GAPS OF SUMMANDS OF THE ZECKENDORF LATTICE

NEELIMA BORADE, DEXTER CAI, DAVID Z. CHANG, BRUCE FANG, ALEX LIANG, STEVEN J. MILLER, AND WANQIAO XU

ABSTRACT. A beautiful theorem of Zeckendorf states that every positive integer has a unique decomposition as a sum of non-adjacent Fibonacci numbers. Such decompositions exist more generally, and much is known about them. First, for any positive linear recurrence $\{G_n\}$ the number of summands in the legal decompositions for integers in $[G_n, G_{n+1})$ converges to a Gaussian distribution. Second, Bower, Insoft, Li, Miller, and Tosteson proved that the probability of a gap between summands in a decomposition which is larger than the recurrence length converges to geometric decay. While most of the literature involves one-dimensional sequences, some recent work by Chen, Guo, Jiang, Miller, Siktar, and Yu have extended these decompositions to d-dimensional lattices, where a legal decomposition is a chain of points such that one moves in all d dimensions to get from one point to the next. They proved that some but not all properties from 1-dimensional sequences still hold. We continue this work and look at the distribution of gaps between terms of legal decompositions, and prove similar to the 1-dimensional cases that when $d = 2$ the gap vectors converge to a bivariate geometric random variable.

1. INTRODUCTION

1.1. **Previous Work.** Zeckendorf's Theorem [Ze] provides an alternative definition of the Fibonacci numbers $\{F_n\}$ (normally defined by $F_1 = 1, F_2 = 2$ and $F_{n+1} = F_n + F_{n-1}$ for all $n \geq 2$): they are the only sequence such that every positive integer can be written uniquely as sum of non-adjacent terms. Such a sum is called the Zeckendorf (or legal) decomposition. Similar results hold for other sequences; see for example [BDEMMTTW, Br, Day, DDKMMV, FGNPT, Fr, GTNP, Ha, Ho, HW, Ke, Lek, LM1, LM2, MW1, MW2, Ste1, Ste2] for a representative sample of results on unique decompositions, as well as on the distribution of the number of summands in these decompositions. Most of the work to date has been on one-dimensional sequences; many of the sequences which at first appear two-dimensional, such as those in [CFHMN2, CFHMNPX], are in fact one-dimensional when viewed properly. In [CCGJMSY], the authors considered generalizations to d-dimensional lattices, where a legal decomposition involved a finite subset of lattice points where each point has all of its coordinates strictly smaller than the previous (thus all motion is both down and to the left). The motivation for their work was to see which properties persist. They were able to show that the distribution of the number of summands is similar, also a Gaussian, but unique decomposition is lost.

In this work we continue these investigations and look at the distribution of gaps between summands in decompositions. For many sequences the resulting distributions converge to geometric decay, with the constant arising from the largest root of the characteristic polynomial of the recurrence relation; see [BBGILMT, BILMT]. The question is more interesting here, as the gaps are now d-dimensional vectors, and there is the possibility of new behavior.

We first recall the construction from [CCGJMSY]; the following description is slightly modified from that work with permission of the authors. A legal decomposition in d dimensions is a finite collection of lattice points for which

(1) each point is used at most once, and
(2) if the point $(i_1, i_2, \ldots, i_d)$ is included then all subsequent points $(i'_1, i'_2, \ldots, i'_d)$ have $i'_j < i_j$ for all $j \in \{1, 2, \ldots, d\}$ (i.e., *all* coordinates must decrease between any two points in the decomposition).

We call these sequences of points on the d-dimensional lattice simple jump paths. One can weaken the second condition and merely require $i'_j \leq i_j$. This restriction was imposed in [CCGJMSY] to simplify the combinatorial analysis, as it led to simple closed form expressions. With additional work we can consider the more general case where it is no longer required that all coordinates decrease; see [FJLLLMSS], where the authors prove similar behavior as in [CCGJMSY].

We now construct our sequence using the above definition of legal decomposition. We concentrate on $d \in \{1, 2\}$ in the main results in this paper; similar results should hold in general, but for small d there are combinatorial identities which simplify the sums that arise and lead to nice closed form expression. Whenever possible we state definitions and ancillary lemmas for the most general case possible.

When $d = 1$ we write $\{y_a\}_{a=0}^{\infty}$ for our sequence, which is defined by

(1) $y_1 := 1$, and
(2) if we have constructed the first k terms of our sequence, the $(k+1)^{\text{st}}$ term is the smallest integer which cannot be written as a sum of terms in the sequence, with each term used at most once.

This case is, not surprisingly, very similar to previous one-dimensional results. A straightforward calculation shows that $a_n = 2^{n-1}$, and the legal decomposition of a number is just its binary expansion.

We now turn to the main object of study in this paper, $d = 2$, and denote the general term of our sequence by $y_{i,j}$. Instead of defining the sequence by traveling along diagonals we could do right angular paths; for the purposes of this paper it does not matter, as we are only concerned with the gaps between chosen summands, and not the values of the summands (and it is the values which are influenced by the mode of construction).

(1) Set $y_{1,1} := 1$.
(2) Iterate through the natural numbers. For each such number, check if any path of numbers in our sequence with a strict leftward and downward movement between each two points sums to the number. If no such path exists, add the number to the sequence so that it is added to the shortest unfilled diagonal moving from the bottom right to the top left.
(3) If a new diagonal must begin to accommodate a new number, set the value $y_{k,1}$ to be that number, where k is minimized so that $y_{k,1}$ has not yet been assigned.

In (1.1) we give the first few diagonals of the 2-dimensional lattice. Note that we no longer have uniqueness of decomposition (for example, 25 has two legal decompositions: $20 + 5$ and

$24 + 1)$.

$$
\begin{matrix}
280 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
157 & 263 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
84 & 155 & 259 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
50 & 82 & 139 & 230 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
28 & 48 & 74 & 123 & 198 & \cdots & \cdots & \cdots & \cdots & \cdots \\
14 & 24 & 40 & 66 & 107 & 184 & \cdots & \cdots & \cdots & \cdots \\
7 & 12 & 20 & 33 & 59 & 100 & 171 & \cdots & \cdots & \cdots \\
3 & 5 & 9 & 17 & 30 & 56 & 93 & 160 & \cdots & \cdots \\
1 & 2 & 4 & 8 & 16 & 29 & 54 & 90 & 159 & \cdots
\end{matrix}
\tag{1.1}
$$

The main result of [CCGJMSY] is that the distribution of the number of summands among all simple jump paths starting at (n,n) and ending at $(0,0)$ converges to a Gaussian as $n \to \infty$ (as all paths must have both a down and a left component, we can add an additional row and an additional column where one of the indices is zero, and require all paths to end at $(0,0)$); a similar result holds for compound paths where each step is either down, left or down and left [FJLLLMSS]. We investigate the distribution of gaps between adjacent summands in legal decompositions. Before stating our results, we first introduce some notation.

1.2. **New Results.** There are several ways to define gaps in these d-dimensional lattice decompositions, leading to slightly different behavior. We give three possibilities here.

Definition 1.1. *For a step from $(x_{m,1}, \ldots, x_{m,d})$ to $(x_{m+1,1}, \ldots, x_{m+1,d})$, its **gap vector** is the difference $(x_{m,1} - x_{m+1,1}, \ldots, x_{m,d} - x_{m+1,d})$. A simple jump path of length k starting at $(a_1, a_2, \ldots, a_d)$ corresponds to the set $\{(x_{i,1} - x_{i+1,1}, \ldots, x_{i,d} - x_{i+1,d})\}_{i=0}^{k-1}$ of k gap vectors, where*

- $(x_{0,1}, \ldots, x_{0,d}) = (a_1, \ldots, a_d)$,
- $(x_{k,1}, \ldots, x_{k,d}) = (0, \ldots, 0)$, *and*
- *for each $i \in \{0, 1, \ldots, k-1\}$ and $j \in \{1, \ldots, d\}$, $x_{i,j} > x_{i+1,j}$.*

Definition 1.2. *Given a gap vector $(x_{i,1} - x_{i+1,1}, \ldots, x_{i,d} - x_{i+1,d})$, its **gap sum** is the sum of the components of the vector: $(x_{i,1} - x_{i+1,1}) + \cdots + (x_{i,d} - x_{i+1,d})$. Similarly the **gap distance** is the length of the gap vector.*

There are three natural quantities we can investigate. We can look at the gap vectors, the gap sums, or the gap distances. The distribution of the gap vectors is the most fundamental quantity, and much of the combinatorics is a natural generalization of previous work for the one-dimensional case [BBGILMT, BILMT]. Knowing the distribution of the gap vectors, we can calculate the distribution of the gap sums by summing over all gap vectors with the same gap sum. The last notion, the gap distance, is harder as this requires summing over a subset of gap vectors to obtain a given gap distance. Note we can interpret the difference between these two perspectives as arising from the norm we use to measure the length of the gap vector; the gap sum comes from using the L^1 norm while the gap distance is from the L^2 norm.

Our main result is that as n goes to infinity, the distribution of the gap vectors in the 2-dimensional lattice converges to a geometric decay, and thus we see similar behavior as in the one-dimensional cases.

Theorem 1.3. *Let n be a positive integer, consider the distribution of gap vectors among all simple jump paths of dimension 2 with starting point $(n+1, n+1)$. For fixed positive integers v_1, v_2, the probability that a gap vector equals (v_1, v_2) converges to $1/2^{v_1 + v_2}$ as $n \to \infty$.*

We prove Theorem 1.3 in §3 through combinatorial identities and Stirling's formula, but for larger d the combinatorial lemmas do not generalize. As an immediate consequence, we obtain the distribution of the gap sums.

Theorem 1.4. *Let n be a positive integer, consider the distribution of gap sums among all simple jump paths of dimension 2 with starting point $(n+1, n+1)$. The probability that a gap sum equals an integer $v \geq 2$ converges to $(v-1)/2^v$ as $n \to \infty$ (the probability of a gap sum of 0 or 1 is zero).*

After reviewing properties of simple jump paths in §2, we prove our main results in §3, and then conclude the paper with questions on alternative definitions of gaps, generalizations to compound paths, and distribution of the longest gap (where we present some partial results). Whenever possible we state and prove results for arbitrary dimensions, to facilitate future research.

2. Properties of Simple Jump Paths

We first recall some notation for our simple jump paths from [CCGJMSY]. As our paper is an extension of [CCGJMSY], the following four paragraphs are reproduced with permission from them.

We have walks in d dimensions starting at some initial point $(a_1, a_2, \ldots, a_d)$ with each $a_j > 0$, and ending at the origin $(0, 0, \ldots, 0)$. Note that our simple jump paths must always have movement in all dimensions at each step. We are just adding one extra point, at the origin, and saying every path must end there. Note that as we always change all of the indices during a step, we never include a point where only some of the coordinates are zero, and thus there is no issue in adding one extra point and requiring all paths to end at the origin.

Our walks are sequences of points on the lattice grid with positive indices or the origin, and we refer to movements between two such consecutive points as **steps**. Thus a simple jump path is a walk where each step has a strict movement in all d-dimensions. More formally, a simple jump path of length k starting at $(a_1, a_2, \ldots, a_d)$ is a sequence of points $\{(x_{i,1}, \ldots, x_{i,d})\}_{i=0}^{k}$ where the following hold:

- $(x_{0,1}, \ldots, x_{0,d}) = (a_1, \ldots, a_d)$,
- $(x_{k,1}, \ldots, x_{k,d}) = (0, \ldots, 0)$, and
- for each $i \in \{0, 1, \ldots, k-1\}$ and $j \in \{1, \ldots, d\}$, $x_{i,j} > x_{i+1,j}$.

For a fixed d and any choice of starting point $(a_1, a_2, \ldots, a_d) \in \mathbb{R}^d$, we let $s_d(a_1, \ldots, a_d)$ denote the number of simple jump paths starting at $(a_1, a_2, \ldots, a_d)$ and ending at $(0, \ldots, 0)$, and $t_d(k; a_1, \ldots, a_d)$ the subset of these paths with length k. In particular, when $a_1 = \cdots = a_d = n$ for a fixed $n \in \mathbb{N}^+$, we let $s_d(n)$ denote the number of simple jump paths from $(n, n, \ldots, n)$ to the origin, and $t_d(k, n)$ the subset of these paths with exactly k steps. As we must reach the origin, every path has at least 1 step, the maximum number of steps is n, and

$$s_d(n) \;=\; \sum_{k=1}^{n} t_d(k, n). \tag{2.1}$$

We now determine $t_d(k, n)$. In one dimension we have $t_d(k, n) = \binom{n-1}{k-1}$, as we must choose exactly $k-1$ of the first $n-1$ terms (we must choose the nth term as well as the origin, and thus choosing $k-1$ additional places ensures there are exactly k steps). Since in higher dimension there is movement in each dimension for each step, for $1 \leq k \leq \min(a_1, \ldots, a_d)$,

$$t_d(k; a_1, \ldots, a_d) \;=\; \binom{a_1 - 1}{k-1}\binom{a_2 - 1}{k-1}\cdots\binom{a_d - 1}{k-1}, \tag{2.2}$$

and

$$s_d(a_1, \ldots, a_d) \;=\; \sum_{k=1}^{\min(a_1,\ldots,a_d)} t_d(k; a_1, \ldots, a_d). \tag{2.3}$$

From the binomial theorem we have $s_1(a_1) = 2^{a_1-1}$. For higher dimensions we need another well-known combinatorial result: Vandermonde's identity. We reproduce the proof here for completeness.

Lemma 2.1 (Vandermonde's Identity). *For $\alpha, \beta \in \mathbb{N}$,*

$$\sum_{k=0}^{\gamma} \binom{\alpha}{\gamma - k} \binom{\beta}{k} \;=\; \binom{\alpha + \beta}{\gamma}. \tag{2.4}$$

Proof. Suppose there are $\alpha + \beta$ balls, with α of them red and β of them blue. We want to select γ balls in total; there are $\binom{\alpha+\beta}{\gamma}$ ways to do this. We can also calculate this by splitting into cases of the number of blue balls chosen. There are $\binom{\alpha}{\gamma-k}\binom{\beta}{k}$ ways to choose k blue balls and $\gamma - k$ red balls. Summing over all cases, it's clear that

$$\sum_{k=0}^{\gamma} \binom{\alpha}{\gamma - k} \binom{\beta}{k} \;=\; \binom{\alpha + \beta}{\gamma}, \tag{2.5}$$

which completes the proof. $\square$

We can now determine the number of simple paths in 2 dimensions. The result below is an extension of results from [CCGJMSY], where only the special case $a_1 = a_2$ has been proved.

Theorem 2.2. *In 2-dimensional lattice,*

$$s_2(a_1, a_2) \;=\; \binom{a_1 + a_2 - 2}{a_1 - 1}. \tag{2.6}$$

Proof. From (2.2) and (2.3), we have

$$s_2(a_1, a_2) \;=\; \sum_{k=1}^{\min(a_1,a_2)} \binom{a_1 - 1}{k - 1} \binom{a_2 - 1}{k - 1}. \tag{2.7}$$

Without loss of generality, assume $\min(a_1, a_2) = a_1$. Then

$$\begin{aligned} s_2(a_1, a_2) \;&=\; \sum_{k=1}^{a_1} \binom{a_1 - 1}{k - 1} \binom{a_2 - 1}{k - 1} \\ &=\; \sum_{k=0}^{a_1 - 1} \binom{a_1 - 1}{a_1 - 1 - k} \binom{a_2 - 1}{k}. \end{aligned} \tag{2.8}$$

Applying Lemma 2.1,

$$s_2(a_1, a_2) \;=\; \binom{a_1 + a_2 - 2}{a_1 - 1} \;=\; \binom{a_1 + a_2 - 2}{a_2 - 1}. \tag{2.9}$$

$\square$

Note that when $a_1 = a_2 = n$, we have

$$s_2(n) = \binom{2n - 2}{n - 1}. \tag{2.10}$$

3. GAPS IN 2-DIMENSIONAL LATTICE

In 1-dimension case, the notion of gaps between adjacent points in a simple jump path is unambiguous; however, for $d \geq 2$, we have several choices. Below we concentrate on the gap vector.

We let $g_d(n)$ denote the number of gap vectors of all simple jump paths from $(n, n, \ldots, n)$ to the origin. We add the origin to the path to facilitate the counting; as all steps must be both down and to the left, the origin is the only point where one of the indices is zero. While this does introduce one extra gap, as $n \to \infty$ the contribution from it is negligible and can thus be safely ignored. This addition now means that each simple jump path of length k contains k gap vectors, and every legal path has at least one and at most n gap vectors. Thus

$$g_d(n) = \sum_{k=1}^{n} k \ t_d(k, n). \tag{3.1}$$

To prove Theorems 1.3 and 1.4, we begin with determining $g(n; (v_1, v_2))$, defined as the number of gap vectors (v_1, v_2) in all simple jump paths starting from (n, n) and ending at $(0, 0)$. Then we find $g_2(n)$, the total number of gap vectors. Due to the presence of $n - 1$ in the formula for $s_2(n)$, we work with $n + 1$ below to simplify some of the algebra.

Lemma 3.1. *Consider all the simple jump paths from $(n+1, n+1)$ to $(0, 0)$ in the 2-dimensional lattice. Let $G((x, y), (x + v_1, y + v_2))$ denote the number of gap vectors (v_1, v_2) starting at $(x + v_1, y + v_2)$ and ending at (x, y) within all simple jump paths from $(n + 1, n + 1)$ to $(0, 0)$, then*

$$G((x, y), (x + v_1, y + v_2)) = \binom{x + y - 2}{x - 1} \binom{2n - v_1 - v_2 - x - y}{n - v_1 - x}. \tag{3.2}$$

Proof. Since each different arrangement of simple jump paths from $(n + 1, n + 1)$ to $(x + v_1, y + v_2)$ and from (x, y) to $(0, 0)$ contributes 1 to the number of gap vectors (v_1, v_2), $G((x, y), (x + v_1, y + v_2))$ is given by the number of simple jump paths from (x, y) to $(0, 0)$ times the number of simple paths from $(n + 1, n + 1)$ to $(x + v_1, y + v_2)$; see Figure 1.

Shifting the coordinates $(x + v_1, y + v_2)$ and $(n + 1, n + 1)$ down to $(0, 0)$ and $(n - x - v_1 + 1, n - y - v_2 + 1)$ respectively, we obtain

$$G((x, y), (x + v_1, y + v_2)) = s_2(x, y) \cdot s_2(n - x - v_1 + 1, n - y - v_2 + 1). \tag{3.3}$$

Applying Theorem 2.2,

$$G((x, y), (x + v_1, y + v_2)) = \binom{x + y - 2}{x - 1} \binom{2n - v_1 - v_2 - x - y}{n - v_1 - x}. \tag{3.4}$$

$\square$

Now we determine the range of x, y in Lemma 3.1. If a gap vector (v_1, v_2) starts at $(x + v_1, y + v_2)$ and ends at (x, y), then it is clear that $x, y \geq 0$. Since we are only considering simple jump paths from $(n + 1, n + 1)$ to $(0, 0)$, the components of $(x + v_1, y + v_2)$ cannot exceed $n + 1$. Thus $x \leq n - v_1 + 1$ and $y \leq n - v_2 + 1$. Combining, we have $0 \leq x \leq n - v_1 + 1$ and $0 \leq y \leq n - v_2 + 1$.

Lemma 3.2. *Recall that $g(n + 1; (v_1, v_2))$ denotes the number of gap vectors (v_1, v_2) in all simple jump paths starting from $(n + 1, n + 1)$ and ending at $(0, 0)$. Then*

$$g(n + 1; (v_1, v_2)) = (2n - v_1 - v_2 - 1)\binom{2n - v_1 - v_2 - 2}{n - v_1 - 1} + 2\binom{2n - v_1 - v_2}{n - v_1}. \tag{3.5}$$
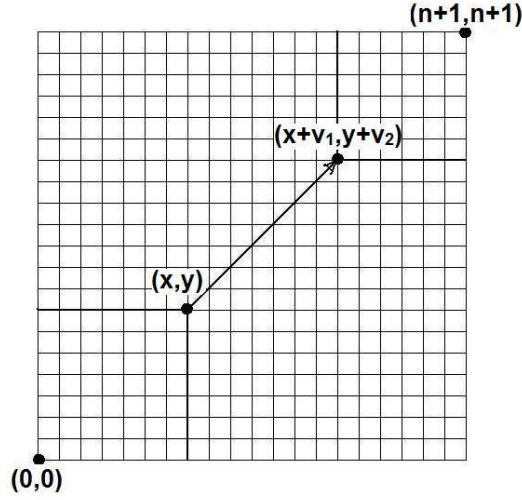
FIGURE 1. Set-up to compute the number of simple jump paths from $(0,0)$ to $(n+1, n+1)$ with a gap of (v_1, v_2) starting at (x, y). Counting the number of such paths is the same as counting the number of paths in the bottom left rectangle and multiplying by the number of paths in the top right.

Proof. For ease of calculation, we break into three cases:

(1) $1 \leq x \leq n - v_1$ and $1 \leq y \leq n - v_2$,
(2) $x = 0$ and $y = 0$,
(3) $x = n - v_1 + 1$ and $y = n - v_2 + 1$.

We first consider Case (1). By Lemma 3.1, the number of gap vectors (v_1, v_2) is given by

$$\sum_{x=1}^{n-v_1} \sum_{y=1}^{n-v_2} G\left((x,y),(x+v_1, y+v_2)\right) \;=\; \sum_{x=1}^{n-v_1} \sum_{y=1}^{n-v_2} \binom{x+y-2}{x-1}\binom{2n-v_1-v_2-x-y}{n-v_1-x}. \quad (3.6)$$

Shifting the index of x, y in the sum, the right hand side of (3.6) becomes

$$\sum_{x=0}^{n-v_1-1} \sum_{y=0}^{n-v_2-1} \binom{x+y}{x}\binom{2n-v_1-v_2-x-y-2}{n-v_1-x-1}. \quad (3.7)$$

Letting $p = n - v_1 - 1$ and $q = n - v_2 - 1$, it is equivalent to calculate

$$\sum_{x=0}^{p}\sum_{y=0}^{q} \binom{x+y}{x}\binom{p+q-(x+y)}{p-x}. \quad (3.8)$$

In the sum, since $0 \leq x \leq p$ and $0 \leq y \leq q$,

$$0 \;\leq\; x+y \;\leq\; p+q, \quad (3.9)$$

so there are $p + q + 1$ different values of $x + y$. Now we bring the problem down to proving that, for each fixed value of $x + y$,

$$\sum_{x=0}^{p} \binom{x+y}{x}\binom{p+q-(x+y)}{p-x} \;=\; \binom{p+q}{p}, \quad (3.10)$$

which is an immediate application of lemma 2.1. Note that not all terms in the left hand side of (3.10) are necessarily non-zero, as x cannot exceed $x + y$.

Thus the number of gap vectors (v_1, v_2) in Case (1) is

$$(p + q + 1)\binom{p+q}{p} = (2n - v_1 - v_2 - 1)\binom{2n - v_1 - v_2 - 2}{n - v_1 - 1}. \tag{3.11}$$

We then consider Case (2). When $x = 0$, $y = 0$, the number of gap vectors (v_1, v_2) equals the number of simple jump paths from $(n + 1, n + 1)$ to (v_1, v_2). Shifting the coordinates $(n + 1, n + 1)$ and (v_1, v_2) down to $(n + 1 - v_1, n + 1 - v_2)$ and $(0, 0)$, respectively, the number of gap vectors in this case (v_1, v_2) is just $s_2(n + 1 - v_1, n + 1 - v_2)$. Applying Theorem 2.2,

$$s_2(n + 1 - v_1, n + 1 - v_2) = \binom{2n - v_1 - v_2}{n - v_1}. \tag{3.12}$$

Similarly, in Case (3), the number of gap vectors (v_1, v_2) is

$$s_2(n + 1 - v_1, n + 1 - v_2) = \binom{2n - v_1 - v_2}{n - v_1}. \tag{3.13}$$

Summing up all three cases,

$$g\left(n + 1; (v_1, v_2)\right) = (2n - v_1 - v_2 - 1)\binom{2n - v_1 - v_2 - 2}{n - v_1 - 1} + 2\binom{2n - v_1 - v_2}{n - v_1}. \tag{3.14}$$

$\square$

Lemma 3.3.

$$g_2\left(n + 1\right) = \left(\frac{n}{2} + 1\right)\binom{2n}{n}. \tag{3.15}$$

We prove the lemma using two different methods, in anticipation that perhaps one might be more useful to future researchers trying to generalize to higher dimensions.

First Proof. Let $\mu_2\left(n + 1\right)$ be the mean for the number of steps of all simple jump paths from $(n + 1, n + 1)$ to $(0, 0)$, and $s_2(n + 1)$ be the total number of simple paths from $(n + 1, n + 1)$ to $(0, 0)$. By Lemma 3.1 in [CCGJMSY],

$$\mu_2\left(n + 1\right) = \frac{n}{2} + 1, \tag{3.16}$$

and

$$s_2(n + 1) = \binom{2n}{n}; \tag{3.17}$$

it is here that we are using $d = 2$, as it is only when $d \leq 2$ that we have simple formulas for $s_d(n + 1)$, though with a more involved analysis similar results shoudl be obtainable for all d.

Since in a simple jump path the number of steps is the same as the number of gap vectors,

$$g_2\left(n + 1\right) = \left(\frac{n}{2} + 1\right)\binom{2n}{n}. \tag{3.18}$$

$\square$

Second Proof. Since a simple jump path from $(n + 1, n + 1)$ to $(0, 0)$ with i points chosen in between (not counting $(n + 1, n + 1)$ and $(0, 0)$) has $(i + 1)$ gap vectors,

$$
\begin{aligned}
g_2(n + 1) &= \sum_{i=0}^{n} (i + 1) \binom{n}{i} \binom{n}{i} \\
&= \sum_{i=0}^{n} i \binom{n}{i} + \sum_{i=0}^{n} \binom{n}{i} \binom{n}{i}.
\end{aligned}
\tag{3.19}
$$

We use the following two standard binomial identities, see for example [CCGJMSY] (again we do not have these for general d):

$$
\begin{aligned}
\sum_{i=0}^{n} \binom{n}{i} \binom{n}{i} &= \binom{2n}{n} \\
\sum_{i=0}^{n} i \binom{n}{i} \binom{n}{i} &= \frac{n}{2} \binom{2n}{n}.
\end{aligned}
\tag{3.20}
$$

Hence

$$
\begin{aligned}
g_2(n + 1) &= \frac{n}{2} \binom{2n}{n} + \binom{2n}{n} \\
&= \left(\frac{n}{2} + 1 \right) \binom{2n}{n}.
\end{aligned}
\tag{3.21}
$$

$\square$

Now we have all the tools to prove Theorem 1.3.

Proof of Theorem 1.3. Let $P(v_1, v_2)$ denote the probability that a given gap vector is (v_1, v_2). By definition,

$$
P(v_1, v_2) = \frac{G(n + 1; (v_1, v_2))}{g_2(n + 1)}.
\tag{3.22}
$$

Using Lemmas 3.2 and 3.3 to simplify,

$$
\begin{aligned}
P(v_1, v_2) &= \frac{(2n - v_1 - v_2 - 1)\binom{2n - v_1 - v_2 - 2}{n - v_1 - 1} + 2\binom{2n - v_1 - v_2}{n - v_1}}{\left(\frac{n}{2} + 1 \right) \binom{2n}{n}} \\
&= \frac{\frac{(2n - v_1 - v_2 - 1)!}{(n - v_1 - 1)!(n - v_2 - 1)!} + 2 \frac{(2n - v_1 - v_2)}{(n - v_1)(n - v_2)} \frac{(2n - v_1 - v_2 - 1)!}{(n - v_1 - 1)!(n - v_2 - 1)!}}{\left(\frac{n}{2} + 1 \right) \frac{(2n)!}{n!n!}} \\
&= \frac{\left(1 + 2 \frac{(2n - v_1 - v_2)}{(n - v_1)(n - v_2)} \right) \frac{(2n - v_1 - v_2 - 1)!}{(n - v_1 - 1)!(n - v_2 - 1)!}}{\left(\frac{n}{2} + 1 \right) \frac{(2n)!}{n!n!}}.
\end{aligned}
\tag{3.23}
$$

As v_1 and v_2 are fixed,

$$
\lim_{n \to \infty} \frac{(2n - v_1 - v_2)}{(n - v_1)(n - v_2)} = 0,
\tag{3.24}
$$

and thus

$$
\lim_{n \to \infty} P(v_1, v_2) = \lim_{n \to \infty} \frac{\frac{(2n - v_1 - v_2 - 1)!}{(n - v_1 - 1)!(n - v_2 - 1)!}}{\left(\frac{n}{2} + 1 \right) \frac{(2n)!}{n!n!}}.
\tag{3.25}
$$

For u large, Stirling's approximation states that $m! \approx \sqrt{2\pi m}(\frac{m}{e})^m$. We approximate the factorials, and can safely drop the lower order error terms as we take the limit as $n \to \infty$ and obtain

$$\lim_{n\to\infty} P(v_1, v_2) = \lim_{n\to\infty} \frac{\left(\frac{2n-v_1-v_2-1}{e}\right)^{2n-v_1-v_2-1}\left(\frac{n}{e}\right)^{2n}}{\left(\frac{n-v_1-1}{e}\right)^{n-v_1-1}\left(\frac{n-v_2-1}{e}\right)^{n-v_2-1}\left(\frac{n}{2}+1\right)\left(\frac{2n}{e}\right)^{2n}}$$
$$\times \frac{(\sqrt{2\pi})^3\sqrt{2n-v_1-v_2-1}(\sqrt{n})^2}{(\sqrt{2\pi})^3\sqrt{n-v_1-1}\sqrt{n-v_2-1}\sqrt{2n}}. \qquad (3.26)$$

Since

$$\lim_{n\to\infty} \frac{\sqrt{2n-v_1-v_2-1}(\sqrt{n})^2}{\sqrt{n-v_1-1}\sqrt{n-v_2-1}\sqrt{2n}} = 1, \qquad (3.27)$$

The right hand side of (3.26) becomes

$$\lim_{n\to\infty} P(v_1, v_2) = \lim_{n\to\infty} \frac{\left(\frac{2n-v_1-v_2-1}{e}\right)^{2n-v_1-v_2-1}\left(\frac{n}{e}\right)^{2n}}{\left(\frac{n-v_1-1}{e}\right)^{n-v_1-1}\left(\frac{n-v_2-1}{e}\right)^{n-v_2-1}\left(\frac{n}{2}+1\right)\left(\frac{2n}{e}\right)^{2n}}$$
$$= \lim_{n\to\infty} \frac{e^{-1}(2n-v_1-v_2-1)^{2n-v_1-v_2-1}}{\left(\frac{n}{2}+1\right)2^{2n}(n-v_1-1)^{n-v_1-1}(n-v_2-1)^{n-v_2-1}}$$
$$= \lim_{n\to\infty} \frac{2n-v_1-v_2-1}{\frac{n}{2}+1}\frac{e^{-1}}{2^{v_1+v_2+2}}\left(\frac{2n-v_1-v_2-1}{2n-2v_1-2}\right)^{n-v_1-1}$$
$$\times \left(\frac{2n-v_1-v_2-1}{2n-2v_2-2}\right)^{n-v_2-1}$$
$$= \lim_{n\to\infty} \frac{2n-v_1-v_2-1}{\frac{n}{2}+1}\frac{e^{-1}}{2^{v_1+v_2+2}}\left(1+\frac{\frac{v_1-v_2+1}{2}}{n-v_1-1}\right)^{n-v_1-1}$$
$$\times \left(1+\frac{\frac{v_2-v_1+1}{2}}{n-v_2-1}\right)^{n-v_2-1}. \qquad (3.28)$$

As

$$\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = e^a, \qquad (3.29)$$

we find

$$\lim_{n\to\infty} P(v_1, v_2) = \lim_{n\to\infty} \frac{2n-v_1-v_2-1}{\frac{n}{2}+1}\frac{e^{-1}}{2^{v_1+v_2+2}}e^{\frac{v_1-v_2+1}{2}}e^{\frac{v_2-v_1+1}{2}}$$
$$= \lim_{n\to\infty} \frac{2n-v_1-v_2-1}{\frac{n}{2}+1}\frac{1}{2^{v_1+v_2+2}}. \qquad (3.30)$$

Since

$$\lim_{n\to\infty} \frac{2n-v_1-v_2-1}{\frac{n}{2}+1} = 4, \qquad (3.31)$$

we obtain

$$\lim_{n\to\infty} P(v_1, v_2) = 4 \times \frac{1}{2^{v_1+v_2+2}} = \frac{1}{2^{v_1+v_2}}, \qquad (3.32)$$

which is clearly a bivariate geometric random variable. $\qquad\square$

We have proved that the distribution of gap vectors converge to a geometric decay as the lattice size goes to infinity. We may now turn to an alternate definition of gaps, namely, the gap sum. Recall Definition 1.2, which states that a gap sum is the sum of components of the corresponding gap vector. The proof for Theorem 1.4 follows immediately from Theorem 1.3.

Proof of Theorem 1.4. For a fixed n, let $P(v)$ denote the probability that a given gap sum equals $v \geq 2$; note $P(0) = P(1) = 0$ as we must have both horizontal and vertical movement in a step, therefore the smallest possible gap sum is 2. By Theorem 1.3, for each value v of gap sum, all pairs (v_1, v_2) with $v_1 + v_2 = v$ contribute equally. As $1 \leq v_1, v_2 \leq v - 1$, there are $v - 1$ such pairs (once v_1 is chosen then v_2 is determined), each pair occurring with probability $1/2^v$. Thus

$$\lim_{n \to \infty} P(v) \ = \ (v - 1)\left(\frac{1}{2}\right)^v, \tag{3.33}$$

completing the proof. $\qquad\qquad\square$

We remark on the difficulty in generalizing the above argument to arbitrary d. The problem is Lemma 2.1; we are not aware of an analogue when $d \geq 3$.

4. Future Work and Concluding Remarks

We end with some problems and comments for future research.

(1) Is there a way to generalize our analysis to the d-dimensional lattice?
(2) Do nice limits for the distribution of gap distances exist as they do for gap vectors and gap sums?
(3) Can we obtain similar results in d-dimensional compound path [FJLLLMSS] with the three definitions of gaps we set forth in this paper?
(4) Can we obtain similar results on the distribution of longest gap in both d-dimensional simple paths and compound paths?

As [CCGJMSY] was able to obtain Gaussian behavior for the distribution of summands for all d, there is reason to be optimistic that a more involved analysis is possible and we could obtain similar extensions for gaps. In that work, however, the simple closed form expressions that exist in two dimensions do not generalize, and combinatorial proofs and analysis are replaced by more involved techniques. We have thus chosen here to concentrate on $d \leq 2$, as this is already enough to see new behavior (i.e., previous problems never looked at the distribution of vector valued quantities for gaps).

For the distribution of gap distances, we need to know not just what numbers are the sums of two squares, but how they are, as the probability of a gap vector (v_1, v_2) is $1/2^{v_1 + v_2}$. Thus for small g we can easily compute all pairs (v_1, v_2) with $v_1^2 + v_2^2 = g$ (it is easier to study the square of the gap distance, as that is always an integer). For large g we would need advanced results from number theory on decompositions, but these values will have negligibly small probabilities of occurring.

As [FJLLLMSS] extended the results from [CCGJMSY] to compound paths through additional book-keeping (especially more involved inclusion-exclusion arguments), with more work it is likely that the gap results can be generalized to the compound setting as well.

Finally, we end with some results on the distribution of the longest gap. The one dimensional case is already known (see for example [GSW, Sch]); it is essentially equivalent to the distribution of the longest run of heads when tossing a fair coin. If we toss n fair coins, the

expected value of the longest run of heads is

$$\log_2 n + \frac{\gamma}{\log 2} - \frac{3}{2} + r_1(n) + \epsilon_1(n), \tag{4.1}$$

where γ is Euler's constant, $|r_1(n)| \leq 0.000016$, and $\epsilon_1(n)$ tends to zero as n tends to infinity. Moreover, the distribution is strongly concentrated about the mean; the variance is

$$\frac{\pi^2}{6 \log^2 2} + \frac{1}{12} + r_2(n) + \epsilon_2(n), \tag{4.2}$$

where $r_2(n) < 0.00006$ and $\epsilon_2(n)$ goes to zero as n tends to infinity. Note the mean is approximately $\log_2 n$ and the variance is *bounded* independent of n.

Interestingly, it is easier to get results in the compound setting, as the freedom to just move in one direction allows us to view the two components of the vectors as independent. In other words, if we wish to look at the length of the longest horizontal or vertical gap, it is essentially the same as in the case of tossing fair coins; the only possible difference is we must end with a 'head', but at worst that increases the length by 1, which is negligible relative to $\log_2 n$. We can thus immediately get decent bounds on the approximate size of the longest gap in the compound case; the horizontal and vertical results give a lower bound, and adding the two (or adding the two and taking a square-root) provides an upper bound. More work is needed to get results for simple paths, because there now *is* a dependence between the two motions (we must have the same number of 'heads' for each), but with some work it is likely that one could obtain results that show $\log_2 n$ is the right order of magnitude for the longest gap (or at least that there is negligible probability of a longest gap of size n^δ for any fixed $\delta > 0$, and that the longest gap is of size $\log \log n$).

References

[BBGILMT]   O. Beckwith, A. Bower, L. Gaudet, R. Insoft, S. Li, S. J. Miller and P. Tosteson, *The Average Gap Distribution for Generalized Zeckendorf Decompositions*, Fibonacci Quarterly **51** (2013), 13–27.

[BM]   I. Ben-Ari, S. Miller, *A Probabilistic Approach to Generalized Zeckendorf Decompositions*, SIAM Journal on Discrete Mathematics, **30** (2016), no. 2, 1302–1332.

[BDEMMTTW]   A. Best, P. Dynes, X. Edelsbrunner, B. McDonald, S. Miller, K. Tor, C. Turnage-Butterbaugh, M. Weinstein, *Gaussian Behavior of the Number of Summands in Zeckendorf Decompositions in Small Intervals*, Fibonacci Quarterly, **52** (2014), no. 5, 47–53.

[BILMT]   A. Bower, R. Insoft, S. Li, S. Miller, P. Tosteson, *The Distribution of Gaps Between Summands in Generalized Zeckendorf Decompositions*, Journal of Combinatorial Theory, **135** (2015), 130–160.

[Br]   J. L. Brown, Jr., *Zeckendorf's Theorem and Some Applications*, The Fibonacci Quarterly, Vol. 2, No. 3 (Oct. 1964), pages 163–168.

[CCGJMSY]   E. Chen, R. Chen, L. Guo, C. Jiang, S. J. Miller, J. M. Siktar and Peter Yu, *Gaussian Behavior in Zeckendorf Decompositions From Lattices*, to appear in the Fibonacci Quarterly.

[CFHMN1]   M. Catral, P. Ford, P. Harris, S. Miller, D. Nelson, *Generalizing Zeckendorf's Theorem: The Kentucky Sequence*, Fibonacci Quarterly, **52** (2014), no. 5, 68–90.

[CFHMN2]   M. Catral, P. Ford, P. E. Harris, S. J. Miller, and D. Nelson, *Legal Decompositions Arising from Non-positive Linear Recurrences*, Fibonacci Quarterly **54** (2016), no. 4, 3448–365.

[CFHMNPX]   M. Catral, P. Ford, P. E. Harris, S. J. Miller, D. Nelson, Z. Pan and H. Xu, *New Behavior in Legal Decompositions Arising from Non-positive Linear Recurrences*, Fibonacci Quarterly **55** (2017), no. 3, 252–275 (expanded arXiv version: http://arxiv.org/pdf/1606.09309).

[Day]   D. E. Daykin, *Representation of Natural Numbers as Sums of Generalized Fibonacci Numbers*, J. London Mathematical Society **35** (1960), 143–160.

[DDKMMV]   P. Demontigny, T. Do, A. Kulkarni, S. Miller, D. Moon, U. Varma, *Generalizing Zeckendorf's Theorem to f-Decompositions*, Journal of Number Theory, **141** (2014), 136–158.

[DFFHMPP]   R. Dorward, P. Ford, E. Fourakis, P. Harris, S. Miller, E. Palsson, H. Paugh, *New Behavior in Legal Decompositions Arising From Non-Positive Linear Recurrences*, Fibonacci Quarterly, **55** (2017), no. 3, 252–275.

[DG]   M. Drmota and J. Gajdosik, *The distribution of the sum-of-digits function*, J. Théor. Nombrés Bordeaux **10** (1998), no. 1, 17–32.

[Eg]   S. Eger, *Stirling's Approximation for Central Extended Binomial Coefficients*, American Mathematical Monthly, **121** (2014), no. 4, 344–349.

[FJLLLMSS]   E. Fang, J. Jenkins, Z. Lee, D. Li, E. Lu, S. J. Miller, D. Salgado and Joshua Siktar *Central Limit Theorems for Compound Paths on the 2-Dimensional Lattice*, preprint 2019. https://arxiv.org/abs/1906.10645.

[FGNPT]   P. Filipponi, P. J. Grabner, I. Nemes, A. Pethö, and R. F. Tichy, *Corrigendum to: "Generalized Zeckendorf expansions"*, Appl. Math. Lett., **7** (1994), no. 6, 25–26.

[Fr]   A. S. Fraenkel, *Systems of Numeration*, Amer. Math. Monthly **92** (1985), no. 2, 105–114.

[GSW]   L. Gordon, M. F. Schilling, and M. S. Waterman, *An extreme value theory for long head runs*, Probability Theory and Related Fields **72** (1986) 279–287.

[GTNP]   P. J. Grabner, R. F. Tichy, I. Nemes, and A. Pethö, *Generalized Zeckendorf expansions*, Appl. Math. Lett. **7** (1994), no. 2, 25–28.

[Ha]   N. Hamlin, *Representing Positive Integers as a Sum of Linear Recurrence Sequences*, Fibonacci Quarterly **50** (2012), no. 2, 99–105.

[Ho]   V. E. Hoggatt, *Generalized Zeckendorf theorem*, Fibonacci Quarterly **10** (1972), no. 1 (special issue on representations), pages 89–93.

[HW]   N. Hamlin and W. A. Webb, *Representing positive integers as a sum of linear recurrence sequences*, Fibonacci Quarterly **50** (2012), no. 2, 99–105.

[Ke]   T. J. Keller, *Generalizations of Zeckendorf's theorem*, Fibonacci Quarterly **10** (1972), no. 1 (special issue on representations), pages 95–102.

[KKMY]   M. Kologlu, G. Kopp, S. Miller, Y. Wang, *On the Number of Summands in Zeckendorf Decompositons*, Journal of Number Theory, **49** (2011), no. 2, 116–130.

[LT]   M. Lamberger and J. M. Thuswaldner, *Distribution properties of digital expansions arising from linear recurrences*, Math. Slovaca **53** (2003), no. 1, 1–20.

[Lek]   C. G. Lekkerkerker, *Voorstelling van natuurlyke getallen door een som van getallen van Fibonacci]*, Simon Stevin 29 (1951-1952), 190–195.

[Len]   T. Lengyel, *A Counting Based Proof of the Generalized Zeckendorf's Theorem*, Fibonacci Quarterly **44** (2006), no. 4, 324–325.

[LM1]   R. Li and S. J. Miller, *A Collection of Central Limit Type results in Generalized Zeckendorf Decompositions*, the 17th International Fibonacci Conference, Fibonacci Quarterly **55** (2017), no. 5, 105–114.

[LM2]   R. Li and S. J. Miller, *Central Limit Theorems for Gaps of Generalized Zeckendorf Decompositions*, to appear in the Fibonacci Quarterly. http://arxiv.org/abs/1606.08110v1.pdf

[Mil]   S. J. Miller, *The Probability Lifesaver*, Princeton University Press, 2017, 752 pages.

[MW1]   S. Miller, Y. Wang, *From Fibonacci Numbers to Central Limit Type Theorems*, Journal of Combinatorial Theory, Series A **119** (2012), no. 7, 1398–1413.

[MW2]   S. Miller, Y. Wang, *Gaussian Behavior in Generalized Zeckendorf Decompositions*, Combinatorial and Additive Number Theory, CANT 2011 and 2012 (Melvyn B. Nathanson, editor), Springer Proceedings in Mathematics & Statistics (2014), 159–173.

[Sch]   M. F. Schilling, *The Longest Run of Heads*, The College Mathematics Journal **21** (1990), no. 3, 196–207.

[Ste1]   W. Steiner, *Parry expansions of polynomial sequences*, Integers, **2** (2002), Paper A14.

[Ste2]   W. Steiner, *The Joint Distribution of Greedy and Lazy Fibonacci Expansions*, Fibonacci Quarterly, **43** (2005), 60–69.

[Ze]   E. Zeckendorf, *Représentation des nombres naturels par une somme des nombres de Fibonacci ou de nombres de Lucas*, Bulletin de la Société Royale des Sciences de Liége **41** (1972), pages 179–182.

MSC2010: 11B02 (primary), 05A02 (secondary)

Department of Mathematics, University of Illinois at Chicago, Chicago, IL 60607
Email address: `nborad2@uic.edu`

Wuhan Britain China School, Wansongyuan Road 48, Jianghan District, Wuhan, China
Email address: `3182935374@qq.com`

High Technology High School, Lincroft, NJ 07738
Email address: `davidchang7636@gmail.com, dachang@ctemc.org`

UWC Changshu China, No. 88 Kun-Cheng-Hu-Xi Road, Changshu, Jiangsu, China 215500
Email address: `bjfang18@uwcchina.org`

Philips Exeter Academy, Exeter, NH 03833
Email address: `yliang@exeter.edu`

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267
Email address: `sjm1@williams.edu`

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109
Email address: `wanqiaox@umich.edu`