Number theory (New York, 1991--1995),Springer, New York, 1996

Some Applications of Probability to Additive Number Theory and Harmonic Analysis

Mihail N. Kolountzakis*

January 1995

Abstract

We present some applications of the probabilistic method in additive number theory and harmonic analysis. We describe two general approaches to the probablistic construction of certain objects. The question of whether one can actually "construct" these is also discussed and several examples of "derandomized" probabilistic proofs are given.

1 Introduction

The term *probabilistic method* (*counting method*) in mathematics refers to the proof of the existence of a certain object by examining the behavior of an appropriate collection of candidates. This collection is equipped with a nonnegative measure of total mass 1 (a *probability measure*). The behavior of the collection, with respect to a certain property of its members, is then examined either on the average or in measure, as will become clear below.

The application of probability to different branches of mathematics in order to prove statements that do not seem to have anything to do with probability has been occuring more and more often since about the middle of this century, pioneered mainly by the work of Erdős. The method is used particularly frequently in combinatorics, number theory and harmonic analysis. In [3] many applications of the method in discrete mathematics and theoretical computer science are given, while [16] is the primary reference for applications in analysis.

In this paper we distinguish and describe two different arguments in which a probability measure can be used to prove the existence of objects with interesting properties. We describe these arguments mainly by giving examples from additive number theory and harmonic analysis. We do not mean to and we cannot exhaustively describe all different ways of using probabilistic ideas to prove theorems in these two branches of mathematics.

The common characteristic of the two kinds of probabilistic arguments is the ad-hoc definition of the underlying probability measure. This definition depends much on the problem at hand but a guiding principle in choosing a probability measure for a certain collection of objects is to ensure at least *good behavior on the average*. Having made our choice of the probability measure the probabilistic arguments that we want to discuss can be described as follows.

1. *The Average Value argument*
   The existence of an object having a certain property follows from the fact that the expected value of the *single* quantity of interest (with respect to the probability measure that we have defined) falls in the desired range.

*School of Mathematics, Institute for Advanced Study, Olden Lane, Princeton, New Jersey 08540. E-mail: kolount@math.ias.edu. Supported by NSF grant DMS 9304580.

2. *The Large Deviation argument*
   When the goodness of an element in our probability space depends on more than one quantity (*random variable*) we cannot establish the existence of a good object by knowing just that these random variables have expected values in the appropriate range. This is so simply because this kind of information does not guarantee the existence of a point in the probability space at which all random variables of interest are *simultaneously* in the proper range. We then need to bound the probability that each of these random variables deviates from its expected value (*the probability of large deviation*) and show that, in total, these probabilities amount to less than 1.

The lack of the power to "construct" the solution to a specific problem is an inherent characteristic of the probabilistic method. Very frequently the probabilistic proof of a theorem is extremely simple compared to a bare hands constructive proof, and that is to be expected since it furnishes less: the mere existence of a solution to a problem rather than the solution itself. Probabilistic proofs are thus almost universally regarded as inferior, if simpler, than "constructive" proofs, and probably with good reason. Yet, one of the points that we want to make here is that, very often, a probabilistic proof can easily be turned into a construction *if one assumes the point of view, that an efficient algorithm is a construction.* We shall give several examples of this so-called *derandomization* of a probabilistic proof.

2 The prototype Average Value argument

The prototype example of the probabilistic method in this form can be considered to be the following obvious statement.

Proposition 2.1 *If* $x_1, \ldots, x_n \in \mathbf{R}$ *and*

$$\frac{x_1 + \cdots + x_n}{n} \geq a \tag{1}$$

then for some j

$$x_j \geq a. \tag{2}$$

The usefulness of the method lies in the fact that the average (1) is often easier to compute than exhibiting a specific x_j for which (2) can be proved to hold.

Let us rephrase Proposition 2.1 in the following more useful form. The measure space Ω is equipped with a nonnegative measure $d\mathbf{Pr}$ of total mass 1, and a real random variable X on Ω is just a measurable function $X : \Omega \to \mathbf{R}$.

Proposition 2.2 *Let* X *be a real random variable on a probability space* $(\Omega, d\mathbf{Pr})$ *whose expected value*

$$\mathbf{E}[X] = \int_\Omega X(\omega) d\mathbf{Pr}(\omega)$$

satisfies

$$\mathbf{E}[X] \geq a.$$

Then there is $\omega \in \Omega$ *such that*

$$X(\omega) \geq a.$$

We remark that because of the obvious linearity property of the expectation of a random variable

$$\mathbf{E}[\alpha X_1 + \beta X_2] = \alpha \mathbf{E}[X_1] + \beta \mathbf{E}[X_2]$$

(whenever the right hand side makes sense), the expected value of quantities of interest are almost always very easy to compute or at least to estimate very well. Notice that no independence is required of the pair X_1, X_2.

We proceed to give some examples.

2.1 An example from graph theory

Let n, m, with $n \geq m \geq 3$, be two positive integers. We denote by K_n the complete graph on n vertices. We want to color the edges of K_n with two colors, say red and blue, so that it contains few monochromatic copies of K_m (m-cliques). Of course it is easy to have many monochromatic K_m's by coloring every edge with the same color. For each subset A of $[n] = \{1, \ldots, n\}$ with $|A| = m$ we define the function of the coloring

$$\chi_A = \begin{cases} 1 & \text{if } A \text{ is monochromatic,} \\ 0 & \text{otherwise.} \end{cases}$$

Then the number X of monochromatic K_m's is

$$X = \sum_{A \subseteq [n],\ |A|=m} \chi_A. \tag{3}$$

We color each edge of K_n red or blue with equal probability $1/2$ and independently of the other edges, i.e. we toss a fair coin for each edge. The expected value of χ_A is then $2(\frac{1}{2})^{\binom{m}{2}}$ and by the linearity of expectation and (3) we get

$$\mathbf{E}[X] = \binom{n}{m} 2^{1-\binom{m}{2}}.$$

We have proved:

Theorem 2.1 *There is a 2-coloring of the edges of the graph K_n which gives rise to no more than $\binom{n}{m}2^{1-\binom{m}{2}}$ monochromatic K_m's.*

2.2 A large sum-free subset of a given set of integers

A subset E of an additive group is called *sum-free* if

$$x + y \neq z, \quad \text{for all } x, y, z \in E. \tag{4}$$

The following theorem of Erdős [10], [2] has a beautiful probabilistic proof. See also [19] for a similar, but computationally more efficient, approach.

Theorem 2.2 *Let $A \subseteq \mathbf{N}$ be a set of N positive integers. Then there is a sum-free subset E of A with*

$$|E| > \frac{1}{3}N.$$

Proof: Let $A = \{n_1 < \cdots < n_N\}$ and choose any prime $p > n_N$ such that $p = 3k + 2$ for some $k \in \mathbf{N}$. View the set A as a subset of the multiplicative group of units of the field $\mathbf{Z}_p$ (the integers mod p). Write

$$S = \{k + 1, \ldots, 2k + 1\}$$

and notice that $|S| > (p-1)/3$ and S is sum-free as a subset of $\mathbf{Z}_p$. Let t be uniformly distributed over $\mathbf{Z}_p^\times = \{1, \ldots, p - 1\}$ and write

$$X = |S \cap (t \cdot A)|,$$

where $t \cdot A = \{t \cdot n_1, \ldots, t \cdot n_N\}$ and the arithmetic is in $\mathbf{Z}_p$. Since

$$X = \sum_{j \in S} \mathbf{1}(t^{-1}j \in A)$$

and

$$\mathbf{E}[\mathbf{1}(t^{-1}j \in A)] = \frac{N}{p-1}, \quad \text{for all } j \in \mathbf{Z}_p^\times$$

($\mathbf{Z}_p^\times$ is a multiplicative group), we have

$$\mathbf{E}[X] = \frac{|S|N}{p-1} > \frac{N}{3}.$$

This implies that there is $t_0 \in \mathbf{Z}_p^\times$ for which $X > N/3$. Define then

$$E = A \cap (t_0^{-1}S).$$

It follows that E is sum-free as a set of integers (even more, it is sum-free mod p) and $|E| > N/3$, as we had to show. □

Open Problem: What is the largest constant that works in place of $1/3$ in the previous theorem? It must be smaller than $12/29$ [2].

Bourgain [5] remarked to the author that a similar result can be proved about infinite sequences of positive integers.

Theorem 2.3 *Let $A = \{a_1, a_2, \ldots\}$ be an infinite sequence of positive integers and write $A(x) = |A \cap [1,x]|$ for the counting function of A. Then there is an infinite sum-free subsequence E of A with*

$$\lim_{x\to\infty} \frac{E(x)}{A(x)} = \frac{1}{3}.$$

Note that Theorem 2.3 does not follow from Theorem 2.2 about sum-free subsets of finite sets.

The proof of Theorem 2.3 follows from the following well known result (see [23, p. 32]) on uniform distribution of sequences mod 1.

Theorem 2.4 *Let $A = \{a_1, a_2, \ldots\}$ be an infinite sequence of positive integers. Then for almost all (Lebesgue) real numbers x the sequence $\{xa_n\}$ is uniformly distributed mod 1, that is for all $\alpha, \beta \in [0,1]$, $\alpha < \beta$,*

$$\lim_{N\to\infty} \frac{|\{(xa_1) \bmod 1, \ldots, (xa_N) \bmod 1\} \cap (\alpha,\beta)|}{N} = \beta - \alpha. \tag{5}$$

Above we denote by $x \bmod 1$ the fractional part of the real number x.

Let then x be one such real number for which (5) holds and define

$$E = \{a \in A : \ xa \bmod 1 \in (1/3, 2/3)\}.$$

Clearly then E is sum-free and

$$\lim_{x\to\infty} \frac{E(x)}{A(x)} = \frac{1}{3}$$

follows from (5).

2.3 Uchiyama's theorem on the L^1 norm of trigonometric polynomials

The following theorem regarding the L^1 norm of trigonometric polynomials was proved by Uchiyama [29]. It is related to the so-called Littlewood Conjecture [20, 25].

Theorem 2.5 *Let $A = \{n_1 < \cdots < n_N\}$ be a set of N positive integers. Then there is a subset $E \subseteq A$ such that*

$$\left\| \sum_{j\in E} e^{ijx} \right\|_1 \ge C\sqrt{N}, \tag{6}$$

where C is a positive constant.

Proof: Let $g(x) = \sum_{j\in A} e^{ijx}$ and

$$f(x) = \sum_{j\in A} \epsilon_j e^{ijx},$$

where $\epsilon_j = \pm 1$ with equal probability and independently. By the triangle inequality it suffices to show that there is an assignment to ϵ_j that makes $\|f\|_1 \gg \sqrt{N}$. To this end we use Hölder's inequality in the form

$$\|f\|_2^2 \le \|f\|_1^{2/3} \|f\|_4^{4/3}.$$

We always have $\|f\|_2^2 = N$ and

$$|f|^4 = \left|\sum_{j,k} \epsilon_j \epsilon_k e^{i(j-k)x}\right|^2.$$

Thus writing (the indices j, k always run through A)

$$r(x) = \sum_{x=j-k} \epsilon_j \epsilon_k$$

we get

$$\|f\|_4^4 = \frac{1}{2\pi}\int_0^{2\pi} |f|^4 = \sum_{x\in\mathbf{N}} r^2(x)$$

and

$$\mathbf{E}[\|f\|_4^4] = \sum_{x\in\mathbf{N}} \mathbf{E}[r^2(x)] = \sum_{j-k=j'-k'} \mathbf{E}[\epsilon_j\epsilon_k\epsilon_{j'}\epsilon_{k'}]$$

and the only terms that will survive are essentially those with $j = j'$, $k = k'$, thus

$$\mathbf{E}[\|f\|_4^4] = N^2 + o(N^2).$$

This implies the existence of an assignment of the ϵ_j such that $\|f\|_4^4 \le N^2 + o(N^2)$. Using Hölder's inequality for this assignment we get $\|f\|_1 \ge N^{1/2} - o(N^{1/2})$ which concludes the proof. □

3 The prototype Large Deviation argument

Often we associate several quantities $X_1, \ldots, X_n$ with a random object (n can be infinite). Typically their averages $\mathbf{E}[X_1], \ldots, \mathbf{E}[X_n]$ will be easy to compute or estimate and their value will be in the desirable range. Our objective is to have the values of the random variables X_j themselves in that range, simultaneously for all j.

Having found a proper distribution of random objects, namely one for which the expected values $\mathbf{E}[X_j]$ are of the desirable magnitude, we still need to bound the probability that *some* X_j deviates too much from its expected value. That is, we want an upper bound on

$$\mathbf{Pr}\,[|X_j - \mathbf{E}[X_j]| > d_j, \text{ for some } j]. \tag{7}$$

The maximum allowed deviations d_j are problem dependent.

It is usually the case that the best upper bound that we know for this probability is

$$\sum_{j=1}^{n} \mathbf{Pr}\,[|X_j - \mathbf{E}[X_j]| > d_j].$$

So we aim for this sum to be strictly less than 1. This implies that with positive probability none of the *bad events*

$$B_j = \{|X_j - \mathbf{E}[X_j]| > d_j\},\ j = 1, \ldots, n,$$

holds. In particular there is an object for which the quantities X_j satisfy

$$\mathbf{E}[X_j] - d_j \ \le\ X_j \ \le\ \mathbf{E}[X_j] + d_j, \text{ for } j = 1, \ldots, n.$$

To achieve this we can use several well known Large Deviation inequalities.

3.1 Inequalities for Large Deviations

The following two inequalities are straightforward to prove.

Proposition 3.1 (Markov's Inequality) *If X is any nonnegative random variable with finite expectation then for all $\alpha > 0$*

$$\mathbf{Pr}\left[X > \alpha \mathbf{E}[X]\right] \le \frac{1}{\alpha}. \tag{8}$$

Proposition 3.2 (Chebyshev's Inequality) *If X is any real random variable with finite* variance $\sigma^2 = \mathbf{E}[(X - \mathbf{E}[X])^2]$ *then for all $\alpha > 0$*

$$\mathbf{Pr}\left[|X - \mathbf{E}[X]| > \alpha\sigma\right] \le \frac{1}{\alpha^2}. \tag{9}$$

The inequalities of Markov and Chebyshev are rather weak in most cases but they are applicable to virtually any random variable and this makes them very useful.

In the following theorems the random variable X is assumed to be of a special form: a sum of independent random variables.

Theorem 3.1 (Chernoff [6], [3, p. 239]) *If $X = X_1 + \cdots + X_k$, and the X_j are independent indicator random variables (that is $X_j \in \{0, 1\}$), then for all $\epsilon > 0$*

$$\mathbf{Pr}\left[|X - \mathbf{E}[X]| > \epsilon \mathbf{E}[X]\right] \le 2e^{-c_\epsilon \mathbf{E}[X]},$$

where $c_\epsilon > 0$ is a function of ϵ alone

$$c_\epsilon = \min\left\{-\log\left(e^\epsilon (1+\epsilon)^{-(1+\epsilon)}\right), \epsilon^2/2\right\}.$$

We call a random variable X which, as above, is the *sum of independent indicator random variables* a SIIRV.

Remarks on Theorem 3.1:

1. Observe that if $X = X' + X''$, where X' and X'' are SIIRV then we have

$$\mathbf{Pr}\left[|X - \mathbf{E}[X]| > \epsilon \mathbf{E}[X]\right] \le 4e^{-c_\epsilon \min\{\mathbf{E}[X'], \mathbf{E}[X'']\}}.$$

2. Since there is no dependence of the bound on k (the number of summands in X), it is easy to prove that the same bound holds for $X = \sum_{j=1}^\infty X_j$, provided that $\sum_{j=1}^\infty \mathbf{E}[X_k] < \infty$.
3. Suppose one has to control n different SIIRV variables, say $Y_1, \ldots, Y_n$. By this we mean that we want to ensure that each Y_j is of the order of its expected value, with high probability. In order then for Theorem 3.1 to be of any use the expectations $\mathbf{E}[Y_j]$ must be large, at least $C \log n$ where the constant C can be made as large as we please. One cannot control with this theorem random variables with sublogarithmic expectations.

Theorem 3.2 [3, p. 236] *Let $p_1, \ldots, p_n \in [0, 1]$ and let the independent zero-mean random variables $X_1, \ldots, X_n$ have the distribution*

$$X_j = \begin{cases} 1 - p_j & \textit{with probability } p_j, \\ -p_j & \textit{with probability } 1 - p_j. \end{cases}$$

If $X = a_1 X_1 + \cdots + a_n X_n$, where $a_1, \ldots, a_n \in \mathbf{C}$, then we have for all $a > 0$

$$\mathbf{Pr}\left[|X| > a\right] \le C_1 \exp\left(-C_2 a^2 / \sum_{j=1}^n |a_j|^2\right),$$

where $C_1, C_2 > 0$ are two absolute constants. In particular, if $|a_j| \le 1$ we have the above probability bounded above by $C_1 \exp\left(-C_2 a^2/n\right)$.

Theorems 3.1 and 3.2 are extremely useful. In the next section we show a nice application of Theorem 3.1 to a problem in additive number theory.

3.2 An asymptotic additive basis with small representation function

A set E of positive integers is called an *asymptotic additive basis of order* 2 if the *representation function*

$$r(x) = r_E(x) = |\{(a,b)\ :\ a,b \in E\ \&\ a \le b\ \&\ x = a+b\}|$$

is strictly positive for all sufficiently large integers x. In other words all sufficiently large x can be expressed as a sum of two elements of E. Examples of asymptotic additive bases are the set $\mathbf{N}$ of natural numbers itself and the set $\{1,2,4,6,8,\ldots\}$. We are interested in bases for which the representation function is small. Notice that in the previous two examples $r(x)$ can be as large as Cx.

We present Erdős's probabilistic proof [8, 9], [14, Ch. 3] that there is an asymptotic basis of order 2 such that

$$c_1 \log x \le r(x) \le c_2 \log x \tag{10}$$

for all sufficiently large x. The ratio of the two absolute constants c_1 and c_2 can be made arbitrarily close to 1.

Define the probabilities

$$p_x = K \cdot \left(\frac{\log x}{x}\right)^{1/2}$$

for the values of x for which the right hand side is in $[0,1]$; otherwise let $p_x = 0$. The constant K will be determined later in the proof. We define a random set E by letting

$$\mathbf{Pr}\left[x \in E\right] = p_x$$

independently for all x. We show that with high probability the random set E has the claimed property (10).

Define the indicator random variables

$$\chi_j = \mathbf{1}(j \in E)$$

with mean values $\mathbf{E}[\chi_j] = p_j$. We then have

$$r(x) = \sum_{j=1}^{\lfloor x/2 \rfloor} \chi_j \chi_{x-j}$$

from which and the independence of χ_j it follows that

$$\mathbf{E}[r(x)] = \sum_{j=1}^{\lfloor x/2 \rfloor} p_j p_{x-j}. \tag{11}$$

Notice also that, for each fixed x, $r(x)$ is a SIIRV. Easy calculations on the right hand side of (11) allow the asymptotic estimate

$$\mathbf{E}[r(x)] \sim IK^2 \log x,$$

where $I = \int_0^{1/2} (s(1-s))^{-1/2} ds$. We now define the bad events

$$A_x = \left\{ |r(x) - \mathbf{E}[r(x)]| > \frac{1}{2}\mathbf{E}[r(x)] \right\},\ \ x = 1,2,3,\ldots.$$

Using Theorem 3.1 we can bound

$$\mathbf{Pr}\left[A_x\right] \le 2\exp\left(-\frac{1}{2} c_{1/2} IK^2 \log x\right) = 2x^{-\alpha}$$

where $\alpha = \frac{1}{2}c_{1/2}IK^2$. All we have to do now is to choose the constant K large enough to have $\alpha > 1$. We deduce that $\sum_x \mathbf{Pr}\left[A_x\right]$ is a convergent series and thus there is $n_0 \in \mathbf{N}$ for which

$$\sum_{x \geq n_0} \mathbf{Pr}\left[A_x\right] < 1,$$

which implies that with positive probability none of the events A_x, $x \geq n_0$, holds. This in turn implies the existence of a set $E \subseteq \mathbf{N}$ such that

$$\frac{1}{2}IK^2 \log x \leq r(x) \leq \frac{3}{2}IK^2 \log x$$

for all $x \geq n_0$, which concludes the proof.

We emphasize the structure of the proof. First we defined an appropriate class of random objects (random subsets of $\mathbf{N}$). We then showed that the quantities of interest (the numbers $r(x)$, $x \in \mathbf{N}$) have expected values of the desired size. The last, and most important, step was to show that, with high probability, none of the quantities of interest deviates much from its expected value.

Open Problem: Is it possible to have $1 \leq r(x) = o(\log x)$? Note that Theorem 3.1 is useless if $\mathbf{E}[r(x)] = o(\log x)$ (see the remarks following that theorem). Is it possible to have $r(x) = C \log x + o(\log x)$? See [14] for problems related to thin additive bases.

3.2.1 Good asymptotic bases of higher order

If one tries to prove a similar theorem for additive bases of order $k \geq 3$, i.e. sets of integers such that any sufficiently large integer can be written as a sum of k of them, one encounters an extra difficulty. Given a set of integers E let us write again

$$r_k(x) = \left|\{(a_1, \ldots, a_k) \in E^k \ :\ a_1 \leq \cdots \leq a_k \ \&\ x = a_1 + \cdots + a_k\}\right|$$

for the number of representations of the integer x as a sum of k elements of E, without taking the order of the summands into account. It is easy to see that the proper class of random sets are those defined by

$$\mathbf{Pr}\left[x \in E\right] = K \frac{\log^{1/k} x}{x^{(k-1)/k}},$$

for $K > 0$ a sufficiently large constant. That is for this class of random sets we have

$$\mathbf{E}[r_k(x)] \sim C_k \log x.$$

All that is missing now in order to prove the existence of sets E for which $C_1(k) \log x \leq r_k(x) \leq C_2(k) \log x$ is a large deviation argument for the random variables $r_k(x)$, which do have the correct order of magnitude. These random variables can, once more, be expressed as

$$r_k(x) = \sum \chi_{a_1} \chi_{a_2} \cdots \chi_{a_k} \tag{12}$$

where the summation extends over all k-tuples $(a_1, \ldots, a_k) \in \mathbf{N}^k$ which satisfy $x = a_1 + \cdots + a_k$. Thus $r_k(x)$ is again a sum of indicator random variables but, already in the case $k = 3$, these are not independent, since a certain χ_j appears in many terms in this representation of $r_k(x)$. Thus the Chernoff bound (Theorem 3.1) is not applicable here.

The problem has been solved recently by Erdős and Tetali [12], [3, p. 108] who proved the following.

Theorem 3.3 *There is an asymptotic basis of order k of the integers such that*

$$c_1 \log x \leq r_k(x) \leq c_2 \log x$$

where $c_1, c_2 > 0$ depend only on k and not on x.

The tools used to tackle the lack of independence in the summands of $r_k(x)$ in (12) were the so called Janson inequalities [3, p. 95], [15] which allow for sparse dependencies in cases like this.

3.3 The density of infinite $B_h[g]$ sets

Let $E \subseteq \mathbf{N}$ and define a corresponding representation function on $\mathbf{N}$

$$r(x) = r_E(x) = |\{(a,b) \ : \ a,b \in E \ \& \ a \le b \ \& \ x = a+b\}|. \tag{13}$$

We say that the set E is in the class B_2 if $r(x) \le 1$ for all $x \in \mathbf{N}$. In other words all sums of the form

$$a+b, \ a,b \in E, \tag{14}$$

are distinct except for permutation of a and b. It is not hard to see that this condition is equivalent to requiring that all differences

$$a-b, \ a,b \in E, \ a \ne b, \tag{15}$$

are distinct. The terminology "Sidon set" is sometimes used to describe B_2 sets but we will avoid it since it has a rather different meaning in harmonic analysis.

Remember the definition

$$r_h(x) = |\{(a_1,\ldots,a_h) \ : \ a_j \in E \ \& \ a_1 \le \cdots \le a_h \ \& \ x = a_1+\cdots+a_h\}|. \tag{16}$$

We call a set E a B_h set if $r_h(x) \le 1$ for all $x \in \mathbf{N}$. We call it a $B_h[g]$ set if $r_h(x) \le g$ for all $x \in \mathbf{N}$. Thus a B_h set is a set of which all sums of the form

$$a_1+\cdots+a_h, \ a_j \in E, \ a_1 \le \cdots \le a_h,$$

are distinct.

While it is possible to have a B_2 subset of $\{1,\ldots,n\}$ with about $\sqrt{n}$ elements, the following theorem of Erdős [14, p. 88], [28] shows that the situation is quite different if we look at infinite B_2 sequences of high lower density.

Theorem 3.4 *If the sequence $\{n_1 < n_2 < \cdots\} \subseteq \mathbf{N}$ is B_2 then we have*

$$\limsup_j \frac{n_j}{j^2 \log j} \ge 0. \tag{17}$$

Thus we cannot have a (finite or infinite – the infinite sequence can be obtained from finite sequences by a diagonal argument) B_2 sequence which satisfies for all j

$$n_j \ll j^2.$$

For a long time the B_2 sequence with the highest lower density known was the one produced by the so called greedy method. Let $n_1 = 1$ and having found $n_1,\ldots,n_k$ choose n_{k+1} to be the smallest positive integer x that is not in the set

$$\{a+b-c \ : \ a,b,c \in \{n_1,\ldots,n_k\}\}.$$

It then follows easily that the sequence n_j is B_2 and that $n_j \le j^3$. The gap between this sequence and Theorem 3.4 still stands except for the following result of Ajtai, Komlós and Szemerédi [1].

Theorem 3.5 *There is a B_2 sequence $\{n_1 < n_2 < \cdots\} \subseteq \mathbf{N}$ such that*

$$n_j \ll \left(\frac{j^3}{\log j}\right).$$

Open Problem: Prove that if the sequence A is of type $B_h[g]$, $g \ge 2$, then

$$\liminf_{x \to \infty} \frac{A(x)}{x^{1/h}} = 0.$$

This is open for all $g \geq 2$. The difficulty in proving this comes from the fact that the only known proof (case of $h = 2, g = 1$) uses the fact that all *differences* (not sums) from A are distinct. This, of course, fails for higher g. Related is the following problem. By just counting the distinct sums of a $B_2[2]$ set $A \subseteq [1,n]$ one gets $|A| \leq 2\sqrt{2}n^{1/2} + o(n^{1/2})$. I believe that this is too high an upper bound but I do not think anything better is known. If the case of $B_2[1]$ sets is suggestive at all then by counting distinct sums we get an upper bound $\sim 2n^{1/2}$ for the size of a $B_2[1]$ set $A \subseteq [1,n]$, by counting distinct differences we get $\sim \sqrt{2}n^{1/2}$, while the true bound is $\sim n^{1/2}$ [7, 13], [14, Ch. 2].

What is the largest constant c for which there is an *infinite* $B_2[1]$ sequence A with

$$\limsup_{x\to\infty} \frac{A(x)}{x^{1/2}} = c?$$

The largest known is $c = 1/\sqrt{2}$ [22], [14, Ch. 2] while by the upper bound for the finite $B_2[1]$ sets one might have c up to 1. The only way known for constructing such sets A with large c is by piecing together dense finite sets of type $B_2[1]$.

The following theorem of Erdős and Rényi deals with dense infinite $B_2[g]$ sequences. The proof is once again probabilistic.

Theorem 3.6 (Erdős and Rényi [11], [14, Ch. 3]) *For every $\delta' > 0$ there is an integer g and a $B_2[g]$ sequence $A = \{a_1 < a_2 < \cdots\}$ such that*

$$a_j \ll j^{2+\delta'}, \tag{18}$$

for all $j > 0$.

Proof: Let $\delta' \in (0,1)$ be given. Let A be a random set with

$$\mathbf{Pr}\,[x \in A] = p_x,$$

independently for all $x \in \mathbf{N}$, where

$$p_x = x^{-1/2-\delta/2}.$$

Then with high probability $A(x) \gg x^{1/2-\delta/2}$ for all x, which implies (18) for $\delta = \delta(\delta')$ properly chosen. Write, as usual, $\chi_j = \mathbf{1}(j \in A)$. Then we have

$$r(x) = \sum_{j=1}^{\lfloor x/2 \rfloor} \chi_j \chi_{x-j}$$

and we can estimate

$$\mathbf{E}[r(x)] \leq Cx^{-\delta},$$

where $C = 2\int_0^{1/2}(s(1-s))^{-1/2-\delta/2}ds$. Define the bad events

$$A_x = \{r(x) > g\} = \left\{ r(x) > (\frac{g}{\mathbf{E}[r(x)]})\mathbf{E}[r(x)] \right\}.$$

We now use the Chernoff bound (Theorem 3.1) with

$$\epsilon = \frac{g}{\mathbf{E}[r(x)]} \geq Cgx^{\delta},$$

observing in Theorem 3.1 that

$$c_\epsilon \sim \epsilon \log \epsilon, \text{ as } \epsilon \to \infty.$$

We get

$$\mathbf{Pr}\,[A_x] \leq 2e^{-2c_\epsilon \mathbf{E}[r(x)]} \leq 2e^{-3(\epsilon \mathbf{E}[r(x)])\log \epsilon} = 2e^{-3g \log \epsilon},$$

and using the estimate on $\mathbf{E}[r(x)]$ we get

$$\mathbf{Pr}\left[A_x\right] \le Ce^{-Cg\delta \log x} = Cx^{-Cg\delta}.$$

Choose now $g = C/\delta$, for large enough C, to get $\mathbf{Pr}\left[A_x\right] \ll x^{-2}$ and thus $\sum_x \mathbf{Pr}\left[A_x\right] < \infty$. So there is $n_0 \in \mathbf{N}$ for which

$$\sum_{x \ge n_0} \mathbf{Pr}\left[A_x\right] < 1,$$

so that with positive probability none of the bad events A_x, $x \ge n_0$, holds. Now discard all elements of the set A up to n_0 to get a $B_2[g]$ set with the desired growth. □

3.4 The Salem-Zygmund theorem for trigonometric polynomials with random coefficients

The following theorem is often used to estimate the maximum of a random trigonometric polynomial

Theorem 3.7 (Salem and Zygmund [26], [16, p. 69]) *Let $f_1(x), \ldots, f_n(x)$, be trigonometric polynomials of degree at most m, and $\xi_1, \ldots, \xi_n$ be independent zero-mean random variables*

$$\xi_j = \begin{cases} 1 - p_j & \textit{with probability } p_j, \\ -p_j & \textit{with probability } 1 - p_j, \end{cases} \tag{19}$$

for some $p_j \in [0,1]$. Write

$$f(x) = \sum_{j=1}^{n} \xi_j f_j(x).$$

Then, for some $C > 0$,

$$\mathbf{Pr}\left[\|f\|_\infty \le C \left(\sum_{j=1}^{n} \|f_j\|_\infty^2 \log m \right)^{1/2} \right] \to 1, \quad \textit{as } m \to \infty.$$

For the proof of the Salem-Zygmund theorem we need the following.

Theorem 3.8 *Let a_{ij}, $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$, be a matrix of complex numbers, such that $|a_{ij}| \le r_j$. Let also $p_1, \ldots, p_{n_2} \in [0,1]$ and the random variables $\xi_1, \ldots, \xi_{n_2}$ be defined as in (19). Then with probability tending to 1 as $n_1 \to \infty$*

$$\left| \sum_{j=1}^{n_2} a_{ij}\xi_j \right| \le C \left(\sum_{j=1}^{n_2} r_j^2 \log n_1 \right)^{1/2}, \quad \textit{for all } i = 1, \ldots, n_1,$$

where C is an absolute constant.

Proof: Define

$$L_i(\xi) = \sum_{j=1}^{n_2} a_{ij}\xi_j.$$

We can clearly work on the real and imaginary parts of the linear forms L_i separately, so we assume $a_{ij} \in \mathbf{R}$. Define the bad events

$$A_i = \left\{ |L_i(\xi)| > C \left(\sum_{j=1}^{n_2} r_j^2 \log n_1 \right)^{1/2} \right\}.$$

Using Theorem 3.2 we get

$$\mathbf{Pr}\left[A_i\right] \le 2\exp\left(-2C^2\sum_{j=1}^{n_2} r_j^2 \log n_1 / \sum_{j=1}^{n_2} r_j^2\right) = C_1 n_1^{-C_2C^2},$$

where the constants $C_1, C_2 > 0$ of Theorem 3.1 are absolute. Now choose the constant $C = (2/C_2)^{1/2}$ to get

$$\mathbf{Pr}\left[\bigcup_{i=1}^{n_1} A_i\right] \le \sum_{i=1}^{n_1} \mathbf{Pr}\left[A_i\right] \le \frac{C_1}{n_1},$$

which concludes the proof. □

To complete the proof of the Salem-Zygmund theorem we note that it is enough to ensure that $f(x_j)$ is small for a sufficiently dense set of points $x_j \in [0, 2\pi)$.

Since f is a trigonometric polynomial of degree at most m we can use Bernstein's inequality [17, p. 12]:

$$\|f'\|_\infty \le m\|f\|_\infty.$$

Define $x_i = i\frac{2\pi}{10m}$ for $i = 1, \ldots, 10m$ and the matrix

$$a_{ij} = f_j(x_i), \quad i = 1, \ldots, 10m, \ j = 1, \ldots, n,$$

for which of course $|a_{ij}| \le \|f_j\|_\infty$. Notice that for all $i = 1, \ldots, 10m$

$$f(x_i) = \sum_{j=1}^{n} \xi_j f_j(x_i) = \sum_{j=1}^{n} \xi_j a_{ij}.$$

From this and Theorem 3.8 follows that

$$\mathbf{Pr}\left[|f(x_i)| \le C\left(\sum_{j=1}^{n} \|f_j\|_\infty{}^2 \log m\right)^{1/2}, \text{ for all } i\right] \to 1 \tag{20}$$

as $m \to \infty$. But the event in (20) implies that $|f(x)| \le C(\sum_{j=1}^{n} \|f_j\|_\infty{}^2 \log m)^{1/2}$ for all $x \in [0, 2\pi)$ and for a larger constant C. For assume that $|f(x_0)| = \|f\|_\infty$ and that

$$|x_k - x_0| \le \frac{2\pi}{10m}.$$

Then, using Bernstein's inequality,

$$|f(x_0) - f(x_k)| \le \frac{2\pi}{10m}\|f'\|_\infty \le \frac{2\pi}{10}|f(x_0)|$$

and, since $2\pi/10 < 1$, we get

$$\|f\|_\infty = |f(x_0)| \le C|f(x_k)| \le C\left(\sum_{j=1}^{n} \|f_j\|_\infty{}^2 \log m\right)^{1/2}.$$

For some applications of the Salem-Zygmund theorem to harmonic analysis see, for example, [4, 16, 20, 21, 24].

4 Are probabilistic proofs constructive?

After a probabilistic proof of the existence of a certain object has been given, it is frequently followed by the remark that " ...although we have proved that almost all elements in this class satisfy our requirements, no one knows how to construct a single one of them". In mathematics the terms *probabilistic proof* and *existential proof* have long been thought of as the former implying the latter. This is a conception that we would like to shake a little bit.

To debate this matter one has to make clear, before the discussion begins, what one means by "constructive". We shall mean the following. We say that we have a constructive proof of the existence of an object if we can give an algorithm, that can run on an ordinary computer, to construct the object and this algorithm takes a reasonable time to finish. By reasonable we usually mean a number of steps bounded by a fixed polynomial of the parameter of the problem (what that parameter is is usually clear – though not always).

As an example of what is *not* acceptable as a construction, suppose that we have proved a theorem stating that, with high probability, a random object that depends on the random variables $X_1, \ldots, X_n$ has a certain property. Assume for simplicity that the X_j's are independent 0-1 random variables, with $\mathbf{Pr}[X_j = 1] = p_j$, the p_j's being fixed given numbers. And as a minimal requirement on the properties that we are trying to ensure our object has we demand that for a given assignment $X_j = 0$ or 1 it is easy to verify (i.e. there exists an efficient algorithm) whether the object that corresponds to this assignment has the desired property. The following simple-minded algorithm is then *not* acceptable as a construction of a good object: check all possible assignments of the X_j's and pick the first that has the property (we know one exists from the proof). Clearly this algorithm may take time that is not bounded by any polynomial in n. On the other hand if we managed to find a good assignment in time $O(n^{10})$ we consider that good enough, if not practical, a construction.

Not everybody agrees with this definition. Many insist that they should be able to actually "see" the object itself rather than the algorithm that will construct it. It is of course impossible to define this rigorously. Nevertheless mathematicians will usually agree that a certain proof fits these aesthetic requirements or not. For the sake of distinguishing this concept from what we have already termed constructive, we call these proofs *explicit*. Thus we think of explicitness as a property more specific than constructibility, but we do not deal with it here. We stick to our definition of what is constructive and proceed to show that many well known proofs of theorems that have usually been called existential are, indeed, easy to turn into "efficient" constructions.

A single method of *derandomization*, a way of turning a probabilistic proof into an efficient algorithm, will be described here. It is the so called *method of conditional probabilities* and is, perhaps, the simplest and most widely applicable derandomization technique. It applies to a great many problems and the requirements for its applicability are easy to state and check. We demonstrate it with a few examples. For a more thorough treatment of this very interesting subject see for example [3, p. 223].

4.1 Coloring a complete graph's edges for few monochromatic cliques

We first describe the method of conditional probabilities on the proof of Theorem 2.1, which stated that there is a 2-coloring of the edges of the complete graph K_n on n vertices such that the number of monochromatic copies of K_m is at most

$$\binom{n}{m} 2^{1-\binom{m}{2}}.$$

Assume that m is fixed and our task is to produce such a coloring of K_n. Trying every possible coloring clearly takes too much time since there are $2^{\binom{n}{2}}$ possible colorings. We describe the derandomization process and arrive at an algorithm for finding such a coloring in time polynomial in n. We keep the same notation as in Section 2.1.

Let $A_1, \ldots, A_k$, where $k = \binom{n}{m}$, be all the copies of K_m in K_n (otherwise known as m-cliques), and enumerate all edges of K_n as $e_1, \ldots, e_{\binom{n}{2}}$. Let the color of edge e_j be the random variable

c_j. We are going to define the colors $a_j \in \{RED,\ BLUE\}$ one by one, for $j = 1, \ldots, \binom{n}{2}$. Define the events

$$R_j = R_j(a_1, \ldots, a_j) = \{(c_1, \ldots, c_{\binom{n}{2}}) : c_1 = a_1, \ldots, c_j = a_j\}.$$

R_0 is the whole probability space. Intuitevely, R_j represents our choices of colors up to the j-th color.

As in Section 2.1 we define the 0-1-valued random variable χ_j to indicate whether A_j is monochromatic or not. We have $X = \sum_{j=1}^k \chi_j$ and we already computed $\mathbf{E}[X] = k 2^{1-\binom{m}{2}}$. We are going to choose the sequence of colors $a_1, \ldots, a_{\binom{n}{2}}$ so that the function of j

$$\mathbf{E}[X \mid R_j]$$

is non-increasing. This is possible for the following general reason (the nature of the variable X is immaterial here):

$$\begin{aligned}&\mathbf{E}[X \mid R_{j-1}(a_1, \ldots, a_{j-1})] = \\ &\quad \frac{1}{2}(\mathbf{E}[X \mid R_j(a_1, \ldots, a_{j-1}, RED)] + \mathbf{E}[X \mid R_j(a_1, \ldots, a_{j-1}, BLUE)]).\end{aligned}$$

This means that at least one of the choices $a_j = RED$ or $a_j = BLUE$ will yield $\mathbf{E}[X \mid R_j] \le \mathbf{E}[X \mid R_{j-1}]$. Which of the two choices works can be decided (here the nature of X plays a role) since we can explicitly compute

$$\mathbf{E}[X \mid R_j(a_1, \ldots, a_j)]$$

for any colors $a_1, \ldots, a_j$. This computation clearly takes time polynomial in n. We proceed like this until the colors of all vertices have been fixed. Then X is completely determined

$$X = \mathbf{E}[X \mid R_{\binom{n}{2}}] \le \cdots \le \mathbf{E}[X \mid R_0] = \mathbf{E}[X] \le \binom{n}{m} 2^{1-\binom{m}{2}}$$

and the coloring $a_1, \ldots, a_{\binom{n}{2}}$ is thus a solution to our problem.

4.2 When can the method be applied?

The very general applicability of the method just described on the example from graph theory should be obvious. The general context is the following. We have n independent random variables $\epsilon_1, \ldots, \epsilon_n$ which, without loss of generality, can be assumed to have the distributions

$$\epsilon_j = \begin{cases} 1 & \text{with probability } p_j, \\ 0 & \text{with probability } 1 - p_j. \end{cases}$$

We also have a certain function $X = X(\epsilon_1, \ldots, \epsilon_n)$ for which we can efficiently compute

$$\mathbf{E}[X \mid \epsilon_1 = v_1, \ldots, \epsilon_m = v_m]$$

for any m and $v_1, \ldots, v_m \in \{0, 1\}$. In particular we can compute $\mathbf{E}[X] = \mu$. We can then, as in the example of the previous section, efficiently compute an assignment

$$\epsilon_1 = v_1, \ldots, \epsilon_n = v_n, \quad v_1, \ldots, v_n \in \{0, 1\},$$

for which $X \le \mu$.

In this formalism we can also fit the following case. Suppose that we have proved for a certain event A that $\mathbf{Pr}[A] < 1$. (This event is in the space defined by the ϵ_j's of the previous. paragraph.) Suppose also that the characteristic function $X(\omega) = \mathbf{1}(\omega \in A)$ satisfies the computability requirements of the previous paragraph. Then we can find an assignment to the ϵ_j's which is not in A.

A good informal description of the method of conditional probabilities is the following. We are fixing the values of the random variables, one by one, taking care to assign them a value that will maximize the probability of our success *given the choices that we have made so far.*

4.3 Constructibility of certain trigonometric polynomials with small maximum

Theorem 3.8 has the following constructive equivalent [3, p. 225]. We give it only in the case $\epsilon_j = \pm 1$, $n_1 = n_2$, but it holds in general.

Theorem 4.1 *Let a_{ij} be a real $n \times n$ matrix, $|a_{ij}| \le 1$. We can then find, in polynomial time in n, signs $\epsilon_1, \ldots, \epsilon_n = \pm 1$ such that for every $i = 1, \ldots, n$ we have*

$$\left| \sum_{j=1}^{n} a_{ij}\epsilon_j \right| \le C(n \log n)^{1/2},$$

where C is an absolute constant.

This of course means that the Salem-Zygmund Theorem 3.7 is equally effective since it is Theorem 3.8 alone that was used in its proof.

Theorem 4.2 *Given trigonometric polynomials $f_1, \ldots, f_n$ of degree at most m and numbers $p_j \in [0,1]$ one can find, in time polynomial in mn, coefficients $\xi_j \in \{1 - p_j, -p_j\}$, $j = 1, \ldots, n$, such that*

$$\left\| \sum_{j=1}^{n} \xi_j f_j \right\|_\infty \le C \left(\sum_{j=1}^{n} \|f_j\|_\infty^2 \log m \right)^{1/2}.$$

As the Salem-Zygmund theorem is the only "random" ingredient used in the proofs of the results in [16, 21, 24] the objects that were proved to exist therein are computable in polynomial time. Namely, in time polynomial in n one can find

1. a trigonometric polynomial $f(x) = \sum_{j=1}^{n} a_j \exp(ijx)$, with $|a_j| = 1$, $j = 1, \ldots, n$, such that
$$\frac{|f(x)|}{\sqrt{n}} \sim 1,$$
as $n \to \infty$, uniformly for all x in $[0, 2\pi]$ (results of [16, 21]).
2. a cosine polynomial $f(x) = M + \sum_{j=1}^{n} a_j \cos(ijx)$ with a_j nonnegative integers, such that $M \le C(s \log s)^{1/3}$, with $s = a_1 + \cdots + a_n$ (result in [24]).

Open Problem: Spencer [3, 27] has proved that, given a complex matrix a_{ij}, with $|a_{ij}| \le 1$, $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$, there are signs $\epsilon_j = \pm 1$, $j = 1, \ldots, n_2$, such that for all i we have $\left| \sum_{j=1}^{n_2} a_{ij}\epsilon_j \right| \le C n_1^{1/2}$. This is, in general, an improvement over the Salem-Zygmund theorem when, say, $n_1 = n_2 = n$, but we do not know how to find the signs ϵ_j in time polynomial in n.

Open Problem: Bourgain [4], [16, p. 78] has proved the existence of integer frequencies $\lambda_1 < \cdots < \lambda_n$ such that $\|\sin \lambda_1 x + \cdots + \sin \lambda_n x\|_\infty \ll n^{2/3}$. The proof involves the Salem-Zygmund theorem on a polynomial of degree super-polynomial in n. Indeed, one can easily see that λ_n must necessarily be super-polynomial in n. Thus the straightforward derandomization of the Salem-Zygmund theorem, as applied in Bourgain's proof, gives rise to an algorithm that does not run in polynomial time. Can one find these frequencies $\lambda_1 < \cdots < \lambda_n$ in time polynomial in n?

4.4 An effective additive basis for the integers

In this section we shall derandomize Erdős's proof of the existence of an additive basis E of the integers of order 2 for which the representation function satisfies

$$C_1 \log x \le r_E(x) \le C_2 \log x \tag{21}$$

for all sufficiently large integers x. Erdős's proof was described in Section 3.2. Our goal will be to give an algorithm for the enumeration of E which enumerates $E \cap [1, n]$ in time polynomial in n. It is not immediately clear that there exists *any* algorithm, whether fast or slow, to enumerate E. Indeed, if one tries to perform any kind of exhaustive search of the probability space, one faces the obstacle that making the decision whether to put a certain integer m in the set E or not affects the values for $r_E(x)$ for all $x > m$, which are of course infinitely many.

This problem can be overcome if one looks at the original, slightly different, proof of Erdős [8], which has been stated using counting arguments and not probability. It uses an existential argument on a finite interval at a time and can thus be readily turned into a construction by examining all possible intersections of E with the interval. But the algorithm which we get this way takes time exponential in n to decide whether n is in E or not.

First we give our modified probabilistic proof. The method of conditional probabilities can then be applied [18, 20].

4.4.1 A modified probabilistic proof

Define the modified representation function $r'(x) = r'_E(x)$ as the number of representations of the nonnegative integer x as a sum $a + b$, with $a, b \in E$, $g(x) \le a \le b$, where $g(x) = (x \log x)^{1/2}$. This is our main difference from Erdős's proof. By doing this modification we have achieved that the presence or absence of a certain number n in our set E affects $r'(x)$ for only a finite number of nonnegative integers x.

Theorem 4.3 *There are positive constants c_1, c_2, c_3, with $c_2 < c_3$, and a set E of positive integers such that*

$$c_2 \log x \le r'(x) \le c_3 \log x$$

and

$$|E \cap [x - g(x), x]| \le c_1 \log x$$

for all large enough $x \in \mathbf{N}$.

Proof: We define the random set E by letting

$$\mathbf{Pr}\left[x \in E\right] = p_x = K \cdot \left(\frac{\log x}{x}\right)^{1/2}$$

independently for all $x \in \mathbf{N}$, where K is a positive constant that will be specified later. Let

$$\mu = \mathbf{E}[r'(x)] = \sum_{t=g(x)}^{x/2} p_t p_{x-t}.$$

Define also

$$s(x) = |E \cap [x - g(x), x]|$$

and

$$\nu = \mathbf{E}[s(x)] = \sum_{t=x-g(x)}^{x} p_t.$$

We can estimate μ and ν for large x to get

$$\mu \sim IK^2 \log x, \quad \nu \sim K \log x,$$

where $I = \int_0^{1/2} (s(1-s))^{-1/2} ds$.

The "bad" events are

$$A_x = \{|r'(x) - \mu| > \epsilon\mu\}, \quad B_x = \{s(x) - \nu > \epsilon\nu\},$$

for a positive constant ϵ, say $\epsilon = 1/2$. Since both $r'(x)$ and $s(x)$ are both SIIRV we can use the Chernoff bound (Theorem 3.1) to get

$$\mathbf{Pr}\left[A_x\right] \le 2e^{-c_\epsilon \mu} \le 2e^{-\frac{1}{2}c_\epsilon I K^2 \log x} = 2x^{-\alpha}$$

and

$$\mathbf{Pr}\left[B_x\right] \le 2e^{-c_\epsilon \nu} \le 2e^{-\frac{1}{2}c_\epsilon K \log x} = 2x^{-\beta}$$

where $\alpha = \frac{1}{2}c_\epsilon I K^2$ and $\beta = \frac{1}{2}c_\epsilon K$. Choose K large enough to make both α and β greater than 1.

Then

$$\sum_{x=1}^{\infty} \mathbf{Pr}\left[A_x\right] + \mathbf{Pr}\left[B_x\right] < \infty$$

which implies the existence of $n_0 \in \mathbf{N}$ such that, with positive probability, none of the events A_x and B_x, $x \ge n_0$, holds. In particular there exists a set E for which

$$\mu/2 \le r'(x) \le 3\mu/2 \text{ and } s(x) \le 3\nu/2,$$

for all $x \ge n_0$. This implies the conclusion of Theorem 4.3 with $c_1 = \frac{1}{2}K$, $c_2 = \frac{1}{2}IK^2$ and $c_3 = \frac{3}{2}IK^2$. □

Observe that $r'(x) \le r(x) \le r'(x) + s(x)$. We deduce that for the set E of Theorem 4.3 we have

$$c_2 \log x \le r(x) \le (c_1 + c_3) \log x$$

so that (21) is true for E.

4.4.2 Derandomization of the modified proof

We showed that the complement of the bad event

$$B = \bigcup_{x \ge n_0} (A_x \cup B_x)$$

has positive probability, by establishing the inequality $\sum_{x \ge n_0} \mathbf{Pr}\left[A_x\right] + \mathbf{Pr}\left[B_x\right] < 1$. This implies the existence of a point E in our probability space $\{0,1\}^{\mathbf{N}}$ which is not in B (there is a natural identification between points in the probability space and subsets of $\mathbf{N}$). We give an algorithm which at the n-th step outputs 0 or 1 to denote the absence or presence of n in our set E.

Denote by $\chi \in \{0,1\}^{\mathbf{N}}$ a generic element in our space and by $R(a_1, \ldots, a_k)$ the event $\chi_1 = a_1, \ldots, \chi_k = a_k$, where $a_1, \ldots, a_k \in \{0,1\}$. It is obvious that for any event $D \subseteq \{0,1\}^{\mathbf{N}}$

$$\begin{aligned} &\mathbf{Pr}\left[D \mid R(a_1, \ldots, a_{n-1})\right] = \\ &\qquad p_n \mathbf{Pr}\left[D \mid R(a_1, \ldots, a_{n-1}, 1)\right] + (1-p_n)\mathbf{Pr}\left[D \mid R(a_1, \ldots, a_{n-1}, 0)\right]. \end{aligned} \tag{22}$$

We are going to define the sequence $a_n \in \{0,1\}$ so that the function

$$b_n = b_n(a_1, \ldots, a_n) = \sum_{x \ge n_0} \mathbf{Pr}\left[A_x \mid R(a_1, \ldots, a_n)\right] + \mathbf{Pr}\left[B_x \mid R(a_1, \ldots, a_n)\right]$$

is non-increasing in n. (Notice that the function $\mathbf{Pr}\left[A_x \mid R(a_1, \ldots, a_n)\right]$ is constant in n when $n > x$, and is equal to either 0 or 1. The same is true for the events B_x.) Since $b_0 = \sum_{x \ge n_0} \mathbf{Pr}\left[A_x\right] + \mathbf{Pr}\left[B_x\right] < 1$, the monotonicity of b_n implies that

$$\sum_{x \ge n_0} \mathbf{Pr}\left[A_x \mid R(a_1, \ldots, a_n, \ldots)\right] + \mathbf{Pr}\left[B_x \mid R(a_1, \ldots, a_n, \ldots)\right] < 1.$$

The probabilities above are either 0 or 1, so they are all 0, and the point $E = (a_1, \ldots, a_n, \ldots)$ is not in B.

So all that remains to be done is to ensure that b_n does not increase. Adding up (22) we get

$$\begin{aligned} & b_{n-1}(a_1, \ldots, a_{n-1}) = \\ & \qquad p_n b_n(a_1, \ldots, a_{n-1}, 1) + (1 - p_n) b_n(a_1, \ldots, a_{n-1}, 0), \end{aligned}$$

which implies that at least one of $b_n(a_1, \ldots, a_{n-1}, 1)$ and $b_n(a_1, \ldots, a_{n-1}, 0)$ is not greater than $b_{n-1}(a_1, \ldots, a_{n-1})$. We let $a_n = 1$ if the first number is smaller than the latter, otherwise we let $a_n = 0$.

Notice that

$$\begin{aligned} \Delta &= b_n(a_1, \ldots, a_{n-1}, 1) - b_n(a_1, \ldots, a_{n-1}, 0) \\ &= \sum_{x=n}^{G(n)} \mathbf{Pr}[A_x \mid R(a_1, \ldots, a_{n-1}, 1)] - \mathbf{Pr}[A_x \mid R(a_1, \ldots, a_{n-1}, 0)] + \\ &\qquad + \mathbf{Pr}[B_x \mid R(a_1, \ldots, a_{n-1}, 1)] - \mathbf{Pr}[B_x \mid R(a_1, \ldots, a_{n-1}, 0)], \end{aligned}$$

where $G(n) = (1 + o(1))n^2/\log n$ is the greatest integer k such that $g(k) \le n$. This is so because the events A_x and B_x, with $x > G(n)$ are independent of $\chi_1, \ldots, \chi_n$ and their probabilities cancel out in the difference above. We have to decide in time polynomial in n whether $\Delta \ge 0$. This is indeed possible since the expression for Δ has $\sim 4n^2/\log n$ terms, each of which can be computed in polynomial time as the following easy Lemma [18, 20] claims.

Lemma 4.1 *Let $X_k = \xi_1 + \cdots + \xi_k$ be a sum of k independent indicator random variables with $\mathbf{Pr}[\xi_j = 1] = p_j$, $j = 1, \ldots, k$. Then the distribution of X_k can be computed in time polynomial in k.*

Thus all probabilities of the form $\mathbf{Pr}[\alpha < X_k < \beta]$ can be efficiently computed. Observe that having fixed $\chi_1 = a_1, \ldots, \chi_n = a_n$ we have

$$r'(x) = \sum_{t=g(x)}^{x/2} \chi_t \chi_{x-t} = \sum_{g(x)}^{n} a_t \chi_{x-t} + \sum_{n+1}^{x/2} \chi_t \chi_{x-t},$$

for $x - g(x) > n$, otherwise $r'(x)$ has already been completely determined by the assigned values of $\chi_1, \ldots, \chi_n$. This means that $r'(x)$ is a SIIRV and so is $s(x)$. Thus the probabilities of A_x and B_x conditioned on $R(a_1, \ldots, a_{n-1}, 1)$ and $R(a_1, \ldots, a_{n-1}, 0)$ can be efficiently computed and $\Delta \ge 0$ can be decided in polynomial time, as we had to show.

One can check that it takes roughly $O(n^7)$ steps to enumerate all of $E \cap [1, n]$.

Open Problem: Is there a basis E that satisfies (21) and an algorithm that can answer the question "Is $n \in E$?" in time polynomial in $\log n$? That would be the next best thing to being able to write down a formula for the elements of E.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, *A dense infinite Sidon sequence*, Europ. J. Comb. 2 (1981), 1-11.

[2] N. Alon and D. J. Kleitman, *Sum-free subsets*, in A. Baker, B. Bollobás, and A. Hajnál, eds., A Tribute to Paul Erdős, Cambridge University Press (1990), 13-26.

[3] N. Alon and J. Spencer, *The probabilistic method*, Wiley Interscience Series in Discrete Mathematics and Optimization, 1992.

[4] J. Bourgain, *Sur les sommes de sinus*, Sém. Anal. Harm., Publ. Math. d'Orsay 84-01 (1984), exp. no 3.

[5] J. Bourgain, *Personal communication*, 1994.

[6] H. Chernoff, *A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Stat. 23 (1952), 493-509.

[7] S. Chowla, *Solution of a problem of Erdős and Turán in additive number theory*, Proc. Nat. Acad. Sci. India 14 (1944) 1-2.

[8] P. Erdős, *On a problem of Sidon in additive number theory*, Acta Sci. Math. (Szeged), 15 (1953-54), 255-259.

[9] P. Erdős, *Problems and results in additive number theory*, Colloque sur la Théorie des Nombres (CBRM, Bruxelles), 127-137.

[10] P. Erdős, *Extremal problems in number theory*, Proceedings of the Symp. Pure Math. VIII, AMS (1965), 181-189.

[11] P. Erdős and A. Rényi, *Additive properties of random sequences of positive integers*, Acta Arith. 6 (1960), 83-110.

[12] P. Erdős and P. Tetali, *Representations of integers as the sum of k terms*, Random Strucrures and Algorithms 1 (1990), 245-261.

[13] P. Erdős and P. Turán, *On a problem of Sidon in additive number theory and some related problems*, J. London Math. Soc. 16 (1941), 212-215; *Addendum* (by P. Erdős), ibid. 19 (1944), 208.

[14] H. Halberstam and K. F. Roth, *Sequences*, Springer-Verlag, New York, 1983.

[15] S. Janson, *Poisson approximation for large deviations*, Random Structures and Algorithms 1 (1990), 221-230.

[16] J.-P. Kahane, *Some random series of functions*, Cambridge Studies in Advanced Mathematics 5, 1985, Second Edition.

[17] Y. Katznelson, *An introduction to Harmonic Analysis*, Wiley, 1968, and Dover, 1976, New York.

[18] M. N. Kolountzakis, *An effective additive basis for the integers*, Discr. Math., to appear. Also in Proc. Symp. On Discrete Algorithms (SODA) 1994.

[19] M. N. Kolountzakis, *Selection of a large sum-free subset in polynomial time*, Inf. Proc. Letters 49 (1994), 255-256.

[20] M. N. Kolountzakis, *Probabilistic and constructive methods in harmonic analysis and additive number theory*, Ph.D. thesis, Department of Mathematics, Stanford University, May 1994.

[21] T. W. Körner, *On a polynomial of J. S. Byrnes*, Bull. London Math. Soc. 12 (1980), 219-224.

[22] F. Krückeberg, *B_2 Folgen und verwandte Zahlenfolgen*, J. Reine Angew. Math. 106 (1961), 53-60.

[23] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley, New York, 1974.

[24] A. M. Odlyzko, *Minima of cosine sums and maxima of polynomials on the unit circle*, J. London Math. Soc. (2) 26 (1982), 412-420.

[25] S. K. Pichorides, *L^p norms of exponential sums*, Sém. Anal. Harm., Publ. Math. d'Orsay 77-73 (1976), 1-65.

[26] R. Salem and A. Zygmund, *Some properties of trigonometric series whose terms have random signs*, Acta Math. 91 (1954), 245-301

[27] J. Spencer, *Six standard deviations suffice*, Trans. Amer. Math. Soc. 289, 2 (1985), 679-706.

[28] A. Stöhr, *Gelöste und ungelöste Fragen über Basen der natürlichen Zahlenreihe II*, J. Reine Angew. Math. 194 (1955), 111-140.

[29] S. Uchiyama, *On the mean modulus of trigonometric polynomials whose coefficients have random signs*, Proc. Amer. Math. Soc. 16 (1965), 1185-1190.