A GENERALIZATION OF ZECKENDORF'S THEOREM VIA CIRCUMSCRIBED m-GONS

ROBERT DORWARD, PARI L. FORD, EVA FOURAKIS, PAMELA E. HARRIS, STEVEN J. MILLER, EYVI PALSSON, AND HANNAH PAUGH

Abstract. Zeckendorf's theorem states that every positive integer can be uniquely decomposed as a sum of nonconsecutive Fibonacci numbers, where the Fibonacci numbers satisfy $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$, $F_1 = 1$ and $F_2 = 2$. The distribution of the number of summands in such decomposition converges to a Gaussian, the gaps between summands converges to geometric decay, and the distribution of the longest gap is similar to that of the longest run of heads in a biased coin; these results also hold more generally, though for technical reasons previous work needed to assume the coefficients in the recurrence relation are non-negative and the first term is positive.

We extend these results by creating an infinite family of integer sequences called the m-gonal sequences arising from a geometric construction using circumscribed m-gons. They satisfy a recurrence where the first $m+1$ leading terms vanish, and thus cannot be handled by existing techniques. We provide a notion of a legal decomposition, and prove that the decompositions exist and are unique. We then examine the distribution of the number of summands used in the decompositions and prove that it displays Gaussian behavior. There is geometric decay in the distribution of gaps, both for gaps taken from all integers in an interval and almost surely in distribution for the individual gap measures associated to each integer in the interval. We end by proving that the distribution of the longest gap between summands is strongly concentrated about its mean, behaving similarly as in the longest run of heads in tosses of a coin.

1. Introduction

The Fibonacci numbers are a heavily studied sequence which arise in many different ways and places. By defining them by $F_1 = 1$, $F_2 = 2$ and $F_{n+1} = F_n + F_{n-1}$, we have the remarkable property that every positive integer can be uniquely written as a sum of non-consecutive Fibonacci numbers; further, this property is equivalent to the Fibonaccis (i.e., if $\{a_n\}$ is a sequence of numbers such that every integer can be written uniquely as a sum of non-adjacent terms in the sequence, then $a_n = F_n$). Zeckendorf proved this in 1939, though he did not publish this result until 1972 [Ze].

Date: December 6, 2015.

2010 *Mathematics Subject Classification.* 11B39, 11B05 (primary) 65Q30, 60B10 (secondary).

Key words and phrases. Zeckendorf decompositions, longest gap.

The authors thank CPT Joseph Pedersen for his programming assistance in the creation of Figure 3. The authors would also like to thank the AIM REUF program, the SMALL REU and Williams College for all of their support and continued funding, which made this work and collaboration possible. Moreover, we thank West Point, in particular the Center for Leadership and Diversity in STEM, for travel funding in support of this work and its' dissemination. The first and third named authors were supported by NSF Grant DMS1347804, and the fifth named author by NSF Grant DMS1265673. This research was performed while the fourth named author held a National Research Council Research Associateship Award at USMA/ARL.

In recent years many have studied generalizations to Zeckendorf's theorem by exploring different notions of decompositions and the properties of the associated sequences, see among others [Al, Day, DDKMMV, DDKMV, DG, FGNPT, GT, GTNP, Ke, Len, MW1, MW2, Ste1, Ste2]. Despite the vast literature in this area, the majority of the research on generalized Zeckendorf decompositions have involved sequences with *positive linear recurrences*. Positive linear recurrence sequences $\{G_n\}$ satisfy a linear recurrence relation where the coefficients are non-negative with the first and last term coefficients being positive[1].

There has been little research which considers cases where the leading coefficient in the recurrence is zero; one such case is found in [CFHMN1]. They studied what they call the Kentucky Sequence, which is defined by the recurrence relation $H_{n+1} = H_{n-1} + 2H_{n-3}$, $H_i = i$ for $i \leq 4$. While the behavior there is similar to the positive linear recurrences, there are sequences with very different behavior. One such is the Fibonacci Quilt, which arises from creating a decomposition rule from the Fibonacci spiral[2] (see [CFHMN2, CFHMN3]), where the number of decompositions is not unique but in fact grows exponentially. This leads to the major motivation of this paper (as well as the motivation for the three papers just mentioned): how important is the assumption that the leading term be positive? The work mentioned above shows that it is not just a technically convenient assumption; markedly different behavior can emerge. Our goal is to try and determine when we have each type of behavior, and thus the purpose of this paper is to explore infinitely many recurrences with leading term absent and see the effect that has on the properties of the decompositions.

Specifically, we consider an infinite family of integer sequences called the *m-gonal sequences*, where $m \geq 3$. These sequences arise from a geometric construction using circumscribed m-gons, and after defining them below we state our results.

1.1. **Definition of m-gonal Sequence.** One interpretation of the Zeckendorf's theorem, which state that every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers, is that we have infinitely many bins with just one number per bin, and if we choose a bin to contribute a summand to a number's decomposition then we cannot choose a summand from an adjacent bin. We can generalize to bins with more elements, as well as disallowing two bins to be used if they are within a given distance (see [CFHMN1, CFHMN2, CFHMN3]). The m-gonal sequences are similar to these constructions, but have a two-dimensional structure arising from circumscribing m-gons about one central m-gon.

Briefly we view the decomposition rule corresponding to the m-gonal sequence, for $m \geq 1$, by saying the sequence is partitioned into bins b_i of length $|b_i|$, where $|b_0| = 1$ and $|b_i| = m$ for all $i \geq 1$. A valid decomposition has no two summands being elements from the same bin. We refer to this decomposition as a *legal m-gonal decomposition of a positive integer z*. We now give details and examples of this construction.

The m-gonal sequence was initially constructed by circumscribing m-gons. For $m \geq 3$ we let M_0 denote a regular m-gon. Circumscribe the m-gon M_1 onto M_0 such that the vertices of M_0 bisect the edges of M_1. Note that this adds m faces to the resulting figure. We continue this process indefinitely, where we circumscribe the m-gon M_i onto M_{i-1}, such that the vertices of M_{i-1} bisect the edges of M_i. At each step we have added an additional m

[1] Thus $G_{n+1} = c_1 G_n + \cdots + c_L G_{n-(L-1)}$ with $c_1 c_L > 0$ and $c_i \geq 0$.

[2] Let $f_n = F_{n-1}$, the standard definition of the Fibonacci numbers. Then the plane can be tiled in a spiral where the dimensions of the n^{th} square is $f_n \times f_n$; we declare a decomposition legal if no two summands used share an edge.

faces to the resulting figure. We depict these initial iterations in Figure 1. Let $\mathcal{M}$ denote all of the faces created through the process of circumscribing m-gons. Then

$$\mathcal{M} \;=\; \{f_0\} \cup \left(\bigcup_{i=1}^{\infty}\{f(i,1), f(i,2), \ldots, f(i,m)\}\right), \tag{1.1}$$

where f_0 is the face of the m-gon M_0 and for $i \geq 1$, $f(i,1), f(i,2), \ldots, f(i,m)$ are the faces added to $\mathcal{M}$ when M_i was circumscribed onto M_{i-1}.

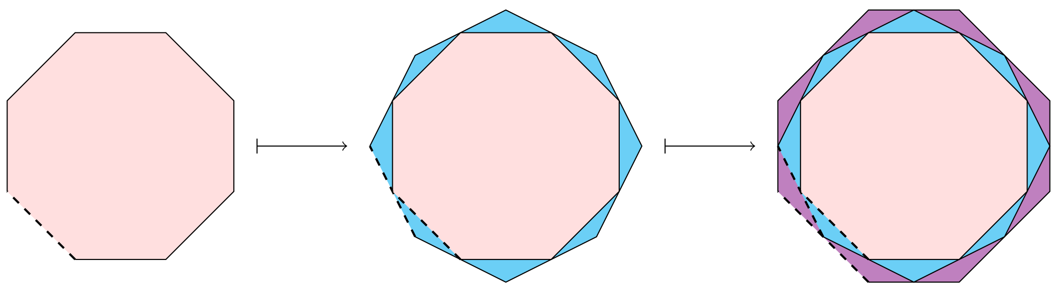


FIGURE 1. Circumscribed m-gons

Fix an integer $m \geq 3$. Suppose $\{a_n\}_{n=0}^{\infty}$ is an increasing sequence of positive integers. We define the following ordered lists, which we refer to as bins, $b_0 = [a_0]$ and for $i \geq 1$, $b_i = [a_{m(i-1)+1},\, a_{m(i-1)+2},\, \ldots,\, a_{mi}]$. Note that for all $i \geq 1$, b_i has size m and b_0 has size one. The integers in bin b_i will correspond directly with the integers which we place on the faces added to $\mathcal{M}$ when M_i was circumscribed onto M_{i-1}. With the elements of our sequence partitioned into bins, we define a *legal m-gonal decomposition* of any positive integer z. If we have

$$z \;=\; a_{\ell_t} + a_{\ell_{t-1}} + \cdots + a_{\ell_2} + a_{\ell_1}, \tag{1.2}$$

where $\ell_1 < \ell_2 < \cdots < \ell_t$ and $\{a_{\ell_j}, a_{\ell_{j+1}}\} \not\subset b_i$ for any $i \geq 0$ and $1 \leq j \leq t-1$, then we call this a legal m-gonal decomposition of z. Namely, a legal m-gonal decomposition cannot use more than one summand from the same bin. And with the generalized construction of the sequence by partitioning the members into bins rather than relying solely on the 2-dimensional circumscribed polygons, we make it a formal definition for $m \geq 1$.

Definition 1.1. *Let an increasing sequence of positive integers $\{a_n\}_{n=0}^{\infty}$ be given and partition the elements into ordered lists that we call bins*

$$b_k \;:=\; [a_{m(k-1)+1}, a_{m(k-1)+2}, \ldots, a_{mk}] \tag{1.3}$$

for $m \geq 1$, $k \geq 1$, and $b_0 = [a_0]$. We declare a decomposition of an integer

$$z \;=\; a_{\ell_t} + a_{\ell_{t-1}} + \cdots + a_{\ell_1} \tag{1.4}$$

where $\ell_1 < \ell_2 < \cdots < \ell_t$ and $\{a_{\ell_j}, a_{\ell_{j+1}}\} \not\subset b_i$ for any i, j to be a legal m-gonal decomposition.

The following definition details the construction of the m-gonal sequence, which is the focus of this paper.

Definition 1.2. *For $m \geq 1$, an increasing sequence of positive integers $\{a_n\}_{n=0}^{\infty}$ is called an* m-gonal sequence *if every a_i ($i \geq 0$) is the smallest positive integer that does not have a legal m-gonal decomposition using the elements $\{a_0, a_1, \ldots, a_{i-1}\}$.*

Example 1.3. *For $m = 1$, all the bins have size 1 and the 1-gonal sequence $\{a_i\}_{i=0}^{\infty}$ is defined by $a_i = 2^i$. This is equivalent to writing an integer in binary. When $m = 2$ we have bins $b_i = [a_{2i-1}, a_{2i}]$ for $i \geq 1$ and $b_0 = [a_0]$. The first few terms of the sequence are*

$$\underbrace{1}_{b_0},\ \underbrace{2,\ 4}_{b_1},\ \underbrace{6,\ 12}_{b_2},\ \underbrace{18,\ 36}_{b_3},\ \underbrace{54, 108}_{b_4},\ \underbrace{162,\ 324}_{b_5},\ \dots.$$

In the case where $m = 3$ the triangle (3-gonal) sequence begins with the terms

$$\underbrace{1}_{b_0},\ \underbrace{2,\ 4,\ 6}_{b_1},\ \underbrace{8,\ 16,\ 24}_{b_2},\ \underbrace{32,\ 64,\ 96}_{b_3},\ \underbrace{128,\ 256,\ 384}_{b_4},\ \underbrace{512,\ 1024,\ 1536}_{b_5},\ \dots.$$

Figure 2 gives a visualization of the beginning of the triangle sequence when the integers are placed in the faces of the circumscribed triangles. Moreover, we note that the triangles used need not be equilateral.

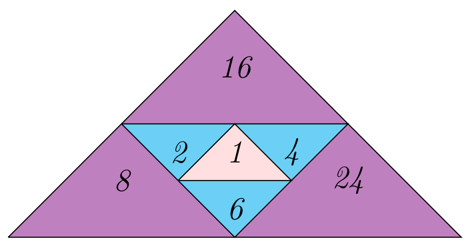


FIGURE 2. Beginning of triangle sequence

Also one can observe that the triangle decomposition of 2015 *is given by*

$$2015 = a_{15} + a_{12} + a_6 + a_3 + a_0 = 1536 + 384 + 64 + 24 + 6 + 1. \tag{1.5}$$

In Section 2 we derive the recurrence relation and explicit closed form expressions for the terms of the the m-gonal sequence, which we state below.

Theorem 1.4. *Let $m \geq 1$. If $\{a_n\}_{n=0}^{\infty}$ is the m-gonal sequence, then*

$$a_n \ = \ \begin{cases} 1 & \text{if } n = 0 \\ 2n & \text{if } 1 \leq n \leq m \\ (m+1)a_{n-m} & \text{if } n > m. \end{cases} \tag{1.6}$$

Then for $n \geq 1$, with $n = km + r$, $k \geq 0$ and $1 \leq r \leq m$

$$a_n \ = \ 2r(m+1)^k\ . \tag{1.7}$$

1.2. **Uniqueness of Decomposition.** Notice that for $m \geq 2$, the recurrence given in Theorem 1.4 is not a positive linear recurrence as the leading coefficients of the first m terms are zero. Therefore past results on positive linear recurrences do not apply to the m-gonal sequence; however, we do still obtain unique decomposition.

Theorem 1.5 (Uniqueness of decompositions)**.** *Fix $m \geq 1$. Every positive integer can be written uniquely as a sum of distinct terms from the m-gonal sequence, where no two summands are in the same bin.*

A proof of Theorem 1.5 is given in Appendix A.

1.3. **Gaussianity.** Previous work with positive linear recurrence sequences proved the number of summands in the decomposition of positive integers converges to a Gaussian (see among others [DDKMMV, MW2]). The same holds for Kentucky decompositions despite the fact that the Kentucky sequence is not a positive linear recurrence [CFHMN1], and also for the m-gonal sequences.

Theorem 1.6 (Gaussian Behavior of Summands)**.** *Let the random variable Y_n denote the number of summands in the (unique) m-gonal decomposition of an integer picked at random from $[0, a_{mn+1})$ with uniform probability.*[3] *Normalize Y_n to $Y_n' = (Y_n - \mu_n)/\sigma_n$, where μ_n and σ_n are the mean and variance of Y_n respectively. Then*

$$\mu_n \;=\; \frac{mn}{m+1} + \frac{1}{2}, \qquad \sigma_n^2 \;=\; \frac{mn}{(m+1)^2} + \frac{1}{4}, \tag{1.8}$$

and Y_n' converges in distribution to the standard normal distribution as $n \to \infty$.

The proof of Theorem 1.6 is given in Section 3.

1.4. **Gaps between summands.** Another property studied of positive linear recurrence sequences is the behavior of the gaps between adjacent summands in decompositions, where, in many instances, it has been shown that there is exponential decay in the distribution of gaps, see [BBGILMT, B-AM, BILMT].[4] Similarly, the Kentucky Sequence displays exponential decay in the distribution of gaps [CFHMN1]. We obtain similar behavior again, though now there is a slight dependence on the residue of gap modulo m (if we split by residue we obtain geometric decay).

Before stating our result we first fix some notation. For the legal m-gonal decomposition

$$z \;=\; a_{\ell_k} + a_{\ell_{k-1}} + \cdots + a_{\ell_1} \quad \text{with} \quad \ell_1 \;<\; \ell_2 \;<\; \cdots \;<\; \ell_k \tag{1.9}$$

and $z \in [0, a_{mn+1})$, we define the multiset of gaps as follows:

$$\text{Gaps}_n(z) \;:=\; \{\ell_2 - \ell_1, \ell_3 - \ell_2, \ldots, \ell_k - \ell_{k-1}\}. \tag{1.10}$$

Observe that we do not consider $\ell_1 - 0$, as a gap. However, doing so would not affect the limiting behavior. For example, notice $z = a_{15} + a_{12} + a_6 + a_3 + a_0$ contributes three gaps of length 3, and one gap of length 6.

Considering all the gaps between summands in legal m-gonal decompositions of all $z \in [0, a_{mn+1})$, we let $P_n(g)$ be the fraction of all these gaps that are of length g. That is, $P_n(g)$ is the probability of a gap of length g among legal m-gonal decompositions of $z \in [0, a_{mn+1})$.

Theorem 1.7 (Average Gap Measure)**.** *Let $g = m\alpha + \beta$, where $\alpha \geq 0$ and $0 \leq \beta < m$. For $P_n(g)$ as defined above, the limit $P(g) := \lim_{n\to\infty} P_n(g)$ exists, and*

$$P(g) \;=\; \begin{cases} \frac{\beta}{m(m+1)} & \text{if } \alpha = 0 \\ \frac{m+1-\beta}{(m+1)^{\alpha+1}} & \text{if } \alpha > 0. \end{cases} \tag{1.11}$$

[3]Using the methods of [BDEMMTTW], these results can be extended to hold almost surely for a sufficiently large sub-interval of $[0, a_{mn+1})$.

[4]The proofs involve technical arguments concerning roots of polynomials associated to the recurrence; in many cases one needs to assume all the recurrence coefficients are positive.

The proof for Theorem 1.7 is given in Section 4.

Via an application of [DFFHMPP, Theorem 1.1] we extract a result on individual gaps for the m-gonal case. In order to state the theorem, we need the following definitions, as were presented in [DFFHMPP], but specialized to the m-gonal case. Let $\{a_n\}$ denote the m-gonal sequence with its unique decomposition as given in Definition 1.1. Let $I_n := [0, a_{mn+1})$ for all $n > 0$ and let $\delta(x-a)$ denotes the Dirac delta functional, assigning a mass of 1 to $x = a$ and 0 otherwise.

- *Spacing gap measure:* We define the spacing gap measure of a $z \in I_n$ with $k(z)$ summands as
$$\nu_{z,n}(x) \;:=\; \frac{1}{k(z)-1}\sum_{j=2}^{k(z)} \delta(x - (\ell_j - \ell_{j-1})). \tag{1.12}$$
- *Average spacing gap measure:* Note that the total number of gaps for all $z \in I_n$ is
$$N_{\text{gaps}}(n) \;:=\; \sum_{z=a_0}^{a_{mn+1}-1} (k(z)-1). \tag{1.13}$$
  The average spacing gap measure for all $z \in I_n$ is
$$\begin{aligned}\nu_n(x) \;&:=\; \frac{1}{N_{\text{gaps}}(n)} \sum_{z=a_0}^{a_{mn+1}-1} \sum_{j=2}^{k(z)} \delta\left(x - (\ell_j - \ell_{j-1})\right)\\ &=\; \frac{1}{N_{\text{gaps}}(n)} \sum_{z=a_0}^{a_{mn+1}-1} (k(z)-1)\,\nu_{z,n}(x).\end{aligned} \tag{1.14}$$
  Letting $P_n(g)$ denote the probability of a gap of length g among all gaps from the decompositions of all $z \in I_n$, we have
$$\nu_n(x) \;=\; \sum_{g=0}^{mn} P_n(g)\delta(x-g). \tag{1.15}$$
- *Limiting average spacing gap measure, limiting gap probabilities:* If the limits exist, we let
$$\nu(x) \;=\; \lim_{n\to\infty} \nu_n(x), \qquad P(k) \;=\; \lim_{n\to\infty} P_n(k). \tag{1.16}$$
- *Indicator function for two gaps:* For $g_1, g_2 \geq 0$
$$X_{j_1,j_1+g_1,j_2,j_2+g_2}(n) \;:=\; \#\left\{ z \in I_n : \substack{a_{j_1},a_{j_1+g_1},a_{j_2},a_{j_2+g_2} \text{ in } z\text{'s decomposition,}\\ \text{but not } a_{j_1+q},a_{j_2+p} \text{ for } 0<q<g_1, 0<p<g_2} \right\}. \tag{1.17}$$
- *Specific gap length probability:* Recall that $P_n(g)$ is the probability
$$P_n(g) \;:=\; \frac{1}{N_{\text{gaps}}(n)} \sum_{i=1}^{mn+1-g} X_{i,i+g}(n). \tag{1.18}$$

Now can now state the result of the individual gap measure for the m-gonal case.

Theorem 1.8 (Individual Gap Measure)**.** *For $z \in I_n$, the individual gap measures $\nu_{z,n}(x)$ converge almost surely in distribution to the limiting gap measure $\nu(x)$.*

We give a proof of Theorem 1.8 in Section 5.

1.5. **Longest Gap.** Another interesting problem is to determine the distribution of the longest gap between summands as $n \to \infty$. The structure of the legal m-gonal decompositions allows us to easily prove the following.

Theorem 1.9 (Distribution of the Longest Gap)**.** *Consider the m-gonal sequence $\{a_n\}$. Then as $n \to \infty$ the mean of the longest gap between summands in legal m-gonal decompositions of integers in $[a_n, a_{n+1})$ is $m \log_2(n/2m) + O_m(1)$, and the variance is $O_m(1)$.*

The proof of Theorem 1.9 is given in Section 6 and bypasses many of the technical arguments used in [BILMT]. There the authors had to deduce properties of somewhat general associated polynomials; the nature of the legal m-gonal decompositions here allows us to immediately convert this problem to a simple generalization of the longest run of heads problem.

2. Recurrence relations and generating functions

Let $m \geq 1$. We can use the division algorithm to observe that the integer a_{mk+r} is the r^{th} integer in the bin b_{k+1} for $mk + r \geq 1$. Hence $1 \leq r \leq m$ denotes the location of the integer within its bin. We let the first bin b_0 contain the element $a_0 = 1$. Then for any $k \geq 0$, we let b_{k+1} denote the set of elements of the $(k+1)^{\text{th}}$ bin. Namely

$$\underbrace{a_0}_{b_0}, \ \underbrace{a_1, a_2, \dots, a_m}_{b_1}, \ \underbrace{a_{m+1}, a_{m+2}, \dots, a_{2m}}_{b_2}, \ \dots, \ \underbrace{a_{mk+1}, a_{mk+2}, \dots, a_{m(k+1)}}_{b_{k+1}}, \ \dots. \tag{2.1}$$

We can now begin our work in describing the terms of this sequence.

The following result, which follows immediately from the definition, is used in many of the proofs in this section. We record it here for easy reference.

Definition 2.1. *Let Ω_n denote the integer with summands from each bin $b_0, b_1, b_2, \dots, b_n$. Then*

$$\Omega_n \ = \ \sum_{i=0}^{n} a_{mi}. \tag{2.2}$$

The first result that makes use of Definition 2.1 is given below.

Lemma 2.2. *Let $m \geq 1$ and $k \geq 1$. If a_{mk+1} is the first entry in bin b_{k+1}, then $a_{mk+1} = a_{mk} + a_{m(k-1)+1}$.*

Proof. We note that since a_{mk+1} and $a_{m(k-1)+1}$ are the first numbers in the bins b_{k+1} and b_k, respectively, by Equation (2.2) we have that

$$a_{mk+1} \ = \ 1 + \Omega_k \ = \ 1 + \sum_{i=0}^{k} a_{mi} \tag{2.3}$$

$$a_{m(k-1)+1} \ = \ 1 + \Omega_{k-1} \ = \ 1 + \sum_{i=0}^{k-1} a_{mi}. \tag{2.4}$$

Then Equations (2.3) and (2.4) yield

$$a_{mk+1} \ = \ 1 + \Omega_k \ = \ 1 + a_{mk} + \Omega_{k-1} \ = \ a_{mk} + (1 + \Omega_{k-1}) \ = \ a_{mk} + a_{m(k-1)+1}, \tag{2.5}$$

as claimed. □

We now prove the more general result.

Lemma 2.3. *If $k \geq 0$ and $1 \leq r \leq m$, then $a_{mk+r} = r \cdot a_{mk+1}$.*

Proof. First consider the bin b_1. As $b_0 = [a_0] = 1$, by construction of the m-gonal sequence it is straightforward to determine that $b_1 = [2, 4, \ldots, 2m]$ and $a_r = r \cdot a_1$ for all $1 \leq r \leq m$.

We proceed for bins b_k with $k \geq 1$ by induction on r, where $1 \leq r \leq m$. The basis case when $r = 1$ clearly holds.
Let $1 \leq x \leq m-1$ and assume that for any $1 \leq r \leq x$, we have that $a_{mk+r} = r \cdot a_{mk+1}$. We want to show that $a_{mk+x+1} = (x+1)a_{mk+1}$. Recall that a_{mk+x+1} is the entry in bin b_{k+1} after a_{mk+x} and by definition a_{mk+x+1} is one more than the largest integer we can create using the elements of bins $b_0, b_1, \ldots, b_k$ along with the element a_{mk+x}. Using Equation (2.2), we have that

$$a_{mk+x+1} \ = \ 1 + a_{mk+x} + \Omega_k \ = \ 1 + a_{mk+x} + \sum_{i=0}^{k} a_{mi}. \tag{2.6}$$

Recalling that $1 + \Omega_k = a_{mk+1}$ and by the use of the induction hypothesis, Equation (2.6) yields

$$a_{mk+x+1} \ = \ a_{mk+x} + a_{mk+1} \ = \ xa_{mk+1} + a_{mk+1} \ = \ (x+1)a_{mk+1}. \tag{2.7}$$

□

We now provide a closed formula for the terms of the m-gonal sequence.

Proposition 2.4. *Let $m \geq 2$, $k \geq 0$, and $1 \leq r \leq m$. Then $a_{mk+r} = 2r(m+1)^k$. For $m = 1$, $a_i = 2^i$.*

Proof. For the case where $m = 1$, each of our bins have size 1 and a legal decomposition has distinct summands. Thus the rule for legal decomposition is precisely a description of writing the positive integers in binary.

We will proceed by induction on k, the subscript on the bin, and r, the location of a_{mk+r} within the bin considered. The basis case $k = 0$ and $r = 1$ gives the expected result, $a_{m\cdot 0+1} = 2(1)(m+1)^0 = 2$. We now assume that for some $k \geq 0$ and some r, $1 \leq r \leq m$, we have

$$a_{mk+r} \ = \ 2r(m+1)^k. \tag{2.8}$$

We need to show that the following two equations hold:

$$a_{mk+r+1} \ = \ 2(r+1)(m+1)^k \tag{2.9}$$

$$a_{m(k+1)+r} \ = \ 2r(m+1)^{k+1}. \tag{2.10}$$

Suppose that $1 \leq r \leq m-1$. To show Equation (2.9) holds it suffices to observe that by Lemma 2.3 and our induction hypothesis we have

$$a_{mk+r+1} \ = \ (r+1)a_{mk+1} \ = \ (r+1) \cdot 2(1)(m+1)^k \ = \ 2(r+1)(m+1)^k. \tag{2.11}$$

When $r = m$, we use Lemma 2.2 and our induction hypothesis to quickly deduce

$$a_{mk+m+1} = 2(m+1)^{k+1}. \tag{2.12}$$

By iterating (2.11) till $r = m-1$, we find that (2.12) holds. Then (2.10) holds by Lemma 2.3. □

The final result gives the recurrence relation stated in Theorem 1.4.

Corollary 2.5. *If $n > m$, then $a_n = (m+1)a_{n-m}$.*

Proof. Let $n > m$ and write $n = mk + r$, where $k \geq 1$ and $1 \leq r \leq m$. By Proposition 2.4, $a_n = a_{mk+r} = 2r(m+1)^k$ and $a_{n-m} = a_{m(k-1)+r} = 2r(m+1)^{k-1}$. So it directly follows that $a_n = (m+1)a_{n-m}$. □

2.1. **Counting integers with exactly k summands.** In [KKMW], Koloğlu, Kopp, Miller and Wang introduced a very useful combinatorial perspective to attack Zeckendorf decomposition problems by partitioning the integers $z \in [F_n, F_{n+1})$ into sets based on the number of summands in their Zeckendorf decomposition. We use a similar technique to prove that the distribution of the average number of summands in the m-gonal decomposition displays Gaussian behavior.

let $p_{n,k}$ denote the number of integers in $I_n := [0, a_{mn+1})$ whose m-gonal decomposition contains exactly k summands, where $k \geq 0$. We begin our analysis with the following result.

Proposition 2.6. *If $n, k \geq 0$, then*

$$p_{n,k} \;=\; \begin{cases} 1 & \text{if } k = 0 \\ m^k\binom{n}{k} + m^{k-1}\binom{n}{k-1} & \text{if } 1 \leq k \leq n+1 \\ 0 & \text{if } k > n+1. \end{cases} \tag{2.13}$$

Proof. Let $n, k \geq 0$. Observe that the unique integer in the interval $I_n = [0, a_{mn+1})$ which has zero summands is zero itself. Thus $p_{n,0} = 1$. Now if k is larger than the number of available bins, it would be impossible to have k summands as one can draw no more than one summand per bin. Therefore $p_{n,k} = 0$, whenever $k > n+1$.

We now show that for $1 \leq k \leq n+1$, $p_{n,k} = m^k\binom{n}{k} + m^{k-1}\binom{n}{k-1}$. There are two cases to consider:

Case 1. One of the k summands is chosen from b_0.
Case 2. None of the k summands are chosen from b_0.

Case 1. Since one of the k summands is coming from b_0 there are $k-1$ available summands to take from the bins $b_1, \dots, b_n$. The number of ways to select $k-1$ bins from n bins is $\binom{n}{k-1}$. As each of the bins $b_1, \dots, b_n$ has exactly m elements and $|b_0| = 1$, once the $(k-1)$ bins are selected, the number of ways to select an element from these bins is m^{k-1}. Thus the number of $z \in I_n$ which have exactly k summands with one summand coming from bin b_0 is $m^{k-1}\binom{n}{k-1}$.

Case 2. We choose k summands from any bin but b_0. Using a similar argument as in Case 1, we can see that the total number of ways to select these k summands is $m^k\binom{n}{k}$.

As the two cases are disjoint, we have shown that the total number of integers in the interval I_n with exactly k summands is

$$p_{n,k} \;=\; m^k\binom{n}{k} + m^{k-1}\binom{n}{k-1}. \tag{2.14}$$

□

We also provide a recursive formula for the value of $p_{n,k}$ as it is used in the proof of Proposition 2.8.

Proposition 2.7. *If $0 < k < n+1$, then $p_{n,k} = mp_{n-1,k-1} + p_{n-1,k}$.*

We omit the proof of Proposition 2.7 as it is a straightforward application of the combinatorial identity $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$. With the recursive formula at hand, we now create a generating function for $p_{n,k}$.

Proposition 2.8. *Let*

$$F(x,y) \;:=\; \sum_{n,k\geq 0} p_{n,k}x^n y^k \tag{2.15}$$

be the generating function of the $p_{n,k}$'s arising from m-gonal decompositions. Then

$$F(x,y) \;=\; \frac{1+y}{1-(my+1)x}. \tag{2.16}$$

Proof. Noting that $p_{n,k} = 0$ if either $n < 0$ or $k < 0$, using explicit values of $p_{n,k}$ and the recurrence relation from Proposition 2.7, after some straightforward algebra we obtain

$$F(x,y) \;=\; mxyF(x,y) + xF(x,y) + 1 + y. \tag{2.17}$$

From this, Equation (2.16) follows. □

3. Gaussian behavior

To motivate this section's main result, we point the reader to the following experimental observations. Taking samples of 200,000 integers from the intervals $[0, 2(4)^{600})$, $[0, 2(5)^{600})$, $[0, 2(6)^{600})$ and $[0, 2(7)^{600})$, in Figure 3 we provide a histogram for the distribution of the number of summands in the m-gonal decomposition of these integers, when $m = 3, 4, 5$ and 6, respectively. Moreover, Figure 3 provides the histograms and Gaussian curves (associated to the respective value of m and n; the interval is $[0, a_{mn+1})$ so $n = 600$ in all experiments). In Table 1 we give the values of the predicted means and variances (as computed using Proposition 3.2), as well as the sample means and variances, for each of the cases considered.

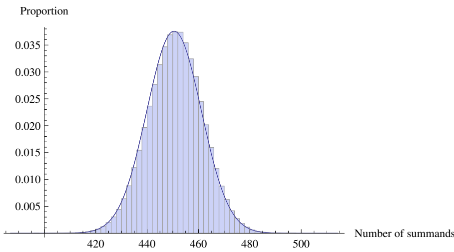
Figure	m	Predicted Mean	Sample Mean	Predicted Variance	Sample Variance
3a	3	450.50	450.49	112.75	112.34
3b	4	480.50	480.52	96.25	95.73
3c	5	500.50	450.49	83.58	83.38
3d	6	514.79	514.76	73.72	73.64

Table 1. Predicted means and variances versus sample means and variances for simulation from Figure 3.

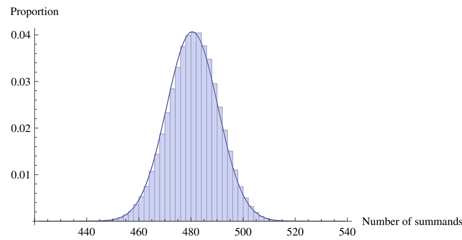
From these observations it is expected that for any $m \geq 1$, the distribution of the number of summands in the m-gonal decompositions of integers in the interval I_n displays Gaussian behavior. This is in fact the statement of Theorem 1.6. We begin by proving a technical result and follow it with the formulas for the mean and variance, which make use of some properties associated with the generating function for the $p_{n,k}$'s.

Proposition 3.1. *If $g_n(y)$ denotes the coefficient of x^n in $F(x,y)$, then*

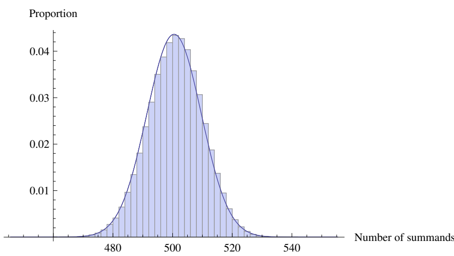
$$g_n(y) \;=\; (1+y)(my+1)^n. \tag{3.1}$$



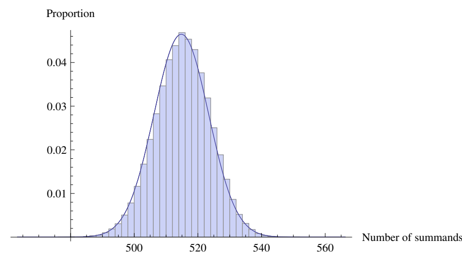
(A) $m = 3$



(B) $m = 4$



(C) $m = 5$



(D) $m = 6$

FIGURE 3. Distributions for the number of summands in the m-gonal decomposition for a random sample with $n = 600$.

Proof. Using the fact that $F(x, y) = \frac{1+y}{1-mxy-x}$ we have by geometric series that

$$F(x, y) \;=\; \sum_{n=0}^{\infty} (1 + y)(my + 1)^n x^n. \tag{3.2}$$

Thus the coefficient of x^n in $F(x, y)$ is $(1 + y)(my + 1)^n$. □

We can now use $g_n(y)$ to find the mean and variance for the number of summands for integers $z \in I_n$.

Proposition 3.2. *Let Y_n be the number of summands in the m-gonal decomposition of a randomly chosen integer in the interval I_n, where each integer has an equal probability of being chosen. Let μ_n and σ_n^2 denote the mean and variance of Y_n. Then*

$$\mu_n \;=\; \frac{nm}{m+1} + \frac{1}{2}, \qquad \sigma_n^2 \;=\; \frac{nm}{(m+1)^2} + \frac{1}{4}. \tag{3.3}$$

Proof. By Propositions 4.7 annd 4.8 in [DDKMMV] the mean and variance of Y_n are

$$\mu_n \;=\; \sum_{i=0}^{n} iP(Y_n = i) \;=\; \sum_{i=0}^{n} i \frac{p_{n,i}}{\sum_{k=0}^{n} p_{n,k}} \;=\; \frac{g_n'(1)}{g_n(1)}, \text{ and} \tag{3.4}$$

$$\sigma_n^2 \;=\; \sum_{i=0}^{n} (i - \mu_n)^2 P(Y_n = i) \;=\; \sum_{i=0}^{n} i^2 \frac{p_{n,i}}{\sum_{k=0}^{n} p_{n,k}} - \mu_n^2 \;=\; \frac{\frac{d}{dy}[yg_n'(y)]|_{y=1}}{g_n(1)} - \mu_n^2. \tag{3.5}$$

Our result follows directly from these formulas and the fact that $g_n(y) = (1+y)(my+1)^n$. □

Normalize Y_n to $Y_n' = \frac{Y_n - \mu_n}{\sigma_n}$, where μ_n and σ_n are the mean and variance of Y_n respectively, as given in Proposition 3.2. We are now ready to prove that Y_n' converges in distribution to the standard normal distribution as $n \to \infty$.

Proof of Theorem 1.6. For convenience we set $r := \frac{t}{\sigma_n}$. Since $\sigma_n = \sqrt{\frac{nm}{(m+1)^2} + \frac{1}{4}}$, we know that $r \to 0$ as $n \to \infty$ for any fixed value of t. Hence we will expand e^r using its power series expansion. We start with

$$M_{Y_n'}(t) \;=\; \frac{g_n(e^{\frac{t}{\sigma_n}})e^{\frac{-t\mu_n}{\sigma_n}}}{g_n(1)}. \tag{3.6}$$

Taking the logarithm of Equation (3.6)

$$\log(M_{Y_n'}(t)) \;=\; \log[g_n(e^r)] - \log[g_n(1)] - \frac{t\mu_n}{\sigma_n}. \tag{3.7}$$

We proceed using Taylor expansions of the exponential and logarithmic functions to expand the following:

$$\begin{aligned}
\log[g_n(e^r)] &= \log(1+e^r) + n\log(me^r+1) \\
&= \log\left(1+\left(1+r+\frac{r^2}{2}\right)\right) + n\log\left(m\left(1+r+\frac{r^2}{2}\right)+1\right) + O(r^3) \\
&= \log(2) + \frac{1}{2}\left(r+\frac{r^2}{2}\right) - \frac{1}{8}\left(r+\frac{r^2}{2}\right)^2 \\
&\quad + n\left[\log(m+1) + \frac{\left(mr+\frac{mr^2}{2}\right)}{m+1} - \frac{\left(mr+\frac{mr^2}{2}\right)^2}{2(m+1)^2}\right] + O(r^3) \\
&= \log(2(m+1)^n) + \frac{r}{2} + \frac{r^2}{8} + \frac{nmr}{m+1} + \frac{nmr^2}{2(m+1)^2} + O(r^3).
\end{aligned} \tag{3.8}$$

From Proposition 3.1 we have that $g_n(1) = 2(m+1)^n$, hence

$$\log[2(m+1)^n] \;=\; \log[g_n(1)]. \tag{3.9}$$

Substituting Equations (3.8) and (3.9) and the values $\mu_n = \frac{nm}{m+1} + \frac{1}{2}$ and $\sigma_n = \sqrt{\frac{nm}{(m+1)^2} + \frac{1}{4}}$ into Equation (3.7) yields

$$\log(M_{Y_n'}(t)) \;=\; \frac{r}{2} + \frac{r^2}{8} + \frac{nmr}{m+1} + \frac{nmr^2}{2(m+1)^2} - \frac{t\left(\frac{nm}{m+1}+\frac{1}{2}\right)}{\sqrt{\frac{nm}{(m+1)^2}+\frac{1}{4}}} + O(r^3). \tag{3.10}$$

After some straightforward algebra we arrive at

$$\log(M_{Y_n'}(t)) \;=\; \frac{t^2}{2} + o(1); \tag{3.11}$$

the moment generating proof of the Central Limit Theorem now yields that the distribution converges to that of the standard normal distribution as $n \to \infty$. □

4. Average Gap Measure

We now turn our attention to our final result in which we determine the behavior of gaps between summands. We begin with some preliminary notation in order to make our approach precise. For a positive integer $z \in I_n = [0, a_{mn+1})$ with m-gonal decomposition

$$z \;=\; a_{\ell_t} + a_{\ell_{t-1}} + \cdots + a_{\ell_1} \tag{4.1}$$

where $\ell_1 < \ell_2 < \cdots < \ell_t$, we define the multiset of gaps of z as

$$Gaps_n(z) \;:=\; \{\ell_2 - \ell_1, \ell_3 - \ell_2, \ldots, \ell_t - \ell_{t-1}\}. \tag{4.2}$$

Our result will average over all $z \in [0, a_{mn+1})$ since we are interested in the average gap measure arising from m-gonal decompositions.

We follow the methods of [BBGILMT, BILMT]. In order to have a gap of length exactly g in the decomposition of z, there must be some index i such that a_i and a_{i+g} occur in z's decomposition, but a_j does not for any j between i and $i+g$. Thus for each i we count how many z have a_i and a_{i+g} but not a_j for $i < j < i+g$; summing this count over i gives the number of occurrences of a gap of length g among all the decompositions of z in our interval of interest. We want to compute the fraction of the gaps (of length g) arising from the decompositions of all $z \in I_n$. This probability is given by

$$P_n(g) \;:=\; \frac{1}{(\mu_n - 1)a_{mn+1}} \sum_{z=0}^{a_{mn+1}-1} \sum_{i=0}^{mn+1-g} X_{i,g}(z), \tag{4.3}$$

where $X_{i,g}(z)$ is the indicator function[5].

We are now ready to prove the result on the exponential decay in the distribution of gaps. The arguments in the proof of our main result (Theorem 1.7) are quite straightforward, however a bit tedious. To simplify our arguments, we write the gap length g as $m\alpha + \beta$, where $\alpha \geq 0$ and $0 \leq \beta < m$.

Proof of Theorem 1.7. Let $g = m\alpha + \beta$, where $\alpha \geq 0$ and $0 \leq \beta < m$. We proceed by considering the following two cases:

Case 1. $\alpha = 0$,
Case 2. $\alpha > 0$.

Case 1. Let $\alpha = 0$. Hence $g = \beta$, where $0 < \beta < m$ and so our gap is less than the size of each bin b_i for $i > 0$ ($\alpha = 0$ and $\beta = 0$ would give us a gap of length 0 which is not m-gonal legal). We first consider gaps of length $g = \beta$ beginning at index 0. If $a_i = a_0$ then the only way to have a gap of length $g = \beta$ is if $a_{i+1} = a_\beta$. Now we are counting integers with m-gonal decompositions of the form $a_{\ell_t} + \cdots + a_{\ell_3} + a_\beta + a_0$. The number of $z \in I_n$ with summands $a_{\ell_3}, \ldots, a_{\ell_t}$ coming from bins $b_2, b_3, \ldots, b_n$ is $(m+1)^{n-1}$.

If $a_i = a_{mk+r}$, $k \geq 0$ and $1 \leq r \leq m$, then $a_{i+g} = a_{mk+r+\beta}$. Notice that the case $r + \beta \leq m$ cannot occur as this would force $\{a_i, a_{i+g}\} \subset b_{k+1}$, which leads to a decomposition that is

[5]For $1 \leq i, g \leq mn$ $X_{i,g}(z)$ denotes whether the decomposition of z has a gap of length g beginning at index i. That is, for $z = a_{\ell_t} + a_{\ell_{t-1}} + \cdots + a_{\ell_1}$,

$$X_{i,g}(z) \;=\; \begin{cases} 1 & \text{if } \exists\, j,\, 1 \leq j \leq t \text{ with } i = \ell_j \text{ and } i + g = \ell_{j+1} \\ 0 & \text{otherwise.} \end{cases} \tag{4.4}$$

not m-gonal legal as only one summand can be taken per bin. Now if $r+\beta > m$, then $a_{i+g} \in b_{k+2}$.

Notice that in this case $a_i = a_{mk+r}$ is one of the largest β entries in bin b_{k+1} and thus there are β many choices for r, namely $r \in \{m-\beta+1, m-\beta+2, \dots, m\}$.

Now we need to count the number of $z \in I_n \setminus \{0\} = [1, a_{mn+1})$ which have summands $a_i \in b_{k+1}$ and $a_{i+g} \in b_{k+2}$. We must have $0 \leq k \leq n-2$.

As we have already used bins b_{k+1} and b_{k+2}, it follows from a straightforward combinatorial counting argument that the total number of integers $z \in I_n$ that can be created with summands $a_i \in b_{k+1}$ and $a_{i+g} \in b_{k+2}$ (with no summands in between) is given by

$$2\beta(m+1)^{n-2}, \tag{4.5}$$

where the factor of β comes from the β possible choices of a_i within the bin b_{k+1}. As we can vary k from 0 to $(n-2)$, we find that the total number of integers $z \in I_n$ which contribute a gap between a_i $(i \neq 0)$ and a_{i+g} (assuming that $r+\beta > m$) is given by

$$2(n-1)\beta(m+1)^{n-2}. \tag{4.6}$$

Observe that Equation (4.6) does not account for the case when $i=0$, which adds an extra factor of $(m+1)^{n-1}$. Therefore

$$\sum_{z=0}^{a_{mn+1}-1} \sum_{i=0}^{mn+1-g} X_{i,g}(z) \;=\; 2\beta(n-1)(m+1)^{n-2} + (m+1)^{n-1}. \tag{4.7}$$

Using Proposition 3.2 and Proposition 2.4 we have that

$$(\mu_n - 1)a_{mn+1} \;=\; \left(\frac{mn}{m+1} - \frac{1}{2}\right)(2(m+1)^n) \;=\; (m+1)^{n-1}(2mn-(m+1)). \tag{4.8}$$

By Equations (4.7) and (4.8), for $g=\beta$, with $0 \leq \beta < m$, we have that

$$P_n(g) \;=\; \frac{2\beta(n-1)}{(m+1)(2mn-m-1)} + \frac{1}{2mn-m-1}. \tag{4.9}$$

Now recall that $P(g) = \lim_{n\to\infty} P_n(g)$, so by letting $n \to \infty$ in Equation (4.9) we have that

$$P(g) \;=\; \frac{\beta}{m(m+1)}, \tag{4.10}$$

whenever $g = \beta$ and $0 \leq \beta < m$. This completes Case 1.

Case 2. Let $g = m\alpha + \beta$, where $\alpha \geq 1$ and $0 \leq \beta < m$. First consider when $a_i = a_0$. If $\beta = 0$, then $a_{i+g} = a_{m\alpha} \in b_\alpha$. Otherwise, when $0 < \beta < m$, $a_{i+g} = a_{m\alpha+\beta} \in b_{\alpha+1}$. In the case of the former, the number of $z \in I_n$ with summands coming from bins $b_{\alpha+1}, b_{\alpha+2}, \dots, b_n$ is $(m+1)^{n-\alpha}$. In the latter case, the number of $z \in I_n$ with summands coming from bins $b_{\alpha+2}, b_{\alpha+3}, \dots, b_n$ is $(m+1)^{n-(\alpha+1)}$.

Now if $a_i = a_{mk+r}$, with $k \geq 0$ and $1 \leq r \leq m$, then $a_{i+g} = a_{m(k+\alpha)+r+\beta}$. Hence $a_i \in b_{k+1}$ and $a_{i+g} \in b_{k+\alpha+1}$ whenever $1 \leq r+\beta \leq m$, or $a_{i+g} \in b_{k+\alpha+2}$ whenever $m < r+\beta < 2m$. Hence we consider the following subcases:

Subcase 1. Let $1 \leq r+\beta \leq m$.

Subcase 2. Let $m < r+\beta < 2m$.

Subcase 1. Let $1 \leq r+\beta \leq m$. In this case $a_i \in b_{k+1}$ and $a_{i+g} \in b_{k+\alpha+1}$. Notice that in this case a_i must be one of the smallest $m-\beta$ entries in bin b_{k+1}. Namely $r = 1, 2, \dots, m-\beta$.

Now we need to count the number of $z \in I_n$ which have summands $a_i \in b_{k+1}$ and $a_{i+g} \in b_{k+\alpha+1}$ (and no summands in between). For the decomposition to only have summands from bins $b_0, \dots, b_{k+1}, b_{k+\alpha+1}, \dots, b_n$, we must have $0 \leq k \leq n-(\alpha+1)$.

In order to have a gap created by $a_i \in b_{k+1}$ and $a_{i+g} \in b_{k+\alpha+1}$, there must be no summands taken from b_j, where $k+1 < j < k+\alpha+1$. Again using a straightforward combinatorial counting argument, the total number of integers $z \in I_n$ which have summands $a_i \in b_{k+1}$ and $a_{i+g} \in b_{k+\alpha+1}$ (with no summands in between) is given by

$$2(m-\beta)(m+1)^{n-\alpha-1}, \tag{4.11}$$

where the factor of $m-\beta$ comes from the $m-\beta$ possible choices of a_i within the bin b_{k+1}.

As we can vary k from 0 to $n-(\alpha+1)$ we find that the total number of integers $z \in I_n$ which contribute a gap between a_i $(i \neq 0)$ and a_{i+g} in this case is

$$2(n-\alpha)(m-\beta)(m+1)^{n-\alpha-1}. \tag{4.12}$$

Subcase 2. Let $m < r+\beta < 2m$. In this case $a_i \in b_{k+1}$ and $a_{i+g} \in b_{k+\alpha+2}$. Notice that in this case a_i can be any of the largest β entries in bin b_{k+1} so $a_i \in b_{k+1}$ and $a_{i+g} \in b_{k+\alpha+2}$. Namely $r = m+1-\beta, m+2-\beta, \dots, m$.

Using the same reasoning as in Subcase 1, we determine that the total number of integers meeting the conditions is

$$2\beta(n-\alpha-1)(m+1)^{n-\alpha-2}. \tag{4.13}$$

This completes Subcase 2.

We still need to account for the number of integers $z \in I_n$ which contribute a gap of length $g = m\alpha+\beta$ ($\alpha \geq 1$ and $0 \leq \beta < m$) beginning at index a_0. Recall we previously computed this quantity to be $(m+1)^{n-\alpha-1}$ when $\beta > 0$ and the quantity is $(m+1)^{n-\alpha}$ when $\beta = 0$.

Therefore we need to sum the values of Equations (4.12), (4.13) along with $(m+1)^{n-\alpha-1}$ when $\beta > 0$ to get that

$$\sum_{z=0}^{a_{mn+1}-1} \sum_{i=0}^{mn+1-g} X_{i,g}(z) = 2(n-\alpha)(m-\beta)(m+1)^{n-\alpha-1} + 2\beta(n-\alpha-1)(m+1)^{n-\alpha-2} + (m+1)^{n-\alpha-1}. \tag{4.14}$$

By Equations (4.14) and (4.8), for $g = m\alpha+\beta$, with $\alpha \geq 1$ and $0 < \beta < m$, we have that

$$P_n(g) = \frac{2(n-\alpha)(m-\beta)}{(m+1)^{\alpha}(2mn-m-1)} + \frac{2\beta(n-\alpha-1)}{(m+1)^{\alpha+1}(2mn-m-1)} + \frac{1}{(m+1)^{\alpha}(2mn-m-1)}. \tag{4.15}$$

Now recall that $P(g) = \lim_{n\to\infty} P_n(g)$, so by letting $n \to \infty$ in Equation (4.15) we have that

$$P(g) = \frac{2(m-\beta)}{(m+1)^{\alpha}(2m)} + \frac{2\beta}{(m+1)^{\alpha+1}(2m)} = \frac{m+1-\beta}{(m+1)^{\alpha+1}}, \tag{4.16}$$

whenever $g = m\alpha+\beta$, $\alpha \geq 1$ and $0 < \beta < m$.

Now for $\beta = 0$ we do not need to consider when $\alpha = 0$ as this would give us a gap $g = 0$. Also, as $1 \le r \le m$, we only meet the conditions of Subcase 1. Thus for $\alpha > 0$ and $\beta = 0$ we need to sum the values of Equations (4.12) along with $(m+1)^{n-\alpha}$ to get

$$\sum_{z=0}^{a_{mn+1}-1} \sum_{i=0}^{mn+1-g} X_{i,g}(z) = 2(n-\alpha)m(m+1)^{n-\alpha-1} + (m+1)^{n-\alpha}. \tag{4.17}$$

By Equations (4.17) and (4.8), for $g = m\alpha$, with $\alpha \ge 1$, we have that

$$P_n(g) \;=\; \frac{2(n-\alpha)m}{(m+1)^{\alpha}(2mn-m-1)} + \frac{1}{(m+1)^{\alpha-1}(2mn-m-1)}. \tag{4.18}$$

Now recall that $P(g) = \lim_{n\to\infty} P_n(g)$, so by letting $n \to \infty$ in Equation (4.18) we have that

$$P(g) \;=\; \frac{1}{(m+1)^{\alpha}} = \frac{m+1}{(m+1)^{\alpha+1}} \tag{4.19}$$

whenever $g = m\alpha + \beta$, $\alpha \ge 1$ and $\beta = 0$. This completes the proof. □

5. Individual Gap Measure

In this section, we prove Theorem 1.8, by checking that the conditions given in [DFFHMPP, Theorem 1.1] are satisfied in the m-gonal case. We restate this theorem below for ease of reference.

Theorem 5.1. [DFFHMPP] *For $z \in I_n$, the individual gap measures $\nu_{z,n}(x)$ converge almost surely in distribution to the average gap measure $\nu(x)$ if the following hold.*

(1) The number of summands for decompositions of $z \in I_n$ converges to a Gaussian with mean $\mu_n = c_{\text{mean}}n + O(1)$ and variance $\sigma_n^2 = c_{\text{var}}n + O(1)$, for constants $c_{\text{mean}}, c_{\text{var}} > 0$, and $k(z) \ll n$ for all $z \in I_n$.

(2) We have the following, with $\lim_{n\to\infty} \sum_{g_1,g_2} \text{error}(n, g_1, g_2) = 0$:

$$\frac{2}{|I_n|\mu_n^2} \sum_{j_1<j_2} X_{j_1,j_1+g_1,j_2,j_2+g_2}(n) \;=\; P(g_1)P(g_2) + \text{error}(n, g_1, g_2). \tag{5.1}$$

(3) The limits in Equation (1.16) exist.

We note that the above theorem is more general than we need, and in our particular case our interval of interest is $I_n = [0, a_{mn+1})$. Now observe that Proposition 3.2 and Theorem 1.6 ensure that the first criterion is met, and $k(z)$ is clearly at most $mn+1$ and thus $k(z) \ll n$. In addition, the exponential decay seen in Theorem 1.7 shows that Condition (3) is met. It remains to show that Condition (2) of Theorem 5.1 holds.

Proposition 5.2. *We have that*

$$\frac{2}{|I_n|\mu_n^2} \sum_{j_1<j_2} X_{j_1,j_1+g_1,j_2,j_2+g_2}(n) \;=\; P(g_1)P(g_2) + \text{error}(n, g_1, g_2) \tag{5.2}$$

and the sum of the error over all pairs (g_1, g_2) goes to zero as $n \to \infty$.

Proof. Let $g_1 = \alpha_1 m + \beta_1$, $g_2 = \alpha_2 m + \beta_2$, $j_1 = k_1 m + r_1$, and $j_2 = k_2 m + r_2$ where $0 \le \beta_1, \beta_2 < m$, $1 \le k_1, k_2 \le m$ and $k_1 < k_2$. Thus a_{j_1} and a_{j_2} are in bins $(k_1 + 1)$ and $(k_2 + 1)$ respectively. There are a number of cases to consider when determining $\sum_{j_1<j_2} X_{j_1,j_1+g_1,j_2,j_2+g_2}(n)$, depending on whether or not $\alpha_1 = 0$ or $\alpha_2 = 0$. We include only the case when both $\alpha_1, \alpha_2 \ge 1$, as the other cases are similar.

There are several subcases to consider. We will first consider the four subcases which contribute to the main term and then bound the remaining cases. In all of the cases contributing to the main term we have $j_1 + g_1 \ne j_2$ and $j_1 \ne 0$ and thus we will suppose these conditions hold below.

Subcase 1: Let $1 \le r_1 + \beta_1 \le m$ and $1 \le r_2 + \beta_2 \le m$. First we determine the possible values of k_1 and k_2. In this case, the gap from g_1 spans $\alpha_1 + 1$ bins and the gap from g_2 spans $\alpha_2 + 1$ bins. Thus $0 \le k_1 \le (n - \alpha_1 - \alpha_2 - 3)$ and $(k_1 + \alpha_1 + 2) \le k_2 \le (n - \alpha_2 - 1)$ and the number of choices for k_1 and k_2 is $n^2/2 + O(n)$. Because of the restrictions of $1 \le r_1 + \beta_1 \le m$ and $1 \le r_2 + \beta_2 \le m$, within each bin there are $(m - \beta_1)$ choices for where to place a_{j_1} and $(m - \beta_2)$ choices for where to place a_{j_2}. Lastly, we determine the number of ways to choose the remaining elements for the decomposition. Because we are spanning $\alpha_1 + 1$ bins for the gap from g_1 and $\alpha_2 + 1$ bins for the gap from g_2, using a straight forward combinatorial counting argument, the number of integers that can be made using what remains is $2(m+1)^{n-\alpha_1-\alpha_2-2}$. Thus the total number of integers that can be made in this case is

$$\begin{aligned} & 2(m-\beta_1)(m-\beta_2)(m+1)^{n-\alpha_1-\alpha_2-2}(n^2/2 + O(n)) \\ &= (m-\beta_1)(m-\beta_2)(m+1)^{n-\alpha_1-\alpha_2-2}(n^2 + O(n)). \end{aligned} \tag{5.3}$$

Through similar arguments we can obtain the remaining three subcases that contribute to the main term.

Subcase 2: Let $1 \le r_1 + \beta_1 \le m$ and $m < r_2 + \beta_2 < 2m$. Then the number of integers that can be made in this case is

$$(m-\beta_1)\beta_2(m+1)^{n-\alpha_1-\alpha_2-3}(n^2 + O(n)). \tag{5.4}$$

Subcase 3: Let $m < r_1 + \beta_1 < 2m$ and $1 \le r_2 + \beta_2 \le m$. Then the number of integers that can be made in this case is

$$(m-\beta_2)\beta_1(m+1)^{n-\alpha_1-\alpha_2-3}(n^2 + O(n)). \tag{5.5}$$

Subcase 4: Let $m < r_1 + \beta_1 < 2m$ and $m < r_2 + \beta_2 < 2m$. Then the number of integers that can be made in this case is

$$\beta_1\beta_2(m+1)^{n-\alpha_1-\alpha_2-4}(n^2 + O(n)). \tag{5.6}$$

The remaining cases occur when $j_1 + g_1 = j_2$ or $j_1 = 0$. In these cases, the number of choices for k_1 and k_2 is on the order of n instead of n^2 and thus the number of integers that can be made in these cases is $O(n(m+1)^{n-\alpha_1-\alpha_2})$. Combining all cases, we have

$$\begin{aligned}\sum_{j_1<j_2} X_{j_1,j_1+g_1,j_2,j_2+g_2}(n) \;&=\; n^2(m-\beta_1)(m-\beta_2)(m+1)^{n-\alpha_1-\alpha_2-2}\\ &+\; n^2(m-\beta_1)\beta_2(m+1)^{n-\alpha_1-\alpha_2-3}\\ &+\; n^2(m-\beta_2)\beta_1(m+1)^{n-\alpha_1-\alpha_2-3}\\ &+\; n^2\beta_1\beta_2(m+1)^{n-\alpha_1-\alpha_2-4}+O(n(m+1)^{n-\alpha_1-\alpha_2}). \qquad (5.7)\end{aligned}$$

Next, recall that by Proposition 3.2 $\mu_n = \frac{nm}{m+1}+\frac{1}{2}$. In addition, $|I_n| = a_{mn+1} = 2(m+1)^n$. Thus in our case we have

$$\begin{aligned}&\frac{2}{|I_n|\mu_n^2}\sum_{j_1<j_2} X_{j_1,j_1+g_1,j_2,j_2+g_2}(n)\\ =\;& \frac{1}{(m+1)^n\left(\frac{nm}{m+1}+\frac{1}{2}\right)^2}\sum_{j_1<j_2} X_{j_1,j_1+g_1,j_2,j_2+g_2}(n)\\ =\;& \left(\frac{1}{(m+1)^n\left(\frac{nm}{m+1}+\frac{1}{2}\right)^2}\right)\Big(n^2(m-\beta_1)(m-\beta_2)(m+1)^{n-\alpha_1-\alpha_2-2}\\ &+\; n^2(m-\beta_1)\beta_2(m+1)^{n-\alpha_1-\alpha_2-3}\\ &+\; n^2(m-\beta_2)\beta_1(m+1)^{n-\alpha_1-\alpha_2-3}\\ &+\; n^2\beta_1\beta_2(m+1)^{n-\alpha_1-\alpha_2-4}+O(n(m+1)^{n-\alpha_1-\alpha_2})\Big)\\ =\;& \left(\frac{1}{(m+1)^{\alpha_1+\alpha_2+2}m^2n^2+O(n)}\right)\Big(n^2(m-\beta_1)(m-\beta_2)(m+1)^2\\ &+\; n^2(m-\beta_1)\beta_2(m+1) \;+\; n^2(m-\beta_2)\beta_1(m+1) \;+\; n^2\beta_1\beta_2\Big)\\ &+\; O\left(\frac{1}{n(m+1)^{\alpha_1+\alpha_2}}\right). \qquad (5.8)\end{aligned}$$

Taking the limit as $n\to\infty$ and rearranging we obtain

$$\frac{m^2(m+1-\beta_1)(m+1-\beta_2)}{m^2(m+1)^{\alpha_1+\alpha_2+2}} \;=\; P(g_1)P(g_2). \qquad (5.9)$$

The fact that the error term decays exponentially in g_1 and g_2 ensures that the error summed over all g_1 and g_2 goes to zero. □

6. Longest Gap

Using the techniques introduced by Bower, Insoft, Li, Miller and Tosteson in [BILMT], we can determine the mean and variance of the distribution of the longest gap between summands in the decomposition of integers in $[a_n, a_{n+1})$ as $n\to\infty$. There are no obstructions to using those methods; however, there are some book-keeping issues due to the nature of our legal m-gonal decomposition. Specifically, we have to worry a little about the residue of the longest gap modulo m. This is a minor issue, as with probability 1 the longest gap is much larger than m and thus we will not have two items in the same bin. Rather than going through the technical argument, we instead give a very short proof that captures the correct main term of the mean of the longest gap, which grows linearly with $\log n$; our error is at the level of the constant term for the mean. We are able to handle the variance similarly, and similarly compute that up to an error that is $O_m(1)$.

Proof of Theorem 1.9. The proof follows immediately from results on the longest run of heads in tosses of a fair coin; we sketch the details below. If a coin has probability p of heads and $q = 1 - p$ of tails, the expected longest run of heads and its variance in n tosses is

$$\mu_n \ = \ \log_{1/p}(nq) - \frac{\gamma}{\log p} - \frac{1}{2} + r_1(n) + \epsilon_1(n), \quad \sigma_n^2 \ = \ \frac{\pi^2}{6\log^2 p} + \frac{1}{12} + r_2(n) + \epsilon_2(n); \quad (6.1)$$

here γ is Euler's constant, the $r_i(n)$ are at most .000016, and the $\epsilon_i(n)$ tend to zero as $n \to \infty$. Note the variance is bounded independently of n (by essentially 3.5); see [Sch] for a proof.

Note that for legal m-gonal decompositions we either have an element in a bin, or we do not. As all decompositions are equally likely, we see that these expansions are equivalent to flipping a coin with probability $1/2$ for each bin, and choosing exactly one of the m possible summands in that bin if we have a tail. As the probability that the longest gap is at the very beginning or very end of a sequence of coin tosses is negligible, we can ignore the fact that the first bin has size 1 and that we may only use part of the last bin if $n+1$ is not a multiple of m. Thus gaps between bins used in a decomposition correspond to strings of consecutive heads.

As our integers lie in $[a_n, a_{n+1})$, we have $\lfloor n/m \rfloor + O(1) = n/m + O(1)$ bins (again, we ignore the presence or absence of the initial bin of length one or a partial bin at the end). We now invoke the results on the length of the longest run for tosses of a fair coin. For us, this translates not to a result on the longest gap between summands, but to a result on the longest number of *bins* between summands. It is trivial to pass from this to our desired result, as all we must do is multiply by m (the error will be at most $O(m)$ coming from the location of where the summands are in the two bins). This completes the proof of our claim on the mean; the variance follows similarly. □

Appendix A. Proof of Theorem 1.5

Proof of Theorem 1.5. Our proof is constructive. We build our sequence by only adjoining terms that ensure that we can *uniquely* decompose a number while never using more than one summand from the same bin. For a fixed $m \geq 1$ the sequence begins:

$$\underbrace{1}_{b_0},\ \underbrace{2,\ 4,\ 6,\ \dots,\ 2m}_{b_1},\ \underbrace{2(m+1),\ 4(m+1),\ 6(m+1),\ \dots, 2m(m+1)}_{b_2},\ \dots \quad \text{(A.1)}$$

Note we would not adjoin 7 because then there would be two legal m-gonal decompositions for 7, one using $7 = 7$ and the other being $7 = 6 + 1$. The next number in the sequence must be the smallest integer which cannot be legally decomposed using the current terms of the sequence.

We can now proceed with proof by induction. Note that the integers $1, 2, 3, \dots, 2m$ have unique decompositions as they are either in the sequence or are the sum of an even number from bin b_1 plus the 1 from bin b_0. The sequence continues:

$$\ldots, \underbrace{a_{m(n-2)+1}, a_{m(n-2)+2}, \ldots, a_{m(n-1)}}_{b_{n-1}}, \underbrace{a_{m(n-1)+1}, \ldots, a_{mn}}_{b_n}, \underbrace{a_{mn+1}, \ldots, a_{m(n+1)}}_{b_{n+1}}, \ldots \quad (A.2)$$

By induction we assume that there exists a unique decomposition for all integers $z \leq a_{mn} + \Omega_{n-1}$, where Ω_{n-1} is the maximum integer that can be legally decomposed using terms in the set $\{a_0, a_1, a_2, a_3, \ldots, a_{m(n-1)}\}$.

By construction we have

$$\Omega_n \;=\; a_{mn} + \Omega_{n-1} \;=\; a_{mn} + a_{m(n-1)+1} - 1.$$

Let x be the maximum integer that can be legally decomposed using terms in the set $\{a_1, a_2, a_3, \ldots, a_{m(n-1)}\}$. Note $x = a_{m(n-1)+1} - 1$ as this is why we include $a_{m(n-1)+1}$ in the sequence.

Claim: $a_{mn+1} = \Omega_n + 1$ and this decomposition is unique.

By induction we know that Ω_n was the largest value that we could legally decompose using only terms in $\{a_0, a_1, a_2, \ldots, a_{mn}\}$. Hence we choose $\Omega_n + 1$ as a_{mn+1} and $\Omega_n + 1$ has a unique decomposition.

Claim: All $N \in [\Omega_n + 1, \Omega_n + 1 + x] = [a_{mn+1}, a_{mn+1} + x]$ have a unique decomposition.

We can legally and uniquely decompose the integers $1, 2, 3, \ldots, x$ using elements in the set $\{a_0, a_1, a_2, \ldots, a_{m(n-1)}\}$. Adding a_{mn+1} to the decomposition of any of these integers would still yield a legal m-gonal decomposition since a_{mn+1} is not in any of the bins $b_0, b_1, b_2, \ldots, b_{n-1}$. The uniqueness of these decompositions follows from the fact that if a_{mn+1} was not included as a summand, then the decomposition does not yield a number big enough to exceed $\Omega_n + 1$.

Claim: $a_{mn+2} = \Omega_n + 1 + x + 1 = a_{mn+1} + x + 1$ and this decomposition is unique.

By construction the largest integer that can be legally decomposed using terms $\{a_0, a_1, a_2, \ldots, a_{mn+1}\}$ is $\Omega_n + 1 + x$.

Claim: All $N \in [a_{mn+2}, a_{mn+2} + x]$ have a unique decomposition.

First note that the decomposition exists as we can legally and uniquely construct $a_{mn+2} + v$, where $0 \leq v \leq x$. For uniqueness, we note that if we do not use a_{mn+2}, then the summation would be too small.

Claim: $a_{mn+2} + x$ is the largest integer that can be legally decomposed using terms $\{a_0, a_1, a_2, \ldots, a_{mn+2}\}$.

This follows from construction. □

References

[Al] H. Alpert, *Differences of multiple Fibonacci numbers*, Integers **9** (2009), 745–749.

[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 Quart. **51** (2013), 13–27.

[B-AM] I. Ben-Ari and S. J. Miller, *A Probabilistic Approach to Generalized Zeckendorf Decompositions*, preprint. http://arxiv.org/pdf/1405.2379.

[BDEMMTTW] A. Best, P. Dynes, X. Edelsbrunner, B. McDonald, S. J. Miller, K. Tor, C. Turnage-Butterbaugh, M. Weinstein, *Gaussian Distribution of Number Summands in Zeckendorf Decompositions in Small Intervals*, preprint.

[BILMT] A. Bower, R. Insoft, S. Li, S. J. Miller and P. Tosteson, *Distribution of Gaps between Summands in Generalized Zeckendorf Decompositions* (and an appendix on *Extensions to Initial Segments* with Iddo Ben-Ari), J. Combin. Theory Ser. A **135** (2015), 130–160.

[CFHMN1] M. Catral, P. Ford, P. E. Harris, S. J. Miller, D. Nelson, *Generalizing Zeckendorf's Theorem: The Kentucky Sequence*, Proc. of the Sixteenth Internat. Conf. on Fibonacci Numbers and Their Appl., Vol. 52, Number 5, pp. 69–91.

[CFHMN2] M. Catral, P. Ford, P. E. Harris, S. J. Miller and D. Nelson, *Legal Decompositions Arising from Non-positive Linear Recurrences*, preprint.

[CFHMN3] M. Catral, P. Ford, P. E. Harris, S. J. Miller and D. Nelson, *New Behavior in Legal Decompositions Arising from Non-positive Linear Recurrences*, preprint.

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

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

[DDKMV] P. Demontigny, T. Do, A. Kulkarni, S. J. Miller and U. Varma, *A Generalization of Fibonacci Far-Difference Representations and Gaussian Behavior*, Fibonacci Quart. **52** (2014), no. 3, 247–273.

[DFFHMPP] R. Dorward, P. Ford, E. Fourakis, P. E. Harris, S. J. Miller, E. Palsson and H. Paugh, *Individual Gap Measures from Generalized Zeckendorf Decompositions* (2015), preprint.

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

[FG] B. E. Fristedt and L. F. Gray, *A modern approach to probability theory*, Birkhäuser, Boston (1996).

[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.

[GT] P. J. Grabner and R. F. Tichy, *Contributions to digit expansions with respect to linear recurrences*, J. Number Theory **36** (1990), no. 2, 160–169.

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

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

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

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

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

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

[KKMW] M. Koloğlu, G. Kopp, S. J. Miller and Y. Wang, *On the number of summands in Zeckendorf decompositions*, Fibonacci Quart. **49** (2011), no. 2, 116–130.

[Kos] T. Koshy, *Fibonacci and Lucas Numbers with Applications*, Wiley-Interscience, New York, 2001.

[MT-B] S. J. Miller and R. Takloo-Bighash, *An Invitation to Modern Number Theory*, Princeton University Press, Princeton, NJ, 2006.

[MW1] S. J. Miller and Y. Wang, *From Fibonacci numbers to Central Limit Type Theorems*, J. Combin. Theory Ser. A **119** (2012), no. 7, 1398–1413.

[MW2] S. J. Miller and 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. .

[Na] M. Nathanson, *Additive Number Theory: The Classical Bases*, Grad. Texts in Math., Springer-Verlag, New York, 1996.

[Sch] M. Schilling, *The longest run of heads*, College Math. J. **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 Quart. **43** (2005), 60–69.

[Ze] E. Zeckendorf, *Représentation des nombres naturels par une somme des nombres de Fibonacci ou de nombres de Lucas*, Bull. Soc. Roy. Sci. Liège **41** (1972), pages 179–182.

E-mail address: rdorward@oberlin.edu

Department of Mathematics, Oberlin College, Oberlin, OH 44074

E-mail address: fordpl@bethanylb.edu

Department of Mathematics and Physics, Bethany College, Lindsborg, KS 67456

E-mail address: erf1@williams.edu

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267

E-mail address: pamela.harris@usma.edu

Department of Mathematical Sciences, United States Military Academy, West Point, NY 10996

E-mail address: sjm1@williams.edu, Steven.Miller.MC.96@aya.yale.edu

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267

E-mail address: eap2@williams.edu

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267

E-mail address: hannah.paugh@usma.edu

Department of Mathematical Sciences, United States Military Academy, West Point, NY 10996