

# THE PROBABILITY THAT A RANDOM TRIPLE OF DICE IS TRANSITIVE

D. H. J. POLYMATH

**ABSTRACT.** An  $n$ -sided die is an  $n$ -tuple of positive integers. We say that a die  $(a_1, \dots, a_n)$  *beats* a die  $(b_1, \dots, b_n)$  if the number of pairs  $(i, j)$  such that  $a_i > b_j$  is greater than the number of pairs  $(i, j)$  such that  $a_i < b_j$ . We show that for a natural model of random  $n$ -sided dice, if  $A, B$  and  $C$  are three random dice then the probability that  $A$  beats  $C$  given that  $A$  beats  $B$  and  $B$  beats  $C$  is approximately  $1/2$ . In other words, the information that  $A$  beats  $B$  and  $B$  beats  $C$  has almost no effect on the probability that  $A$  beats  $C$ . This proves a statement that was conjectured by Conrey, Gabbard, Grant, Liu and Morrison for a different model.

## 1. INTRODUCTION

It is an amusing fact, first observed by Bradley Efron in the 1960s, that there can be four dice  $A_0, A_1, A_2, A_3$ , each with six sides but with non-standard numberings such that if they are all rolled, then each of the four events “ $A_i$  shows a higher number than  $A_{i+1}$ ” occurs with probability  $2/3$  (where  $i + 1$  is interpreted mod 4). Thus, if we say that one die *beats* another if it has a better-than-50% chance of showing a higher number, then the relation “beats” is not transitive.

Much more recently, Conrey, Gabbard, Grant, Liu and Morrison decided to investigate how common a phenomenon intransitivity is. They defined a notion of random  $n$ -sided dice, defined a suitable relation “beats” for such dice, and did some computer experiments that indicated, to their surprise, that if  $A, B$  and  $C$  are three random dice, then the probability that  $A$  beats  $C$  given that  $A$  beats  $B$  and  $B$  beats  $C$  is, for large  $n$ , approximately equal to  $1/2$ . That is, the information that  $A$  beats  $B$  and  $B$  beats  $C$  gives almost no clue about whether  $A$  beats  $C$ .

The definition they gave of a random  $n$ -sided die, which we shall refer to as the *multiset model*, is as follows: they define an  $n$ -sided die to be a multiset with  $n$  elements that add up to  $n(n + 1)/2$  (or equivalently average  $(n + 1)/2$ ), and a random  $n$ -sided die is simply an  $n$ -sided die chosen uniformly at random. An equivalent definition is that a random  $n$ -sided die is a random non-decreasing sequence  $(a_1, \dots, a_n)$  of positive integers between 1 and  $n$  that add up to  $n(n + 1)/2$ . For example, the 4-sided dice are  $(1, 1, 4, 4)$ ,  $(1, 2, 3, 4)$ ,  $(1, 3, 3, 3)$ ,  $(2, 2, 2, 4)$ , and  $(2, 2, 3, 3)$ .

Given two random  $n$ -sided dice  $A = (a_1, \dots, a_n)$  and  $B = (b_1, \dots, b_n)$ , we say that  $A$  *beats*  $B$  if the number of pairs  $(i, j)$  such that  $a_i > b_j$  is greater than the number of pairs  $(i, j)$  such that  $a_i < b_j$ . For example, the die  $A = (1, 1, 4, 4)$  beats the die  $B = (1, 3, 3, 3)$  because there are eight

pairs  $(i, j)$  with  $a_i > b_j$  and only six with  $a_i < b_j$ . If the two numbers are equal we say that  $A$  *ties* with  $B$ .

Conrey, Gabbard, Grant, Liu and Morrison made the following two conjectures.

**Conjecture 1.1.** *Let  $n$  be a positive integer and let  $A$  and  $B$  be independent random  $n$ -sided dice in the multiset model. Then the probability that  $A$  ties with  $B$  is  $o(1)$ .*

**Conjecture 1.2.** *Let  $n$  be a positive integer and let  $A, B$  and  $C$  be independent random  $n$ -sided dice in the multiset model. Then the probability that  $A$  beats  $C$  given that  $A$  beats  $B$  and  $B$  beats  $C$  is  $\frac{1}{2} + o(1)$ .*

They also conjectured a strengthening of Conjecture 1.2, which is the following. Recall that a *tournament* is a complete graph for which every edge is given a direction. We shall regard it as a set  $T$  of ordered pairs of distinct elements of a set  $V$  such that for any two distinct elements  $v, w$  of  $V$  exactly one of  $(v, w)$  or  $(w, v)$  belongs to  $T$ .

**Conjecture 1.3.** *Let  $T$  be a tournament with vertices  $1, 2, \dots, k$ . Then if  $A_1, \dots, A_k$  are independent random  $n$ -sided dice in the multiset model, then the probability that for each  $1 \leq i < j \leq k$  we have that  $A_i$  beats  $A_j$  if and only if  $(i, j) \in T$  is  $2^{-\binom{k}{2}} + o(1)$ .*

The conclusion about the tournament  $T$  is stating that it is *quasirandom* in a sense introduced by Chung and Graham [?]. It turns out to be equivalent to the statement that for all but a fraction  $o(1)$  of the pairs of vertices  $x, y$ , the fraction of vertices  $z$  such that either  $(x, z)$  and  $(y, z)$  belong to  $T$  or  $(z, x)$  and  $(z, y)$  belong to  $T$  is  $\frac{1}{2} + o(1)$ . There are many different equivalent conditions for quasirandomness: the one conjectured to hold by Conrey, Gabbard, Grant, Liu and Morrison states that all small tournaments occur in  $T$  with approximately the frequency one would expect in a random tournament.

Conrey, Gabbard, Grant, Liu and Morrison also looked at other models, and the experimental evidence was surprisingly sensitive to the model chosen, with Conjecture 1.2 (and hence also Conjecture 1.3) appearing to be false for most of them. However, there was one other model for which it seemed to be true, which we shall refer to as the *balanced sequences model*. Here an  $n$ -sided die is simply a sequence  $(a_1, \dots, a_n)$  of elements of  $\{1, 2, \dots, n\}$  that adds up to  $n(n+1)/2$  and a random  $n$ -sided die is an  $n$ -sided die chosen uniformly at random. Note that permuting a sequence does not affect which other sequences it beats. If we say that two dice that are permutations of one another are *equivalent*, then the difference between the balanced sequences model and the multiset model is that the multiset model gives the same weight to each equivalence class, while the balanced sequences model gives the same weight to each individual sequence.

The main results of this paper are that Conjectures 1.1 and 1.2 are true for the balanced sequences model. We also report on experimental evidence that suggests that the stronger conjecture, Conjecture 1.3 is false for both models.

The method of proof can be summarized as follows. We begin by showing that Conjecture 1.2 is equivalent to the statement that almost every random die beats approximately half the other dice and is beaten by approximately half the other dice (and therefore ties with almost no dice). We then argue that unless a die  $A$  has a very “atypical” distribution, then it is indeed the case that it beats approximately half the other dice and is beaten by approximately half the other dice. We can regard this last statement as the claim that if  $(b_1, \dots, b_n)$  is a random sequence of elements of  $\{1, 2, \dots, n\}$ , then the probability that it is beaten by  $A$  given that it sums to  $n(n+1)/2$  is approximately  $1/2$ , as is the probability that it beats  $A$  given the same condition. It turns out that this is true provided that a certain sum of independent random variables with values in  $\mathbb{Z}$  is sufficiently close to a discrete Gaussian distribution. In order to prove this, we need a rather explicit quantitative local central limit theorem, which we prove by standard Fourier-analytic means, using probabilistic arguments to prove that the behaviour we need the Fourier transform (or characteristic function) to satisfy holds for almost all dice  $A$ .

This paper is the result of an open online collaboration between several authors. A complete record of the discussion that led to its existence can be found in a series of five consecutive blog posts and comments on them, of which the first is <https://gowers.wordpress.com/2017/04/28/a-potential-new-polymath-project-intransitive-dice/>. The posts belong to a category entitled polymath13.

## 2. THE PRELIMINARY REDUCTION

We begin with a lemma about tournaments, or rather about near tournaments, by which we mean directed graphs with  $n$  vertices and  $(1 - o(1))\binom{n}{2}$  edges. Given a triple of vertices  $(x, y, z)$ , we shall call it *intransitive* if the subgraph induced by the three vertices is a directed cycle of length 3, and *transitive* if it is a triangle but not a directed 3-cycle. The *out-degree*  $d_+(x)$  of a vertex  $x$  is the number of vertices  $y$  such that  $(x, y)$  is an edge, and the *in-degree*  $d_-(x)$  is the number of vertices  $y$  such that  $(y, x)$  is an edge.

**Lemma 2.1.** *Let  $T$  be a directed graph with  $n$  vertices and  $(1 - o(1))\binom{n}{2}$  edges. Then the following two statements are equivalent.*

- (1) *The probability that a random triple of vertices is intransitive is  $\frac{1}{4} + o(1)$ .*
- (2) *If  $x$  is a random vertex, then with probability  $1 - o(1)$   $d_+(x) = (\frac{1}{2} + o(1))n$ .*

{equivaler

*Proof.* Write  $x \rightarrow y$  if  $(x, y)$  is an edge of  $T$ . First let us count the number of triples  $(x, y, z)$  such that  $x \rightarrow y \rightarrow z$ . A directed triangle  $xyz$  in  $T$  gives rise to three such triples, namely  $(x, y, z)$ ,  $(y, z, x)$  and  $(z, x, y)$ . Any other triangle gives rise to just one: for example, if  $x \rightarrow y$ ,  $x \rightarrow z$  and  $y \rightarrow z$ , then the only triple we obtain is  $(x, y, z)$ . Since the number of triangles is  $(1 - o(1))\binom{n}{3}$ , we find that the number of triples  $(x, y, z)$  such that  $x \rightarrow y \rightarrow z$  is  $(1 - o(1))\binom{n}{3}$  plus twice the number of directed triangles. Note that  $\binom{n}{3} = (1 + o(1))n^3/6$ .

But the number of such triples is also  $\sum_y d_+(y)d_-(y)$ . Since the number of edges is  $(1 - o(1))\binom{n}{2}$ , this is equal to  $(1 + o(1)) \sum_y d_+(y)(n - d_+(y))$ . Also,  $\sum_y d_+(y) = (1 - o(1))\binom{n}{2} = (1 - o(1))n^2/2$ , so this is  $(1 + o(1))(n^3/2 - \sum_y d_+(y)^2)$ . Therefore, twice the number of directed triangles is  $(1 + o(1))(n^3/3 - \sum_y d_+(y)^2)$ .

If a random triple of vertices has a probability  $\frac{1}{4} + o(1)$  of being intransitive, then twice the number of directed triangles is also  $(\frac{1}{2} + o(1))\binom{n}{3} = (\frac{1}{12} + o(1))n^3$ . It follows that  $\sum_y d_+(y)^2 = (\frac{1}{4} + o(1))n^3$ , and therefore that  $\mathbb{E}_y d_+(y)^2 = (\frac{1}{4} + o(1))n^2$ . But  $\mathbb{E}_y d_+(y) = (\frac{1}{2} + o(1))n$ , so  $\text{var}(d_+(y)) = o(n^2)$ , which implies that  $d_+(y) = (\frac{1}{2} + o(1))n$  with probability  $1 - o(1)$ .

The steps in the previous paragraph can also be reversed, so the lemma is proved.  $\square$

### 3. A RANDOM VARIABLE RELATED TO AN $n$ -SIDED DIE AND A SECOND REDUCTION

Write  $[n]$  for the set  $\{1, 2, \dots, n\}$ . Given an  $n$ -sided die  $A = (a_1, \dots, a_n)$  (in fact the definition we are about to give applies to any sequence in  $[n]^n$ ) we define a cumulative distribution function  $f_A$  by

$$f_A(j) = |\{i \in [n] : a_i < j\}| + \frac{1}{2}|\{i \in [n] : a_i = j\}|.$$

We typically expect  $f_A(j)$  to be around  $j - \frac{1}{2}$ , so it is convenient also to define a function  $g_A$  by  $g_A(j) = f_A(j) - j + \frac{1}{2}$ . For a fixed  $A$ , we shall be interested in the random variable  $(g_A(j), j - (n + 1)/2)$ , which is defined on  $[n]$ . More precisely, we choose  $j$  uniformly from  $[n]$  and evaluate the pair  $(g_A(j), j - (n + 1)/2)$ .

To see why this is useful to look at, let us do a few simple calculations.

First of all,

$$\begin{aligned}
 \sum_j f_A(j) &= \sum_j \sum_i (\mathbb{1}_{[a_i < j]} + \frac{1}{2} \mathbb{1}_{[a_i = j]}) \\
 &= \sum_i \sum_j (\mathbb{1}_{[a_i < j]} + \frac{1}{2} \mathbb{1}_{[a_i = j]}) \\
 &= \sum_i (n - a_i + 1/2) \\
 &= n^2/2,
 \end{aligned}$$

where the last equality follows from the fact that  $A$  is an  $n$ -sided die and therefore  $\sum_i a_i = n(n+1)/2$ . This gives us that

$$\sum_j g_A(j) = n^2/2 - \sum_j (j - 1/2) = n^2/2 - n(n+1)/2 + n/2 = 0,$$

and therefore that the mean of the random variable  $(g_A(j), j)$  is  $(0, 0)$ .

Next, let  $B = (b_1, \dots, b_n)$  be another  $n$ -sided die. Then

$$\sum_j f_A(b_j) = \sum_j \sum_i (\mathbb{1}_{[a_i < b_j]} + \frac{1}{2} \mathbb{1}_{[a_i = b_j]}) = |\{(i, j) : a_i < b_j\}| + \frac{1}{2} |\{(i, j) : a_i = b_j\}|.$$

But

$$\sum_j g_A(b_j) = \sum_j (f_A(b_j) - b_j + 1/2) = \sum_j f_A(b_j) - n^2/2,$$

where the last inequality follows from the fact that  $\sum_j b_j = n(n+1)/2$ . It follows that

$$\sum_j g_A(b_j) = |\{(i, j) : a_i < b_j\}| + \frac{1}{2} |\{(i, j) : a_i = b_j\}| - n^2/2.$$

Since there are  $n^2$  pairs  $(i, j)$ , this tells us that  $\sum_j g_A(b_j) > 0$  if and only if

$$|\{(i, j) : a_i < b_j\}| + \frac{1}{2} |\{(i, j) : a_i = b_j\}| > |\{(i, j) : a_i > b_j\}| + \frac{1}{2} |\{(i, j) : a_i = b_j\}|,$$

which is true if and only if  $B$  beats  $A$ . Similarly  $A$  beats  $B$  if and only if  $\sum_j g_A(b_j) < 0$ .

We will therefore be done if we can prove the following claim.

{main}

**Claim 3.1.** *If  $A$  is a random  $n$ -sided die, then with probability  $1 - o(1)$  we have that the proportion of  $n$ -sided dice  $B = (b_1, \dots, b_n)$  with  $\sum_j g_A(b_j) > 0$  is  $\frac{1}{2} + o(1)$ .*

The proof that this claim is sufficient requires one small observation. Given a die  $A = (a_1, \dots, a_n)$ , define the *complementary die*  $\bar{A}$  to be the sequence  $(n+1-a_1, \dots, n+1-a_n)$ . Then  $A$  beats  $B$  if and only if  $\bar{B}$  beats  $\bar{A}$ . So if the claim is true, then with probability  $1 - o(1)$ , the proportion of  $B$  such that  $A$  beats  $B$  is  $\frac{1}{2} + o(1)$  and the proportion of  $B$  such that  $\bar{A}$  beats  $\bar{B}$  is also  $\frac{1}{2} + o(1)$ , which implies that the proportion of  $B$  such that  $A$  ties with  $B$  is  $o(1)$ .

#### 4. A HEURISTIC ARGUMENT FOR CLAIM 3.1

We begin by explaining why one would expect Claim 3.1 to be true. Once we have done that, we shall turn our heuristic argument into a rigorous one. It is at that point that we shall need to prove a local central limit theorem with sufficiently explicit bounds.

Let  $(b_1, \dots, b_n)$  be a purely random sequence belonging to  $[n]^n$  – that is, one where the  $b_i$  are chosen uniformly and independently from  $[n]$  and there is no restriction on the sum. Then to prove Claim 3.1 for a fixed  $A$  we need to show that

$$\mathbb{P}\left[\sum_j g_A(b_j) > 0 \mid \sum_j b_j = n(n+1)/2\right] = \frac{1}{2} + o(1),$$

which is equivalent to the assertion that

$$\mathbb{P}\left[\sum_j g_A(b_j) > 0 \mid \sum_j (b_j - (n+1)/2) = 0\right] = \frac{1}{2} + o(1).$$

But  $b_1, \dots, b_n$  are uniformly and independently chosen from  $[n]$ . Thus, if we write  $(X_j, Y_j)$  for the random variable  $(g_A(b_j), b_j - (n+1)/2)$ , then  $(X_1, Y_1), \dots, (X_n, Y_n)$  are  $n$  independent copies of the random variable  $(g_A(j), j - (n+1)/2)$  mentioned earlier, and we are concerned with the sum  $\sum_{j=1}^n (X_j, Y_j)$ , which we shall write as  $(X, Y)$ .

The central limit theorem suggests that the distribution of this sum will be approximately Gaussian, and since each  $(X_i, Y_i)$  has mean  $(0, 0)$  we would in particular expect that the distribution would be approximately symmetric about the origin. Also, we would expect a typical value of  $g_A(j)$  to have magnitude around  $\sqrt{n}$ , so the standard deviation of  $X$  ought to be around  $n$ . Also  $Y$  has standard deviation of order  $n^{3/2}$  and the two random variables, though correlated, will probably not be too heavily correlated.

If all these heuristics are correct, then the probability that  $X = 0$  given that  $Y = 0$  should be of order  $n^{-1}$ , and certainly  $o(1)$ . The symmetry should imply that  $\mathbb{P}[X > 0 | Y = 0] \approx \mathbb{P}[X < 0 | Y = 0]$ , and these statements taken together would give us that  $\mathbb{P}[X > 0 | Y = 0] = \frac{1}{2} + o(1)$  and  $\mathbb{P}[X < 0 | Y = 0] = \frac{1}{2} + o(1)$ , which is equivalent, as we have seen, to the statement that the proportion of dice that beat  $A$  is  $\frac{1}{2} + o(1)$  and the proportion of dice that  $A$  beats is  $\frac{1}{2} + o(1)$ .

The reason this heuristic argument cannot immediately be turned into a proof is that the central limit theorem is too blunt a tool. There are two reasons for this. The first is that although it tells us that a sum of i.i.d. random variables will converge to a Gaussian, it does not tell us how fast that convergence will occur, and we need it to have occurred (to within a small error) when we take a sum of  $n$  copies of  $(g_A(j), j)$ . And we cannot just let  $n$  tend to infinity because the random variables themselves depend on  $n$ . This second problem applies not just to the central limit theorem but also to the Berry-Esseen theorem, which gives a rate of convergence in the central limit theorem, but with a constant that (necessarily) depends on the random variable.

A second problem is that the notion of convergence in the central limit theorem and the Berry-Esseen theorem is not suitable for our purposes. We need to be able to estimate the probability that  $(X, Y)$  belongs to the positive x-axis, which is a “probability zero event” from the point of view of the central limit theorem and Berry-Esseen theorem. Instead, we need a *local central limit theorem*, the name given to versions of the central limit theorem that can give us estimates for the density function at individual values. Unfortunately, the local central limit theorems that appear in the literature tend still to involve inexplicit constants that depend on the random variable, again necessarily. (We did find an exception to this, but it proved a one-dimensional theorem where we need a two-dimensional one [?].)

In the end, we have proved for ourselves a local central limit theorem that is tailored to our application. It is not hard to prove using Fourier analysis, which is one of the standard methods for proving such results, but it requires the random variable to have certain properties, as we shall explain later, in order for us to be able to make the implied constant explicit. So the rest of the proof splits into two parts: first we shall prove that the random variable  $(U, V) = (g_A(j), j - (n + 1)/2)$  has certain properties with high probability (when  $A$  is a random  $n$ -sided die). Then we shall use those properties to establish a suitable local central limit theorem, after which the argument will essentially be finished.

## 5. PROPERTIES OF THE RANDOM VARIABLE $(U, V)$

We begin with an almost standard fact (Lemma 5.2 below), but for convenience we provide a complete proof. (The fact and its proof could be thought of as a weakening of a very special case of a one-dimensional local central limit theorem.) First we prove an even more basic lemma.

**Lemma 5.1.** *Let  $I_n$  be the set  $\{-(n-1)/2, -(n-3)/2, \dots, (n-3)/2, (n-1)/2\}$  and let  $f$  be defined on  $\frac{1}{2}\mathbb{Z}$  by taking  $f(x) = n^{-1}$  if  $x \in I_n$  and  $f(x) = 0$  otherwise. (Thus,  $f(x) = \mathbb{P}[V = x]$ .) Then the  $k$ -fold convolution  $f^{*k}$  of  $f$  is supported on  $\mathbb{Z}$  except if  $k$  is odd and  $n$  is even, in which case it is*

supported on  $\mathbb{Z} + \frac{1}{2}$ . In all cases,  $f^{*k}$  is an even function, and its non-zero values increase when  $x < 0$  and decrease when  $x > 0$ .

*Proof.* The statements about the support and the symmetry are trivial. To prove the increasing and decreasing properties, we note that they follow easily by induction. Indeed, let  $g$  be any even function supported on an arithmetic progression of common difference 1 that increases towards the middle, and let  $x \geq 0$ . Then the inner product of  $g$  with  $I_n + x$  is greater than or equal to the inner product of  $g$  with  $I_n + x + 1$ , since  $g(x - (n - 1)/2) \geq g(x + 1 + (n - 1)/2)$ . Therefore,  $g * f$  decreases when  $x$  is positive, and by symmetry it increases when  $x$  is negative (when we restrict to appropriate supports).  $\square$

What we care about here is that when  $k = n$ , the maximum of  $f^{*n}$  is attained at zero. And all we really need from the next lemma is that the probability that  $Y = 0$  is not tiny.

**Lemma 5.2.** *Let  $(a_1, \dots, a_n)$  be an element of  $[n]^n$  chosen uniformly at random. Then the probability that  $\sum_i a_i = n(n + 1)/2$  is at least  $n^{-3/2}/4$ .*

*Proof.* The probability that  $\sum_i a_i = n(n + 1)/2$  is  $f^{*n}(0)$ . Equivalently, it is the probability that  $Y = 0$ , where  $Y$  is the sum of  $n$  independent copies of  $V$ . The variance of  $V$  is at most  $n^2/4$ , so the variance of  $Y$  is at most  $n^3/4$ . Therefore, by Chebyshev's inequality, the probability that  $|Y| \geq n^{3/2}$  is at most  $1/4$ , which implies that the probability that  $|Y| \leq n^{3/2}$  is at least  $3/4$ . Since 0 is the most likely value of  $Y$ , it follows that  $Y = 0$  with probability at least  $3/8(n^{3/2} + 1) \geq n^{-3/2}/4$ .  $\square$

We shall now obtain an upper bound for  $\|U\|_\infty$ . Our method is to obtain an upper bound that holds with such high probability for a purely random element of  $[n]^n$  that it continues to hold with high probability even when we condition on the sum being  $n(n + 1)/2$ .

**Lemma 5.3.** *Let  $A$  be a random  $n$ -sided die. Then with probability  $1 - o(1)$  we have that  $\max_j |g_A(j)| \leq 6\sqrt{n \log n}$ .*

*Proof.* Let  $(a_1, \dots, a_n)$  be a purely random sequence – that is, an element of  $[n]^n$  chosen uniformly at random. For each  $j$ , let  $n_A(j)$  be the number of  $i$  such that  $a_i \leq j$ . Note that  $f_A(j)$  is the average of  $n_A(j - 1)$  and  $n_A(j)$ .

Now  $n_A(j)$  is a sum of  $n$  independent Bernoulli random variables of mean  $j/n$ . By Chernoff's bounds, the probability that  $|n_A(j) - j| \geq m$  is at most  $2 \exp(-m^2/6n)$ . Therefore, the probability that there exists  $j$  such that  $|n_A(j) - j| \geq m$  is at most  $2n \exp(-m^2/6n)$ . Setting  $m = 6\sqrt{n \log n}$ , this is at most  $2 \exp(-6 \log n) = 2n^{-6}$ . By Lemma 5.2, if we now condition on the event that  $\sum_i a_i = n(n + 1)/2$ , then this probability rises to at most  $8n^{-9/2}$ .



If no such  $j$  exists, then for every  $j$  we have that

$$|(n_A(j-1) + n_A(j))/2 - (j-1 + j)/2| \leq 6\sqrt{n \log n},$$

by the triangle inequality. The left-hand side of this inequality is  $|g_A(j)|$ . □

Our next aim is to prove that with high probability  $\sum_j g_A(j)^2$  is not too small.

**\*\*\*\*We know that a typical order of magnitude of  $g_A(j)$  is  $\sqrt{n}$ , so we would expect  $\mathbb{E}_j g_A(j)^2$  to be at least  $n/100$  with high probability. However, because the different values of  $g_A(j)$  are far from independent, it seems to be tricky to prove this. I don't yet have a nice argument. I think I can probably hack something out that would be just about sufficient, but I'd much rather do it properly.\*\*\*\***

**\*\*\*\*There will also be a lemma that says that  $\hat{f}(\alpha, \beta)$  is bounded away from 1 in magnitude by enough for  $\hat{f}(\alpha, \beta)^n$  to be very small, except if  $\alpha$  and  $\beta$  are so small that the Gaussian approximation to  $\hat{f}(\alpha, \beta)^n$  is accurate.\*\*\*\***

## 6. AN EXPLICIT TWO-DIMENSIONAL LOCAL CENTRAL LIMIT THEOREM

In this section we shall use a small amount of Fourier analysis. Recall that if  $f : \mathbb{Z}^2 \rightarrow \mathbb{C}$  then  $\hat{f} : \mathbb{T}^2 \rightarrow \mathbb{C}$  is defined by the formula

$$\hat{f}(\alpha, \beta) = \sum_{x,y} f(x, y) e(\alpha x + \beta y),$$

where  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$  and  $e(\theta)$  is shorthand for  $\exp(2\pi i \theta)$ . When  $f$  is a random variable  $(U, V)$  on  $\mathbb{Z}^2$  (in other words,  $f(x, y)$  is the probability that  $(U, V) = (x, y)$ ), then  $\hat{f}$  is often called the *characteristic function* of  $(U, V)$ . The characteristic function of  $(U, V)$  relates in a simple way to its moments.

We have that

$$\frac{\partial^{r+s}}{\partial^r \alpha \partial^s \beta} \hat{f}(\alpha, \beta) = (2\pi i)^{r+s} \sum_{x,y} x^r y^s f(x, y) e(\alpha x + \beta y),$$

and evaluating this at zero we get  $\sum_{x,y} x^r y^s f(x, y) = (2\pi i)^{r+s} \mathbb{E}(U^r V^s)$ .

Writing  $\partial_1$  and  $\partial_2$  for the operators of partially differentiating with respect to the first and second variables, respectively, we shall use the following estimate, which follows from Taylor's theorem and the observation about the partial derivatives.

**Lemma 6.1.** *Let  $f$  be as above. Then*

$$\hat{f}(\alpha, \beta) = 1 - 2\pi^2(\alpha^2 \mathbb{E}U^2 + 2\alpha\beta \mathbb{E}UV + \beta^2 \mathbb{E}V^2) + R(\alpha, \beta),$$

{taylor}

where  $|R(\alpha, \beta)| \leq \frac{4\pi^3}{3}(|\alpha|||U||_\infty + |\beta|||V||_\infty)^3$ .

{1c1t}

**Theorem 6.2.** *Let  $(U, V)$  be a random variable with mean  $(0, 0)$  taking values in  $\mathbb{Z}^2$ , let  $(X_1, Y_1), \dots, (X_n, Y_n)$  be independent copies of  $(U, V)$  and let  $(X, Y) = \sum_i (X_i, Y_i)$ . Suppose that  $\|U\|_\infty^2 \leq \rho_1 \mathbb{E}U^2$ , that  $\|V\|_\infty^2 \leq \rho_2 \mathbb{E}V^2$ , and that  $|\mathbb{E}UV| \leq \gamma(\mathbb{E}U^2)^{1/2}(\mathbb{E}V^2)^{1/2}$ . \*\*\*\***Precise statement yet to be formulated.**\*\*\*\**

*Proof.* For each  $(x, y) \in \mathbb{Z}^2$ , let  $f(m, n) = \mathbb{P}[(U, V) = (x, y)]$ . If we add  $n$  independent copies of  $(U, V)$ , then the resulting random variable is the  $n$ -fold convolution of  $f$  with itself. The convolution law from Fourier analysis thus tells us that the characteristic function of  $(X, Y)$  is  $\hat{f}^n$ . The inversion formula then gives us a formula for  $\mathbb{P}[(X, Y) = (x, y)]$ , namely

$$\mathbb{P}[(X, Y) = (x, y)] = \int_{\mathbb{T}^2} \hat{f}(\alpha, \beta)^n e(-\alpha x - \beta y) d\alpha d\beta.$$

As is usual for such theorems, the rough idea of the proof is to look at the Taylor expansion of  $\hat{f}$  and deduce that near zero  $\hat{f}$  can be approximated by a function  $1 - q(\alpha, \beta)$  for some positive definite quadratic form  $q$ . That tells us that  $\hat{f}(\alpha, \beta)^n$  is close to  $\exp(-nq(\alpha, \beta))$  near zero, which is the formula for a Gaussian. Provided  $\hat{f}(\alpha, \beta)$  is not too close to 1 when  $(\alpha, \beta)$  is not close to zero, that implies that  $\mathbb{P}[(X, Y) = (x, y)]$  is approximately given by the same formula as that of the Fourier transform of a Gaussian on  $\mathbb{R}^2$ .

Our hypothesis tells us that the remainder term  $R(\alpha, \beta)$  in Lemma 6.1 is bounded above by  $\frac{4\pi^3}{3}(|\alpha|(\rho_1 \mathbb{E}U^2)^{1/2} + |\beta|(\rho_2 \mathbb{E}V^2)^{1/2})^3$ .

\*\*\*\***Proof not yet finished.**\*\*\*\*

□

## 7. COMPUTATIONAL EVIDENCE THAT CONJECTURE 1.3 IS FALSE

\*\*\*\***Section needs to be planned and written.**\*\*\*\*