

GENERALIZING ZECKENDORF'S THEOREM TO f -DECOMPOSITIONS

PHILIPPE DEMONTIGNY, THAO DO, ARCHIT KULKARNI, STEVEN J. MILLER, DAVID MOON,
AND UMANG VARMA

ABSTRACT. A beautiful theorem of Zeckendorf states that every positive integer can be uniquely decomposed as a sum of non-consecutive Fibonacci numbers $\{F_n\}$, where $F_1 = 1$, $F_2 = 2$ and $F_{n+1} = F_n + F_{n-1}$. For general recurrences $\{G_n\}$ with non-negative coefficients, there is a notion of a legal decomposition which again leads to a unique representation, and the number of summands in the representations of uniformly randomly chosen $m \in [G_n, G_{n+1})$ converges to a normal distribution as $n \rightarrow \infty$.

We consider the converse question: given a notion of legal decomposition, is it possible to construct a sequence $\{a_n\}$ such that every positive integer can be decomposed as a sum of terms from the sequence? We encode a notion of legal decomposition as a function $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ and say that if a_n is in an “ f -decomposition”, then the decomposition cannot contain the $f(n)$ terms immediately before a_n in the sequence; special choices of f yield many well known decompositions (including base- b , Zeckendorf and factorial). We prove that for any $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$, there exists a sequence $\{a_n\}_{n=0}^\infty$ such that every positive integer has a unique f -decomposition using $\{a_n\}$. Further, if f is periodic, then the unique increasing sequence $\{a_n\}$ that corresponds to f satisfies a linear recurrence relation. Previous research only handled recurrence relations with no negative coefficients. We find a function f that yields a sequence that cannot be described by such a recurrence relation. Finally, for a class of functions f , we prove that the number of summands in the f -decomposition of integers between two consecutive terms of the sequence converges to a normal distribution.

CONTENTS

1. Introduction	2
2. Constructing f -Sequences	4
2.1. Existence and Uniqueness Results	4
2.2. Linear Recurrences and Periodic f	5
3. Radix Representation and the Factorial Number System	6
4. b -Bin Decompositions	8
4.1. Zeckendorf's Theorem for b -Bin Decompositions	8
4.2. Generating Function	9
4.3. Computing The Mean and Variance	11
4.4. Gaussian Behavior	12
5. Conclusion and Future Questions	14
Appendix A. Example of Linear Recurrence	14

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

Key words and phrases. Zeckendorf decompositions, recurrence relations, Stirling numbers of the first kind, Gaussian behavior.

This research was conducted as part of the 2013 SMALL REU program at Williams College and was partially supported funded by NSF grant DMS0850577 and Williams College; the fourth named author was also partially supported by NSF grant DMS1265673. We would like to thank our colleagues from the Williams College 2013 SMALL REU program for helpful discussions, especially Francisc Bozgan, Taylor Corcoran, Joseph R Iafrate, Jaclyn Porfilio and Jirapat Samranvedhya, as well as Kevin O'Bryant for helpful conversations.

A.1. Subsequence $\{a_{0,n}\}$	15
A.2. Common recurrence relation	16
Appendix B. Negative Coefficients in Linear Recurrence	16
References	17

1. INTRODUCTION

The Fibonacci numbers are a very well known sequence, whose properties have fascinated mathematicians for centuries. Zeckendorf [Ze] proved an elegant theorem stating that every positive integer can be written uniquely as the sum of non-consecutive Fibonacci numbers $\{F_n\}$, where¹ $F_1 = 1$, $F_2 = 2$ and $F_n = F_{n-1} + F_{n-2}$. More is true, as the number of summands for integers in $[F_n, F_{n+1})$ converges to a normal distribution as $n \rightarrow \infty$. These results have been generalized to Positive Linear Recurrence Relations of the form

$$G_{n+1} = c_1 G_n + \cdots + c_L G_{n+1-L}, \quad (1.1)$$

where L, c_1, \dots, c_L are non-negative and L, c_1 and c_L are positive. For every such recurrence relation there is a notion of “legal decomposition” with which all positive integers have a unique decomposition as a non-negative integer linear combination of terms from the sequence, and the distribution of the number of summands of integers in $[G_n, G_{n+1})$ converges to a Gaussian. There is an extensive literature for this subject; see [Al, BCCSW, Day, GT, Ha, Ho, Ke, Len, MW1, MW2] for results on uniqueness of decomposition, [DG, FGNPT, GTNP, KKMW, Lek, LamTh, MW1, St] for Gaussian behavior, and [BBGILMT] for recent work on the distribution of gaps between summands.

An alternative definition of the Fibonacci sequence can be framed in terms of the Zeckendorf non-consecutive condition: The Fibonacci sequence (beginning $F_1 = 1$, $F_2 = 2$) is the unique increasing sequence of natural numbers such that every positive integer can be written uniquely as a sum of non-consecutive terms from the sequence. This is a special case of our results (described below). The condition that no two terms in the decomposition may be consecutive is the notion of legal decomposition in the case of Zeckendorf decompositions. In this paper, we encode notions of legal decomposition by a function $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$.

Definition 1.1. *Given a function $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$, a sum $x = \sum_{i=0}^k a_{n_i}$ of terms of $\{a_n\}$ is an f -decomposition of x using $\{a_n\}$ if for every a_{n_i} in the f -decomposition, the previous $f(n_i)$ terms $(a_{n_i-f(n_i)}, a_{n_i-f(n_i)+1}, \dots, a_{n_i-1})$ are not in the f -decomposition.*

We prove the following theorems about f -decompositions.

Theorem 1.2. *For any $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$, there exists a sequence of natural numbers $\{a_n\}$ such that every positive integer has a unique legal f -decomposition in $\{a_n\}$.*

Theorem 1.3. *Let $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ be a function, and let $\{a_n\}$ be a sequence where $a_n = 1$ for $n < 1$ and $a_n = a_{n-1} + a_{n-1-f(n-1)}$ for $n \geq 1$. Then $\{a_n\}_{n=0}^\infty$ is the only increasing sequence of natural numbers in which every positive integer has a unique legal f -decomposition.*

As we study various such sequences, it is useful to give them a name.

¹We don’t start the Fibonacci numbers $F_1 = 1$, $F_2 = 1$, $F_3 = 2$ because doing so would lead to multiple decompositions for some positive integers.

Definition 1.4. Given $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$, the associated f -sequence is the unique increasing sequence of natural numbers $\{a_n\}$ such that every positive integer has a unique f -decomposition.

If we let f be the constant function $f(n) = 1$ for all $n \in \mathbb{N}_0$, we get the Zeckendorf condition² that consecutive terms of the sequence may not be used. Hence the Fibonacci numbers are the f -sequence associated with the constant function $f(n) = 1$ for all $n \in \mathbb{N}_0$.

The Fibonacci sequence is a solution to a recurrence relation. For certain f , we can prove similar connections between f -sequences and linear recurrence relations.

Theorem 1.5. If $f(n)$ is periodic, then the associated f -sequence $\{a_n\}$ is described by a linear recurrence relation.

In Sections 3 and 4 we consider various kinds of f -decompositions and study the distribution of the number of summands in the f -decomposition of integers picked in an interval. We find that these distributions converge in distribution³ to a normal distribution for suitable growing intervals, which we now describe.

Definition 1.6. Let $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ be the function defined as

$$\{f(n)\} = \{0, 0, 1, 0, 1, 2, 0, 1, 2, 3, \dots\}, \quad (1.2)$$

where each “bin” is one term wider than the previous, each bin begins with 0, and f increases by exactly 1 within bins. We say that the f -decomposition of $x \in \mathbb{N}$ for the function f above is the Factorial Number System Representation of the natural number x .

Theorem 1.7. Let the random variable X_n denote the number of summands in the Factorial Number System Representation of an integer picked randomly from $[0, (n+1)!)$ with uniform probability. If we normalize X_n as X'_n , so that X'_n has mean 0 and variance 1, then X'_n converge in distribution to the standard normal distribution as $n \rightarrow \infty$.

The above theorem immediately implies the well-known result that the Stirling numbers of the first kind are asymptotically normally distributed (see Corollary 3.2).

Definition 1.8. Let $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ be a periodic function defined by⁴

$$\{f(n)\} = \{1, 1, 2, \dots, b-1, 1, 1, 2, \dots, b-1, \dots\}. \quad (1.3)$$

Let $\{a_n\}$ be the f -sequence that corresponds to this function f . We say that the f -decomposition of $x \in \mathbb{N}$ for the function f above is the b -bin representation of x .

Theorem 1.9. Let the random variable X_n denote the number of summands in the b -bin representation of an integer picked at random from $[0, a_{bn})$ with uniform probability. Normalize X_n as $Y_n = (X_n - \mu_n)/\sigma_n$, where μ_n and σ_n are the mean and variance of X_n respectively. If $b \geq 3$, Y_n converges in distribution to the standard normal distribution as $n \rightarrow \infty$.

²We say $f(0) = 1$ only for notational convenience. Note that it is redundant as there are no terms before a_0 in the sequence.

³While we work with moment generating functions, since these functions converge well we could multiply the arguments by $i = \sqrt{-1}$ and obtain convergence results about the characteristic functions. By showing the moment generating functions converge pointwise, by the Lévy continuity theorem we obtain convergence in distribution.

⁴Again, $f(0) = 1$ for convenience.

2. CONSTRUCTING f -SEQUENCES

2.1. Existence and Uniqueness Results. Our proof of Theorem 1.2 is constructive, and for any $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ gives a sequence $\{a_n\}$ such that every positive integer has a unique f -decomposition using $\{a_n\}$. This is an analogue to Zeckendorf's Theorem [Ze].

Proof of Theorem 1.2. Let $a_0 = 1$. For $n \geq 1$, define

$$a_n = a_{n-1} + a_{n-1-f(n-1)}, \quad (2.1)$$

where a_n may be assumed to be 1 when $n < 0$. Notice that $\{a_n\}$ is a strictly increasing sequence. We use this definition of the sequence to show that all integers in $[a_m, a_{m+1})$ have an f -decomposition in $\{a_n\}_{n=0}^m$ for all $m \in \mathbb{N}$. We proceed by induction.

The sequence always begins $a_0 = 1, a_1 = 2$. Thus all integers in $[a_0, a_1)$ can be legally decomposed in $\{a_n\}$. For $m > 0$, recall that $a_{m+1} = a_m + a_{m-f(m)}$.

Case I: If $m - f(m) < 0$, we have $a_{m+1} = a_m + 1$. Therefore $[a_m, a_{m+1}) = \{a_m\}$ and a_m has an f -decomposition.

Case II: If $m - f(m) \geq 0$, consider any $x \in [a_m, a_{m+1}) = [a_m, a_m + a_{m-f(m)})$. Therefore $x - a_m \in [0, a_{m-f(m)})$. By the induction hypothesis, $x - a_m$ has an f -decomposition in $\{a_n\}_{n=0}^{m-f(m)-1}$. Since no terms from $\{a_n\}_{n=m-f(m)}^{m-1}$ are in the f -decomposition of $x - a_m$, we may use this f -decomposition and add a_m to get x .

We now prove uniqueness of f -decompositions. Let $S = \sum_{i=1}^k a_{n_i}$ be an f -decomposition with $n_1 > n_2 > \dots > n_k$. We first show that $S \in [a_{n_1}, a_{n_1+1})$. It is clear that $S \geq a_{n_1}$. We show by induction on n_1 that $S < a_{n_1+1}$. If $n_1 = 0$, then $S = a_0 = 1$. If $n_1 \geq 1$, then $S = a_{n_1} + \sum_{i=2}^k a_{n_i}$. By the induction hypothesis, $\sum_{i=2}^k a_{n_i} < a_{n_2+1}$. We know from our notion of f -decomposition that $n_2 \leq n_1 - f(n_1) - 1$. Since $\{a_n\}$ is increasing, $\sum_{i=2}^k a_{n_i} < a_{n_1-f(n_1)}$. This gives us $S = \sum_{i=1}^k a_{n_i} < a_{n_1} + a_{n_1-f(n_1)} = a_{n_1+1}$.

Consider two f -decompositions $x = \sum_{i=1}^k a_{n_i} = \sum_{j=1}^l a_{m_j}$ for the same positive integer x with $n_1 > n_2 > \dots > n_k$ and $m_1 > m_2 > \dots > m_l$. Assume for the sake of contradiction that $\{n_1, n_2, \dots, n_k\} \neq \{m_1, m_2, \dots, m_l\}$. Let h be the smallest natural number such that $n_h \neq m_h$ (it is clear that such an h exists with $h \leq k$ and $h \leq l$). We have $\sum_{i=1}^{h-1} a_{n_i} = \sum_{j=1}^{h-1} a_{m_j}$. Thus $\sum_{i=h}^k a_{n_i} = \sum_{j=h}^l a_{m_j}$. However, as shown above, $\sum_{i=h}^k a_{n_i} \in [a_{n_h}, a_{n_h+1})$ and $\sum_{j=h}^l a_{m_j} \in [a_{m_h}, a_{m_h+1})$, which are disjoint intervals. Therefore $\sum_{i=1}^k a_{n_i} \neq \sum_{j=1}^l a_{m_j}$, a contradiction. \square

Theorem 1.2 gives us a construction for $\{a_n\}$. Theorem 1.3 tells us that $\{a_n\}$ is the only increasing sequence of natural numbers for a given $f(n)$. As mentioned in Definition 1.4, we call this sequence the f -sequence.

Proof of Theorem 1.3. We proceed by induction. Let $\{a'_n\}$ be an increasing sequence such that every positive has a unique f -decomposition using $\{a'_n\}$. Since 1 cannot be written as a sum of other positive integers, we require $a'_0 = a_0 = 1$. Now suppose that $a'_i = a_i$ for each $0 \leq i \leq m-1$. As shown in the proof of Theorem 1.2, each $x \in [0, a_m)$ has a unique f -decomposition in $\{a'_n\}_{n=0}^{m-1}$. Thus, $a'_m \geq a_m$; otherwise, it would not have a unique f -decomposition. On the other hand, if $a'_m > a_m$, then the integer with value a_m does not have an f -decomposition. Thus, $a'_m = a_m$. \square

2.2. Linear Recurrences and Periodic f . Now that we have the unique increasing sequence for a given $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$, we prove Theorem 1.5, which says that this sequence satisfies a linear recurrence relation if f is periodic. The following lemma is a key ingredient in the proof; see for example [LaTa] for a proof.

Lemma 2.1. *Let $a_{1,n}, a_{2,n}, \dots, a_{b,n}$ be linearly recurrent sequences, whose recurrence relations need not be equal. Then the sequence $\{a_n\}_{n=0}^\infty$ constructed by interlacing the sequences as $a_{1,1}, a_{2,1}, \dots, a_{b,1}, a_{1,2}, a_{2,2}, \dots$ is also linearly recurrent.*

To see why this is true, note that if $f_i(x)$ is the characteristic polynomial of the sequence $\{a_{i,n}\}_{n=0}^\infty$, then each subsequence $\{a_{i,n}\}$ satisfies the recurrence relation whose characteristic polynomial is $\prod_{i=1}^b f_i(x)$ and the interlaced recurrence relation satisfies the recurrence relation whose characteristic polynomial is $\prod_{i=1}^b f_i(x^b)$.

Proof of Theorem 1.5. Let p be the period of f and let b the smallest integer multiple of p such that $b \geq f(n) + 1$ for all $n \in \mathbb{N}_0$. When $n \equiv r \pmod{b}$ for some $r \in \mathbb{N}_0$, we have $f(n) = f(r)$ and we have a linear expression (i.e., $a_{n+1} - a_n - a_{n-f(r)} = 0$) for a_{n+1} in terms of the previous b terms. We can write this as a vector of sufficient dimension ($b^2 + 1$ suffices):

$$\vec{v}_0 = [1 \ -1 \ \underbrace{0 \ 0 \ 0 \ \dots \ 0}_{f(r)-1 \text{ times}} \ -1 \ 0 \ \dots]. \quad (2.2)$$

Such a recurrence relation exists for any a_{n-i} where $i \in \mathbb{N}_0$ and $n \equiv r \pmod{b}$. When $i \leq b^2 - b$ (i.e., when there is room in the vector to fit the recurrence relation), we can write the corresponding vector as

$$\vec{v}_i = [\underbrace{0 \ 0 \ \dots \ 0}_i \ 1 \ -1 \ \underbrace{0 \ 0 \ 0 \ 0 \ \dots \ 0}_{f(n-i)-1 \text{ times}} \ -1 \ 0 \ \dots]. \quad (2.3)$$

Our goal is to find a recurrence relation that holds for all $n \equiv r \pmod{b}$ and only uses terms whose indices are from the same residue class modulo b . We begin by finding a recurrence relation for each residue class, not necessarily the same recurrence relation. More precisely, when $n \equiv r \pmod{b}$, we claim that a_n satisfies a recurrence relation of the form

$$a_n = \sum_{i=1}^{b+1} c_i a_{n-bi}. \quad (2.4)$$

Our proof of the above claim is algorithmic. In the proof, we index the coordinates of vectors starting at 0. We define *bad coordinates* to be non-zero coordinates whose indices are not multiples of b . A vector without bad coordinates corresponds to a recurrence relation of the form (2.4).

Let $\vec{w}_0 = \vec{v}_0$. Let \vec{u}_0 be a truncated copy of \vec{w}_0 containing only the coordinates between coordinate 0 and coordinate b , both exclusive. Hence \vec{u}_0 is of dimension $b - 1$. We iteratively find $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_{b-1}$ and $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_{b-1}$ through the following algorithm, looping from $i = 1$ to $i = b - 1$.

- All bad coordinates have index between $(i-1)b$ and ib . Use $\vec{v}_{(i-1)b+1}, \vec{v}_{(i-1)b+2}, \dots, \vec{v}_{ib-1}$ to cancel all coordinates with index from $(i-1)b + 1$ to $ib - 1$ so that they are all zero. This yields the vector \vec{w}_i . All bad coordinates of \vec{w}_i have index between ib and $(i+1)b$ because the degree of the recurrence relation corresponding to vectors $\vec{v}_0, \vec{v}_1, \dots, \vec{v}_{b^2-b}$ are all at most b .
- If \vec{w}_i has zeros in all coordinates with index from $ib + 1$ and $(i+1)b - 1$, we have a recurrence relation of the form (2.4), where the only terms with non-zero coefficients are of the form a_{n-ib} for $i \in \mathbb{N}_0$. We are done. Otherwise, we continue.

- Let u_i be a truncation of \vec{w}_i , containing only coordinates with indices from $ib + 1$ to $(i + 1)b - 1$ (i.e., the bad coordinates), so \vec{w}_i is a vector of dimension $(b - 1)$.

Let $U = \{\vec{u}_0, \vec{u}_1, \dots, \vec{u}_{b-1}\}$. Notice U has b vectors, each of dimension $b - 1$. Since U cannot be a linearly independent set, we have a non-trivial solution to $\lambda_0 \vec{u}_0 + \lambda_1 \vec{u}_1 + \dots + \lambda_{b-1} \vec{u}_{b-1} = \vec{0}$. We want to find a non-trivial linear combination of shifted versions of the vectors $\vec{w}_0, \vec{w}_1, \dots, \vec{w}_{b-1}$ so that it contains no bad coordinates. Let $T_b : \mathbb{R}^{b^2+1} \rightarrow \mathbb{R}^{b^2+1}$ shift the coordinates of a vector to the right by b coordinates. Since each coordinate remains in the same residue class modulo b , the relation corresponding to a vector with shifted coordinates still holds. All the bad coordinates in $T_b^{b-i-1}(\vec{w}_i)$ have index between $b^2 - b + 1$ and $b^2 - 1$. Thus $\sum_{i=0}^{b-1} \lambda_i T_b^{b-i-1}(\vec{w}_i)$ is of the form (2.4) because the bad coordinates cancel to give 0.

Notice that this sum is not the zero vector because the first coordinate of \vec{w}_i is always 1 and for the largest i such that $\lambda_i \neq 0$, the first non-zero coordinate is further left than the first non-zero coordinate of any other vector in the sum (because the other vectors are shifted even further to the right).

For each residue class modulo b , we have a recurrence relation that describes the subsequence of $\{a_n\}$ with indices from that residue class. Thus we can write $\{a_n\}$ as b linearly recurrent sequences interlaced together. By Lemma 2.1, $\{a_n\}$ is also a linearly recurrent sequence. \square

As the above proof is algorithmic, a detailed step-by-step example is provided in Appendix A for illustrative purposes.

3. RADIX REPRESENTATION AND THE FACTORIAL NUMBER SYSTEM

Numerous radix representations can be interpreted as f -decompositions. The most basic are base b representations, which can be interpreted as f -decompositions where $f(n) = n \bmod (b-1)$.

Example 3.1. Consider the following function f :

$$\{f(n)\} = \{0, 1, 2, 3, 0, 1, 2, 3, 0, 1, 2, 3, \dots\}. \quad (3.1)$$

The associated f -sequence is

$$\{a_n\} = \{1, 2, 3, 4, 5, 10, 15, 20, 25, 50, 75, 100, \dots\}. \quad (3.2)$$

An f -decomposition permits at most one summand from each “bin”. In the base 5 representation of any natural number, the i^{th} digit denotes one of $1 \cdot 5^{i-1}$, $2 \cdot 5^{i-1}$, $3 \cdot 5^{i-1}$, $4 \cdot 5^{i-1}$, or 0. This highlights the relationship between these f -decompositions and base 5 representations.

The Factorial Number System is a mixed radix numeral system, where the i^{th} radix corresponds to a place value of $(i - 1)!$ and the digits go from 0 to $i - 1$. Like base b decompositions, this can be interpreted as f -decompositions, where

$$\{a_n\} = \{1, 2, 4, 6, 12, 18, 24, 48, 72, 96, 120, 240, 360, 480, 600\} \quad (3.3)$$

and

$$\{f(n)\} = \{0, 0, 1, 0, 1, 2, 0, 1, 2, 3, 0, 1, 2, 3, 4, \dots\}. \quad (3.4)$$

Here, too, at most one term from each bin may be chosen in an f -decomposition.

Let the random variable X_n denote the number of summands in the Factorial Number System Representation of an integer picked randomly from $[0, (n + 1)!]$ with uniform probability. We now prove the distribution of X_n converges in distribution to a normal distribution.

Proof of Theorem 1.7. We can write X_n as a sum of random variables Y_1, Y_2, \dots, Y_n , where Y_i represents the number of summands from the i^{th} bin. For any integer $i \in [1, n]$, there are an equal number of integers in the interval $[n!, (n+1)!)$ with a given term from the i^{th} bin in the integer's decomposition as there are integers with no terms from the i^{th} bin in its decomposition. Therefore each digit from 0 to i has equal probability in the i^{th} bin, independent of the terms in other bins.

This proves that each Y_i is independent from the other Y_j random variables, and $P(Y_i = 1) = \frac{i-1}{i}$, and $P(Y_i = 0) = \frac{1}{i}$. We apply the Lyapunov Central Limit Theorem (see for example [Bi, pp. 371]) to show that X_n converges in distribution to a Gaussian distribution as $n \rightarrow \infty$; note we cannot apply the standard version as our random variables are not identically distributed.

Let $s_n^2 = \sum_{i=1}^n \sigma_i^2$. The Lyapunov Central Limit Theorem states that if there exists some $\delta > 0$ such that

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^n \mathbb{E}(|Y_i - \mu_i|^{2+\delta}) = 0, \quad (3.5)$$

then the Central Limit Theorem for Y_i holds. That is, $\frac{1}{N} \sum_{i=1}^N Y_i$ converges in distribution to a normal distribution.

We let $\mu_i = \mathbb{E}(Y_i)$ and $\sigma_i^2 = \mathbb{E}(Y_i^2 - \mu_i^2)$. From our definition of Y_i above, we can see

$$\begin{aligned} \mu_i &= \mathbb{E}(Y_i) = \frac{i-1}{i} \\ \sigma_i^2 &= \mathbb{E}(Y_i^2 - \mu_i^2) = \frac{i-1}{i} - \frac{(i-1)^2}{i^2} = \frac{i-1}{i^2}. \end{aligned} \quad (3.6)$$

We show we may take $\delta = 2$.

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^n \mathbb{E}(|Y_i - \mu_i|^{2+\delta}) &= \lim_{n \rightarrow \infty} \frac{1}{\left(\sum_{i=1}^n \frac{i-1}{i^2}\right)^2} \sum_{i=1}^n \left(\frac{i-1}{i} \left(\frac{1}{i}\right)^4 + \frac{1}{i} \left(\frac{i-1}{i}\right)^4 \right) \\ &= \lim_{n \rightarrow \infty} \frac{1}{\left(\sum_{i=1}^n \frac{i-1}{i^2}\right)^2} \sum_{i=1}^n \frac{i^4 - 4i^3 + 6i^2 - 3i}{i^5} \\ &= \lim_{n \rightarrow \infty} \frac{O(\log n)}{\log^2(n) + o(\log^2 n)} = 0. \end{aligned} \quad (3.7)$$

Thus the Lyapunov Central Limit Theorem conditions are met, and X'_n , the normalization of X_n , converges in distribution to the standard normal distribution as $n \rightarrow \infty$. \square

Similar results for base b representations follow trivially from the classical Central Limit Theorem. As a corollary to Theorem 1.7, we can show the following previously known result (see [FS]).

Corollary 3.2. *The unsigned Stirling numbers of the first kind $\begin{bmatrix} n \\ k \end{bmatrix}$ are asymptotically normally distributed.*

Proof. Let $p_{n,k}$ be the number of integers in $[0, (n+1)!)$ whose f -decomposition contains exactly k summands. It is clear that these summands all come from the first n bins. To count the number of ways to choose k summands from n bins, we select k of the n bins. For each bin chosen (say we choose the i^{th} bin), we can choose any of the i elements in the bin.

Let $\mathcal{I} = \{1, 2, \dots, n\}$. We can write $p_{n,k}$ as

$$p_{n,k} = \sum_{S \subset \mathcal{I}, |S|=k} \prod_{s \in S} s. \quad (3.8)$$

Notice that defined in this way, $p_{n,k}$ is the coefficient of x^{n-k+1} in the expansion of $x(x+1)(x+2) \cdots (x+n)$. This allows us to write $p_{n,k}$ as the Stirling number $\begin{bmatrix} n+1 \\ n-k+1 \end{bmatrix}$. Therefore, by Theorem 1.7 the unsigned Stirling numbers of the first kind are asymptotically normally distributed. \square

4. b -BIN DECOMPOSITIONS

4.1. Zeckendorf's Theorem for b -Bin Decompositions. Previous work on generalizing Zeckendorf's Theorem (see for example [MW1, MW2, St]) only handled linear recurrence relations with non-negative coefficients. Not only do some types of f -decompositions have notions of legal decomposition that do not result from any Positive Linear Recurrence Relations [MW1, MW2] (called G -ary digital expansions in [St]), but we can also find functions f that correspond to recurrence relations with some negative coefficients, which is beyond the scope of previous work. In this section we explore a class of f which include some of these new cases.

Let $b \geq 3$ be an integer. We partition non-negative integers as “bins” of b consecutive integers (i.e., $\{0, 1, 2, \dots, b-1\}, \{b, b+1, \dots, 2b-1\}, \dots$).

Definition 4.1 (*b -bin Decomposition*). *Let $b \geq 3$ be an integer. A b -bin decomposition of a positive integer is legal if the following two conditions hold.*

- (1) *No two distinct terms a_i, a_j in a decomposition can have indices i, j from the same bin.*
- (2) *No consecutive terms from the sequence may be in the decomposition.*

We can interpret this as an f -decomposition, where

$$f(n) = \max\{1, n \bmod b\}. \quad (4.1)$$

Example 4.2. *When $b = 3$, the resulting sequence is*

$$\{a_n\} = \underbrace{1, 2, 3, 4, 7, 11, 15, 26, 41, 56, 97, 153, 209, 362, 571, \dots}_{(4.2)}$$

Notice the similarity to base- b representation discussed in Section 3.

We begin by finding a recurrence relation for the sequence $\{a_n\}_{n=0}^{\infty}$ resulting from constant-width bins.

Proposition 4.3. *If $f(n) = \max\{1, n \bmod b\}$ (and thus we have a b -bin decomposition), then the associated f -sequence $\{a_n\}_{n=0}^{\infty}$ satisfies $a_n = (b+1)a_{n-b} - a_{n-2b}$.*

Proof. The proof follows by induction. We prove two base cases and show that the recurrence relation holds for the remaining terms by applying the relation (2.1) that was used to generate the sequence, regrouping terms, and finding the desired recurrence.

We begin by finding the initial $2b+1$ terms of the sequence.

- For $n \leq b$, we have $a_n = n+1$.
- For $n = b+1$, we have $b - f(b) = b-1$. Thus $a_n = a_b + a_{b-1} = 2b+1$.
- For $b+2 \leq n \leq 2b$, we have $n-1 - f(n-1) = b$. Therefore $a_n = a_{n-1} + a_b$, which gives us $a_n = 2b+1 + (n-b-1)(b+1)$.
- For $n = 2b+1$, we have $n-1 - f(n-1) = 2b-1$. Therefore, $a_n = a_{2b} + a_{2b-1} = 2b+1 + (b-1)(b+1) + 2b+1 + (b-2)(b+1) = 2b^2 + 3b - 1$.

First, we prove the base cases and show that the recurrence relation holds for a_{2b} and a_{2b+1} . We have

$$a_{2b} = b^2 + 2b = (b+1)(b+1) - 1 = (b+1)a_{2b-b} - a_{2b-2b} \quad (4.3)$$

and

$$a_{2b+1} = 2b^2 + 3b - 1 = (b+1)(2b+1) - 2 = (b+1)a_{(2b+1)-b} - a_{(2b+1)-2b}. \quad (4.4)$$

For any $n' > 2b+1$, our induction hypothesis states that $a_n = (b+1)a_{n-b} - a_{n-2b}$ for all $n \in [2b, n']$. We use $n' - 1 - f(n' - 1) \geq 2b$ for $n' > 2b+1$.

We constructed the sequence as $a_{n'} = a_{n'-1} + a_{n'-1-f(n'-1)}$. Therefore

$$\begin{aligned} a_{n'} &= ((b+1)a_{n'-1-b} - a_{n'-1-2b}) + ((b+1)a_{n'-1-f(n'-1)-b} - a_{n'-1-f(n'-1)-2b}) \\ &= (b+1)(a_{n'-1-b} + a_{n'-1-f(n'-1)-b}) - (a_{n'-1-2b} + a_{n'-1-f(n'-1)-2b}). \end{aligned} \quad (4.5)$$

As f is periodic, $f(n' - 1) = f(n' - b - 1)$. Thus

$$\begin{aligned} a_{n'} &= (b+1)(a_{n'-b-1} + a_{n'-b-1-f(n'-b-1)}) - (a_{n'-2b-1} + a_{n'-2b-1-f(n'-2b-1)}) \\ &= (b+1)a_{n'-b} - a_{n'-2b}. \end{aligned} \quad (4.6)$$

□

When $b = 3$, the corresponding recurrence relation is $a_n = 4a_{n-3} - a_{n-6}$. In Appendix B we prove that there is no linear recurrence relation with non-negative coefficients that this sequence satisfies. Thus this f -sequence is a new sequence whose behavior cannot be analyzed by previous methods.

While previous techniques cannot handle this sequence, it is still natural to ask if we obtain Gaussian behavior, as this sequence has similar properties to previously studied sequences in terms of uniqueness of decomposition. The answer is yes, and we prove below that the number of summands for integers chosen from $[0, a_{bn})$ converges in distribution to being normally distributed as $n \rightarrow \infty$. In §4.2 we calculate the generating function for the number of summands, then compute the mean and variance in §4.3, and finally prove Gaussianity in §4.4.

4.2. Generating Function. We take a combinatorial approach to finding the distribution of the number of summands for integers in $[0, a_{bn})$. We begin by finding a two-dimensional sequence for the number of integers that can be written as the sum of k summands chosen from the first n bins.

Proposition 4.4. *Let $p_{n,k}$ be the number of integers that are the sum of k summands from n consecutive bins. We have*

$$p_{n,k} = p_{n-1,k} + bp_{n-1,k-1} - p_{n-2,k-2}. \quad (4.7)$$

Proof. We count all possible ways to legally choose k summands from the first n bins. We call the bin containing $\{a_{bn-(b-1)}, a_{bn-(b-2)}, \dots, a_{bn}\}$ the first bin, the bin containing $\{a_{b(n-1)-(b-1)}, \dots, a_{b(n-1)}\}$ the second bin, and so on. We derive a recurrence relation for the $\{p_{n,k}\}$ by counting how many valid choices there are.

In the arguments below we assume $n \geq 2$ so that there are at least two bins. There are two ways to have a contribution to $p_{n,k}$.

- (1) We may choose no summands from the first bin. Therefore we need k summands from $n - 1$ bins and there are exactly $p_{n-1,k}$ ways of doing so.

- (2) We may choose any of the b summands in the first bin, leaving $k - 1$ terms to choose from $n - 1$ bins. There are $bp_{n-1,k-1}$ ways of doing that. However, this argument counts some illegal decompositions. We are not allowed to choose the last element of the first bin and the first element of the second bin. There are exactly $p_{n-2,k-2}$ such decompositions because after choosing these two terms, $k - 2$ terms remain to be chosen from the $n - 2$ remaining bins, and there are no restrictions on our choice from the remaining bins.

Hence for $n \geq 2$ we have the following recurrence relation (in two variables):

$$p_{n,k} = p_{n-1,k} + bp_{n-1,k-1} - p_{n-2,k-2}. \quad (4.8)$$

□

This recurrence relation allows us to compute a closed form expression for $F(x, y)$, the generating function of the $p_{n,k}$'s.

Proposition 4.5. *Let $F(x, y) = \sum_{n,k \geq 0} p_{n,k} x^n y^k$. The closed form expression of $F(x, y)$ is*

$$F(x, y) = \frac{1}{1 - x - bxy + x^2 y^2} \quad (4.9)$$

Proof. We have $p_{0,0} = 1$, $p_{0,k} = 0$ if $k > 0$, and $p_{n,k} = 0$ if $n < 0$. Using the recurrence relation (4.7), after some algebra we find

$$F(x, y) = xF(x, y) + bxyF(x, y) - x^2 y^2 F(x, y) + 1, \quad (4.10)$$

which yields (4.9). □

We now find the coefficient of x^n in $F(x, y)$, which we denote by $g_n(y)$.

Proposition 4.6. *Let $g_n(y) = \sum_{k \geq 0} p_{n,k} y^k$, which is the coefficient of x^n in the generating function of the $p_{n,k}$'s. For $b > 2$, the closed form expression of $g_n(y)$ is*

$$g_n(y) = \frac{\left(by + 1 + \sqrt{(b^2 - 4)y^2 + 2by + 1} \right)^{n+1} - \left(by + 1 - \sqrt{(b^2 - 4)y^2 + 2by + 1} \right)^{n+1}}{2^{n+1} \sqrt{(b^2 - 4)y^2 + 2by + 1}}. \quad (4.11)$$

Proof. Let $x_1(y), x_2(y)$ be the solutions for x of $1 - x - bxy + x^2 y^2 = 0$. It is clear that

$$y^2(x - x_1(y))(x - x_2(y)) = 1 - x - bxy + x^2 y^2. \quad (4.12)$$

It is easy to verify that for all $b > 2$ and non-negative y , $x_1(y)$ and $x_2(y)$ are unequal. Thus

$$F(x, y) = \frac{1}{y^2(x_1(y) - x_2(y))} \left[\frac{1}{x - x_1(y)} - \frac{1}{x - x_2(y)} \right]. \quad (4.13)$$

We use the geometric series formula to expand (4.13), obtaining

$$\begin{aligned} F(x, y) &= \frac{1}{y^2(x_2(y) - x_1(y))} \left[\frac{1/x_1(y)}{1 - x/x_1(y)} - \frac{1/x_2(y)}{1 - x/x_2(y)} \right] \\ &= \frac{1}{y^2(x_2(y) - x_1(y))} \sum_{i=0}^{\infty} \left[\frac{1}{x_1} \left(\frac{x}{x_1} \right)^i - \frac{1}{x_2} \left(\frac{x}{x_2} \right)^i \right]. \end{aligned} \quad (4.14)$$

Using the quadratic formula, we find the roots of the quadratic equation in (4.12) are

$$\begin{aligned} x_1(y) &= \frac{by + 1 - \sqrt{(b^2 - 4)y^2 + 2by + 1}}{2y^2} \\ x_2(y) &= \frac{by + 1 + \sqrt{(b^2 - 4)y^2 + 2by + 1}}{2y^2}. \end{aligned} \quad (4.15)$$

If we write $F(x, y)$ as a power series of x and define $g_n(y)$ to be the coefficient of x^n in $F(x, y)$, then (4.14) implies that

$$g_n(y) = \frac{1}{y^2(x_2(y) - x_1(y))} \left(\frac{x_2(y)^{n+1} - x_1(y)^{n+1}}{(x_1(y)x_2(y))^{n+1}} \right). \quad (4.16)$$

Using (4.15) we find

$$\begin{aligned} g_n(y) &= \frac{y^{2n+2}(x_2(y)^{n+1} - x_1(y)^{n+1})}{\sqrt{5y^2 + 6y + 1}} \\ &= \frac{\left(by + 1 + \sqrt{(b^2 - 4)y^2 + 2by + 1} \right)^{n+1} - \left(by + 1 - \sqrt{(b^2 - 4)y^2 + 2by + 1} \right)^{n+1}}{2^{n+1} \sqrt{(b^2 - 4)y^2 + 2by + 1}}, \end{aligned} \quad (4.17)$$

which completes the proof. \square

4.3. Computing The Mean and Variance. Let X_n be the random variable denoting the number of summands in the unique b -bin decomposition of an integer chosen uniformly from $[0, a_{bn})$. The integers in $[0, a_{bn})$ are exactly those integers whose unique decomposition contains only terms from the first n bins.

Proposition 4.7. *The mean number of summands of b -bin decompositions, μ_n , for integers in $[0, a_{bn})$ is*

$$\mu_n = \frac{(b^2 + b - 4 + b\sqrt{b^2 + 2b - 3})n}{\sqrt{b^2 + 2b - 3}(1 + b + \sqrt{b^2 + 2b - 3})} + O(1). \quad (4.18)$$

Proof. The mean value μ_n of X_n is

$$\mu_n = \sum_{i=0}^n iP(X_n = i) = \sum_{i=0}^n i \frac{p_{n,i}}{\sum_{k=0}^n p_{n,k}} = \frac{g'_n(1)}{g_n(1)}. \quad (4.19)$$

Computing the value of $g'_n(1)/g_n(1)$, we get

$$\frac{g'_n(1)}{g_n(1)} = \frac{4\sqrt{b^2 + 2b - 3} + (b^2 + 2b - 3)(b^2 + b - 4 + b\sqrt{b^2 + 2b - 3})n}{(b^2 + 2b - 3)^{3/2}(1 + b + \sqrt{b^2 + 2b - 3})}, \quad (4.20)$$

which gives us (4.18). \square

Proposition 4.8. *The variance σ_n^2 of X_n is*

$$\sigma_n^2 := \frac{(b^2 + b - 4)n}{(b^2 + 2b - 3)^{3/2}} + O(1). \quad (4.21)$$

Proof. Similar to the computation of the mean, the variance σ_n of X_n can be computed as

$$\sigma_n^2 = \sum_{i=0}^n (i - \mu_n)^2 P(X_n = i) = \sum_{i=0}^n i^2 \frac{p_{n,i}}{\sum_{k=0}^n p_{n,k}} - \mu_n^2 = \frac{\frac{d}{dy} [y g'_n(y)] \Big|_{y=1}}{g(1)} - \mu_n^2. \quad (4.22)$$

Computing (4.22), we get

$$\begin{aligned} \sigma_n^2 = & \left(2(b^6 n + b^5(5 + \sqrt{-3 + 2b + b^2})n + b^4(-2 + 4\sqrt{-3 + 2b + b^2})n \right. \\ & + 4(-9 + \sqrt{-3 + 2b + b^2} + 3(-1 + \sqrt{-3 + 2b + b^2})n) - 2b^3(1 + \sqrt{-3 + 2b + b^2} \\ & + (12 + \sqrt{-3 + 2b + b^2})n) + b(2(9 + 7\sqrt{-3 + 2b + b^2}) + (35 + \sqrt{-3 + 2b + b^2})n) \\ & \left. + b^2(22 - (5 + 16\sqrt{-3 + 2b + b^2})n) \right) / \left((-3 + 2b + b^2)^{5/2} (1 + b + \sqrt{-3 + 2b + b^2})^2 \right), \end{aligned} \quad (4.23)$$

which gives us (4.21). \square

4.4. Gaussian Behavior. In Section 4.2 we found a closed form expression for the generating function $g_n(y)$ of the sequence $p_{n,k}$ for any n . Our expansion of the generating function $g_n(y)$ allows us to explicitly find the moment generating function of X_n , which converges in distribution to a Gaussian and thus proves Theorem 1.9 (that the distribution of the number of summands converges to a Gaussian). Before we prove Theorem 1.9, we need a lemma.

Lemma 4.9. *The moment generating function $M_{Y_n}(t)$ of Y_n is*

$$M_{Y_n}(t) = \mathbb{E}(e^{tY_n}) = \frac{g_n(e^{t/\sigma_n})e^{-t\mu_n/\sigma_n}}{g_n(1)}. \quad (4.24)$$

Proof. Our goal is to study the distribution of X_n as $n \rightarrow \infty$, where

$$P(X_n = k) = \frac{p_{n,k}}{\sum_{k \geq 0} p_{n,k}}. \quad (4.25)$$

Observe that

$$\frac{g_n(e^t)}{g_n(1)} = \sum_{k \geq 0} \frac{p_{n,k} e^{tk}}{\sum_{k \geq 0} p_{n,k}} = \mathbb{E}(e^{tX}). \quad (4.26)$$

If we let μ_n and σ_n be the mean and standard deviation of X_n respectively, then we can normalize the generating function by letting $Y_n = \frac{X_n - \mu_n}{\sigma_n}$, which yields

$$M_{Y_n}(t) = \mathbb{E}(e^{tY_n}) = \sum_{k \geq 0} \frac{p_{n,k} e^{t \frac{(k - \mu_n)}{\sigma_n}}}{\sum_{k \geq 0} p_{n,k}} = \frac{g_n(e^{t/\sigma_n})e^{-t\mu_n/\sigma_n}}{g_n(1)}. \quad (4.27)$$

\square

We now prove Theorem 1.9, which says that Y_n converges in distribution to a normal distribution, by showing that the moment generating function $M_{Y_n}(t)$ of Y_n converges pointwise to the standard normal as $n \rightarrow \infty$ (see for example [Wi]).

Proof of Theorem 1.9. For convenience we set $r := t/\sigma_n$. Since $\sigma_n = c\sqrt{n} + o(\sqrt{n})$, where c is some positive constant, we know that $r \rightarrow 0$ as $n \rightarrow \infty$ for a fixed value of t . Thus we may expand functions of r , such as e^r , using their power series representations. We begin by manipulating $g_n(e^r)$:

$$g_n(e^r) = \frac{(be^r + 1 + \sqrt{(b^2 - 4)e^{2r} + 2be^r + 1})^{n+1} - (be^r + 1 - \sqrt{(b^2 - 4)e^{2r} + 2be^r + 1})^{n+1}}{2^{n+1}\sqrt{(b^2 - 4)e^{2r} + 2be^r + 1}}. \quad (4.28)$$

Let $\delta_1 = (be^r + 1 - \sqrt{(b^2 - 4)e^{2r} + 2be^r + 1})/2$. For large n , we have $e^r = 1 + o(1)$. We can write δ_1 as $(b + 1 - \sqrt{(b + 1)^2 - 4})/2 + o(1)$. Now, let $\delta'_1 = (b + 1 + \sqrt{(b + 1)^2 - 4})/2$. Notice $\delta_1\delta'_1 = 1 + o(1)$. Since $\delta'_1 > 1$, it is clear that $0 < \delta_1 < 1$. Hence δ_1^{n+1} is $o(1)$, and therefore

$$g_n(e^r) = \frac{\left(\frac{b}{2}e^r + \frac{1}{2} + \frac{1}{2}\sqrt{(b^2 - 4)e^{2r} + 2be^r + 1}\right)^{n+1} + o(1)}{\sqrt{(b^2 - 4)e^{2r} + 2be^r + 1}}. \quad (4.29)$$

To focus on individual parts of this equation, we define

$$\beta_1(r) = \sqrt{b^2 + 2be^{-r} + e^{-2r} - 4} \quad (4.30)$$

$$\beta_2(r) = \frac{b}{2} + \frac{e^{-r}}{2} + \frac{\beta_1(r)}{2}. \quad (4.31)$$

We can now write $g_n(y)$ as

$$g_n(e^r) = \frac{e^{rn}\beta_2(r)^{n+1}}{\beta_1(r)} + o(1). \quad (4.32)$$

We are evaluating $M_{Y_n}(t)$ and require the value of $\log(\beta_2(r))$. We can expand β_2 and find

$$\begin{aligned} \beta_2(r) &= \frac{1}{2} \left(\sqrt{b^2 + 2b - 3} + b + 1 \right) + \left(\frac{-b - 1}{2\sqrt{b^2 + 2b - 3}} - \frac{1}{2} \right) r \\ &\quad + \left(\frac{b^3 + 3b^2 - b - 7}{4(b^2 + 2b - 3)^{3/2}} + \frac{1}{4} \right) r^2 + O(r^3) \\ \log(\beta_2(r)) &= \log \left(\frac{1}{2} \left(\sqrt{b^2 + 2b - 3} + b + 1 \right) \right) - \frac{r}{\sqrt{b^2 + 2b - 3}} \\ &\quad + \frac{(b^2 + b - 4)r^2}{2(b^2 + 2b - 3)^{3/2}} + O(r^3). \end{aligned} \quad (4.33)$$

From Lemma 4.9, we have

$$\begin{aligned} M_{Y_n}(t) &= \frac{g_n(e^{t/\sigma_n})e^{-t\mu_n/\sigma_n}}{g_n(1)} \\ \log(M_{Y_n}(t)) &= \log(g_n(e^{t/\sigma_n})) - t\mu_n/\sigma_n - \log(g_n(1)). \end{aligned} \quad (4.34)$$

Because $e^{rn}\beta_2(r)^{n+1}\beta_1(r)^{-1} > 0$, we may move the error term in $g_n(e^r)$ outside the logarithm and simplify:

$$\begin{aligned}\log(M_{Y_n}(t)) &= tn/\sigma_n + (n+1)\log(\beta_2(t/\sigma_n)) - \log(\beta_1(t/\sigma_n)) - t\mu_n/\sigma_n - \log(g_n(1)) + o(1) \\ &= tn/\sigma_n + n\log(\beta_2(t/\sigma_n)) + \log(\beta_2(0) + o(1)) - \log(\beta_1(0) + o(1)) \\ &\quad - t\mu_n/\sigma_n - \log(g_n(1)) + o(1).\end{aligned}\tag{4.35}$$

Plugging in our values of $\beta_1(r)$ from equation (4.30), $\beta_2(r)$ from equation (4.31), μ_n from Proposition 4.7, σ_n from Proposition 4.8, $g_n(1)$ from equation (4.28), and recalling $r = t/\sigma_n$, we get

$$\log(M_{Y_n}(t)) = \frac{t^2}{2} + O\left(n\left(\frac{t}{\sigma_n}\right)^3\right) = \frac{t^2}{2} + o(1).\tag{4.36}$$

The moment generating function for a normal distribution is $e^{t\mu + \frac{1}{2}\sigma^2 t^2}$. Thus, $M_{Y_n}(t)$ pointwise converges to the moment generating function of the standard normal distribution as $n \rightarrow \infty$, which from standard probability machinery implies the densities converge to a standard normal. \square

5. CONCLUSION AND FUTURE QUESTIONS

Encoding notions of legal decomposition as functions provides a new approach to decomposition problems. We were able to generalize Zeckendorf's (and other) theorems to new classes of sequences which were not amenable to previous techniques. Our work leads to further questions that we hope to return to at a later date. These include:

- (1) Our functions f encode notions of legal decomposition where the forbidden terms associated with any a_n are contiguous on the left of a_n . Are there weaker conditions on the notion of legal decomposition under which it is possible to construct a sequence $\{a_n\}$ so that all positive integers have unique decompositions using $\{a_n\}$?
- (2) The distribution of gaps between summands for Zeckendorf decompositions were studied in [BBGILMT]. What is the distribution of gaps (i.e., the difference in indices, $n_j - n_{j+1}$ in decomposition of $x = \sum_{i=1}^k a_{n_i}$, where $\{n_i\}$ is a decreasing sequence) for the Factorial Number System and b -bin decompositions? Can this be studied in general for all f -decompositions?
- (3) The b -bin decompositions are examples of f -decompositions with periodic functions f . Is it true for all periodic f that the number of summands in f -decompositions tend to a normal distribution for integers picked from an appropriate growing interval? Under what conditions on f does Gaussian behavior occur?
- (4) One could attempt to prove results about the number of summands by considering Markov processes where the transition probabilities are related to the f -legal decompositions. Such an approach quickly leads to a concentration result for the number of summands (this follows from standard stationarity results), but not to Gaussian behavior. As this method does not weigh all numbers uniformly, however, we do not pursue those investigations here.

APPENDIX A. EXAMPLE OF LINEAR RECURRENCE

Theorem 1.5 uses an algorithm to prove that the f -sequences associated to periodic functions satisfy linear recurrence relations; we go through that algorithm for an example below.

Consider the case of 3-bin decompositions. We have a periodic function $f : \mathbb{N}_0 \rightarrow \mathbb{N}_0$ defined by

$$f(n) = \begin{cases} 1, & \text{if } n \equiv 0 \pmod{3} \\ 1, & \text{if } n \equiv 1 \pmod{3} \\ 2, & \text{if } n \equiv 2 \pmod{3}. \end{cases} \quad (\text{A.1})$$

The associated f -sequence is

$$\{a_n\}_{n=0}^\infty = \{1, 2, 3, 4, 7, 11, 15, 26, 41, 56, 97, 153, 209, 362, 571, 780, 1351, 2131, 2911, \dots\}. \quad (\text{A.2})$$

The subsequences $\{a_{i,n}\}_{n=0}^\infty = \{a_{3n+i}\}_{n=0}^\infty$ for $i \in \{0, 1, 2\}$ are

$$\{a_{0,n}\} = \{1, 4, 15, 56, 209, 780, 2911, \dots\} \quad (\text{A.3})$$

$$\{a_{1,n}\} = \{2, 7, 26, 97, 362, 1351, 5042, \dots\} \quad (\text{A.4})$$

$$\{a_{2,n}\} = \{3, 11, 41, 153, 571, 2131, 7953, \dots\}. \quad (\text{A.5})$$

A.1. Subsequence $\{a_{0,n}\}$. For any $n \equiv 0 \pmod{3}$, we have

$$\begin{aligned} a_n &= a_{n-1} + a_{n-3} \\ a_{n-1} &= a_{n-2} + a_{n-3} \\ a_{n-3} &= a_{n-4} + a_{n-5}. \end{aligned} \quad (\text{A.6})$$

These expressions follow a periodic pattern. Their vector representation is

$$\begin{array}{l} \vec{v}_0 = \begin{bmatrix} * & & & & & & & & & & \\ 1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \vec{v}_1 = \begin{bmatrix} & * & & & & & & & & & \\ 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \vec{v}_2 = \begin{bmatrix} & & * & & & & & & & & \\ 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \vec{v}_3 = \begin{bmatrix} & & & * & & & & & & & \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \vec{v}_4 = \begin{bmatrix} & & & & * & & & & & & \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \vec{v}_5 = \begin{bmatrix} & & & & & * & & & & & \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 \end{bmatrix}. \end{array} \quad (\text{A.7})$$

Recall from Theorem 1.5 that we 0-index coordinates of vectors in this algorithm (thus the first coordinate has index 0, the second has index 1 and so on). The stars (*) indicate indices that are multiples of $b = 3$. All non-zero coordinates that do not fall under these indices are considered “bad coordinates.”

We begin with $\vec{w}_0 = \vec{v}_0 = [1 \ -1 \ 0 \ -1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$. We use the vectors \vec{u}_i to keep track of the coordinates between those whose indices are multiples of b . In \vec{w}_0 , such coordinates are only between indices 0 and $b = 3$ (exclusive). Thus $\vec{u}_0 = [-1 \ 0]$.

Using other vectors \vec{v}_i , we move the bad coordinates in \vec{w}_0 to the right, so that the indices between 0 and $b = 3$ (exclusive) are all zero. This yields $\vec{w}_1 = \vec{v}_0 + \vec{v}_1 + \vec{v}_2 = [1 \ 0 \ 0 \ -3 \ -1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$. The bad coordinates are only between b and $2b$ (exclusive). Hence $\vec{u}_1 = [-1 \ 0]$.

We move the bad coordinates to the right once more:

$$\vec{w}_2 = \vec{w}_1 + \vec{v}_4 + \vec{v}_5 = [1 \ 0 \ 0 \ -3 \ 0 \ 0 \ -2 \ -1 \ 0 \ 0 \ 0].$$

Here we have $\vec{u}_2 = [-1 \ 0]$.

We have three vectors (\vec{u}_0 , \vec{u}_1 , and \vec{u}_2) of dimension 2. Therefore, there exists a non-trivial solution to $\sum_{i=0}^2 \lambda_i \vec{u}_i = 0$. One such solution is $\lambda_0 = -1$, $\lambda_1 = 1$, $\lambda_2 = 0$.

Shifting the \vec{w}_i vectors using $T_b : \mathbb{R}^{10} \rightarrow \mathbb{R}^{10}$ so that the bad coordinates line up, we obtain

$$\begin{aligned}
T_b^2(\vec{w}_0) &= \begin{bmatrix} * & & & * & & & * & & & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 \end{bmatrix} \\
T_b(\vec{w}_1) &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & -3 & -1 & 0 & 0 \end{bmatrix} \\
\vec{w}_2 &= \begin{bmatrix} 1 & 0 & 0 & -3 & 0 & 0 & -2 & -1 & 0 & 0 \end{bmatrix}.
\end{aligned} \tag{A.8}$$

Now $\sum_{i=0}^b \lambda_i T_b^{b-1-i}(\vec{w}_i) = -T_b^2(\vec{w}_0) + T_b(\vec{w}_1) = [0 \ 0 \ 0 \ 1 \ 0 \ 0 \ -4 \ 0 \ 0 \ 1]$. Thus for all $n \equiv 0 \pmod 3$, we have $a_n = 4a_{n-3} - a_{n-6}$. The subsequence $\{a_{0,n}\}$ satisfies the recurrence relation $a_{0,n} = 4a_{0,n-1} - a_{0,n-2}$.

A.2. Common recurrence relation. Following the same algorithm for $n \equiv 1 \pmod 3$ and $n \equiv 2 \pmod 3$, we find the same relations (i.e., $a_{1,n} = 4a_{1,n-1} - a_{1,n-2}$ and $a_{2,n} = 4a_{2,n-1} - a_{2,n-2}$) as we did for $n \equiv 0 \pmod 3$.

In this case, we do not require Lemma 2.1 to find a common recurrence relation because we have found the same recurrence relation for all subsequences.

Since all subsequences satisfy $s_n = 4s_{n-1} - s_{n-2}$, the interlaced sequence $\{a_n\}$ (i.e., the f -sequence) satisfies $s_n = 4s_{n-3} - s_{n-6}$.

APPENDIX B. NEGATIVE COEFFICIENTS IN LINEAR RECURRENCE

In §4.1 we claimed that the f -sequence associated to 3-bin decompositions does not satisfy any linear homogeneous recurrence relation with constant non-negative coefficients. We now give the proof.

First, we prove that each linearly recurrent sequence has a “minimal linear recurrence relation”. This is a key step in proving that 3-bin decompositions fall outside the scope of previously studied sequences with legal decompositions (the Positive Linear Recurrence Relation Decompositions of [MW1, MW2] or the G -ary Representations in [St]).

Lemma B.1. *For any linearly recurrent sequence $\{a_n\}_{n=0}^\infty$, there exists a linear recurrence relation $\sum_{i=0}^k c_i s_{n-i} = 0$ so that the characteristic polynomial of any linear recurrence relation that $\{a_n\}$ satisfies is multiple of the characteristic polynomial of $\sum_{i=0}^k c_i s_{n-i}$.*

Proof. Let $\{s_n\}$ be a placeholder sequence. Let \mathcal{R} be the set of all recurrence relations that $\{a_n\}$ satisfies. Let $L : \mathcal{R} \rightarrow \mathbb{N}$ be a function, where $L(r)$ is the degree of the recurrence relation r for any $r \in \mathcal{R}$. By the Well-Ordering Principle, $L(\mathcal{R})$ contains its minimum. Let $k = \min L(\mathcal{R})$ and let $\sum_{i=0}^k c_i s_{n-i} = 0$ be a recurrence relation in \mathcal{R} of degree k . Notice $k > 0$.

Consider any recurrence relation $\sum_{i=0}^l p_i s_{n-i}$ in \mathcal{R} . Let $p(x) = \sum_{i=0}^l p_i x^{l-i}$ be the characteristic polynomial of this recurrence relation. Let $c(x) = \sum_{i=0}^k c_i x^{k-i}$ be the characteristic polynomial of $\sum_{i=0}^k c_i s_{n-i} = 0$. By the division algorithm for polynomials, there exist polynomials $q(x), r(x)$, where $r(x)$ has degree less than k , so that $p(x) = q(x)c(x) + r(x)$.

Assume for contradiction that $r(x) \not\equiv 0$. We know $\{a_n\}$ satisfies the recurrence relation whose characteristic polynomial is $p(x)$. The recurrence relation whose characteristic polynomial is $q(x)c(x)$ is a linear combination of index-shifted versions of $\sum_{i=0}^k c_i s_{n-i} = 0$ and hence $\{a_n\}$ satisfies the recurrence relation whose characteristic polynomial is $q(x)c(x)$. Since $\{a_n\}$ satisfies both these recurrence relations, it has to satisfy their difference, whose characteristic polynomial is $r(x)$. However, $r(x)$ has degree less than k , which contradicts that $\min L(\mathcal{R}) = k$. \square

We are now ready to prove our claim about 3-bin decompositions.

Proposition B.2. *The f -sequence associated with 3-bin decompositions satisfies no linear homogeneous recurrence relation with non-negative integer coefficients.*

Proof. The characteristic polynomial of $s_n = 4s_{n-3} - s_{n-6}$ is $x^6 - 4x^3 + 1$, which is irreducible in $\mathbb{Q}[x]$. Hence the minimal linear recurrence relation for the 3-bin sequence (see Appendix A) is $s_n = 4s_{n-3} - s_{n-6}$. We now need to show that no polynomial multiple of $x^6 - 4x^3 + 1$ can be written as $x^h - \sum_{i=1}^h d_i x^{h-i}$ where all d_i are non-negative integers.

Consider any multiple $\sum_{i=0}^{k+6} c_i x^i = (\sum_{j=0}^k p_j x^j)(x^6 - 4x^3 + 1)$, where $p_k \neq 0$. This corresponds to a linear recurrence with non-negative coefficients if and only if $c_i \leq 0$ for all $i < k + 6$ and $c_{k+6} > 0$. Assume for contradiction that $c_{k+6} > 0$ and $c_i \leq 0$ for all $i < k + 6$.

By expanding, we find

$$\sum_{i=0}^{k+6} c_i x^i = \left(\sum_{j=0}^k p_j x^j \right) (x^6 - 4x^3 + 1) = \sum_{i=0}^{k+6} (p_i - 4p_{i-3} + p_{i-6}) x^i \quad (\text{B.1})$$

Let t be the smallest non-negative integer so that $p_t \neq 0$.

We claim that for all $n \in \mathbb{N}$ with $t + 3n < k + 6$, we have $p_{t+3n} \leq 3p_{t+3n-3}$ and $p_{t+3n} < 0$. The proof follows by induction. In the arguments below we frequently use $c_{k+6} > 0$ and $c_i \leq 0$ for all $i < k + 6$ (which we are assuming to show a contradiction follows).

We have $c_t = p_t$ because $p_i = 0$ for all $i < t$. Since $t \leq k$, we know $p_t = c_t < 0$. Hence $p_t \leq p_{t-3}$. Further, $c_{t+3} = p_{t+3} - 4p_t + p_{t-3}$. Since $c_{t+3} \leq 0$ and $p_{t-3} = 0$, we have $p_{t+3} \leq 4p_t < 0$. This proves the base cases $n = 0$ and $n = 1$.

For any n such that $t + 3n < k + 6$, we know $c_{t+3n} \leq 0$. This gives us $p_{t+3n} - 4p_{t+3n-3} + p_{t+3n-6} \leq 0$, from which we have $p_{t+3n} \leq 4p_{t+3n-3} - p_{t+3n-6}$. We know $p_{t+3n-3} - p_{t+3n-6} \leq 0$ by the induction hypothesis and hence $p_{t+3n} \leq 3p_{t+3n-3} < 0$.

By induction, we have $p_{t+3n} \leq 3p_{t+3n-3}$ and $p_{t+3n} < 0$ for all $n \in \mathbb{N}$ with $t + 3n < k + 6$.

Choose n' so that $k < t + 3n' < k + 6$. By the above claim, $p_{t+3n'} < 0$. However, we know $t + 3n' > k$ and $p_i = 0$ for all $i > k$. This is a contradiction. Hence the recurrence relation corresponding to any polynomial multiple of $x^6 - 4x^3 + 1$ has at least one negative coefficient d_i when written as $s_n = \sum_{i=1}^k d_i s_{n-i}$. \square

REFERENCES

- [Al] H. Alpert, *Differences of multiple Fibonacci numbers*, Integers: Electronic Journal of Combinatorial Number Theory **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*, The Fibonacci Quarterly **51** (2013), 13–27.
- [Bi] Billingsley, P. *Probability and Measure*, 2nd ed. New York: Wiley, 1986.
- [BCCSW] E. Burger, D. C. Clyde, C. H. Colbert, G. H. Shin and Z. Wang, *A Generalization of a Theorem of Lekkerkerker to Ostrowski's Decomposition of Natural Numbers*, Acta Arith. **153** (2012), 217–249.
- [Day] D. E. Daykin, *Representation of Natural Numbers as Sums of Generalized Fibonacci Numbers*, J. London Mathematical Society **35** (1960), 143–160.
- [DG] M. Drmota and J. Gajdosik, *The distribution of the sum-of-digits function*, J. Théor. Nombres Bordeaux **10** (1998), no. 1, 17–32.
- [FGNPT] P. Filipponi, P. J. Grabner, I. Nemes, A. Pethő, and R. F. Tichy, *Corrigendum to: "Generalized Zeckendorf expansions"*, Appl. Math. Lett., **7** (1994), no. 6, 25–26.
- [FS] P. Flajolet, and M. Soria, *Gaussian limiting distributions for the number of components in combinatorial structures*, Journal of Combinatorial Theory, Series A, **53** (1990), 165–182.
- [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. Pethő, *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*, Abstracts of Talks, Fourteenth International Conference on Fibonacci Numbers and Their Applications (2010), pages 2–3.
- [Ho] V. E. Hoggatt, *Generalized Zeckendorf theorem*, Fibonacci Quarterly **10** (1972), no. 1 (special issue on representations), pages 89–93.
- [Ke] T. J. Keller, *Generalizations of Zeckendorf’s theorem*, Fibonacci Quarterly **10** (1972), no. 1 (special issue on representations), pages 95–102.
- [KKMW] Kologlu, Kopp, Miller and Wang: *Gaussianity for Fibonacci case*, Fibonacci Quarterly **49** (2011), no. 2, 116–130.
- [Kos] T. Koshy, *Fibonacci and Lucas Numbers with Applications*, Wiley-Interscience, New York, 2001.
- [LamTh] M. Lamberger and J. M. Thuswaldner, *Distribution properties of digital expansions arising from linear recurrences*, Math. Slovaca **53** (2003), no. 1, 1–20.
- [LaTa] R. G. Larson and E. J. Taft, *The algebraic structure of linearly recursive sequences under Hadamard product*, Israel J. Math. **72** (1990), issue 1-2, pages 118–132.
- [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.
- [MW1] S. J. Miller and Y. Wang, *From Fibonacci Numbers to Central Limit Type Theorems*, Journal of Combinatorial Theory, Series A **119** (2012), no. 7, 1398–1413.
- [MW2] S. J. Miller and Y. Wang, *Gaussian Behavior in Generalized Zeckendorf Decompositions*, to appear in the conference proceedings of the 2011 Combinatorial and Additive Number Theory Conference. <http://arxiv.org/abs/1107.2718>.
- [St] W. Steiner, *Parry expansions of polynomial sequences*, Integers **2** (2002), Paper A14.
- [Wi] S. S. Wilks, *Mathematical Statistics*, John Wiley & Sons, New York, 1962.
- [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 Liège **41** (1972), pages 179–182.

E-mail address: Philippe.P.Demontigny@williams.edu

DEPARTMENT OF MATHEMATICS & STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

E-mail address: thao.do@stonybrook.edu

MATHEMATICS DEPARTMENT, STONY BROOK UNIVERSITY, STONY BROOK, NY, 11794

E-mail address: auk@andrew.cmu.edu

DEPARTMENT OF MATHEMATICAL SCIENCES, CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PA 15213

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

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

E-mail address: Dong.Hwan.Moon@williams.edu

DEPARTMENT OF MATHEMATICS & STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

E-mail address: Umang.Varma10@kzoo.edu

DEPARTMENT OF MATHEMATICS & COMPUTER SCIENCE, KALAMAZOO COLLEGE, KALAMAZOO, MI, 49006