,$$

so that by Corollary 4.8,

$$\sum_{|j| \geq m'} |\hat{f}(j)|^4 \geq B(\Re\hat{f}(1)) - o(1).$$

By hypothesis $\|f * f\|_\infty \leq 1.229837$, so that by Lemma 4.11,

$$\Re\hat{f}(1) \leq \sqrt{\frac{\|f * f\|_\infty}{\pi} \sin\left(\frac{\pi}{\|f * f\|_\infty}\right)} < 0.466.$$

But $B(x_1)$ is a decreasing function for $x_1 \in [0, 0.47]$, so that

$$\begin{aligned} B(\Re\hat{f}(1)) &\geq B\left(\sqrt{\tfrac{\|f*f\|_\infty}{\pi} \sin\left(\tfrac{\pi}{\|f*f\|_\infty}\right)}\right) \\ &> 21.922911 - 33.711941\|f * f\|_\infty + 13.676987\|f * f\|_\infty^2 \end{aligned}$$

after a straightforward computation. ■

4.7 The Kernel Problem

Let $\mathbb{K}$ be the class of functions $K \in L^2$ satisfying $K(x) \geq 1$ on $[-\frac{1}{4}, \frac{1}{4}]$. Proposition 4.2 suggests the problem of computing

$$\inf_{K \in \mathbb{K}} \|\hat{K}\|_p = \inf_{K \in \mathbb{K}} \left(\sum_{j=-\infty}^{\infty} |\hat{K}(j)|^p \right)^{1/p}.$$

In Proposition 4.2 the case $p = \frac{4}{3}$ arose, but using the Hausdorff-Young inequality in place of Parseval's identity we are led to consider $1 < p \leq \frac{4}{3}$. Also, we assumed in Proposition 4.2 that K was continuous, but this assumption can be removed by taking the pointwise limit of continuous functions.

As similar problems occur in [CRT] and in [Gre01], we feel it is worthwhile to detail the thoughts and experiments that led to the kernel functions chosen in Corollaries 4.3 and 4.8.

4.7.1 *Ad hoc* Observations

Our first observation is that if $G \in \mathbb{K}$, then so is $K(x) := \frac{1}{2}(G(x) + G(-x))$, and since $|\hat{K}(j)| = |\Re \hat{G}(j)| \leq |\hat{G}(j)|$ we know that $\|\hat{K}\|_p \leq \|\hat{G}\|_p$. Thus, we may restrict our attention to the *even* functions in $\mathbb{K}$.

We also observe that $|\hat{K}(j)|$ decays more rapidly if many derivatives of K are continuous. This suggests that we should restrict our attention to continuous K, perhaps even to infinitely differentiable K. However, computations suggest that the best functions K are continuous but *not* differentiable at $x = \frac{1}{4}$ (see in particular Section 4.7.4 and Figure 4).

In the argument of Proposition 4.2 we used the inequality $\int f \leq \int fK$, which is an equality if we take K to be equal to 1 on $[-\frac{1}{4}, \frac{1}{4}]$, instead of merely at least 1. In light of this, we should not be surprised if the optimal functions in $\mathbb{K}$ are exactly 1 on $[-\frac{1}{4}, \frac{1}{4}]$. This is supported by our computations.

Finally, we note that if $K_i \in \mathbb{K}$, and $\alpha_i > 0$ with $\sum_i \alpha_i = 1$, then $\sum_i \alpha_i K_i(x) \in \mathbb{K}$ also. This is particularly useful with $K_1(x) := 1$. Specifically, given any $K_2 \in \mathbb{K}$ with known $\|\hat{K_2}\|_p$ (we stipulate $\|K_2\|_1 = \hat{K}(0) =\leq 1$ to avoid technicalities), we may easily compute the $\alpha \in [0, 1]$ for which $\|\hat{K}\|_p$ is minimized, where $K(x) := \alpha K_1(x) + (1 - \alpha)K_2(x)$. We have

$$\|\hat{K}\|_p^p = (\alpha + (1-\alpha)\hat{K}_2(0))^p + (1-\alpha)^p {}_1\|\hat{K}\|_p^p = (1 - (1-\alpha)M)^p + (1-\alpha)^p N, \tag{14}$$

where we have set $M := 1 - \hat{K}_2(0)$ and $N := {}_1\|\hat{K}\|_p^p$. Taking the derivative with respect to α, we obtain

$$p(1-\alpha)^{p-1}\left(M\Big(\frac{1}{1-\alpha} - M\Big)^{p-1} - N\right),$$

the only root of which is $\alpha = 1 - \frac{M^{q/p}}{M^q + N^{q/p}}$ (where $\frac{1}{p} + \frac{1}{q} = 1$). It is straightforward (albeit tedious) to check by substituting α into the second derivative of the expression (14) that this value of α yields a local maximum for $\|\hat{K}\|_p^p$. The maximum value attained is then calculated to equal $N\big(M^q + N^{q/p}\big)^{1-p}$, which is easily computed from the known function K_2.

Notice that when $p=\frac{4}{3}$ (so $q=4$), applying Proposition 4.2 with our optimal function K yields

$$\|f*f\|_2^2 \ge \|\hat{K}\|_{4/3}^{-4} = \Big(N\big(M^4+N^3\big)^{-1/3}\Big)^{-3} = \frac{M^4+N^3}{N^3} = 1+\frac{(1-\hat{K}_2(0))^4}{{}_1\|\hat{K}\|_{4/3}^4},$$

whereupon we recover the conclusion of Proposition 4.6 with $m=1$.

4.7.2 Trigonometric Polynomials

We wish to identify families of functions that are at least 1 on $[-\frac{1}{4},\frac{1}{4}]$ and whose Fourier coefficients have small ℓ^p norm. Natural candidates are functions which have many Fourier coefficients equal to 0. In this section we consider trigonometric polynomials $K(x)=\sum_{j=-m}^m \hat{K}(j)e^{2\pi ijx}$ of degree m.

Montgomery [Mon94, Chapter 1] defines the Selberg polynomials $S_m^+(\alpha,\beta,x)$ and shows that $S_m^+(\alpha,\beta,x)\ge\chi_{[\alpha,\beta]}(x)$ for all x, provided that $\alpha\le\beta\le\alpha+1$; moreover, these functions are (in some senses) optimal L^1 majorants for $\chi_{[\alpha,\beta]}(x)$ among all trigonometric polynomials of bounded degree. These provide a natural family for investigation.

We are concerned with $[\alpha,\beta]=[-\frac{1}{4},\frac{1}{4}]$. We have for instance

$$S_2^+(-\tfrac{1}{4},\tfrac{1}{4},x)=\frac{5}{6}+\big(\frac{4}{9\sqrt{3}}+\frac{2}{3\pi}\big)\cos(2\pi x)-\frac{2}{9}\cos(4\pi x),$$

which satisfies $S_2^+(-\frac{1}{4},\frac{1}{4},x)-\frac{1}{18}\ge 1$ when $x\in[-\frac{1}{4},\frac{1}{4}]$, and

$$\inf_{0\le\alpha\le1}\|(\alpha+(1-\alpha)(S_2^+(-\tfrac{1}{4},\tfrac{1}{4},x)-\frac{1}{18})^\wedge\|_{4/3}>0.990.$$

However, $L(x):=2\cos(2\pi x)-\cos(4\pi x)\in\mathbb{K}$, and

$$\inf_{0\le\alpha\le1}\|(\alpha+(1-\alpha)(L(x))^\wedge\|_{4/3}<0.989.$$

Thus, even among trigonometric polynomials of degree 2, the Selberg polynomials are not optimal. In general, we have been unable to identify the degree-m trigonometric polynomial $K(x)$ that is in $\mathbb{K}$ and for which $\sum_{j=-m}^m|\hat{K}(j)|^{4/3}$ is minimized.

4.7.3 Wavelets

We can give an exact, finite expression for the p-norm of the Fourier coefficients of some large classes of functions. Sums of Haar wavelets give the simplest theoretical instance and the largest class of functions.

Define

$$\psi(x):=\begin{cases}1 & 0\le x<\frac{1}{2},\\ -1 & \frac{1}{2}\le x<1,\\ 0 & \text{otherwise.}\end{cases}$$

and $\psi_{m,n}(x) := 2^{-m/2}\psi(2^m x - n)$. It is well-known that the Haar Wavelets $\{\psi_{m,n} \colon m, n \in \mathbb{Z}\}$ form an orthonormal basis of the subspace of $L^2(\mathbb{R})$ consisting of functions that have integral 0. By the comments in Section 4.7.1,

$$\inf_{K \in \mathbb{K}} \|\hat{K}\|_{4/3} = \inf_{\substack{K \in \mathbb{K} \\ \int K = 0}} \left(1 + \frac{1}{{}_1\|\hat{K}\|_{4/3}^4}\right)^{-1/4},$$

so that the $\int K = 0$ restriction is not a substantial restriction.

It follows that every even function $K \in \mathbb{K}$ with $\int K = 0$ and $x \in [-\frac{1}{4}, \frac{1}{4}) \Rightarrow K(x) = 1$ can be written in the form

$$K(x) = \sqrt{2}\,(\psi_{1,0}(x) + \psi_{1,0}(-x)) + \sum_{n=1}^{\infty} \alpha_n \left(\psi_{2+\lfloor \log_2 n \rfloor, n}(x) + \psi_{2+\lfloor \log_2 n \rfloor, n}(-x)\right).$$

Since this expression is linear, we can give the Fourier coefficients of K in terms of the easily computable Fourier coefficients of the $\psi_{m,n}$ and in terms of the parameters α_n. Truncating the infinite sum at N, we obtain a reasonably large family of possible functions K for which $\|\hat{K}\|_p$ can be computed quickly enough to numerically optimize $\alpha_1, \ldots, \alpha_N$. Graphs of the optimal $K(x)$ for $p = \frac{4}{3}$ and various values of N are displayed in Figure 4. Note that if these wavelet-based functions are converging to some limit function in $\mathbb{K}$, that limit function certainly does not seem to be differentiable at $\pm\frac{1}{4}$.

While the exact form of this expression for $\|\hat{K}\|_p$ is not difficult to compute, we suppress the details as they are very similar to the expressions computed in the next section.

4.7.4 Piecewise-Linear Functions

More useful computationally is the class of continuous piecewise-linear even functions whose vertices all have abscissae with a given denominator. Let $\zeta(s, a) := \sum_{k=0}^{\infty} (k + a)^{-s}$ denote the Hurwitz zeta function. If $\mathbf{v}$ is a vector, define $\Lambda_p(\mathbf{v})$ to be the vector whose coordinates are the pth powers of the absolute values of the corresponding coordinates of $\mathbf{v}$.

Proposition 4.14. *Let T be a positive integer, n a nonnegative integer, and $p \geq 1$ a real number. For each integer $0 \leq t \leq T$, define $x_t := \frac{1}{4} + \frac{t}{4T}$, and let y_t be an arbitrary real number, except that $y_0 = 1$. Let $K(x)$ be the even function on $\mathbb{T}$ that is linear on $[0, \frac{1}{4}]$ and each of the intervals $[x_{t-1}, x_t]$ $(1 \leq t \leq T)$, satisfying $K(0) = 1$ and $K(x_t) = y_t$ $(0 \leq t \leq T)$. Then*

$${}_n\|\hat{K}\|_p = (2\Lambda_p(\mathbf{d}A) \cdot \mathbf{z})^{1/p},$$

where $\mathbf{d}$ is the T-dimensional vector $\mathbf{d} = (y_1 - y_0, y_2 - y_1, \ldots, y_T - y_{T-1})$, A is the $T \times 4T$ matrix whose (t, k)-th component is

$$A_{tk} = \cos(2\pi(n + k - 1)x_t) - \cos(2\pi(n + k - 1)x_{t-1}),$$

and $\mathbf{z}$ is the $4T$-dimensional vector

$$\mathbf{z} = (8T\pi^2)^{-p}\big(\zeta(2p, \tfrac{j}{4T}), \zeta(2p, \tfrac{j+1}{4T}), \ldots, \zeta(2p, \tfrac{j+4T-1}{4T})\big).$$

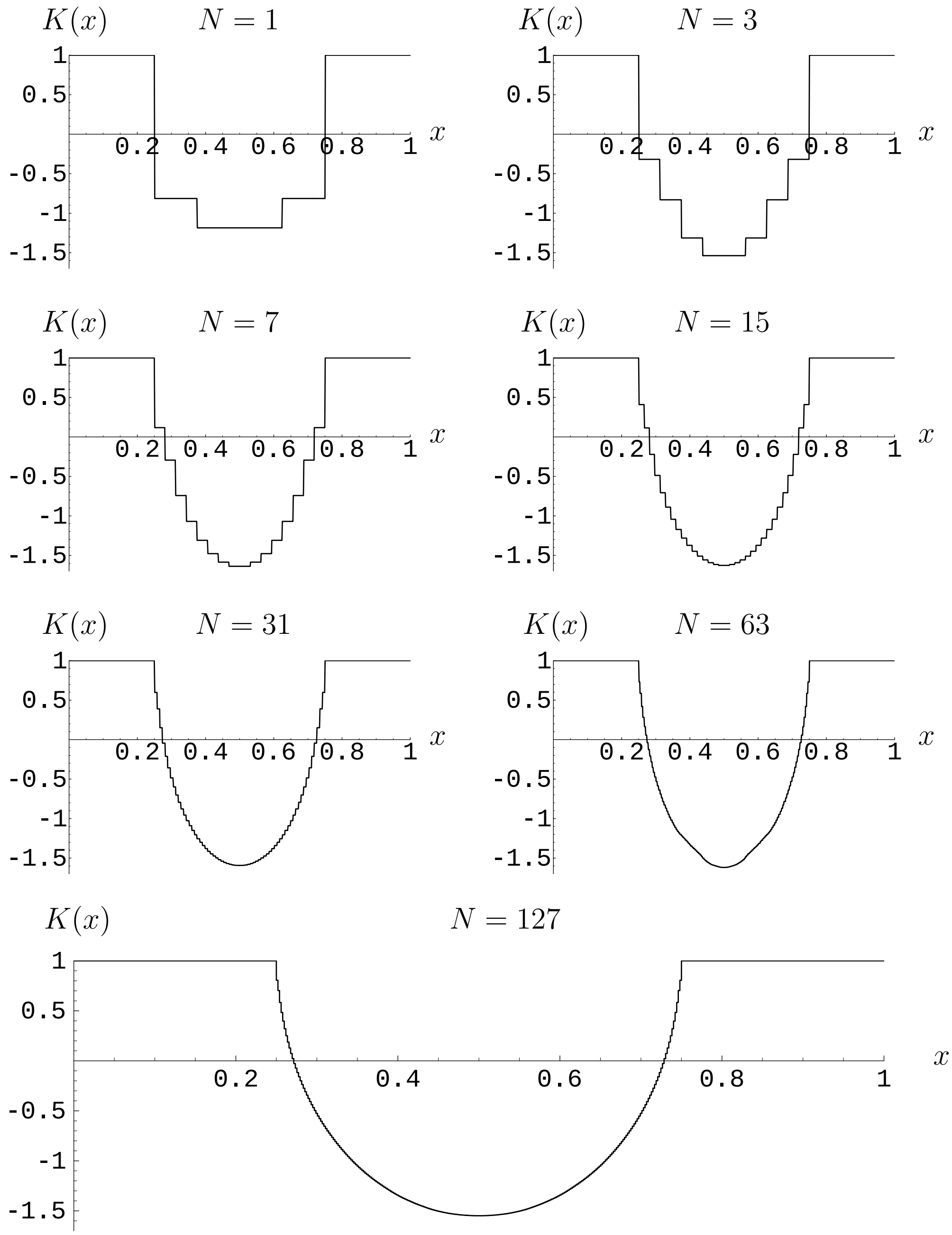

Figure 4: Optimal kernels generated by Haar wavelets

Proof. Note that

$$
\begin{aligned}
\hat{K}(-j) = \hat{K}(j) &= \int_{-1/2}^{1/2} K(u)\cos(2\pi j u)\,du \\
&= 2\int_{0}^{1/4} \cos(2\pi j u)\,du + 2\sum_{t=1}^{T}\int_{x_{t-1}}^{x_t} (m_t u + b_t)\cos(2\pi j u)\,du,
\end{aligned}
$$

where m_t and b_t are the slope and y-intercept of the line going through (x_{t-1}, y_{t-1}) and (x_t, y_t). If we define $C(j) := \frac{\pi^2 j^2}{2T}\hat{K}(j)$, then integrating by parts we have

$$
\begin{aligned}
C(j) = \left(\frac{\pi^2 j^2}{T}\left(\frac{1}{2\pi j}\sin(2\pi j u)\bigg|_0^{1/4}\right) + \frac{\pi^2 j^2}{T}\sum_{t=1}^{T}\left(\frac{m_t u + b_t}{2\pi j}\sin(2\pi j u)\bigg|_{x_{t-1}}^{x_t}\right)\right) \\
+ \left(\frac{\pi^2 j^2}{T}\sum_{t=1}^{T}\frac{m_t}{(2\pi j)^2}\cos(2\pi j u)\bigg|_{x_{t-1}}^{x_t}\right). \qquad (15)
\end{aligned}
$$

The first term of this expression is

$$
\begin{aligned}
&\frac{\pi j}{2T}\left(\sin\left(\pi\tfrac{j}{2}\right) + \sum_{t=1}^{T}\big((m_t x_t + b_t)\sin(2\pi j x_t) - (m_t x_{t-1} + b_t)\sin(2\pi j x_{t-1})\big)\right) \\
&\quad = \frac{\pi j}{2T}\left(\sin\left(\pi\tfrac{j}{2}\right) + \sum_{t=1}^{T}(m_t x_t + b_t)\sin(2\pi j x_t) - \sum_{t=0}^{T-1}(m_{t+1}x_t + b_{t+1})\sin(2\pi j x_t)\right).
\end{aligned}
$$

Since $m_{t+1}x_t + b_{t+1} = y_t = m_t x_t + b_t$ and $x_0 = \frac{1}{4}$, $x_T = \frac{1}{2}$, this entire expression is a telescoping sum whose value is zero. Eq. (15) thus becomes

$$
\begin{aligned}
C(j) &= \frac{\pi^2 j^2}{T}\sum_{t=1}^{T}\frac{m_t}{(2\pi j)^2}\cos(2\pi j u)\bigg|_{x_{t-1}}^{x_t} \\
&= \sum_{t=1}^{T}(y_t - y_{t-1})\left(\cos(2\pi j x_t) - \cos(2\pi j x_{t-1})\right) \qquad (16)
\end{aligned}
$$

using $m_t = \frac{y_t - y_{t-1}}{x_t - x_{t-1}} = 4T(y_t - y_{t-1})$. Each x_t is rational and can be written with denominator $4T$, so we see that the sequence of normalized Fourier coefficients $C(j)$ is periodic with period $4T$.

We proceed to compute ${}_n\|\hat{K}\|_p$ with n positive and $p \geq 1$.

$$
\begin{aligned}
\left({}_n\|\hat{K}\|_p\right)^p &= \sum_{|j|\geq n}|\hat{K}(j)|^p = 2\sum_{j=n}^{\infty}|\hat{K}(j)|^p = 2\sum_{j=n}^{\infty}\left|C(j)\frac{2T}{\pi^2 j^2}\right|^p \\
&= 2\left(\frac{2T}{\pi^2}\right)^p\sum_{j=n}^{\infty}\frac{|C(j)|^p}{j^{2p}}.
\end{aligned}
$$

Because of the periodicity of $C(j)$, we may write this as

$$\begin{aligned}\left({}_n\|\hat{K}\|_p\right)^p &= 2\left(\frac{2T}{\pi^2}\right)^p\left(\sum_{j=n}^{n+4T-1}|C(j)|^p\sum_{r=0}^{\infty}(4Tr+j)^{-2p}\right)\\ &= 2\left(\frac{2T}{(4T)^2\pi^2}\right)^p\left(\sum_{j=n}^{n+4T-1}|C(j)|^p\zeta\left(2p,\tfrac{j}{4T}\right)\right), \qquad (17)\end{aligned}$$

which concludes the proof. ∎

Proposition 4.14 is useful in two ways. The first is that only $\mathbf{d}$ depends on the chosen values y_t. That is, the vector $\mathbf{z}$ and the matrix A may be precomputed (assuming T is reasonably small), enabling us to compute ${}_n\|\hat{K}\|_p$ quickly enough as a function of $\mathbf{d}$ to numerically optimize the y_t. The second use is through Eq. (17). For a given K, we set $y_t = K(x_t)$, whereupon $C(j)$ is computed for each j using the formula in Eq. (16). Thus we can use Eq. (17) to compute ${}_n\|\hat{K_1}\|_p$ with arbitrary accuracy, where K_1 is almost equal to K. We have found that with $T = 10000$ one can generally compute ${}_n\|\hat{K_1}\|_p$ quickly.

In performing these numerical optimizations, we have found that "good" kernels $K(x) \in \mathbb{K}$ have a very negative slope at $x = \frac{1}{4}^+$ (e.g., see Figure 4). Viewing graphs of these numerically optimized kernels suggests that functions of the form

$$K_{d_1,d_2}(x) = \begin{cases}1 & |x| \le \frac{1}{4},\\ 1-(1-(4(\frac{1}{2}-x))^{d_1})^{d_2} & \frac{1}{4} < |x| \le \frac{1}{2},\end{cases}$$

which have slope $-\infty$ at $x = \frac{1}{4}^+$, may be very good. (Note that the graph of $K_{2,1/2}(x)$ between $\frac{1}{4}$ and $\frac{3}{4}$ is the lower half of an ellipse.) More good candidates are functions of the form

$$K_{e_1,e_2,e_3}(x) = \begin{cases}1 & |x| \le \frac{1}{4},\\ \left(\frac{2}{\pi}\tan^{-1}\left(\frac{(1-2x)^{e_1}}{(4x-1)^{e_2}}\right)\right)^{e_3} & \frac{1}{4} < |x| \le \frac{1}{2},\end{cases}$$

where e_1, e_2, and e_3 are positive. We have used a function of the form K_{d_1,d_2} in the proof of Corollary 4.8 and a function of the form K_{e_1,e_2,e_3} in the proof of Corollary 4.3.

4.8 A Lower Bound for $\Delta(\frac{1}{2})$

We begin with a fundamental relationship between $\Re\hat{f}(1)$ and $\Re\hat{f}(2)$.

Lemma 4.15. *Let f be a pdf supported on $[-\frac{1}{4}, \frac{1}{4}]$. Then*

$$2\big(\Re\hat{f}(1)\big)^2 - 1 \le \Re\hat{f}(2) \le 2(\Re\hat{f}(1)) - 1.$$

Proof. Since $L_2(x) := 2\cos(2\pi x) - \cos(4\pi x)$ is at least 1 for $-\frac{1}{4} \le x \le \frac{1}{4}$, we have

$$1 \le \int f(x)L(x)\,dx = \sum_{j=-2}^{2}\hat{f}(j)\hat{L}(-j) = 2(\Re\hat{f}(1)) - \Re\hat{f}(2).$$

Rearranging, we arrive at $\Re\hat{f}(2) \le 2(\Re\hat{f}(1)) - 1$.

We give two proofs of the inequality $2\big(\Re\hat{f}(1)\big)^2 - 1 \le \Re(\hat{f}(2))$, each with its own advantages.

First proof: Since $f(x)$ and $\cos(2\pi x)$ are both nonnegative on $[-\frac{1}{4}, \frac{1}{4}]$, the Cauchy-Schwartz inequality gives

$$\begin{aligned}
2(\Re\hat{f}(1))^2 - 1 &= 2\left(\int_{-1/4}^{1/4} f(x)\cos(2\pi x)\,dx\right)^2 - \int_{-1/4}^{1/4} f(x)\,dx \\
&\le 2\left(\int_{-1/4}^{1/4} f(x)\,dx\right)\left(\int_{-1/4}^{1/4} f(x)\left(\cos(2\pi x)\right)^2\,dx\right) - \int_{-1/4}^{1/4} f(x)\,dx \\
&= 2\int_{-1/4}^{1/4} f(x)\left(\cos(2\pi x)\right)^2\,dx - \int_{-1/4}^{1/4} f(x)\,dx \\
&= \int_{-1/4}^{1/4} f(x)\left(2\cos^2(2\pi x) - 1\right)\,dx \\
&= \int_{-1/4}^{1/4} f(x)\cos(4\pi x)\,dx = \Re\hat{f}(2).
\end{aligned}$$

Second proof: Set $L_b(x) = b\cos(2\pi x) - \cos(4\pi x)$ (with $b \ge 0$) and observe that for $-\frac{1}{4} \le x \le \frac{1}{4}$, we have $L_b(x) \le 1 + \frac{b^2}{8}$. Thus

$$1 + \tfrac{b^2}{8} \ge \int f(x)L_b(x)\,dx = \sum_{j=-2}^{2} \hat{f}(j)\hat{L_b}(-j) = b\Re\hat{f}(1) - \Re\hat{f}(2).$$

Rearranging, we arrive at $\Re\hat{f}(2) \ge b(\Re\hat{f}(1)) - 1 - \frac{b^2}{8}$. Setting $b = 4\Re\hat{f}(1)$, we find that $\Re\hat{f}(2) \ge 2(\Re\hat{f}(1))^2 - 1$. ∎

The first proof may be adapted to also give $4(\Re\hat{f}(1))^3 - 3\Re\hat{f}(1) \le \Re\hat{f}(3)$. The proof does not immediately extend to higher coefficients. The second proof can be strengthened with the additional hypothesis that f be an nif. We take advantage of this in Proposition 4.16.

From the inequality $\Re\hat{f}(2) \le 2\Re\hat{f}(1) - 1$ (Lemma 4.15) one easily computes that $\max\{|\hat{f}(1)|, |\hat{f}(2)|\} \ge \frac{1}{3}$, and with Lemma 4.11 this gives

$$\frac{1}{9} \le \frac{\|f * f\|_\infty}{\pi}\sin\left(\frac{\pi}{\|f * f\|_\infty}\right).$$

This yields $\|f * f\|_\infty \ge 1.11$, a non-trivial bound. If one assumes that f is an nif supported on a subset of $[-\frac{1}{4}, \frac{1}{4}]$ with large measure, then one can do much better than Lemma 4.15. The following proposition establishes the lower bound on $\Delta(\varepsilon)$ given in Theorem 1.1(iii).

Proposition 4.16. *Let f be an nif supported on a subset of $[-\frac{1}{4}, \frac{1}{4}]$ with measure $\varepsilon/2$. Then*

$$\|f * f\|_\infty \ge 1.1092 + 0.176158\,\varepsilon$$

and consequently

$$\Delta(\varepsilon) \ge 0.5546\varepsilon^2 + 0.088079\varepsilon^3.$$

Proof. For $\varepsilon \geq \frac{5}{8}$, this proposition is weaker than Lemma 3.1, and for $\varepsilon \leq \frac{3}{8}$ it is weaker than Proposition 4.12, so we restrict our attention to $\frac{3}{8} < \varepsilon < \frac{5}{8}$.

Let $b > -1$ be a parameter and set $L_b(x) := \cos(4\pi x) - b\cos(2\pi x)$. If we define $F := \max\{\Re \hat{f}(1), -\Re \hat{f}(2)\}$, then

$$\int f(x) L_b(x)\, dx = \Re \hat{f}(2) - b \Re \hat{f}(1) \geq -(b+1)F$$

on the one hand, and

$$\int f(x) L_b(x)\, dx \leq \int f^{\mathrm{sdr}}(x) {L_b}^{\mathrm{sdr}}(x)\, dx = \int_{-\varepsilon/4}^{\varepsilon/4} \tfrac{2}{\varepsilon}\, {L_b}^{\mathrm{sdr}}(x)\, dx$$

on the other, where ${L_b}^{\mathrm{sdr}}(x)$ is the symmetric decreasing rearrangement of $K_b(x)$ on the interval $[-\frac{1}{4}, \frac{1}{4}]$. Thus

$$F \geq \frac{-1}{b+1} \frac{2}{\varepsilon} \int_{-\varepsilon/4}^{\varepsilon/4} {L_b}^{\mathrm{sdr}}(x)\, dx.$$

The right-hand side may be computed explicitly as a function of ε and b and then the value of b chosen in terms of ε to maximize the resulting expression. One finds that for $\varepsilon < \frac{5}{8}$, the optimal choice of b lies in the interval $2 < b < 4$, and the resulting lower bound for F is

$$F \geq \frac{3\cos(\frac{\pi\varepsilon}{4}) + \sin(\frac{\pi\varepsilon}{4}) - \sqrt{3 + 4\cos(\frac{\pi\varepsilon}{2}) + 2\cos(\pi\varepsilon) - \sin(\frac{\pi\varepsilon}{2})}}{\pi\varepsilon \cos(\frac{\pi\varepsilon}{4}) + \pi\varepsilon \sin(\frac{\pi\varepsilon}{4})}.$$

From Lemma 4.11 we know that $F^2 \leq \frac{\|f*f\|_\infty}{\pi} \sin\left(\frac{\pi}{\|f*f\|_\infty}\right)$. We compare these bounds on F to conclude the proof. Specifically,

$$F^2 \leq \frac{\|f*f\|_\infty}{\pi} \sin\left(\frac{\pi}{\|f*f\|_\infty}\right) \leq \frac{3}{5\pi} + \frac{\left(6 + 5\sqrt{3}\pi\right)\left(\|f*f\|_\infty - \frac{6}{5}\right)}{12\pi}, \tag{18}$$

where the expression on the right-hand side of this equation is from the Taylor expansion of $\frac{x}{\pi}\sin(\frac{\pi}{x})$ at $x_0 = \frac{6}{5}$, and

$$\begin{aligned} F^2 &\geq \left(\frac{3\cos(\frac{\pi\varepsilon}{4}) + \sin(\frac{\pi\varepsilon}{4}) - \sqrt{3 + 4\cos(\frac{\pi\varepsilon}{2}) + 2\cos(\pi\varepsilon) - \sin(\frac{\pi\varepsilon}{2})}}{\pi\varepsilon \cos(\frac{\pi\varepsilon}{4}) + \pi\varepsilon \sin(\frac{\pi\varepsilon}{4})}\right)^2 \\ &\geq \frac{-8\left(-3 - \sqrt{2} + \sqrt{3} + \sqrt{6}\right)}{\pi^2} \\ &\quad + \frac{\left(96\left(-3 - \sqrt{2} + \sqrt{3} + \sqrt{6}\right) - 4\left(9\sqrt{2} - 10\sqrt{3} + \sqrt{6}\right)\pi\right)\left(\varepsilon - \frac{1}{2}\right)}{3\pi^2}, \end{aligned} \tag{19}$$

where the expression on the right-hand side is from the Taylor expansion of the middle expression at $\varepsilon_0 = \frac{1}{2}$. Comparing Eqs. (18) and (19) gives a lower bound on $\|f*f\|_\infty$, say $\|f*f\|_\infty \geq c_1 + c_2\varepsilon$ with certain constants c_1, c_2. It is easily checked that $c_1 > 1.1092$ and $c_2 > 0.176158$, concluding the proof of the first asserted inequality. The second inequality then follows from Lemma 4.1. ∎

5 Upper Bounds for $R(g, n)$ Arising from $\Delta(\varepsilon)$

5.1 Inequalities Relating $\Delta(\varepsilon)$ and $R(g, n)$

A symmetric set consists of pairs (x, y) all with a fixed midpoint $c = \frac{x+y}{2}$. If there are few pairs in $E \times E$ with a given sum $2c$, then there will be no large symmetric subset of E with center c. We take advantage of the constructions of large integer sets whose pairwise sums repeat at most g times to construct large real subsets of $[0, 1)$ with no large symmetric subsets. Recall the definition (5) of the function $R(g, n)$.

Proposition 5.1. *For any integers $n \geq g \geq 1$, we have $\Delta(\frac{R(g,n)}{n}) \leq \frac{g}{n}$.*

Proof. Let $S \subseteq \{1, 2, \ldots, n\}$ be a $B^*[g]$ set with $|S| = R(g, n)$. Let

$$A(S) := \bigcup_{s \in S} [\frac{s-1}{n}, \frac{s}{n})$$

as in Eq. (4); it suffices to show that the largest symmetric subset of $A(S)$ has measure at most $\frac{g}{n}$. Notice that the set $A(S)$ is a finite union of intervals, and so the function $\lambda\big(A(S) \cap (2c - A(S))\big)$, which gives the measure of the largest symmetric subset of $A(S)$ with center c, is piecewise linear. (Figure 5.1 contains a typical example of the set $A(S)$ portrayed in dark gray below the c-axis, together with the function $\lambda\big(A(S) \cap (2c - A(S))\big)$ shown as the upper boundary of the light gray region above the c-axis, for $S = \{1, 2, 3, 5, 8, 13\}$.) Without loss of generality, therefore, we may restrict our attention to those symmetric subsets of $A(S)$ whose center c is the midpoint of endpoints of any two intervals $\big(\frac{s-1}{n}, \frac{s}{n}\big)$. In other words, we may assume that $2nc \in \mathbb{Z}$.

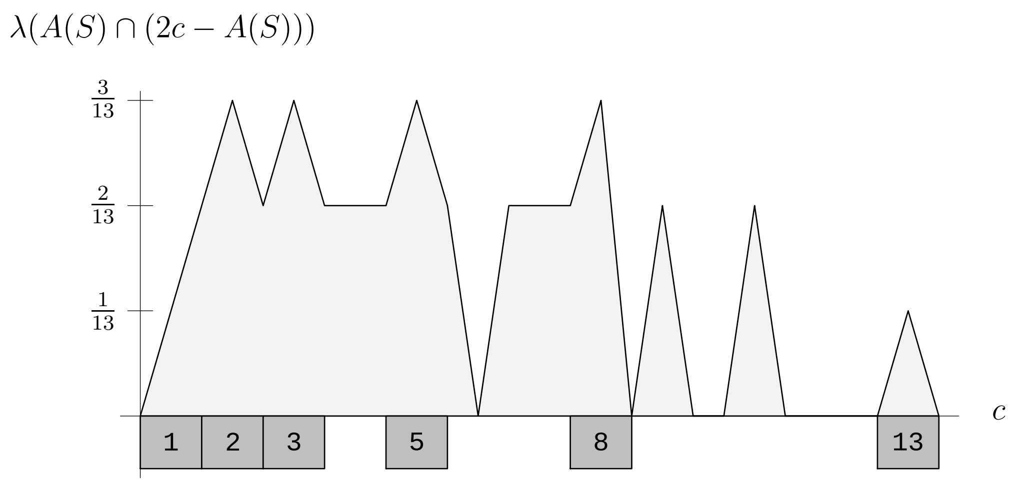

Figure 5: $A(S)$, and the function $\lambda(A(S) \cap (2c - A(S)))$, with $S = \{1, 2, 3, 5, 8, 13\}$

Suppose u and v are elements of $A(S)$ such that $\frac{u+v}{2} = c$. Write $u = \frac{s_1}{n} - \frac{1}{2n} + x$ and $v = \frac{s_2}{n} - \frac{1}{2n} + y$ for integers $s_1, s_2 \in S$ and real numbers x, y satisfying $|x|, |y| < \frac{1}{2n}$. (We may ignore the possibility that nu or nv is an integer, since this is a measure-zero event for any fixed c.) Then $2nc = n(u+v) = s_1 + s_2 - 1 + n(x+y)$, and since $2nc$, s_1, and s_2 are all integers, we see that $n(x+y)$ is also an integer. But $|n(x+y)| < 1$, so $x + y = 0$ and $s_1 + s_2 = 2nc + 1$.

Since S is a $B^*[g]$ set, there are at most g solutions (s_1, s_2) to the equation $s_1 + s_2 = 2nc+1$. If it happens that $s_1 = s_2$, the interval $\left(\frac{s_1-1}{n}, \frac{s_1}{n}\right)$ (a set of measure $\frac{1}{n}$) is contributed to the symmetric subset with center c. Otherwise, the set $\left(\frac{s_1-1}{n}, \frac{s_1}{n}\right) \cup \left(\frac{s_2-1}{n}, \frac{s_2}{n}\right)$ (a set of measure $\frac{2}{n}$) is contributed to the symmetric subset with center c, but this counts for the two solutions (s_1, s_2) and (s_2, s_1). In total, then, the largest symmetric subset having center c has measure at most $\frac{g}{n}$. This establishes the theorem. ∎

Using Proposition 5.1, we can translate lower bounds on $\Delta(\varepsilon)$ into upper bounds on $R(g,n)$, as in Corollary 5.2.

Corollary 5.2. *If $\delta \le \inf_{0<\varepsilon<1} \Delta(\varepsilon)/\varepsilon^2$, then $R(g,n) \le \delta^{-1/2}\sqrt{gn}$ for all $n \ge g \ge 1$.*

We remark that we may take $\delta = 0.591389$ by Proposition 4.6, and so this corollary implies that $R(g,n) \le 1.30036\sqrt{gn}$, which is one of the assertions of Theorem 1.2.

Proof. Combining the hypothesized lower bound $\Delta(\varepsilon) \ge \delta\varepsilon^2$ with Proposition 5.1, we find that

$$\delta\left(\frac{R(g,n)}{n}\right)^2 \le \Delta\left(\frac{R(g,n)}{n}\right) \le \frac{g}{n}$$

which is equivalent to $R(g,n) \le \delta^{-1/2}\sqrt{gn}$. ∎

We have been unable to prove or disprove that

$$\lim_{g\to\infty} \lim_{n\to\infty} \frac{R(g,n)}{\sqrt{gn}} = \left(\inf_{0<\varepsilon<1} \frac{\Delta(\varepsilon)}{\varepsilon^2}\right)^{-1/2},$$

i.e., that Corollary 5.2 is best possible as $g \to \infty$. At any rate, for small g it is possible to do better by taking advantage of the shape of the set $A(S)$ used in the proof of Proposition 5.1. This is the subject of Section 5.2.

Proposition 5.1 provides a one-sided inequality linking $\Delta(\varepsilon)$ and $R(g,n)$. It will also be useful for us to prove a theoretical result showing that the problems of determining the asymptotics of the two functions are, in a weak sense, equivalent. In particular, the following proposition implies that the trivial lower bound $\Delta(\varepsilon) \ge \frac{1}{2}\varepsilon^2$ and the trivial upper bound $R(g,n) \le \sqrt{2gn}$ are actually equivalent. Further, any nontrivial lower bound on $\Delta(\varepsilon)$ gives a nontrivial upper bound on $R(g,n)$, and vice versa.

Proposition 5.3. $\Delta(\varepsilon) = \inf\{\frac{g}{n} \colon n \ge g \ge 1,\ \frac{R(g,n)}{n} \ge \varepsilon\}$ *for all* $0 \le \varepsilon \le 1$.

Proof. That $\Delta(\varepsilon)$ is bounded above by the right-hand side follows immediately from Proposition 5.1 and the fact that Δ is an increasing function. For the complementary inequality, let $S \subseteq [0,1)$ with $\lambda(S) = \varepsilon$. There exists a finite union T of open intervals such that $\lambda(S \diamond T) < \eta$, and it is easily seen that T can be chosen to meet the following criteria: $T \subseteq [0,1)$, the endpoints of the finitely many intervals comprising T are rational, and $\lambda(T) > \varepsilon$. Choosing a common denominator n for the endpoints of the intervals comprising

T, we may write $T = \bigcup_{m\in M} \left[\frac{m-1}{n}, \frac{m}{n}\right)$ (up to a finite set of points) for some set of integers $M \subseteq \{1, \ldots, n\}$; most likely we have greatly increased the number of intervals comprising T by writing it in this manner, and M contains many consecutive integers. Let g be the maximal number of solutions $(m_1, m_2) \in M \times M$ to $m_1 + m_2 = k$ as k varies over all integers, so that M is a $B^*[g]$ set and thus $|M| \leq R(g, n)$ by the definition of R. It follows that $\varepsilon < \lambda(T) = |M|/n \leq R(g, n)/n$. Now T is exactly the set $A(M)$ as defined in Eq. (4); hence $D(T) = \frac{g}{n}$ as we saw in the proof of Proposition 5.1. Therefore by Lemma 3.4,

$$D(S) \geq D(T) - 2\eta = \tfrac{g}{n} - 2\eta \geq \inf\{\tfrac{g}{n} \colon n \geq g \geq 1,\ \tfrac{R(g,n)}{n} \geq \varepsilon\} - 2\eta.$$

Taking the infimum over appropriate sets S and noting that $\eta > 0$ was arbitrary, we derive the desired inequality $\Delta(\varepsilon) \geq \inf\{\frac{g}{n} \colon n \geq g \geq 1,\ \frac{R(g,n)}{n} \geq \varepsilon\}$. ■

5.2 Upper Bounds on $R(g, n)$ and $\overline{\rho}(g)$

The bulk of Theorem 1.2 follows immediately from Proposition 5.4, which is proved in this section.

Let $S \subseteq \{1, 2, \ldots, n\}$ be a $B^*[g]$ set. We call the function

$$f(x) := \begin{cases} \frac{2n}{|S|}, & \frac{s-1}{2n} - \frac{1}{4} \leq x < \frac{s}{2n} - \frac{1}{4} \text{ for some } s \in S, \\ 0, & \text{otherwise} \end{cases}$$

the nif corresponding to S. We think of S, and consequently f, as depending on n (many readers may prefer that we write S_n and f_n, but this is neither customary nor sufficiently brief). For example, when we write "Let $S \subseteq \{1, 2, \ldots, n\}$ be a $B^*[g]$ set with $|S| > 0.7\sqrt{gn}$", we mean "For each $n \in \mathbb{Z}^+$, let $S = S_n \subseteq \{1, 2, \ldots, n\}$ be a $B^*[g]$ set with $|S| = |S_n| > 0.7\sqrt{gn}$."

Note that f is piecewise constant, so that $f * f$ is piecewise linear. Furthermore, the corners of $f * f$ are only at the points $\frac{1}{4n}\mathbb{Z}$, and if we define

$$r(t) := \#\left\{(x, y) \in S^2 \colon x + y = t\right\}, \tag{20}$$

then $f * f(\frac{k}{4n}) = \frac{2nr(k+n+1)}{|S|^2}$. In particular, S is a $B^*[g]$ set if and only if $\|f * f\|_\infty \leq \frac{2gn}{|S|^2}$.

Proposition 5.4. *Let* $\phi := 1.14915$ *be the constant appearing in Corollary 4.7, and let* $\theta_0 := 21.922911$, $\theta_1 := -33.711941$, *and* $\theta_2 := 13.676987$ *be the constants appearing in Lemma 4.13. Then:*

i. $R(g, n) \leq \sqrt{2/\phi}\sqrt{(g-1)n} + \frac{1}{3}$ *for* $n \geq g \geq 2$ *and* g *odd;*

ii. $\overline{\rho}(g) \leq \sqrt{2/\phi}\left(1 - \frac{1}{g\sqrt{3}}\right)^{1/2}$ *for* $g \geq 2$, *and if moreover* $\overline{\rho}(g) \geq 1.275237$ *then*

$$\overline{\rho}(g) \leq \frac{2\sqrt{3\theta_2 g}}{\left(3(1-\theta_1)g - \sqrt{3} - \sqrt{\left(3(1-\theta_1)g - \sqrt{3}\right)^2 - 36\theta_0\theta_2 g^2}\right)^{1/2}};$$

iii. If $g \geq 2$ is odd, then $\overline{\rho}(g) \leq \sqrt{2/\phi}\left(1-\frac{1+1/\sqrt{3}}{g}\right)^{1/2}$, and if moreover $\overline{\rho}(g) \geq 1.275237$, then

$$\overline{\rho}(g) \leq \frac{2\sqrt{3\theta_2 g}}{\left(3\left(1-\theta_1\right)g-\sqrt{3}-3-\sqrt{\left(3\left(1-\theta_1\right)g-\sqrt{3}-3\right)^2-36\theta_0\theta_2 g^2}\right)^{1/2}}.$$

We begin with a simple lemma.

Lemma 5.5. *If h, p, q are nonnegative functions with $h = p + q$ and $\|h\|_\infty \geq \|p\|_\infty + \|q\|_\infty$, then*

$$\|h\|_\infty \geq \frac{\|h\|_2^2 - \|p\|_2^2}{\|h\|_1 + \|p\|_1} + \|p\|_\infty.$$

Proof. We have

$$\begin{aligned}
\|h\|_2^2 &= \|p+q\|_2^2 \\
&= \|p\|_2^2 + \|(2p+q)q\|_1 \\
&\leq \|p\|_2^2 + \|2p+q\|_1\|q\|_\infty \\
&= \|p\|_2^2 + (\|h\|_1 + \|p\|_1)\|q\|_\infty \\
&\leq \|p\|_2^2 + (\|h\|_1 + \|p\|_1)(\|h\|_\infty - \|p\|_\infty).
\end{aligned}$$

∎

We will use this lemma with $h = f * f$ (f a pdf), and p chosen so that $\|p\|_1 \to 0$, $\|p\|_2 \to 0$, $\|p\|_\infty \not\to 0$. In this case, we have the inequality

$$\|f * f\|_\infty \gtrsim \|f * f\|_2^2 + \|p\|_\infty$$

which is stronger than Hölder's Inequality: $\|f * f\|_\infty \geq \|f * f\|_2^2$.

Proof of Proposition 5.4(i). The idea of the proof is that even though $f * f$ might take values near $\frac{2gn}{|S|^2}$, it does so only on a set of small measure, and away from that small set it is bounded by $\frac{2(g-1)n}{|S|^2}$. If the pair (s_1, s_2) contributes to $r(k)$, then so does (s_2, s_1), and therefore $r(k)$ is odd if and only if $k \in \{2s \colon s \in S\}$. There are only $|S|$ such integers k, and no two are consecutive. Consequently, if $f * f(\frac{k}{4n}) = \frac{2gn}{|S|^2}$, then $f * f(\frac{k-1}{4n}) \leq \frac{2(g-1)n}{|S|^2}$ and $f * f(\frac{k+1}{4n}) \leq \frac{2(g-1)n}{|S|^2}$.

We put this idea into effect by writing $f * f(x) = p(x) + q(x)$, where p represents the small contribution to $r(k)$ of pairs (s, s) and q is the remaining majority of $f * f$. More precisely, let $p(x) := \sum_{s \in S} T(x - \frac{s-1}{n} - \frac{1}{2})$, where $T(x)$ is the "tent function"

$$T(x) := \begin{cases} \frac{2n}{|S|^2}(1 - 2n|x|), & \text{if } |x| \leq \frac{1}{2n}, \\ 0, & \text{otherwise,} \end{cases}$$

and let $q(x) := f * f(x) - p(x)$. (See Figure 6 for an illustration of a typical example, again using the $B^*[3]$ set $S = \{1, 2, 3, 5, 8, 13\}$.) Because of our judicious choice of the peaks of p,

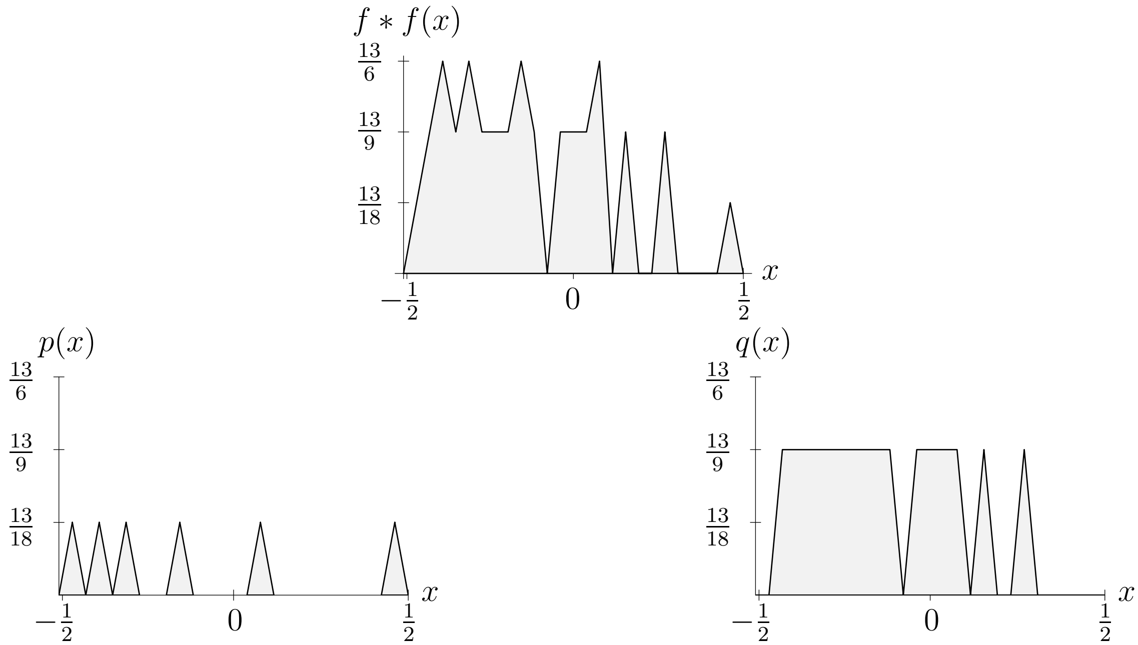

Figure 6: The decomposition $f * f(x) = p(x) + q(x)$, where f is the nif corresponding to $S = \{1, 2, 3, 5, 8, 13\}$

both p and q are nonnegative and $\|f * f\|_\infty \geq \|q\|_\infty + \|p\|_\infty$. We compute directly from the definition of p that $\|p\|_1 = |S| \|T\|_1 = \frac{1}{2|S|}$, $\|p\|_2^2 = |S| \|T\|_2^2 = \frac{2n/3}{|S|^3}$, and $\|p\|_\infty = \|T\|_\infty = \frac{2n}{|S|^2}$.

Lemma 5.5 with $h = f * f$ gives

$$\|f * f\|_\infty \geq \frac{\|f * f\|_2^2 - \|p\|_2^2}{1 + \|p\|_1} + \|p\|_\infty. \tag{21}$$

Using the inequality $\|f * f\|_2^2 \geq \phi$ from Corollary 4.7, the inequality $\|f * f\|_\infty \leq \frac{2gn}{|S|^2}$, and the above computations of $\|p\|_1$, $\|p\|_2$, and $\|p\|_\infty$, we can deduce from Eq. (21) that

$$|S|^2 \leq \frac{2(g-1)n}{\phi} + \frac{(g-1)n + 2n/3}{\phi |S|} \leq \frac{2(g-1)n}{\phi} + \frac{gn}{\phi |S|},$$

from which it follows that $|S| \leq \sqrt{\frac{2(g-1)n}{\phi}} + \frac{1}{3}$. ■

Lemma 5.6. *If f is the nif corresponding to $S \subseteq \{1, 2, \ldots, n\}$, and $m' \to \infty$ with $m' = o\left(\frac{n}{|S|}\right)$, then*

$$\sum_{|j| > m'} |\hat{f}(j)|^4 \geq \frac{2}{\sqrt{3}} \frac{n}{|S|^2} - o(1).$$

Proof. Note that for $j \neq 0$,

$$\begin{aligned}
j\hat{f}(j) &= j \int f(x) e^{-2\pi i j x}\, dx \\
&= \sum_{s \in S} \|f\|_\infty\, j \int_{\frac{s-1}{2n}-\frac{1}{4}}^{\frac{s}{2n}-\frac{1}{4}} e^{-2\pi i j x}\, dx \\
&= \sum_{s \in S} \|f\|_\infty \frac{j}{-2\pi i j} \left(e^{-2\pi i j (\frac{s}{2n}-\frac{1}{4})} - e^{-2\pi i j (\frac{s-1}{2n}-\frac{1}{4})} \right) \\
&= \frac{-\|f\|_\infty}{2\pi i} \sum_{s \in S} \left(e^{-2\pi i j (\frac{s}{2n}-\frac{1}{4})} - e^{-2\pi i j (\frac{s-1}{2n}-\frac{1}{4})} \right).
\end{aligned}$$

In particular, if we set $c(j) := j\,|\hat{f}(j)|$, then $c(j) = c(j+2n)$ (provided neither j nor $j+2n$ is zero) and $c(j) = c(-j)$. Exploiting this periodicity is the heart of this lemma.

We define

$$\begin{aligned}
J :=& \{m'+1, m'+2, \ldots, n\} \cup \{2n-m', 2n-m'+1, \ldots, 2n-1\} \\
J' :=& [m'+1, m'+2n] \setminus J,
\end{aligned}$$

and for $j \in J$ we set

$$j' = \begin{cases} 2n-j, & m'+1 \leq j < n; \\ 2n, & j = n; \\ 4n-j, & 2n-m' \leq j < 2n. \end{cases}$$

Observe that $c(j) = c(j')$ and $J' = \{j' \colon j \in J\}$. Also, for almost all $j \in J$ we have $j' = 2n-j$, since $m' = o(n)$.

For all $p > 1$,

$$\begin{aligned}
\sum_{|j|>m'} |\hat{f}(j)|^p &= 2 \sum_{j=m'+1}^{\infty} |\hat{f}(j)|^p \\
&= 2 \sum_{j=m'+1}^{\infty} \left(\frac{c(j)}{j} \right)^p \\
&= 2 \sum_{j=m'+1}^{m'+2n} \sum_{k=0}^{\infty} \left(\frac{c(j)}{j + 2n \cdot k} \right)^p \\
&= 2 \sum_{j=m'+1}^{m'+2n} \left(\frac{c(j)}{2n} \right)^p \zeta(p, \tfrac{j}{2n}) \\
&= 2 \sum_{j \in J} \left(\frac{c(j)}{2n} \right)^p \left(\zeta(p, \tfrac{j}{2n}) + \zeta(p, \tfrac{j'}{2n}) \right),
\end{aligned} \tag{22}$$

where $\zeta(s,a) := \sum_{k=0}^{\infty} (k+a)^{-s}$ is the Hurwitz zeta function.

We shall use Cauchy's inequality in the form $\sum a_j^2 \geq \frac{(\sum a_j b_j)^2}{\sum b_j^2}$, with

$$a_j = \left(\frac{c(j)}{2n} \right)^2 \left(\zeta(4, \tfrac{j}{2n}) + \zeta(4, \tfrac{j'}{2n}) \right)^{1/2}$$

and

$$b_j = \frac{\zeta(2, \frac{j}{2n}) + \zeta(2, \frac{j'}{2n})}{\Big(\zeta(4, \frac{j}{2n}) + \zeta(4, \frac{j'}{2n})\Big)^{1/2}}.$$

For notational convenience set

$$Z(m', n) := \frac{1}{2n} \sum_{j \in J} b_j^2 = \frac{1}{2n} \sum_{j \in J} \frac{\Big(\zeta(2, \frac{j}{2n}) + \zeta(2, \frac{j'}{2n})\Big)^2}{\zeta(4, \frac{j}{2n}) + \zeta(4, \frac{j'}{2n})},$$

and note that since $m' = o(n)$,

$$\begin{aligned}
\lim_{n\to\infty} Z(m', n) &= \lim_{n\to\infty} \frac{1}{2n} \sum_{j=o(n)}^{n} \frac{\big(\zeta(2, \frac{j}{2n}) + \zeta(2, \frac{2n-j}{2n})\big)^2}{\zeta(4, \frac{j}{2n}) + \zeta(4, \frac{2n-j}{2n})} \\
&= \int_0^{1/2} \frac{(\zeta(2, a) + \zeta(2, 1-a))^2}{\zeta(4, a) + \zeta(4, 1-a)}\, da = \int_0^{1/2} \frac{3}{2 + \cos(2\pi a)}\, da = \frac{\sqrt{3}}{2}. \qquad (23)
\end{aligned}$$

Therefore from Eq. (22) with $q = 4$ and Cauchy's inequality,

$$\begin{aligned}
\sum_{|j|>m'} |\hat{f}(j)|^4 &= 2 \sum_{j \in J} \left(\frac{c(j)}{2n}\right)^4 \left(\zeta(4, \frac{j}{2n}) + \zeta(4, \frac{j'}{2n})\right) \\
&\geq 2 \frac{\left(\sum_{j\in J} \left(\frac{c(j)}{2n}\right)^2 \left(\zeta(2, \frac{j}{2n}) + \zeta(2, \frac{j'}{2n})\right)\right)^2}{\sum_{j\in J} \big(\zeta(2, \frac{j}{2n}) + \zeta(2, \frac{j'}{2n})\big)^2 \big(\zeta(4, \frac{j}{2n}) + \zeta(4, \frac{j'}{2n})\big)^{-1}} \\
&= 2 \frac{\left(\frac{1}{2} \sum_{|j|>m'} |\hat{f}(j)|^2\right)^2}{2n\, Z(m', n)} \\
&= \frac{\left(\|f\|_2^2 - \sum_{|j|\leq m'} |\hat{f}(j)|^2\right)^2}{4n\, Z(m', n)}, \qquad (24)
\end{aligned}$$

the last equality following from Parseval's formula. Trivially, $|\hat{f}(j)| \leq \hat{f}(0) = 1$, and since f is the nif corresponding to S, we have $\|f\|_2^2 = \|f\|_\infty^2 \lambda(\mathrm{supp}(f)) = \|f\|_\infty = \frac{2n}{|S|}$. Since $m' = o(n/|S|)$, Eq. (24) becomes

$$\sum_{|j|>m'} |\hat{f}(j)|^4 \geq \frac{1}{4n} \frac{\left(\frac{2n}{|S|} - 2m' - 1\right)^2}{Z(m', n)} = \frac{2}{\sqrt{3}} \frac{n}{|S|^2} - o(1).$$

■

Proof of Proposition 5.4(ii). Let $S \subseteq \{1, 2, \ldots, n\}$ be a $B^*[g]$ set with $|S| = R(g, n)$ and let f be the corresponding nif. We shall use the inequality

$$\|f * f\|_\infty \geq \|f * f\|_2^2 = \sum_j |\hat{f}(j)|^4 = \sum_{|j|\leq m'} |\hat{f}(j)|^4 + \sum_{|j|>m'} |\hat{f}(j)|^4.$$

We use Proposition 4.7 and Proposition 4.6 to bound the sum over small $|j|$, and we use Lemma 5.6 to bound the sum over large $|j|$.

We are now prepared to prove Proposition 5.4(ii). We have

$$\frac{2}{\overline{\rho}^2(g)} = \liminf_{n\to\infty} \frac{2gn}{|S|^2} \geq \liminf_{n\to\infty} \|f * f\|_\infty \geq \liminf_{n\to\infty} \|f * f\|_2^2$$
$$= \liminf_{n\to\infty} \sum_j |\hat{f}(j)|^4 = \liminf_{n\to\infty} \Bigg(\sum_{|j|\leq m'} |\hat{f}(j)|^4 + \sum_{|j|>m'} |\hat{f}(j)|^4 \Bigg).$$

We see from Corollary 4.7 and Lemma 5.6 that

$$\frac{2}{\overline{\rho}^2(g)} \geq \liminf_{n\to\infty} \Big(\phi - o(1) + \frac{2}{\sqrt{3}} \frac{n}{|S|^2} - o(1) \Big) = \phi + \frac{2}{\sqrt{3}} \frac{1}{\overline{\rho}^2(g) g}$$

Solving this inequality for $\overline{\rho}(g)$ yields

$$\overline{\rho}(g) \leq \sqrt{\frac{2}{\phi}} \left(1 - \frac{1}{g\sqrt{3}}\right)^{1/2}.$$

If $\overline{\rho}(g) \geq 1.275237$, then $\|f * f\|_\infty \leq 1.229837$ for infinitely many n, and we can use Lemma 4.13 instead of Corollary 4.7. Setting

$$F := \liminf_{n\to\infty} \|f * f\|_\infty \leq \liminf_{n\to\infty} \frac{2gn}{|S|^2},$$

we have

$$\begin{aligned}
F &\geq \Bigg(\liminf_{n\to\infty} \sum_{|j|\leq m'} |\hat{f}(j)|^4 + \sum_{|j|>m'} |\hat{f}(j)|^4 \Bigg) \\
&\geq \liminf_{n\to\infty} \Big(\theta_0 + \theta_1 \|f * f\|_\infty + \theta_2 \|f * f\|_\infty^2 - o(1) + \frac{2}{\sqrt{3}} \frac{n}{|S|^2} - o(1) \Big) \\
&\geq \theta_0 + \theta_1 F + \theta_2 F^2 + \frac{F}{g\sqrt{3}},
\end{aligned}$$

whence

$$\overline{\rho}(g) \leq \sqrt{\frac{2}{F}} \leq \frac{2\sqrt{3\theta_2 g}}{\left(3(1-\theta_1)g - \sqrt{3} - \sqrt{\left(3(1-\theta_1)g - \sqrt{3}\right)^2 - 36\theta_0\theta_2 g^2} \right)^{1/2}}.$$

■

Proof of Proposition 5.4(iii). We combine the ideas of parts (i) and (ii). As in the proof of Proposition 5.4(ii), we let $S \subseteq \{1, 2, \ldots, n\}$ be a $B^*[g]$ set with $|S| = R(g, n)$ and f be the corresponding nif.

If g is odd, then (defining p and q as in the proof of Proposition 5.4(i)) Eq. (21) is valid, and $\|p\|_1 \to 0, \|p\|_2 \to 0$, and $\|p\|_\infty = \frac{2n}{|S|^2} \to \frac{2}{\overline{\rho}^2(g)\, g}$.

We find

$$\begin{aligned}
\frac{2}{\overline{\rho}^2(g)} &= \lim_{n\to\infty} \frac{2gn}{|S|^2} \\
&= \lim_{n\to\infty} \|f*f\|_\infty \\
&\geq \lim_{n\to\infty} \frac{\|f*f\|_2^2 - \|p\|_2^2}{1+\|p\|_1} + \|p\|_\infty \\
&= \lim_{n\to\infty} \frac{\left(\sum_{|j|\leq m'} |\hat{f}(j)|^4 + \sum_{|j|>m'} |\hat{f}(j)|^4\right) - \|p\|_2^2}{1+\|p\|_1} + \|p\|_\infty \\
&= \lim_{n\to\infty} \frac{\left(\phi - o(1) + \frac{2}{\sqrt{3}}\frac{n}{|S|^2} - o(1)\right) - \|p\|_2^2}{1+\|p\|_1} + \|p\|_\infty \\
&= \phi + \frac{2}{\sqrt{3}}\frac{1}{\overline{\rho}^2(g)g} + \frac{2}{\overline{\rho}^2(g)g} \\
&= \phi + 2\frac{1+1/\sqrt{3}}{\overline{\rho}^2(g)g}.
\end{aligned}$$

Solving this inequality for $\overline{\rho}^2(g)$ yields

$$\overline{\rho}(g) \leq \sqrt{\frac{2}{\phi}}\left(1 - \frac{1+1/\sqrt{3}}{g}\right)^{1/2}.$$

If $\overline{\rho}(g) \geq 1.275237$, then $\|f*f\|_\infty \leq 1.229837$ for infinitely many n, and we can use Lemma 4.13. We have:

$$\begin{aligned}
\|p\|_1 &= \frac{1}{2|S|} \to 0 \\
\|p\|_2^2 &= \frac{2n/3}{|S|^3} \to 0 \\
\|p\|_\infty &= \frac{2n}{|S|^2} = \frac{1}{g}\|f*f\|_\infty \\
\sum_{|j|\leq m'} |\hat{f}(j)|^4 &\geq \theta_0 + \theta_1\|f*f\|_\infty + \theta_2\|f*f\|_\infty^2 - o(1) \\
\sum_{|j|>m'} |\hat{f}(j)|^4 &= \frac{2}{\sqrt{3}}\frac{n}{|S|^2} - o(1) \geq \frac{\|f*f\|_\infty}{g\sqrt{3}} - o(1).
\end{aligned}$$

Thus, again writing $F = \liminf_{n\to\infty}\|f*f\|_\infty$, we now know

$$\begin{aligned}
F &:= \liminf_{n\to\infty} \|f*f\|_\infty \\
&\geq \liminf_{n\to\infty} \frac{\left(\sum_{|j|\leq m'} |\hat{f}(j)|^4 + \sum_{|j|>m'} |\hat{f}(j)|^4\right) - \|p\|_2^2}{1+\|p\|_1} + \|p\|_\infty \\
&\geq \theta_0 + \theta_1 F + \theta_2 F^2 + \frac{F}{g\sqrt{3}} + \frac{F}{g}.
\end{aligned}$$

Isolating F, we obtain

$$F \geq \frac{\theta_1 + \frac{1+\sqrt{3}}{g\sqrt{3}} - 1 + \sqrt{\left(\theta_1 + \frac{1+\sqrt{3}}{g\sqrt{3}} - 1\right)^2 - 4\theta_0\theta_2}}{-2\theta_2},$$

and so

$$\overline{\rho}(g) \leq \sqrt{\frac{2}{F}} \leq \frac{2\sqrt{3\theta_2 g}}{\left(3\left(1-\theta_1\right)g - \sqrt{3} - 3 - \sqrt{\left(3\left(1-\theta_1\right)g - \sqrt{3} - 3\right)^2 - 36\theta_0\theta_2 g^2}\right)^{1/2}},$$

which concludes the proof. ■

5.3 Ubiquity of Repeated Sums in $B^*[g]$ Sets

The method of proof of Proposition 5.4(i) can be adapted to yield more information about the number of representations of integers as sums of pairs of elements from $B^*[g]$ sets. The following theorem gives a quantitative statement of the fact that, if S is a dense enough $B^*[g]$ set, then there is a substantial number of integers t such that $r(t)$ is large, where $r(t)$ is defined in Eq. (20).

Theorem 5.7. *Let $S \subseteq \{1, \ldots, n\}$ be a $B^*[g]$ set, and let $0 < L < g$ be a real number. The number of integers t such that $r(t) > L$ is at least*

$$\frac{\phi|S|^4 - 2Ln|S|^2}{n(g-L)(g+2L)},$$

where $\phi := 1.14915$ is the constant appearing in Corollary 4.7.

We give an illustration of this theorem after the proof.

Proof. Let f be the nif corresponding to S. We note again that $\|f * f\|_1 = \|f\|_1^2 = 1$ and $\|f * f\|_2^2 \geq \phi$ by Corollary 4.7.

Now let Q be the number of integers t between 1 and $2n$ such that $r(t) > L$. Our aim is to show that Q is large if $|S|$ is large enough and L is small enough. Define the functions $p(x)$ and $q(x)$ by

$$q(x) = \min\left\{f * f(x), \frac{2Ln}{|S|^2}\right\}, \quad p(x) = f * f(x) - q(x).$$

We have

$$\begin{aligned}\phi \leq \|f * f\|_2^2 = \|p + q\|_2^2 &= \|p\|_2^2 + \|(2p+q)q\|_1 \\ &\leq \|p\|_2^2 + \|2p+q\|_1\|q\|_\infty \\ &= \|p\|_2^2 + (1 + \|p\|_1)\|q\|_\infty, \qquad (25)\end{aligned}$$

where the latter inequality is by Hölder. By its definition, we have $\|q\|_\infty \le \frac{2Ln}{|S|^2}$. To bound $\|p\|_1$ and $\|p\|_2$, we look more closely at the function p.

Suppose $r(t+1), r(t+2), \ldots, r(t+K)$ are all greater than L, but that $r(t)$ and $r(t+K+1)$ are at most L. Then the graph of $f*f$ between $\frac{t-n-1}{4n}$ and $\frac{t+K-n}{4n}$ lies under the trapezoid with vertices $\left(\frac{t-n-1}{4n}, \frac{2Ln}{|S|^2}\right)$, $\left(\frac{t-n}{4n}, \frac{2gn}{|S|^2}\right)$, $\left(\frac{t+K-n-1}{4n}, \frac{2gn}{|S|^2}\right)$, and $\left(\frac{t+K-n}{4n}, \frac{2Ln}{|S|^2}\right)$. The area of this trapezoid, which bounds the contribution to $\|p\|_1$ from x in the interval $\left[\frac{t-n-1}{4n}, \frac{t+K-n}{4n}\right]$, equals $\frac{K(g-L)}{2|S|^2}$. Since the numbers K from all such intervals sum to Q, we have established the bound $\|p\|_1 \le \frac{Q(g-L)}{2|S|^2}$.

Similarly, the graph of $(f*f)^2$ between $\frac{t-n-1}{4n}$ and $\frac{t+K-n}{4n}$ lies under the trapezoid with vertices $\left(\frac{t-n-1}{4n}, \frac{4L^2n^2}{|S|^4}\right)$, $\left(\frac{t-n}{4n}, \frac{4g^2n^2}{|S|^4}\right)$, $\left(\frac{t+K-n-1}{4n}, \frac{4g^2n^2}{|S|^4}\right)$, and $\left(\frac{t+K-n}{4n}, \frac{4L^2n^2}{|S|^4}\right)$ (since the graph of $f*f$ is piecewise convex). The area of this trapezoid, which bounds the contribution to $\|p\|_2^2$ from x in the interval $\left[\frac{t-n-1}{4n}, \frac{t+K-n}{4n}\right]$, equals $\frac{Kn(g^2-L^2)}{|S|^4}$. Since the numbers K from all such intervals sum to Q, we have established the bound $\|p\|_2^2 \le \frac{Qn(g^2-L^2)}{|S|^4}$.

Inserting these bounds into Eq. (25), we conclude that

$$\phi \le \frac{Qn(g^2-L^2)}{|S|^4} + \left(1 + \frac{Q(g-L)}{2|S|^2}\right)\frac{2Ln}{|S|^2}$$

or equivalently

$$Q \ge \frac{\phi|S|^4 - 2Ln|S|^2}{n(g-L)(g+2L)} \tag{26}$$

as desired. ∎

We remark that, using the techniques from the proof of Proposition 5.4(ii), the constant 1.14915 in the statement of Theorem 5.7 can be replaced with 1.182778 (provided one is concerned only with the case $n \to \infty$).

To give an illustration of this theorem, define $\gamma = |S|/\sqrt{gn}$, $\alpha = L/g$, and $\kappa = Q/2n$. Then the inequality (26) becomes

$$\kappa \ge \frac{\gamma^2(0.574575\gamma^2 - \alpha)}{(1-\alpha)(1+2\alpha)}. \tag{27}$$

For example, take $\gamma = 0.7$ and $\alpha = 0.25$, so that the right-hand side of this inequality is greater than $0.0137382 > \frac{1}{73}$. Theorem 1.3 tells us that for every $g \ge 2$ except $g = 3$ and possibly $g = 5$ and $g = 7$, there exists a $B^*[g]$ set S contained in $\{1, \ldots, n\}$ with at least $0.7\sqrt{gn}$ elements, at least when n is large. For every such set S, the inequality (27) asserts that at least $\frac{2n}{73}$ of the integers t between 1 and $2n$ have at least $\frac{g}{4}$ representations $t = s_1 + s_2$ with $s_1, s_2 \in S$.

To give a basis for comparison for Theorem 5.7, we note that the simple argument

$$|S|^2 = \sum_t r(t) = \sum_{t\colon r(t)\le L} r(t) + \sum_{t\colon r(t)>L} r(t) \le (2n-Q)L + Qg$$

gives $Q \ge \frac{|S|^2 - 2nL}{g-L}$, which translates in the above notation into

$$\kappa \ge \frac{\gamma^2 - 2\alpha}{2-2\alpha}. \tag{28}$$

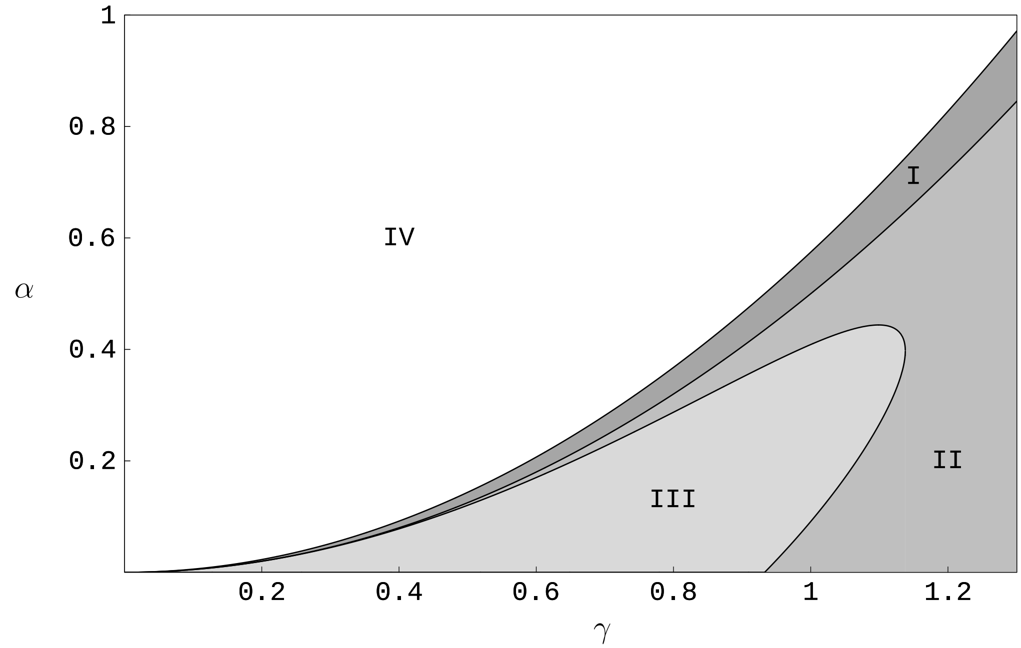

Figure 7: Comparing the bounds of Eq. (28) and Eq. (27)

In Figure 7, Region IV corresponds to the pairs (γ, α) for which neither Eq. (27) nor Eq. (28) gives a nontrivial bound on κ (the trivial bound is $\kappa \geq 0$). In Region I, which contains our example point $(0.7, 0.25)$, Eq. (27) gives a nontrivial bound while Eq. (28) does not. Eq. (28) gives a nontrivial bound in Regions II and III, but in Region II the bound in Eq. (27) is better. That is, the simple bound in Eq. (28) is better than Eq. (27) only in Region III.

5.4 The Uniform Distribution Hypothesis

We say that a sequence $(S_n)_{n=1}^{\infty}$ of sets of positive integers *becomes uniformly distributed* if the discrepancy of S_n goes to 0 as $n \to \infty$. That is, $(S_n)_{n=1}^{\infty}$ becomes uniformly distributed if

$$\limsup_{n\to\infty} \sup_{0\leq\alpha<\beta\leq 1} \left| \frac{|S_n \cap [\alpha M_n, \beta M_n]|}{|S_n|} - (\beta - \alpha) \right| = 0,$$

where M_n denotes the largest element of S_n.

It has long been known [Erd44, Cho44] that $R(2,n) \sim \sqrt{n}$. In 1991, Erdős and Freud [EF91] proved that if $S_n \subseteq \{1,\ldots,n\}$ is a sequence of $B^*[2]$ sets with $|S_n| \sim \sqrt{n}$, then (S_n) becomes uniformly distributed. We are led to make the following conjecture.

Conjecture 5.8. *Let $g \geq 2$ be an integer. Suppose that $S_n \subseteq \{1,\ldots,n\}$ is a sequence of $B^*[g]$ sets with $|S_n| \sim R(g,n)$. Then (S_n) becomes uniformly distributed.*

Independent of the conjecture, we are able to prove a strong result on the cardinality of uniformly distributed $B^*[g]$ sets.

Theorem 5.9. *Let $g \geq 2$ be an integer. Let $S_n \subseteq \{1,\ldots,n\}$ be a sequence of $B^*[g]$ sets that becomes uniformly distributed as $n \to \infty$. Then $|S_n| \lesssim 1.15988\sqrt{gn}$.*

Proof. Let f_n be the nif corresponding to S_n. Since the S_n become uniformly distributed, the functions f_n converge in measure to the function that is identically 2 on $[-\frac{1}{4}, \frac{1}{4}]$. Therefore

$$\hat{f}_n(1) \to \int_{-1/4}^{1/4} 2e^{-2\pi i x}\, dx = \frac{2}{\pi}.$$

Now, apply Lemma 4.11 to find $\liminf_n \|f_n * f_n\|_\infty \geq 1.486634$. Since $\|f_n * f_n\|_\infty = \frac{g}{2n}\left(\frac{2n}{|S_n|}\right)^2$, we have shown $|S_n| \lesssim 1.15988\sqrt{gn}$. ∎

In light of Theorem 5.9, we see that Conjecture 5.8 implies that $\overline{\rho}(g) \leq 1.15988$, improving Theorem 1.2 for all sufficiently large values of g.

6 Constructions of $B^*[g]$ Sets and Lower Bounds for $R(g, n)$

We begin by considering a modular version of $B^*[g]$ sets. A set S is a $B^*[g] \pmod n$ set if for any given m there are at most g ordered pairs $(s_1, s_2) \in S \times S$ with $s_1 + s_2 \equiv m \pmod n$ (equivalently, if the coefficients of the least-degree representative of $\left(\sum_{n \in S} z^n\right)^2 \pmod{z^n - 1}$ are bounded by g). For example, the set $\{0, 1, 2, 4\}$ is a $B^*[3] \pmod 7$ set, and $\{0, 1, 3, 7\}$ is a $B^*[2] \pmod{12}$ set. Note that $7 + 7 \equiv 1 + 1 \pmod{12}$, so that $\{0, 1, 3, 7\}$ is not a "modular Sidon set" as defined by some authors, e.g., [GS80] or [Guy94, Problem C10]. It is, however, the natural companion to the study of $B^*[g]$ sets, as evidenced by the clean forms of Propositions 6.4 and 6.9 below.

Just as we defined $R(g, n)$ to be the largest possible cardinality of a $B^*[g]$ set contained in $[0, n)$, we define $C(g, n)$ to be the largest possible cardinality of a $B^*[g] \pmod n$ set. The mnemonic is "R" for the $\mathbb{R}$eal problem and "C" for the Circular problem. After exhibiting some basic bounds for this new function $C(g, n)$ in the next section, we construct some explicit families of large $B^*[g] \pmod n$ sets in Section 6.2, which are used in turn to construct large $B^*[g]$ sets themselves in Section 6.4. We also demonstrate the existence of large $B^*[g] \pmod n$ sets via a probabilistic construction in Section 6.3, and we give a similar probabilistic construction of large $B^*[g]$ sets in Section 6.5. These constructions will yield lower bounds on the size of $B^*[g]$ sets, which we collect in Section 6.6.

6.1 Upper Bounds on $C(g, n)$

Since $B^*[g] \pmod n$ sets have not been rigorously developed in the literature, we begin this section by giving some simple *upper* bounds on $C(g, n)$.

We can obtain

$$C(g, n) \leq \begin{cases} \sqrt{gn} & g \text{ even}, \\ \sqrt{1 - \frac{1}{g}}\sqrt{gn} + 1 & g \text{ odd}, \end{cases}$$

from the pigeonhole principle as follows. There are $|S|^2$ pairs of elements from S, and there are just n possible values for the sum of two elements, and a possible value is realized at most g times. Thus $|S| \leq \sqrt{gn}$. The only way a sum can occur an odd number of times is if it is twice an element of S, so for odd g, $|S|^2 \leq (g - 1)n + |S|$.

For $g = 2$ we can be more precise: $\binom{C(2,n)}{2} \leq \left\lfloor \frac{n}{2} \right\rfloor$. To establish this, let $S \subseteq [0, n)$ be a witness and note that there are $\binom{|S|}{2}$ pairs of distinct elements from S, and each such pair s_1, s_2 leads to a pair of differences $\{s_1 - s_2, s_2 - s_1\} \in \{\{i, n - i\} \colon 1 \leq i < n\}$. If each of (s_1, s_2) and (s_3, s_4) is a pair of incongruent elements and if $s_1 - s_2 \equiv s_3 - s_4 \pmod n$, then $s_1 + s_4 \equiv s_4 + s_1 \equiv s_2 + s_3 \equiv s_3 + s_2 \pmod n$. The fact that S is a $B^*[2] \pmod n$ set forces $\{s_1, s_4\} \equiv \{s_2, s_3\} \pmod n$, and since $s_1 \not\equiv s_2 \pmod n$ by assumption we conclude that $s_1 \equiv s_3 \pmod n$ and $s_2 \equiv s_4 \pmod n$. Therefore distinct pairs of incongruent elements lead to distinct sets of differences, of which there are at most $\lfloor n/2 \rfloor$, establishing $\binom{C(2,n)}{2} \leq \left\lfloor \frac{n}{2} \right\rfloor$. This bound is actually achieved for $n = p^2 + p + 1$ when p is prime (see Proposition 6.1(iii) below) and implies that $C(2, n) < \sqrt{n} + 1$.

6.2 Explicit Constructions of $B^*[g] \pmod{n}$ Sets

We turn now to the problem of constructing large $B^*[g] \pmod{n}$ sets. The literature contains several examples of families of $B^*[2] \pmod{n}$ sets which we can generalize to families of $B^*[g] \pmod{n}$ sets. The following proposition collects several lower bounds for $C(g,x)$ corresponding to various constructions given in the proofs thereafter.

Proposition 6.1. *Let p be a prime power, and let $1 \le k < p$.*

i. If p is a prime, then $C(2k^2, p^2 - p) \ge k(p-1)$;

ii. $C(2k^2, p^2 - 1) \ge kp$;

iii. $C(2k^2, p^2 + p + 1) \ge kp + 1$.

The $k = 1$ cases of Proposition 6.1(i), (ii), and (iii) are due to Ruzsa [Ruz93], Bose [Bos42], and Singer [Sin38], respectively. Part (i) uses the existence of a primitive root modulo a prime p and the Chinese Remainder Theorem (the fact that the orders of the additive and multiplicative groups modulo p are relatively prime is important, which is why the construction does not generalize to all finite fields). Part (ii) uses the existence of the finite fields $\mathrm{GF}(p^t)$ with cyclic multiplicative groups and their vector space structures. The third part also uses a fact (Lemma 6.3) connecting the multiplicative group of a finite field to its vector space structure.

All three constructions make use of the following lemma. Although the lemma is a special case of unique factorization, we give here a simple, elementary proof of the special case that we require.

Lemma 6.2. *In any field, there are at most 2 ordered pairs (a,b) of solutions to the polynomial equation $x^2 - c_1x + c_2 = (x-a)(x-b)$.*

Proof. Suppose that there are three solutions (a_m, b_m), $1 \le m \le 3$. Obviously we have $a_1 + b_1 = c_1 = a_2 + b_2$ and $a_1b_1 = c_2 = a_2b_2$. This leads to $0 = a_1(a_1 + b_1 - a_2 - b_2) = a_1^2 + a_1b_1 - a_1a_2 - a_1b_2 = a_1^2 + a_2b_2 - a_1a_2 - a_1b_2 = (a_1 - a_2)(a_1 - b_2)$, whence $a_1 \in \{a_2, b_2\}$. If $a_1 = a_2$, then $b_1 = b_2$ and the three solutions are not distinct. Otherwise $a_1 = b_2$, and likewise $a_1 = b_3$, whence $b_2 = b_3$, again a contradiction. ■

Proof of Proposition 6.1(i). Let g be a primitive root modulo p. Using the Chinese Remainder Theorem, define $a_{t,i}$ for $1 \le t < p$ and $1 \le i \le k$ by the pair of congruences

$$a_{t,i} \equiv t \pmod{p-1} \quad \text{and} \quad a_{t,i} \equiv ig^t \pmod{p}. \tag{29}$$

Set $S_i := \{a_{t,i} : 1 \le t < p\}$; clearly $|S_i| = p - 1$. we shall show that $\bigcup_{i=1}^k S_i$ witnesses $C(2k^2, p(p-1)) \ge k(p-1)$.

Suppose $a_{m_1,i} = a_{m_2,j}$, with $m_1, m_2 \in [1,p)$. We have $m_1 \equiv a_{m_1,i} = a_{m_2,j} \equiv m_2 \pmod{p-1}$, so $m_1 = m_2$. Reducing the equation $a_{m_1,i} = a_{m_2,j}$ modulo p, we find $ig^{m_1} \equiv jg^{m_2} = jg^{m_1} \pmod{p}$, so $i = j$. Thus S_i and S_j are distinct for distinct i, j, and $\left|\bigcup_{i=1}^k S_i\right| = k(p-1)$.

First assume that, given any $i, j \in [1,k]$ and any $n \in [1, p(p-1)]$, there are at most two pairs $(r,v) \in S_i \times S_j$ with $r + v \equiv n \pmod{p(p-1)}$. Since there are k^2 such choices of i, j, and each such choice leads to at most two pairs with a given sum, this shows that each n arises as the sum of at most $2k^2$ pairs, concluding the proof.

Now suppose if possible that there are three pairs $(a_{r_m,i}, a_{v_m,j}) \in S_i \times S_j$ satisfying $a_{r_m,i} + a_{v_m,j} \equiv n \pmod{p(p-1)}$. Then obviously $a_{r_m,i} + a_{v_m,j} \equiv n \pmod{p}$. Also, since $a_{r_m,i} + a_{v_m,j} \equiv n \pmod{p-1}$, we have $a_{r_m,i} \cdot a_{v_m,j} = ig^{r_m} jg^{v_m} = ijg^{r_m+v_m} \equiv ijg^n \pmod{p}$. Thus the pairs $(a_{r_m,i}, a_{v_m,j})$ are solutions to $x^2 - nx + ijg^n \pmod{p}$. Applying Lemma 6.2, we find that two of the pairs $(a_{r_m,i}, a_{v_m,j})$ are equal, say $a_{r_1,i} \equiv a_{r_2,i} \pmod{p}$. By Eq. (29), $ig^{r_1} \equiv a_{r_1,i} \equiv a_{r_2,i} \equiv ig^{r_2} \pmod{p}$. Since g has order $p-1$, this tells us that $a_{r_1,i} \equiv r_1 \equiv r_2 \equiv a_{r_2,i} \pmod{p-1}$. Since $a_{r_1,i} \equiv a_{r_2,i} \pmod{p}$, $a_{r_1,i} \equiv a_{r_2,i} \pmod{p-1}$, and $a_{t,i} \in [1, p(p-1))$ (by definition), we see that $a_{r_1,i} = a_{r_2,i}$, whence the three pairs are not distinct. ■

Proof of Proposition 6.1(ii). Let θ generate the multiplicative group of the finite field $\mathrm{GF}(p^2)$, and observe that $\{1, \theta\}$ is a basis of $\mathrm{GF}(p^2)$ as a vector space over $\mathrm{GF}(p)$. For $i \in \mathrm{GF}(p)$, define

$$S_i := \{s' \in [1, p^2-1] : \theta^{s'} = i\theta + s, \quad s \in \mathrm{GF}(p)\}.$$

Bose [Bos42] showed that each S_i is a Sidon set. we shall show that $S = \bigcup_{i=1}^k S_i$ witnesses $C(2k^2, p^2-1) \geq kp$.

First, note that for each $s \in \mathrm{GF}(p)$, there is an integer $s' \in [1, p^2-1]$ with $\theta^{s'} = i\theta + s$, so that $|S_i| = p$. Since $1, \theta$ is a basis for $\mathrm{GF}(p^2)$ over $\mathrm{GF}(p)$, we know that $i\theta + s_1 = j\theta + s_2$ implies that $i = j$ and $s_1 = s_2$. In particular, if $i \neq j$, then $S_i \cap S_j = \emptyset$. Thus $|\bigcup_{i=1}^k S_i| = kp$.

It is sufficient to show that for each n and any i, j (for which there are k^2 choices) there are at most 2 pairs (r'_m, v'_m) in $S_i \times S_j$ with $r'_m + v'_m \equiv n \pmod{p^2-1}$. Define $c_1, c_2 \in \mathrm{GF}(p)$ by $(ij)^{-1}\theta^n - \theta^2 = c_1\theta + c_2$. Since $r'_m + v'_m \equiv n \pmod{p^2-1}$, we have

$$\begin{aligned} c_1\theta + c_2 = (ij)^{-1}\theta^n - \theta^2 = (ij)^{-1}\theta^{r'_m+v'_m} - \theta^2 = (ij)^{-1}\theta^{r'_m}\theta^{v'_m} - \theta^2 = \\ (ij)^{-1}(i\theta + r_m)(j\theta + v_m) - \theta^2 = (i^{-1}r_m + j^{-1}v_m)\theta + i^{-1}r_m j^{-1}v_m. \end{aligned}$$

This means that $(a, b) = (i^{-1}r_m, j^{-1}v_m)$ is a solution to $x^2 - c_1x + c_2 = (x-a)(x-b)$. By Lemma 6.2, there are at most two such pairs. ■

To extend Singer's [Sin38] construction, we shall need the following lemma.

Lemma 6.3. *Let p be a prime power, and let $\theta \in GF(p^3)$ generate the multiplicative group. Then θ^a and θ^b are linearly dependent over $GF(p)$ iff $a \equiv b \pmod{p^2+p+1}$.*

Proof. The multiplicative group of $\mathrm{GF}(p)$ is a subgroup of the multiplicative group of $\mathrm{GF}(p^3)$, i.e., $\mathrm{GF}(p) = \{\theta^k : k \equiv 0 \pmod{\frac{p^3-1}{p-1}}\}$. Since two elements of $\mathrm{GF}(p^3)$ are linearly dependent over $\mathrm{GF}(p)$ exactly if their ratio is in $\mathrm{GF}(p)$, we see that θ^a and θ^b are linearly dependent exactly if $a - b \equiv 0 \pmod{\frac{p^3-1}{p-1}}$. ■

Proof of Proposition 6.1(iii). Let $\theta \in \mathrm{GF}(p^3)$ generate the multiplicative group. Since θ is algebraic with degree 3 over $\mathrm{GF}(p)$, the elements $1, \theta, \theta^2$ are a basis for $\mathrm{GF}(p^3)$ over $\mathrm{GF}(p)$. Define

$$T_i := \{0\} \cup \{s' \in [1, p^3-1] : \theta^{s'} = i\theta + s, \quad s \in \mathrm{GF}(p)\}$$

for $1 \leq i \leq k$, and define S_i to be the set of congruence classes modulo $q := \frac{p^3-1}{p-1} = p^2 + p + 1$ which intersect T_i.

Since each nonzero $s' \in T_i$ corresponds to a unique $s \in \mathrm{GF}(p)$, and for each $s \in \mathrm{GF}(p)$ there is an $s' \in [1, p^3-1]$ with $\theta^{s'} = i\theta + s$, we see that $|T_i| = |\mathrm{GF}(p)| + 1 = p+1$, and so by virtue of Lemma 6.3, $|S_i| = p+1$. Also, each $s' \neq 0$ occurs in at most one of the T_i, so that $|\bigcup_{i=1}^k S_i| = kp+1$.

We wish to show that, given any n, there are at most two pairs $(r', v') \in T_h \times T_j$ with $r' + v' \equiv n \bmod q$. Since there are k^2 choices of h, j, this will establish that $\bigcup_{i=1}^k S_i$ witnesses $C(2k^2, p^2+p+1) \geq kp+1$. Define $L_0 = 1$, and for each $x' \in T_i$ set $L_{x'} = \theta^{x'} = i\theta + x$.

Suppose that n is fixed, and $(r'_m, v'_m) \in T_h \times T_j$ $(1 \leq m \leq 3)$ are distinct pairs with $r'_m + v'_m \equiv n \bmod q$ for $1 \leq m \leq 3$. It follows from Lemma 6.3 that each pair of $L_{r'_1}L_{v'_1}, L_{r'_2}L_{v'_2}, L_{r'_3}L_{v'_3}$ are linearly dependent, i.e., they are multiples of each other. If $r'_1 = 0$, then $L_{r'_1}L_{v'_1}$ is linear. This means that $L_{r'_2}L_{v'_2}$ is also linear, and so one of r'_2, v'_2 is zero and the other is equal to $v'_1 \in T_j$. If $r'_2 = 0$, then $(r'_1, v'_1) = (r'_2, v'_2)$, contradicting distinctness, and if $v'_2 = 0$, then $v'_1 \equiv r'_2 \pmod q$, which is only possible if $v'_1 = r'_2 = 0$, whence again $(r'_1, v'_1) = (r'_2, v'_2)$.

Thus each $L_{r'_m}L_{v'_m}$ is a quadratic in θ. The coefficient of θ^2 in each $\theta^n = L_{r'_m}L_{v'_m} = (h\theta + r_m)(j\theta + v_m) = hj\theta^2 + (r_m j + v_m h)\theta + r_m v_m$ is hj. Since the $L_{r'_m}L_{v'_m}$ are multiples of each other with the same lead coefficients, they must in fact be equal. Set c_1, c_2 by $(hj)^{-1}\theta^n - \theta^2 = c_1\theta + c_2$, and observe that $(a, b) = (h^{-1}r_m, j^{-1}v_m)$ is a solution to $x^2 - c_1 x + c_2 = (x-a)(x-b)$; hence, by Lemma 6.2, there are only two such pairs. ∎

We now show how to combine two $B^*[g] \pmod n$ sets to construct another.

Proposition 6.4. *Let* $\gcd(x, y) = 1$*, and let* S *be a* $B^*[g] \pmod x$ *set and* M *be a* $B^*[h] \pmod y$ *set. Then the set* $M + yS := \{m + ys \colon m \in M, s \in S\}$ *is a* $B^*[gh] \pmod{xy}$ *set. In particular,* $C(gh, xy) \geq C(g, x)C(h, y)$ *for any positive integers* g, h, x, y *with* $\gcd(x, y) = 1$.

Proof. Consider $m_i, n_i \in M$ and $s_i, t_i \in S$ with

$$(m_1 + ys_1) + (n_1 + yt_1) \equiv \cdots \equiv (m_{gh+1} + ys_{gh+1}) + (n_{gh+1} + yt_{gh+1}) \pmod{xy}. \quad (30)$$

We need to show that $m_i = m_j$, $s_i = s_j$, $n_i = n_j$, and $t_i = t_j$, for some i, j. Reducing these Eq. (30) modulo y, we see that $m_1 + n_1 \equiv m_2 + n_2 \equiv \cdots \equiv m_{gh+1} + n_{gh+1} \pmod y$. Since M is a $B^*[h] \pmod y$ set, we can reorder the m_i, n_i, s_i, t_i so that $m_1 = m_2 = \cdots = m_{g+1}$ and $n_1 = n_2 = \cdots = n_{g+1}$. Reducing Eq. (30) modulo x we arrive at

$$ys_1 + yt_1 \equiv ys_2 + yt_2 \equiv \cdots \equiv ys_{g+1} + yt_{g+1} \pmod x$$

whence, since $\gcd(x, y) = 1$,

$$s_1 + t_1 \equiv s_2 + t_2 \equiv \cdots \equiv s_{g+1} + t_{g+1} \pmod x.$$

The s_i and t_i are from a $B^*[g] \pmod x$ set, so that for some i, j, $s_i = s_j$ and $t_i = t_j$. ∎

We have computed $C(g, n)$ for small g and n by exhaustive search. The results are summarized in Table 1. The entry for $(k, g) = (10, 5)$ is 28; this means that $C(5, 28) \geq 10$ (a witness is $\{0, 1, 2, 4, 5, 8, 12, 15, 23, 24\}$), while there is no $n < 28$ for which $C(5, n) \geq 10$.

It is straightforward to verify that $C(5, 28) \geq 10$: one simply verifies that the witness has 10 elements and is indeed a $B^*[5] \pmod{28}$ set. It is not straightforward, however, to verify that $C(5, n) < 10$ for $n < 28$. We have made these verifications for each entry given in the table by a long computer search using *Mathematica*.

k \ g	2	3	4	5	6	7	8	9	10	11
3	6									
4	12	7								
5	21	11	8							
6	31	19	11	9						
7	48	29	14	13	10					
8	57	43	22	17	12	11				
9	73	57	28	19	16	13	12			
10	91		36	28	19	17	14	13		
11				35	22	21	18	15	14	
12					30	23	21	19	16	15
13						31	24	22	19	17
14							28	25		20

Table 1: $\min\{n\colon C(g,n) \geq k\}$

6.3 Probabilistic Constructions of $B^*[g] \pmod n$ Sets

The algebraic methods of the previous section provide effective and completely explicit constructions of large $B^*[g] \pmod n$ sets. However, we can establish the existence of even larger $B^*[g] \pmod n$ sets using a probabilistic construction. We rely upon the following two lemmas, which are quantitative statements of the Law of Large Numbers for sums of many independent random variables.

Lemma 6.5. *Let $p_1, \dots, p_n$ be real numbers in the range $[0,1]$, and set $p = (p_1 + \cdots + p_n)/n$. Define mutually independent random variables $X_1, \dots, X_n$ such that X_i takes the value $1-p_i$ with probability p_i and the value $-p_i$ with probability $1-p_i$ (so that the expectation of each X_i is zero), and define $X = X_1 + \cdots + X_n$. Then for any positive number a,*

$$\Pr[X > a] < \exp\left(\frac{-a^2}{2pn} + \frac{a^3}{2p^2n^2}\right) \quad \textit{and} \quad \Pr[X < -a] < \exp\left(\frac{-a^2}{2pn}\right).$$

Proof. These assertions are Theorems A.11 and A.13 of [AS00]. ∎

Lemma 6.6. *Let $p_1, \dots, p_n$ be real numbers in the range $[0,1]$, and set $E = p_1 + \cdots + p_n$. Define mutually independent random variables $Y_1, \dots, Y_n$ such that Y_i takes the value 1 with probability p_i and the value 0 with probability $1-p_i$, and define $Y = Y_1 + \cdots + Y_n$ (so that the expectation of Y equals E). Then $\Pr[Y > E + a] < \exp\left(\frac{-a^2}{3E}\right)$ for any real number $0 < a < E/3$, and $\Pr[Y < E - a] < \exp\left(\frac{-a^2}{2E}\right)$ for any positive real number a.*

Proof. This follows immediately from Lemma 6.5 upon defining $X_i = Y_i - p_i$ for each i and noting that $E = pn$ and that $\frac{a^3}{2E^2} < \frac{a^2}{6E}$ under the assumption $0 < a < E/3$. ∎

We now give the probabilistic construction of large $B^*[g] \pmod n$ sets.

Proposition 6.7. *For every $0 < \varepsilon \leq 1$, there is a sequence of ordered pairs (n_j, g_j) of positive integers such that $\frac{C(g_j,n_j)}{n_j} \gtrsim \varepsilon$ and $\frac{g_j}{n_j} \lesssim \varepsilon^2$.*

Proof. Let n be an odd integer. We define a random subset S of $\{1, \ldots, n\}$ as follows: for every $1 \le i \le n$, let Y_i be 1 with probability ε and 0 with probability $1-\varepsilon$ with the Y_i mutually independent, and let $S := \{i : Y_i = 1\}$. We see that $|S| = \sum_{i=1}^n Y_i$ has expectation $E = \varepsilon n$. Setting $a = \sqrt{\varepsilon n \log 4}$, Lemma 6.6 gives

$$\Pr\left[|S| < \varepsilon n - \sqrt{\varepsilon n \log 4}\right] < \frac{1}{2}.$$

Now for any integer k, define the random variable

$$\begin{aligned} R_k &:= \#\{1 \le c, d \le n : c + d \equiv k \pmod{n},\ Y_c = Y_d = 1\} \\ &= \sum_{c+d \equiv k \pmod{n}} Y_c Y_d, \end{aligned}$$

so that R_k is the number of representations of $k \pmod{n}$ as the sum of two elements of S. Observe that R_k is the sum of $n-1$ random variables taking the value 1 with probability ε^2 and the value 0 otherwise, plus one random variable (corresponding to $c \equiv d \equiv 2^{-1}k \pmod{n}$) taking the value 1 with probability ε and the value 0 otherwise. Therefore the expectation of R_k is $E = (n-1)\varepsilon^2 + \varepsilon$. Setting $a = \sqrt{3((n-1)\varepsilon^2 + \varepsilon)\log 2n}$, and noting that $a < E/3$ when n is sufficiently large in terms of ε, Lemma 6.6 gives

$$\Pr\left[R_k > (n-1)\varepsilon^2 + \varepsilon - \sqrt{3((n-1)\varepsilon^2 + \varepsilon)\log 2n}\right] < \frac{1}{2n}$$

for each $1 \le k \le n$. Therefore, there exists a $B^*[g] \pmod{n}$ set $S \subseteq \{1, \ldots, n\}$, with $g \le (n-1)\varepsilon^2 + \varepsilon - \sqrt{3((n-1)\varepsilon^2 + \varepsilon)\log 2n} \lesssim \varepsilon^2 n$, such that $|S| \ge \varepsilon n - \sqrt{\varepsilon n \log 4} \gtrsim \varepsilon n$. This establishes the proposition. ■

Define $\Delta_{\mathbb{T}}(\varepsilon)$ to be the supremum of those real numbers δ such that every subset of $\mathbb{T}$ with measure ε has a subset with measure δ that is fixed by a reflection $t \mapsto c - t$. The function $\Delta_{\mathbb{T}}(\varepsilon)$ stands in relation to $C(g, n)$ as $\Delta(\varepsilon)$ stands to $R(g, n)$. However, it turns out that $\Delta_{\mathbb{T}}$ is much easier to understand:

Corollary 6.8. *Every subset of $\mathbb{T}$ with measure ε contains a symmetric subset with measure ε^2, and this is best possible for every ε. In particular, $\Delta_{\mathbb{T}}(\varepsilon) = \varepsilon^2$ for all $0 \le \varepsilon \le 1$.*

Proof. In the proof of the trivial lower bound for $\Delta(\varepsilon)$ (Lemma 3.2), we saw that every subset of $[0, 1]$ with measure ε contains a symmetric subset with measure at least $\frac{1}{2}\varepsilon^2$. The proof is easily modified to show that every subset of $\mathbb{T}$ with measure ε contains a symmetric subset with measure ε^2. This shows that $\Delta_{\mathbb{T}}(\varepsilon) \ge \varepsilon^2$ for all ε. On the other hand, the proof of Proposition 5.1 is also easily modified to show that $\Delta_{\mathbb{T}}\left(\frac{C(g,n)}{n}\right) \le \frac{g}{n}$, as is the proof of Lemma 3.5 to show that $\Delta_{\mathbb{T}}$ is continuous. Then, by virtue of Proposition 6.7 and the monotonicity of $\Delta_{\mathbb{T}}$, we have $\Delta_{\mathbb{T}}(\varepsilon) \le \varepsilon^2$. ■

6.4 Explicit Constructions of $B^*[g]$ Sets

If $S \subseteq [0, 100)$ is a $B^*[g] \pmod{200}$ set, then the modular sums are the same as the real sums, and so S is a $B^*[g]$ set as well. This observation is the fundamental idea behind using the method of Proposition 6.4 to construct $B^*[g]$ sets.

Proposition 6.9. *Let g, h, x, y be positive integers. Then*

$$R(gh, xy) \geq R(gh, xy+1-\lceil \tfrac{y}{C(h,y)} \rceil) \geq R(g,x)C(h,y).$$

Proof. Let $M \subseteq [1,m] \subseteq [1,y]$ witness the value of $C(h,y)$, and let $S \subseteq [0,s)$ witness the value of $R(g,s)$. Take $x > 2s$, and relatively prime to y, and note that S is a $B^*[g] \pmod{x}$ set. By Proposition 6.4, the set $M+yS$ is a $B^*[gh] \pmod{xy}$ set. But by taking x to be sufficiently large, we see that $M+yS$ is actually a $B^*[gh]$ set.

We now compute the smallest and largest element of $M+yS$. Clearly the smallest element is 1, and the largest is $m+y(s-1)$. Since M is a $B^*[h] \pmod{y}$ set, we may shift it modulo y so as to minimize m. M has $C(h,y)$ elements, so there must be two consecutive elements whose difference $\pmod{y}$ is at least $\lceil \frac{y}{C(h,y)} \rceil$, i.e., we may take $m \leq y+1-\lceil \frac{y}{C(h,y)} \rceil$. Thus $M+yS \subseteq [1, y+1-\lceil \frac{y}{C(h,y)} \rceil + y(s-1)] = [1, ys+1-\lceil \frac{y}{C(h,y)} \rceil]$, and $|M+yS| = R(g,s)C(h,y)$. ■

The reader might feel that the part of the argument concerning the largest gap in M is more trouble than it is worth. We include this for two reasons. First, Erdős [Guy94, Problem C9] offered \$500 for an answer to the question, "Is $R(2,n) = \sqrt{n} + O(1)$?" This question would be answered in the negative if one could show, for example, that the $B^*[2] \pmod{p^2-1}$ sets constructed by Bose (the $k=1$ case of Proposition 6.1(ii)) contain a gap which is not $O(p)$, as seems likely from the experiments of Zhang [Zha94] and Lindström [Lin98]. Second, there is some literature (e.g., [ESS95] and [Ruz96]) concerning the possible size of the largest gap in a maximal Sidon set contained in $\{1, \ldots, n\}$. In short, we include this argument because there is some reason to believe that this is a significant source of the error term in at least one case, and because there is some reason to believe that improvement is possible.

Our plan is to employ the inequality of Proposition 6.9 when y is large, $h=2$, and $x \approx \frac{8}{3}g$. In other words, we need nontrivial lower bounds for $C(2,n)$ for $n \to \infty$ and for $R(g,n)$ for values of n that are not much larger than g. The first need is filled by Proposition 6.1, while the second need is filled by the following lemma.

Lemma 6.10. *For all $g \geq 1$ we have $R(g, 3g - \lfloor g/3 \rfloor + 1) \geq g + 2\lfloor g/3 \rfloor + \lfloor g/6 \rfloor$.*

Proof. One can verify that a witness is

$$\left[0, \left\lfloor \frac{g}{3} \right\rfloor\right) \cup \left\{ g - \left\lfloor \frac{g}{3} \right\rfloor + 2\left[0, \left\lfloor \frac{g}{6} \right\rfloor\right) \right\} \cup \left[g, g + \left\lfloor \frac{g}{3} \right\rfloor\right) \cup \left(2g - \left\lfloor \frac{g}{3} \right\rfloor, 3g - \left\lfloor \frac{g}{3} \right\rfloor\right].$$

■

We remark that this family of examples was motivated by the finite sequence $S = \{1, 0, \frac{1}{2}, 1, 0, 1, 1, 1\}$, which has the property that its autocorrelations are small relative to the sum of its entries. In other words, the ratio of the ℓ^∞-norm of $S * S$ to the ℓ^1-norm of S itself is small. If we could find a finite sequence of rational numbers for which the corresponding ratio were smaller, we could convert it directly into a family of examples that would improve the lower bound for $\underline{\rho}(2g)$ in Theorem 1.3 for large g (see the proof of the theorem in Section 6.6).

In addition to these parametric results, we have established by direct (exhaustive) computation the exact value of $R(g,n)$ for small values of g and n. Table 2 records, for given values

k \ g	2	3	4	5	6	7	8	9	10	11
3	4									
4	7	5								
5	12	8	6							
6	18	13	8	7						
7	26	19	11	9	8					
8	35	25	14	12	10	9				
9	45	35	18	15	12	11	10			
10	56	46	22	19	14	13	12	11		
11	73	58	27	24	17	15	14	13	12	
12	≤ 92	≤ 72	31	29	20	18	16	15	14	13
13	≤ 143	≤ 101	37	34	24	21	18	17	16	15
14		≤ 128	44	40	28	26	21	19	18	17
15			≤ 52	≤ 47	32	29	24	22	20	19
16					36	34	27	24	22	21
17					≤ 42	≤ 38	30	28	24	23
18							34	32	27	25
19							≤ 38	≤ 36	30	28
20									33	31
21									≤ 37	35
21										≤ 38

Table 2: $\min\{n\colon R(g,n) \ge k\}$

of g and k, the smallest possible value of $\max S$ among all $B^*[g]$ sets S consisting of exactly k positive integers; in other words, the entry corresponding to k and g is $\min\{n\colon R(g,n) \ge k\}$. For example, the $(k,g) = (8,2)$ entry records the fact that there exists an 8-element Sidon set of integers from $[1,35]$ but no 8-element Sidon set of integers from $[1,34]$.

To show that $R(2,35) \ge 8$, for instance, it is only necessary to observe that the witness $\{1,3,13,20,26,31,34,35\}$ has 8 elements and is a $B^*[2]$ set. To show that $R(2,35) \le 8$, however, seems to require an extensive search.

6.5 Probabilistic Constructions of $B^*[g]$ Sets

We can use the probabilistic methods employed in Section 6.3 to construct large $B^*[g]$ sets in $\mathbb{Z}$. The proof is more complicated because it is to our advantage to endow different integers with different probabilities of belonging to our random set. Although all of the constants in the proof could be made explicit, we are content with inequalities having error terms involving big-O notation.

Proposition 6.11. *Let $\gamma \ge \pi$ be a real number and $n \ge \gamma$ be an integer. There exists a $B^*[g]$ set $S \subseteq \{1,\dots,n\}$, where $g = \gamma + O(\sqrt{\gamma \log n})$, with $|S| \ge 2\sqrt{\frac{\gamma n}{\pi}} + O(\gamma + (\gamma n)^{1/4})$.*

Proof. Define mutually independent random variables Y_k, taking only the values 0 and 1, by

g	x	$R(g,x)$	Witness	$R(g,x)/\sqrt{2gx}$
2	7	4	$\{1,2,5,7\}$	$\frac{2}{\sqrt{7}} \approx 0.756$
3	5	4	$\{1,2,3,5\}$	$\frac{2\sqrt{2}}{\sqrt{15}} \approx 0.730$
4	31	12	$\{1,2,4,10,11,12,14,19,25,26,30,31\}$	$\frac{2}{\sqrt{7}} \approx 0.756$
5	9	7	$\{1,2,3,4,5,7,9\}$	$\frac{7}{3\sqrt{10}} \approx 0.738$
6	20	12	$\{1,2,3,4,5,6,9,10,13,15,19,20\}$	$\frac{\sqrt{3}}{\sqrt{5}} \approx 0.775$
7	15	11	$\{1,2,3,7,8,9,10,11,12,13,15\}$	$\frac{11}{\sqrt{210}} \approx 0.759$
8	30	17	$\{1,2,5,7,8,9,11,12,13,14,16,18,26,27,28,29,30\}$	$\frac{17}{4\sqrt{30}} \approx 0.776$
9	24	16	$\{1,2,3,4,5,6,7,8,9,13,14,15,17,22,23,24\}$	$\frac{4}{3\sqrt{3}} \approx 0.770$
10	33	20	$\{1,2,4,5,6,7,8,9,10,11,13,15,20,21,22,23,30,31,32,33\}$	$\frac{2\sqrt{5}}{\sqrt{33}} \approx 0.778$
11	25	18	$\{1,2,3,4,5,11,12,13,14,15,16,17,18,19,20,21,23,25\}$	$\frac{18}{5\sqrt{22}} \approx 0.768$

Table 3: Important values of $R(g,x)$ and witnesses

$$\Pr\{Y_k = 1\} = p_k := \begin{cases} 1, & 1 \le k < \frac{\gamma}{\pi}, \\ \sqrt{\frac{\gamma}{\pi k}}, & \frac{\gamma}{\pi} \le k \le n, \\ 0, & k > n. \end{cases} \tag{31}$$

(Notice that $p_k \le \sqrt{\frac{\gamma}{\pi k}}$ for all $k \ge 1$.) These random variables define a random subset $S = \{k \colon Y_k = 1\}$ of the integers from 1 to n. We shall show that, with positive probability, S is a large $B^*[g]$ set with g not much bigger than γ.

The expected size of S is

$$\begin{aligned} E_0 := \sum_{1 \le j \le n} p_j &= \sum_{1 \le j < \gamma/\pi} 1 + \sum_{\gamma/\pi \le j \le n} \sqrt{\frac{\gamma}{\pi j}} \\ &= \frac{\gamma}{\pi} + \int_{\gamma/\pi}^{n} \sqrt{\frac{\gamma}{\pi t}}\, dt + O(1) = 2\sqrt{\frac{\gamma n}{\pi}} - \frac{\gamma}{\pi} + O(1). \end{aligned} \tag{32}$$

If we set $a_0 := \sqrt{2E_0 \log 3}$, then Lemma 6.6 tells us that

$$\Pr[|S| < E_0 - a_0] < \exp\left(\frac{-a_0^2}{2E_0}\right) = \frac{1}{3}.$$

Now for any integer $k \in [\gamma, 2n]$, let

$$R_k := \sum_{1 \le j \le n} Y_j Y_{k-j} = 2 \sum_{1 \le j < k/2} Y_j Y_{k-j} + Y_{k/2},$$

the number of representations of k as $k = s_1 + s_2$ with $s_1, s_2 \in S$. (Here we adopt the convention that $Y_{k/2} = p_{k/2} = 0$ if k is odd). Notice that in this latter sum, $Y_{k/2}$ and the $Y_j Y_{k-j}$ are mutually independent random variables taking only values 0 and 1, with

$\Pr[Y_j Y_{k-j} = 1] = p_j p_{k-j}$. Thus the expectation of R_k is

$$\begin{aligned} E_k := 2 \sum_{1 \le j < k/2} p_j p_{k-j} + p_{k/2} &\le 2 \sum_{1 \le j < k/2} \sqrt{\frac{\gamma}{\pi j}} \sqrt{\frac{\gamma}{\pi(k-j)}} + \sqrt{\frac{\gamma}{\pi k/2}} \\ &\le \frac{2\gamma}{\pi} \int_0^{k/2} \sqrt{\frac{1}{t(k-t)}}\, dt + \sqrt{\frac{2\gamma}{\pi k}} = \gamma + \sqrt{\frac{2\gamma}{\pi k}} < \gamma + 1 \qquad (33) \end{aligned}$$

using the inequalities $p_k \le \sqrt{\frac{\gamma}{\pi k}}$ and $k \ge \gamma$.

If we set $a = \sqrt{3(\gamma+1)\log 3n}$, then Lemma 6.6 tells us that

$$\Pr[R_k > \gamma + 1 + a] < \Pr[R_k > E_k + a] < \exp\left(\frac{-a^2}{3E_k}\right) < \exp\left(\frac{-a^2}{3(\gamma+1)}\right) = \frac{1}{3n}$$

for every k in the range $\gamma \le k \le 2n$. Note that $R_k \le \gamma$ trivially for k in the range $1 \le k \le \gamma$. Therefore, with probability at least $1 - \frac{1}{3} - (2n-\gamma)\frac{1}{3n} = \frac{\gamma}{3n} > 0$, the set S has at least $E_0 - a_0 = 2\sqrt{\frac{\gamma n}{\pi}} + O(\gamma + (\gamma n)^{1/4})$ elements and satisfies $R_k \le \gamma + 1 + a$ for all $1 \le k \le 2n$. Setting $g := \gamma + 1 + a = \gamma + O(\sqrt{\gamma \log n})$, we conclude that any such set S is a $B^*[g]$ set. This establishes the proposition. ∎

Schinzel conjectured that among all pdfs supported on $[0, \frac{1}{2}]$, the function

$$f(x) = \begin{cases} \frac{1}{\sqrt{2x}}, & x \in [0, \frac{1}{2}], \\ 0, & \text{otherwise} \end{cases}$$

has the property that $\|f * f\|_\infty$ is minimal. We have

$$f * f(x) = \begin{cases} \frac{\pi}{2}, & x \in [0, \frac{1}{2}], \\ \frac{\pi}{2} - 2\tan^{-1}\sqrt{2x-1}, & x \in [\frac{1}{2}, 1], \\ 0, & \text{otherwise} \end{cases}$$

and so $\|f*f\|_\infty = \frac{\pi}{2}$. We have adapted this function in our definition (31) of the probabilities p_k; the constant $\frac{\pi}{2}$ appears as the value of the last integral in Eq. (33). If Schinzel's conjecture were false, then we could immediately incorporate any better function f into the proof of Proposition 6.11 and improve the lower bound on $|S|$. Indeed, Schinzel's conjecture is one of the motivations for our Conjecture 5.8, which by the above discussion is logically stronger.

Theorem 1.4. *For any $\delta > 0$, we have $R(g,n) > \left(\frac{2}{\sqrt{\pi}} - \delta\right)\sqrt{gn}$ if both $\frac{g}{\log n}$ and $\frac{n}{g}$ are sufficiently large in terms of δ.*

Proof. In the proof of Proposition 6.11, we saw that $\gamma \le g$ and $g = \gamma + O(\sqrt{\gamma \log n})$; this implies that $\gamma = g + O(\sqrt{g \log n}) = g\bigl(1 + O\bigl(\sqrt{\frac{\log n}{g}}\bigr)\bigr)$. Therefore the size of the constructed set S was at least

$$\begin{aligned} 2\sqrt{\tfrac{\gamma n}{\pi}} + O(\gamma + (\gamma n)^{1/4}) &= 2\sqrt{\tfrac{gn}{\pi}\Bigl(1 + O\Bigl(\sqrt{\tfrac{\log n}{g}}\Bigr)\Bigr)} + O(g + (gn)^{1/4}) \\ &= 2\sqrt{\tfrac{gn}{\pi}}\Bigl(1 + O\Bigl(\sqrt{\tfrac{\log n}{g}} + \sqrt{\tfrac{g}{n}}\Bigr)\Bigr). \end{aligned}$$

This establishes the theorem. ∎

6.6 Lower Bounds on $R(g, n)$

We are now ready to prove Theorem 1.3, which we restate here for the reader's convenience. Recall that $\underline{\rho}(g) = \liminf_{n\to\infty} \frac{R(g,n)}{\sqrt{gn}}$.

Theorem 1.3. *We have*

$$\begin{aligned}
\underline{\rho}(4) &\geq \tfrac{2}{\sqrt{7}} > 0.755, & \underline{\rho}(14) &\geq \tfrac{11}{\sqrt{210}} > 0.759,\\
\underline{\rho}(6) &\geq \tfrac{2\sqrt{2}}{\sqrt{15}} > 0.730, & \underline{\rho}(16) &\geq \tfrac{17}{4\sqrt{30}} > 0.775,\\
\underline{\rho}(8) &\geq \tfrac{2}{\sqrt{7}} > 0.755, & \underline{\rho}(18) &\geq \tfrac{4}{3\sqrt{3}} > 0.769,\\
\underline{\rho}(10) &\geq \tfrac{7}{3\sqrt{10}} > 0.737, & \underline{\rho}(20) &\geq \tfrac{2\sqrt{5}}{\sqrt{33}} > 0.778,\\
\underline{\rho}(12) &\geq \tfrac{\sqrt{3}}{\sqrt{5}} > 0.774, & \underline{\rho}(22) &\geq \tfrac{18}{5\sqrt{22}} > 0.767,
\end{aligned}$$

and for any $g \geq 12$,

$$\underline{\rho}(2g) \geq \frac{g + 2\lfloor g/3\rfloor + \lfloor g/6\rfloor}{\sqrt{6g^2 - 2g\lfloor g/3\rfloor + 2g}}.$$

In particular, for any $\delta > 0$ we have $R(g, n) > (\frac{11}{8\sqrt{3}} - \delta)\sqrt{gn}$ if both g and $\frac{n}{g}$ are sufficiently large in terms of δ.

Proof. For any positive integers x and $m \leq \sqrt{n/x}$, the monotonicity of R in the second variable gives $R(2g, n) \geq R(2g, x(m^2 - 1)) \geq R(g, x)C(2, m^2 - 1)$ by Proposition 6.9. If we choose m to be the largest prime not exceeding $\sqrt{n/x}$ (so that $m \gtrsim \sqrt{n/x}$ by the Prime Number Theorem), then Proposition 6.9 gives $R(2g, n) \geq R(g, x) \cdot m \gtrsim R(g, x)\sqrt{\frac{n}{x}}$ for any fixed positive integer g, and hence

$$\underline{\rho}(2g) = \liminf_{n\to\infty} \frac{R(2g, n)}{\sqrt{2gn}} \geq \frac{R(g, x)}{\sqrt{2gx}}.$$

The problem now is to choose x so as to make $\frac{R(g,x)}{\sqrt{2gx}}$ as large as we can for each g. For $g = 2, 3, \ldots, 11$, we use Table 2 to choose $x = 7, 5, 31, 9, 20, 15, 30, 24, 33, 25$, respectively (see Table 3 for witnesses to the values claimed for $R(g, x)$). This yields the first group of assertions in Theorem 1.3. For $g \geq 12$, we set $x = 3g - \lfloor g/3\rfloor + 1$ and appeal to Theorem 6.10, giving the second assertion of Theorem 1.3.

We remark that the above proof gives the more refined result

$$R(2g, n) \geq \frac{11}{8\sqrt{3}}\sqrt{2gn}\left(1 + O\left(g^{-1} + \left(\frac{n}{g}\right)^{(\alpha-1)/2}\right)\right)$$

as $\frac{n}{g}$ and g both go to infinity, where $\alpha < 1$ is any number such that for sufficiently large y, there is always a prime between $y - y^\alpha$ and y. For instance, we can take $\alpha = 0.525$ by [BHP01]. This clarification implies the final assertion of the theorem for even g, and the obvious inequality $R(2g + 1, n) \geq R(2g, n)$ implies the final assertion for odd g as well. ■

Habsieger and Plagne [HP] have proven that $R(2, x)/\sqrt{4x}$ is maximized at $x = 7$. For $g > 2$, we have chosen x based solely on the computations reported in Table 2. For general g, it appears that $R(g, x)/\sqrt{2gx}$ is actually maximized at a fairly small value of x, suggesting that this construction suffers from "edge effects" and is not best possible.

7 Upper Bounds for $\Delta(\varepsilon)$

7.1 Upper Bounds Derived from Constructions of $B^*[g]$ Sets

In Section 5 we used the connection between $B^*[g]$ sets and measurable sets with small symmetric subsets to deduce upper bounds for $R(g,n)$ from lower bounds for $\Delta(\varepsilon)$. In this section we exploit this relationship in the opposite direction, converting the lower bounds on $R(g,n)$ established in Section 6 into upper bounds for $\Delta(\varepsilon)$. Our first proposition verifies the statement of Theorem 1.1(i).

Proposition 7.1. *$\Delta(\varepsilon) = 2\varepsilon - 1$ for $\frac{11}{16} \le \varepsilon \le 1$, and $\Delta(\varepsilon) \ge 2\varepsilon - 1$ for all $0 < \varepsilon \le 1$.*

Proof. Recall from Lemma 3.5 that the function Δ satisfies the Lipschitz condition $|\Delta(x) - \Delta(y)| \le 2|x-y|$. Therefore the inequality $\Delta(\varepsilon) \ge 2\varepsilon - 1$ for all $0 < \varepsilon \le 1$ follows easily from the trivial value $\Delta(1) = 1$. To prove that $\Delta(\varepsilon) = 2\varepsilon - 1$ for $\frac{11}{16} \le \varepsilon \le 1$, then, it suffices to prove that $\Delta(\varepsilon) \le 2\varepsilon - 1$ in that range; and again by the Lipschitz condition, it suffices to prove simply that $\Delta\left(\frac{11}{16}\right) \le \frac{3}{8}$.

For any positive integer g, we combine Proposition 5.1 and Lemma 6.10 and the monotonicity of Δ to see that

$$\frac{g}{3g - \lfloor g/3 \rfloor + 1} \ge \Delta\left(\frac{R(g, 3g - \lfloor g/3 \rfloor + 1)}{3g - \lfloor g/3 \rfloor + 1}\right) \ge \Delta\left(\frac{g + 2\lfloor g/3 \rfloor + \lfloor g/6 \rfloor}{3g - \lfloor g/3 \rfloor + 1}\right).$$

Since Δ is continuous by Lemma 3.5, we may take the limit of both sides as $g \to \infty$ to obtain $\Delta\left(\frac{11}{16}\right) \le \frac{3}{8}$ as desired. ∎

Proposition 7.2. *The function $\frac{\Delta(\varepsilon)}{\varepsilon^2}$ is increasing on $(0,1]$.*

Proof. Choose $0 < \varepsilon < \varepsilon_0$. By Proposition 6.9, we have

$$\frac{R(g,x)}{x}\frac{C(h,y)}{y} \le \frac{R(gh,xy)}{xy}.$$

With the monotonicity of $\Delta(\varepsilon)$ and Proposition 5.1, this gives

$$\Delta\left(\frac{R(g,x)}{x}\frac{C(h,y)}{y}\right) \le \Delta\left(\frac{R(gh,xy)}{xy}\right) \le \frac{gh}{xy}.$$

Let g_i, x_i be such that $\frac{R(g_i,x_i)}{x_i} \to \varepsilon_0$ and $\frac{g_i}{x_i} \to \Delta(\varepsilon_0)$, which is possible by Proposition 5.3. By Proposition 6.7, we may choose sequences of integers h_j and y_j such that $\frac{C(h_j,y_j)}{y_j} \gtrsim \frac{\varepsilon}{\varepsilon_0}$ and $\frac{h_j}{y_j} \lesssim \left(\frac{\varepsilon}{\varepsilon_0}\right)^2$ as $j \to \infty$. This implies

$$\frac{R(g_i,x_i)}{x}\frac{C(h_j,y_j)}{y_j} \gtrsim \varepsilon \quad \text{and} \quad \frac{g_i}{x_i}\frac{h_j}{y_j} \lesssim \Delta(\varepsilon_0)\left(\frac{\varepsilon}{\varepsilon_0}\right)^2,$$

so that, again using the monotonicity and continuity of Δ,

$$\Delta(\varepsilon_0)\frac{\varepsilon^2}{\varepsilon_0^2} \gtrsim \frac{g_i h_j}{x_i y_j} \ge \Delta\left(\frac{R(g_i,x_i)}{x_i}\frac{C(h_j,y_j)}{y_j}\right) \gtrsim \Delta(\varepsilon)$$

as $j \to \infty$. This shows that $\frac{\Delta(\varepsilon)}{\varepsilon^2} \le \frac{\Delta(\varepsilon_0)}{\varepsilon_0^2}$ as desired. ∎

We can immediately deduce two nice consequences of this proposition.

Corollary 7.3. $\lim_{\varepsilon\to 0^+} \frac{\Delta(\varepsilon)}{\varepsilon^2}$ *exists.*

Proof. This follows from the fact that the function $\frac{\Delta(\varepsilon)}{\varepsilon^2}$ is increasing and bounded below by $\frac{1}{2}$ on $(0,1]$ by the trivial lower bound (Lemma 3.2). ∎

Corollary 7.4. $\Delta(\varepsilon) \le \frac{96}{121}\varepsilon^2$ *for* $0 \le \varepsilon \le \frac{11}{16}$.

Proof. This follows from the value $\Delta\big(\frac{11}{16}\big) = \frac{3}{8}$ calculated in Proposition 7.1 and the fact that the function $\frac{\Delta(\varepsilon)}{\varepsilon^2}$ is increasing. ∎

The corollary above proves part (iv) of Theorem 1.1, leaving only part (v) yet to be established. The following proposition finishes the proof of Theorem 1.1.

Proposition 7.5. $\frac{\Delta(\varepsilon)}{\varepsilon^2} \le \frac{\pi}{(1+\sqrt{1-\varepsilon})^2}$ *for all* $0 < \varepsilon \le 1$.

Proof. Define $\alpha := 1 - \sqrt{1-\varepsilon}$, so that $2\alpha - \alpha^2 = \varepsilon$. If we set $\gamma = \pi\alpha^2 n$ in the proof of Proposition 6.11, then the sets constructed are $B^*[g]$ sets with $g = \pi\alpha^2 n + O(\sqrt{n\log n})$ and have size at least

$$E_0 - a_0 = 2\sqrt{\frac{\pi\alpha^2 n^2}{\pi}} - \frac{\pi\alpha^2 n}{\pi} + O(1 + a_0) = (2\alpha - \alpha^2)n + O((\gamma n)^{1/4}) = \varepsilon n + O(\sqrt{n})$$

from Eq. (32).

Therefore, for these values of g and n,

$$\Delta\big(\frac{R(g,n)}{n}\big) \ge \Delta\big(\frac{\varepsilon n + O(\sqrt{n})}{n}\big) \to \Delta(\varepsilon)$$

as n goes to infinity, by the continuity of Δ. On the other hand, we see by Proposition 5.1 that

$$\varepsilon^{-2}\Delta\big(\frac{R(g,n)}{n}\big) \le \frac{\varepsilon^{-2} g}{n} = \frac{\pi\alpha^2 n + O(\sqrt{n\log n})}{\varepsilon^2 n}$$
$$= \frac{\pi\alpha^2}{(2\alpha-\alpha^2)^2} + O\Big(\sqrt{\frac{\log n}{\varepsilon^2 n}}\Big) = \frac{\pi}{(2-\alpha)^2} + o(1) \to \frac{\pi}{(1+\sqrt{1-\varepsilon})^2}$$

as n goes to infinity. Combining these two inequalities yields $\frac{\Delta(\varepsilon)}{\varepsilon^2} \le \frac{\pi}{(1+\sqrt{1-\varepsilon})^2}$ as desired. ∎

7.2 Upper Bounds Derived from Finite Unions of Intervals

Another way to approach bounding $\Delta(\varepsilon)$ is to compute precisely $\Delta_k(\varepsilon)$, the supremum of those real numbers δ such that every subset of $[0,1)$ with measure ε that is the union of k intervals has a symmetric subset with measure δ. From the definition it is easy to see that $\Delta_k(\varepsilon)$ is a decreasing function of k. In fact, it follows directly from the proof of Proposition 5.3 that $\Delta(\varepsilon) = \inf_k \Delta_k(\varepsilon) = \lim_{k\to\infty} \Delta_k(\varepsilon)$. We also have the following formula.

Lemma 7.6. *Let $E = \bigcup_{i=1}^{k} (\alpha_i, \beta_i) \subseteq [0,1)$ be the union of k disjoint intervals. The largest measure of a symmetric subset of E is*

$$D(E) = \max_{0 \le c \le 1} \left\{ \sum_{i=1}^{k} \sum_{j=1}^{k} \max\{0, \min\{c-\alpha_i, \beta_j - c\} - \max\{c - \beta_i, \alpha_j - c\}\} \right\}.$$

Proof. This follows immediately from the fact that, for any real numbers $0 \le \alpha_i < \beta_i \le 1$ and c, the measure of the set $\{x \colon c - x \in (\alpha_1, \beta_1),\ c + x \in (\alpha_2, \beta_2)\}$ is $\max\big\{0, \min\{c - \alpha_1, \beta_2 - c\}$ $- \max\{c - \beta_1, \alpha_2 - c\}\big\}$. ■

Since the maximum over c is achieved for some c that is the midpoint of endpoints of the intervals (i.e., $c = (\alpha_i + \beta_j)/2$, $c = (\alpha_i + \alpha_j)/2$, or $c = (\beta_i + \beta_j)/2$), this theorem provides an effective method for the computation of $D(E)$ for any particular union of k intervals. Indeed, this formula reduces finding a particular value $\Delta_k(\varepsilon)$ to a finite number of linear programming problems, though in practice this computation becomes unmanageably large even for small values of k. In principal, the entire function Δ_k could be calculated by solving these linear programming problems with the constant ε remaining unspecified, branching finitely many times depending on various simple inequalities for ε.

The case $k = 2$ is simple enough to deal with directly; we state the result in Proposition 7.7 but omit the proof. We have shown computationally that the graphs of $\Delta_3(\varepsilon)$ and $\Delta_4(\varepsilon)$ lie on or below the polygonal paths described in Conjecture 7.8 (see Figure 8), but we have not verified that these upper bounds are in fact sharp.

Proposition 7.7. *The graph of the function $(\varepsilon, \Delta_2(\varepsilon))$ is the polygonal path connecting $(0,0)$, $(3/4, 1/2)$, and $(1,1)$.*

Conjecture 7.8. *The graph of the function $\Delta_3(\varepsilon)$ is the polygonal path connecting $(0,0)$, $(\frac{4}{7}, \frac{2}{7})$, $(\frac{7}{11}, \frac{4}{11})$, $(\frac{5}{7}, \frac{3}{7})$, and $(1,1)$. The graph of the function $\Delta_4(\varepsilon)$ is the polygonal path connecting $(0,0)$, $(\frac{5}{12}, \frac{1}{6})$, $(\frac{9}{19}, \frac{4}{19})$, $(\frac{1}{2}, \frac{2}{9})$, $(\frac{2}{3}, \frac{10}{27})$, $(\frac{17}{24}, \frac{5}{12})$, and $(1,1)$.*

It seems likely that the graph of the function $\Delta_k(\varepsilon)$ always contains the line segment connecting $(0,0)$ and $(\frac{k+1}{n}, \frac{2}{n})$, where n is the least integer for which $R(2,n) = k+1$. It is easy to show that for every $k \ge 2$, the graph of the function $\Delta_k(\varepsilon)$ contains the line segment connecting the points $(\frac{3}{4}, \frac{1}{2})$ and $(1,1)$.

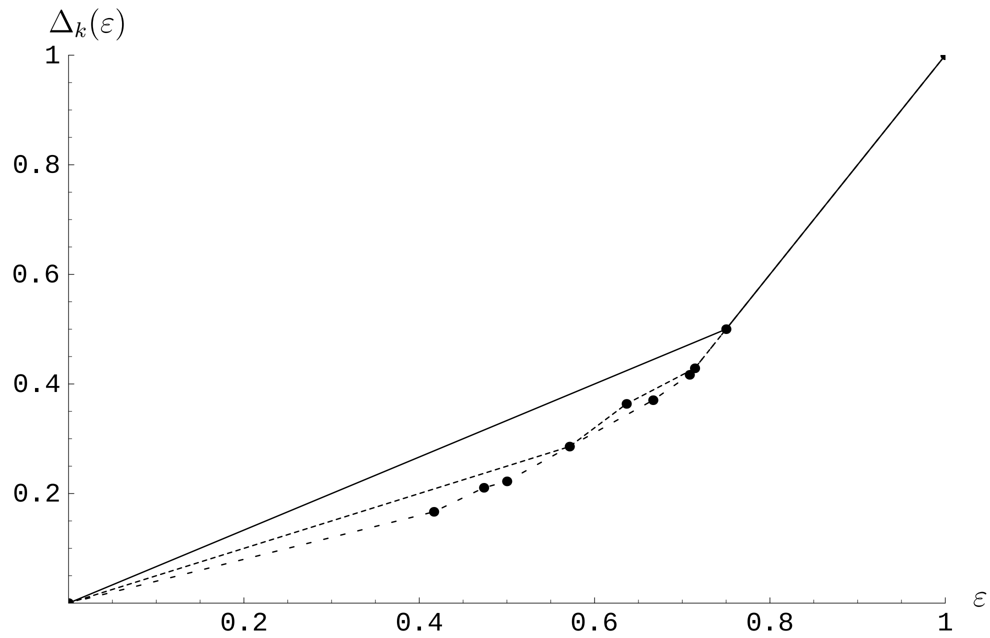

Figure 8: The graph of $\Delta_2(\varepsilon)$ and the conjectured graphs of $\Delta_3(\varepsilon)$ and $\Delta_4(\varepsilon)$

8 Some Remaining Questions

We group the problems in this section into three categories, although some problems do not fit clearly into any of the categories and others fit into more than one. (We also refer the reader to the Conjectures 4.4, 5.8 and 7.8 already propounded.)

8.1 Properties of the Function $\Delta(\varepsilon)$

The first open problem on the list must of course be the exact determination of $\Delta(\varepsilon)$ for all values $0 \le \varepsilon \le 1$. In the course of our investigations, we have come to believe the following assertion.

Conjecture 8.1. *$\Delta(\varepsilon) = \max\{2\varepsilon - 1, \frac{\pi}{4}\varepsilon^2\}$ for all $0 \le \varepsilon \le 1$.*

Notice that the upper bounds given in Theorem 1.1 are not too far from this conjecture, the difference between the constants $\frac{96}{121} = 0.7934$ and $\frac{\pi}{4} = 0.7854$ in the middle range for ε being the only discrepancy. In fact, we believe it might be possible to prove that the expression in Conjecture 8.1 is indeed an upper bound for $\Delta(\varepsilon)$ by a more refined application of the probabilistic method employed in Section 6.5. The key would be to show that the various events $R_k > \gamma + 1 + a$ are more or less independent of one another (as it stands we have to assume the worst—that they are all mutually exclusive—in obtaining our bound for the probability of obtaining a "bad" set), so that we could omit the $\log 3n$ factor from the chosen value of a.

There are some intermediate qualitative results about the function $\Delta(\varepsilon)$ that might be easier to resolve. It seems likely that $\Delta(\varepsilon)$ is convex, for example, but we have not been able to prove this. A first step towards clarifying the nature of $\Delta(\varepsilon)$ might be to prove that

$$\frac{|\Delta(x)-\Delta(y)|}{|x-y|} \ll \max\{x,y\}.$$

Also, we would not be surprised to see accomplished an exact computation of $\Delta(\frac{1}{2})$, but we have been unable to make this computation ourselves.

We do not believe that there is always a set with measure ε whose largest symmetric subset has measure $\Delta(\varepsilon)$. In fact, we do not believe that there is a set with measure $\varepsilon_0 := \inf\{\varepsilon\colon \Delta(\varepsilon) = 2\varepsilon - 1\}$ whose largest symmetric subset has measure $\Delta(\varepsilon_0)$, but we do not even know the value of ε_0. In Proposition 7.1, we showed that $\varepsilon_0 \le \frac{11}{16}$, but this was found by rather limited computations and is unlikely to be sharp. The quantity $\frac{11}{8\sqrt{3}}$ in Theorem 1.3 is of the form $\frac{\varepsilon_0}{\sqrt{4\varepsilon_0-2}}$, and similarly the quantity $\frac{96}{121}$ in Theorem 1.1(iv) is of the form $\frac{2\varepsilon_0-1}{\varepsilon_0^2}$. Thus any improvement in the bound $\varepsilon_0 \le \frac{11}{16}$ would immediately result in improvements to Theorem 1.3 and Theorem 1.1(iv). We remark that Conjecture 8.1 implies that $\varepsilon_0 = \frac{2}{2+\sqrt{4-\pi}} = 0.6834$, which in turn would allow us to replace the constant $\frac{11}{8\sqrt{3}}$ in Theorem 1.3 by $\sqrt{\frac{2}{\pi}}$.

8.2 Artifacts of our Proof

Let $\mathbb{K}$ be the class of functions $K \in L^2(\mathbb{T})$ satisfying $K(x) \ge 1$ on $[-\frac{1}{4}, \frac{1}{4}]$. How small can we make $\|\hat{K}\|_p$ for $1 \le p \le 2$? We are especially interested in $p = \frac{4}{3}$, but a solution for any p may be enlightening.

To give some perspective to this problem, note that a trivial upper bound for $\inf_{K\in\mathbb{K}} \|\hat{K}\|_p$ can be found by taking K to be identically equal to 1, which yields $\|\hat{K}\|_p = 1$. One can find functions that improve upon this trivial choice; for example, the function K defined in Eq. (8) is an example where $\|\hat{K}\|_{4/3} = 0.96585$. On the other hand, since the ℓ^p-norm of a sequence is a decreasing function of p, Parseval's identity immediately gives us the lower bound $\|\hat{K}\|_p \ge \|\hat{K}\|_2 = \|K\|_2 \ge \left(\int_{-1/4}^{1/4} 1^2\,dt\right)^{1/2} = \frac{1}{\sqrt{2}} = 0.707107$, and of course this is the exact minimum for $p = 2$.

We remark that Proposition 4.2 and the function $b(x)$ defined after the proof of Corollary 4.3 provide a stronger lower bound for $1 \le p \le \frac{4}{3}$. By direct computation we have $1.14939 > \|b*b\|_2^2$, and by Proposition 4.2 we have $\|b*b\|_2^2 \ge \|\hat{K}\|_{4/3}^{-4}$ for any $K \in \mathbb{K}$. Together these imply that $\|\hat{K}\|_p \ge \|\hat{K}\|_{4/3} > 0.96579$. In particular, for $p = \frac{4}{3}$ we know the value of $\inf_{K\in\mathbb{K}} \|\hat{K}\|_{4/3}$ to within one part in ten thousand. The problem of determining the actual infimum for $1 < p < 2$ seems quite mysterious. We remark that Green [Gre01] considered the discrete version of a similar optimization problem, namely the minimization of $\|\hat{K}\|_p$ over all pdfs K supported on $[-\frac{1}{4}, \frac{1}{4}]$.

As mentioned at the end of Section 4.2, we used the inequality $\|g\|_2^2 \le \|g\|_\infty \|g\|_1$ which is exact when g takes on one non-zero value, i.e., when g is an nif. We apply this inequality when $g = f*f$ with f supported on an interval of length $\frac{1}{2}$, which usually looks very different from an nif. In this circumstance, the inequality does not seem to be best possible, although the corresponding inequality in the exponential sums approach of [CRT] and in the discrete

Fourier approach of [Gre01] clearly is best possible. Specifically, we ask for a lower bound on

$$\sup_{\substack{f \geq 0 \\ \operatorname{supp}(f) \subseteq [-\frac{1}{4}, \frac{1}{4}]}} \frac{\|f * f\|_\infty \|f * f\|_1}{\|f * f\|_2^2}$$

that is strictly greater than 1.

8.3 The Analogous Problem for Other Sets

More generally, for any subset E of an abelian group endowed with a measure, we can define $\Delta_E(\varepsilon) := \inf\{D(A)\colon A \subseteq E,\, \lambda(A) = \varepsilon\}$, where $D(A)$ is defined in the same way as in Eq. (2). For example, $\Delta_{[0,1]}(\varepsilon)$ is the function $\Delta(\varepsilon)$ we have been considering throughout this paper, and $\Delta_{\mathbb{T}}(\varepsilon)$ was considered in Section 6.3.

We believe that for each $E \subseteq \mathbb{R}$, there is a positive constant c such that $\Delta(\varepsilon) \geq \Delta_E(\varepsilon) \geq c\varepsilon^2$ for $0 < \varepsilon \leq 1$. The work of Abbott [Abb90] seems relevant. If we are concerned only with $\varepsilon \to 0$, and we normalize by considering only sets E for which $\lambda(E) = 1$, then it may be possible to take an absolute constant. In other words, is it true that

$$\inf_{\substack{E \subseteq \mathbb{R} \\ \lambda(E)=1}} \liminf_{\varepsilon \to 0} \frac{\Delta_E(\varepsilon)}{\varepsilon^2} > 0\,?$$

Most of the work in this paper generalizes easily from $E = [0,1]$ to $E = [0,1]^d$. We have had difficulties, however, in finding good kernel functions in higher dimensions. That is, we need functions $K(\bar{x})$ such that

$$\sum_{\bar{j} \in \mathbb{Z}^d} \left|\hat{K}(\bar{j})\right|^{4/3}$$

is as small as possible, while $K(\bar{x}) \geq 1$ if all components of $\bar{x}$ are less than $\frac{1}{4}$ in absolute value. This restricts K on one-half of the space in 1 dimension, one-quarter of the space in 2 dimensions, and only 2^{-d} of the space in d dimensions. For this reason one might expect that *better* kernels exist in higher dimensions, but the computational difficulties have prevented us from finding them.

Acknowledgements. The authors thank Heini Halberstam for thoughtful readings of this manuscript and helpful suggestions. The first author was supported in part by grants from the Natural Sciences and Engineering Research Council.

References

[Abb90] H. L. Abbott. Sidon sets. *Canad. Math. Bull.*, 33(3):335–341, 1990.

[AS00] Noga Alon and Joel H. Spencer. *The probabilistic method.* Wiley-Interscience [John Wiley & Sons], New York, second edition, 2000. With an appendix on the life and work of Paul Erdős.

[Bec75] W. Beckner. Inequalities in Fourier analysis. *Ann. of Math.*, 102:159–182, 1975.

[BHP01] R. C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes. II. *Proc. London Math. Soc. (3)*, 83(3):532–562, 2001.

[Bos42] R. C. Bose. An affine analogue of Singer's theorem. *J. Indian Math. Soc. (N.S.)*, 6:1–15, 1942.

[BVV00] T. Banakh, O. Verbitsky, and Ya. Vorobets. A Ramsey treatment of symmetry. *Electron. J. Combin.*, 7(1):Research Paper 52, 25 pp. (electronic), 2000.

[CEG] Fan Chung, Paul Erdős, and Ronald Graham. On sparse sets hitting linear forms. *J. Number Theory.* to appear.

[Cho44] S. Chowla. Solution of a problem of Erdös and Turan in additive-number theory. *Proc. Nat. Acad. Sci. India. Sect. A.*, 14:1–2, 1944.

[CRT] J. Cilleruelo, I. Ruzsa, and C. Trujillo. Upper and lower bounds for finite $B_h[g]$ sequences, $g > 1$. to appear.

[EF91] P. Erdős and R. Freud. On sums of a Sidon-sequence. *J. Number Theory*, 38(2):196–205, 1991.

[Erd44] P. Erdös. On a problem of Sidon in additive number theory and on some related problems. Addendum. *J. London Math. Soc.*, 19:208, 1944.

[ESS95] P. Erdős, A. Sárközy, and V. T. Sós. On sum sets of Sidon sets. II. *Israel J. Math.*, 90(1-3):221–233, 1995.

[ET41] P. Erdös and P. Turán. On a problem of Sidon in additive number theory, and on some related problems. *J. London Math. Soc.*, 16:212–215, 1941.

[Fol84] Gerald B. Folland. *Real analysis.* John Wiley & Sons Inc., New York, 1984. Modern techniques and their applications, A Wiley-Interscience Publication.

[Gre01] Ben Green. The number of squares and $B_h[g]$ sets. *Acta Arithmetica*, 100(4):365–390, 2001.

[GS80] R. L. Graham and N. J. A. Sloane. On additive bases and harmonious graphs. *SIAM J. Algebraic Discrete Methods*, 1(4):382–404, 1980.

[Guy94] Richard K. Guy. *Unsolved problems in number theory.* Springer-Verlag, New York, second edition, 1994. Unsolved Problems in Intuitive Mathematics, I.

[HLP88] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities.* Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition.

[HP] Laurent Habsieger and Alain Plagne. Ensembles $B_2[2]$: l'étau se resserre. Preprint.

[Lin98] Bernt Lindström. Finding finite B_2-sequences faster. *Math. Comp.*, 67(223):1173–1178, 1998.

[Mon94] Hugh L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1994.

[Ruz93] Imre Z. Ruzsa. Solving a linear equation in a set of integers. I. *Acta Arith.*, 65(3):259–282, 1993.

[Ruz96] Imre Z. Ruzsa. Sumsets of Sidon sets. *Acta Arith.*, 77(4):353–359, 1996.

[Sin38] James Singer. A theorem in finite projective geometry and some applications to number theory. *Trans. Amer. Math. Soc.*, 43(3):377–385, 1938.

[Świ58] S. Świerczkowski. On the intersection of a linear set with the translation of its complement. *Colloq. Math.*, 5:185–197, 1958.

[Zha94] Zhen Xiang Zhang. Finding finite B_2-sequences with larger $m-a_m^{1/2}$. *Math. Comp.*, 63(207):403–414, 1994.