GAUSSIAN BEHAVIOR OF THE NUMBER OF SUMMANDS IN ZECKENDORF DECOMPOSITIONS IN SMALL INTERVALS

ANDREW BEST, PATRICK DYNES, XIXI EDELSBRUNNER, BRIAN MCDONALD, STEVEN J. MILLER, KIMSY TOR, CAROLINE TURNAGE-BUTTERBAUGH, AND MADELEINE WEINSTEIN

Abstract. Zeckendorf’s theorem states that every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers $F_n$, with initial terms $F_1 = 1$, $F_2 = 2$. We consider the distribution of the number of summands involved in such decompositions. Previous work proved that as $n \to \infty$ the distribution of the number of summands in the Zeckendorf decompositions of $m \in [F_n, F_{n+1})$, appropriately normalized, converges to the standard normal. The proofs crucially used the fact that all integers in $[F_n, F_{n+1})$ share the same potential summands.

We generalize these results to subintervals of $[F_n, F_{n+1})$ as $n \to \infty$; the analysis is significantly more involved here as different integers have different sets of potential summands. Explicitly, fix an integer sequence $\alpha(n) \to \infty$. As $n \to \infty$, for almost all $m \in [F_n, F_{n+1})$ the distribution of the number of summands in the Zeckendorf decompositions of integers in the subintervals $[m, m + F_{\alpha(n)})$, appropriately normalized, converges to the standard normal. The proof follows by showing that, with probability tending to 1, $m$ has at least one appropriately located large gap between indices in its decomposition. We then use a correspondence between this interval and $[0, F_{\alpha(n)})$ to obtain the result, since the summands are known to have Gaussian behavior in the latter interval. We also prove the same result for more general linear recurrences.

Contents

1. Introduction 47
   1.1. History 47
   1.2. Main Result 48
   2. Preliminaries 49
   3. Proof of Theorem 1.1 51
   4. Conclusion and Future Work 52
   References 52

1. Introduction

1.1. History. Let $\{F_n\}$ denote the Fibonacci numbers, normalized so that $F_1 = 1$, $F_2 = 2^1$, and $F_{n+1} = F_n + F_{n-1}$. An interesting equivalent definition of the Fibonacci numbers is that they are the unique sequence of positive integers such that every positive integer has a unique decomposition into non-consecutive Fibonacci numbers.

This research was conducted as part of the 2014 SMALL REU program at Williams College and was supported by NSF grant DMS1347804 and DMS1265673, Williams College, and the Clare Boothe Luce Program of the Henry Luce Foundation. It is a pleasure to thank the participants of the SMALL REU and the 16th International Conference on Fibonacci Numbers and their Applications for helpful discussions.

1We define the sequence this way to retain uniqueness in our decompositions.
For the Fibonacci numbers, Lekkerkerker [19] proved that the average number of summands needed in the Zeckendorf decomposition of an integer \( m \in [F_n, F_{n+1}) \) is 
\[
\frac{n}{\phi^2 + 1} + O(1),
\]
where \( \phi = \frac{1 + \sqrt{5}}{2} \), the golden mean, is the largest root of the Fibonacci recurrence. This has been extended to other positive linear recurrence sequences, and much more is known. Namely, the distribution of the number of summands converges to a Gaussian as \( n \to \infty \). There are several different methods of proof, from continued fractions to combinatorial perspectives to Markov processes. See [9, 11, 12, 13, 14, 15, 16, 20, 17, 18, 21, 22, 23, 24] for a sampling of results and methods along these lines, [1, 6, 7, 8, 10, 11] for generalizations to other types of representations, and [2, 5] for related questions on the distribution of gaps between summands.

The analysis in much of the previous work was carried out for \( m \in [F_n, F_{n+1}) \). The advantage of such a localization\(^2\) is that each \( m \) has the same candidate set of summands and is of roughly the same size. The purpose of this work is to explore some of the above questions on a significantly smaller scale and determine when and how often we obtain Gaussian behavior.

Note that we cannot expect such behavior to hold for all sub-intervals of \( [F_n, F_{n+1}) \), even if we require the size to grow with \( n \). To see this, consider the interval
\[
[F_{2n} + F_n + F_{n-2} + \cdots + F_{\lceil n/4 \rceil}, F_{2n} + F_{n+1} + F_{\lceil n/4 \rceil}).
\]

The integers in the above interval that are less than \( F_{2n} + F_{n+1} \) have on the order of \( n/2 \) summands, while those that are larger have at most on the order of \( n^{1/4} \) summands. Thus the behavior cannot be Gaussian.\(^3\)

1.2. Main Result.

We first introduce some notation before stating our main result. Fix any increasing positive integer valued function \( \alpha(n) \) with
\[
\lim_{n \to \infty} \alpha(n) = \lim_{n \to \infty} (n - \alpha(n)) = \infty. \tag{1.2}
\]
Fix a non-decreasing positive function \( q(n) < n - \alpha(n) \) taking on even integer values with the restrictions that \( q(n) \to \infty \) and \( q(n) = o(\sqrt{n}) \). For \( m \in [F_n, F_{n+1}) \) with decomposition
\[
m = \sum_{j=1}^{n} a_F j,
\]
define
\[
C_1(m) := (a_1, a_2, \ldots, a_{\alpha(n)}),
\]
\[
C_2(m) := (a_{\alpha(n)+1}, \ldots, a_{\alpha(n)+q(n)}), \quad \text{and}
\]
\[
C_3(m) := (a_{\alpha(n)+q(n)+1}, \ldots, a_n). \tag{1.4}
\]
Note that each \( a_i \in \{0, 1\} \) for all \( 1 \leq i \leq n \). Let \( s(m) \) be the number of summands in the decomposition of \( m \). That is, let
\[
s(m) := \sum_{j=1}^{n} a_j. \tag{1.5}
\]

\(^2\)As the sequence \( \{F_n\} \) is exponentially growing, it is easy to pass from \( m \) in this interval to \( m \in [0, F_n) \).

\(^3\)Though in this situation it would be interesting to investigate separately the behavior on both sides.
GAUSSIAN BEHAVIOR OF ZECKENDORF DECOMPOSITIONS

Similarly, let \( s_1(m), s_2(m), \) and \( s_3(m) \) be the number of summands contributed by \( C_1(m), C_2(m), \) and \( C_3(m) \) respectively. Note that no two consecutive \( a_j \)'s equal 1.

Our main result, given in the following theorem, extends the Gaussian behavior of the number of summands in Zeckendorf decompositions to smaller intervals. Note that requiring \( m \) to be in \([F_n, F_{n+1})\) is not a significant restriction because given any \( m \), there is always an \( n \) such that this holds.

**Theorem 1.1** (Gaussianity on small intervals). For \( \alpha(n) \) satisfying (1.2), the distribution of the number of summands in the decompositions of integers in the interval \([m, m + F_\alpha(n))\) converges to a Gaussian distribution when appropriately normalized for almost all \( m \in [F_n, F_{n+1}). \) Specifically, this means that using the notation from equations (1.3) and (1.4), the Gaussian behavior holds for all \( m \) where there is a gap of length at least 2 in the \( C_2(m) \) (and \( q(n) = o(\sqrt{n}) \) is an increasing even function that diverges to infinity).

2. Preliminaries

In order to prove Theorem 1.1, we establish a correspondence between the decompositions of integers in the interval \([m, m + F_\alpha(n))\) and those in \([0, F_\alpha(n)). \)

**Lemma 2.1.** Let \( x \in [m, m + F_\alpha(n))\). If there are at least two consecutive 0's in \( C_2(m) \), then \( C_3(x) \) is constant, and hence \( s_3(x) \) is constant as well.

**Proof.** Assume there are at least two consecutive 0’s in \( C_2(m) \). Then for some \( k \in [\alpha(n) + 2, \alpha(n) + q(n)) \), we have \( a_{k-1} = a_k = 0 \). Let \( m' \) denote the integer obtained by truncating the decomposition of \( m \) at \( a_{k-2}F_k \). (Note that if \( a_{k-2} = 1 \), we include \( F_{k-2} \) in the truncated decomposition, and if \( a_{k-2} = 0 \) we do not.) Then \( m' < F_{k-1} \). Since \( F_\alpha(n) \leq F_{k-2} \), it follows that for any \( h < F_\alpha(n) \) we have

\[
m' + h < F_{k-1} + F_{k-2} = F_k,
\]

and thus the decomposition of \( m' + h \) has largest summand no greater than \( F_{k-1} \). Therefore, the Zeckendorf decomposition of \( m + h \) is obtained simply by concatenating the decompositions for \( m - m' \) and \( m' + h \). Hence \( C_3(m + h) = C_3(m - m') = C_3(m) \). \( \square \)

With this lemma, we see that the distribution of the number of summands involved in the decomposition of \( x \in [m, m + F_\alpha(n)) \) depends (up to a shift) only on what happens in \( C_1(x) \) and \( C_2(x) \), provided that there is a gap between summands of length at least two somewhere in \( C_2(m) \). In light of this stipulation, we will show the following items in order to prove our main theorem.

- With high probability, \( m \) is of the desired form (i.e., there is a gap between summands of length at least two in \( C_2(m) \)).

- When \( m \) is of the desired form, the distribution of the number of summands involved in \( C_1(x) \) for \( x \in [m, m + F_\alpha(n)) \) converges to Gaussian when appropriately normalized.

- The summands involved in \( C_2(x) \) produce a negligible error term (i.e., there are significantly fewer summands from \( C_2(x) \) than there are from \( C_1(m) \)).

We address the first point with the following lemma.

**Lemma 2.2.** With probability \( 1 + o(1) \), there are at least 2 consecutive 0’s in \( C_2(m) \) if \( m \) is chosen uniformly at random from the integers in \([F_n, F_{n+1}). \)
THE FIBONACCI QUARTERLY

Proof. Suppose $m$ is not of the desired form. Recalling that $q(n)$ takes on even integer values, it follows that either $C_2(m) = (1, 0, 1, 0, \ldots, 1, 0)$ or $C_2(m) = (0, 1, 0, 1, \ldots, 0, 1)$. For each of these two cases, we now count the total number of ways to choose the coefficients for $C_3(m)$ and $C_1(m)$.

In the former case, we have $a_{\alpha(n)+q(n)} = 0$. Thus the number of ways to choose the coefficients for $C_3(m)$ is equal to the number of ways to legally construct

$$\sum_{j=\alpha(n)+q(n)+1}^{n} a_j F_j$$

with no nonzero consecutive coefficients and $a_n = 1$ (since $m \in [F_n, F_{n+1})$ we must select $F_n$). There are $F_n-a_{\alpha(n)}-q(n)-1$ ways to make such a construction, so we conclude that the number of ways to choose the coefficients for $C_3(m)$ is equal to $F_n-a_{\alpha(n)}-q(n)-1$. To see this, we argue as in [2, 5]. By shifting indices, the number of legal constructions here is the same as the number of legal ways to choose the coefficients in

$$\sum_{j=1}^{n-a_{\alpha(n)}-q(n)} \bar{a}_j F_j$$

where we must choose the final summand. By Zeckendorf’s theorem, this is equivalent to counting the number of elements in $[F_{-\alpha(n)}-q(n), F_{-\alpha(n)}-q(n)+1)$, which by the Fibonacci recurrence is just $F_{n-a_{\alpha(n)}-q(n)-1}$. Thus the number of ways to choose the coefficients for $C_3(m)$ is equal to $F_{n-a_{\alpha(n)}-q(n)-1}$. Similarly, since $a_{\alpha(n)} = 0$ the number of ways to choose the coefficients for $C_1(m)$ is equal to $F_{\alpha(n)}$. Thus, if $C_2(m) = (1, 0, 1, 0, \ldots, 1, 0)$, there are $F_{n-3-a_{\alpha(n)}-q(n)} F_{\alpha(n)}$ ways to choose the coefficients for $C_3(m)$ and $C_1(m)$.

A similar counting argument shows that if $C_2(m) = (0, 1, 0, 1, \ldots, 0, 1)$, then the coefficients for $C_3(m)$ and $C_1(m)$ can be chosen in $F_{n-a_{\alpha(n)}-q(n)-2} F_{\alpha(n)+1}$ different ways. Therefore, since $q(n) \to \infty$ as $n \to \infty$, the probability of $m$ not being of the desired form is

$$\frac{F_{n-a_{\alpha(n)}-q(n)-1} F_{\alpha(n)} + F_{n-a_{\alpha(n)}-q(n)-2} F_{\alpha(n)+1}}{F_{n-1}} \sim \frac{2}{\sqrt{5}} q^{-n} = o(1).$$

Assuming $m$ is of the desired form, we now consider the distribution of $s(x)$ for $x \in [m, m + F_{\alpha(n)}]$.

Lemma 2.3. If $m$ has at least 2 consecutive 0’s in $C_2(m)$, then for all $x \in [m, m + F_{\alpha(n)})$, we have

$$0 \leq s(x) - s_3(m) - s(t(x)) < q(n),$$

where $t(x)$ denotes some bijection

$$t : \mathbb{Z} \cap [m, m + F_{\alpha(n)}) \to \mathbb{Z} \cap [0, F_{\alpha(n)}).$$

Proof. First, note that the number of summands in the decomposition of $x$ with indices $i \in [\alpha(n), \alpha(n) + q(n))$ must be less than $q(n)$. Next, let $m_0$ be the sum of the terms in the decomposition of $x$ truncated at $a_{\alpha(n)-1} F_{\alpha(n)-1}$. Define the bijection $t$ by

$$t(m+h) := \begin{cases} m_0 + h, & \text{if } m_0 + h < F_{\alpha(n)} \\ m_0 + h - F_{\alpha(n)} & \text{if } m_0 + h \geq F_{\alpha(n)}. \end{cases}$$

VOLUME 52, NUMBER 5
GAUSSIAN BEHAVIOR OF ZECKENDORF DECOMPOSITIONS

For any \( x \in [m, m + F_{\alpha(n)}] \), the decompositions of \( t(x) \) and \( x \) agree for the terms with index less than \( \alpha(n) \). Furthermore, the decompositions of \( x \) and \( m \) agree for terms with index greater than \( \alpha(n) + q(n) \). Therefore, the number of summands in the decomposition of \( x \) with indices \( i \in [\alpha(n), \alpha(n) + q(n)] \) is equal to \( s(x) - s_3(m) - s(t(x)) \). Combining this with our initial observation, the lemma now follows. \( \square \)

As a result of this lemma, the distribution of \( s(x) \) over the integers in \([m, m + F_{\alpha(n)}]\) is a shift of its distribution over \([0, F_{\alpha(n)}]\), up to an error bounded by \( q(n) \). With this fact, we are now ready to prove the main theorem.

3. Proof of Theorem 1.1

We now prove our main result. The key idea is that with probability approaching 1, we have a gap of length at least 2 in the middle summands of our decompositions, and this allows us to use our bijection to reduce questions on the distribution of the number of summands in \([m, m + F_{\alpha(n)}]\) to similar statements on \([0, F_{\alpha(n)}]\). In doing so, the fluctuations in the difference between the two quantities is bounded by \( q(n) \), which is a free parameter in our splitting of the decomposition, and can therefore be taken to be sufficiently small.

Proof. For a fixed \( m \in [F_n, F_{n+1}] \) with two consecutive 0’s somewhere in \( C_2(m) \), we define random variables \( X_n \) and \( Y_n \) by

\[
X_n := s(X), \quad Y_n := s(Y),
\]

where \( X \) is chosen uniformly at random from \( \mathbb{Z} \cap [m, m + F_{\alpha(n)}] \) and \( Y \) is chosen uniformly at random from \( \mathbb{Z} \cap [0, F_{\alpha(n)}] \). Let

\[
X_n' := \frac{1}{\sigma_x(n)}(X_n - \mathbb{E}[X_n]),
\]

and

\[
Y_n' := \frac{1}{\sigma_y(n)}(Y_n - \mathbb{E}[Y_n]),
\]

where \( \sigma_x(n) \) and \( \sigma_y(n) \) are the standard deviations of \( X_n \) and \( Y_n \), respectively, so that \( X_n \) and \( Y_n \) are normalized with mean 0 and variance 1. It is known that the densities of \( Y_n' \) converge to the density of the standard normal\(^4\), and we claim that \( X_n' \) converges to the standard normal as well. Though we only need the order of magnitude of \( \sigma_y(n) \), for completeness we remark that the mean of \( Y_n \) is \( \frac{n}{\varphi + 2} + O(1) \) and the variance \( \sigma_y(n)^2 \) is \( \frac{\varphi n}{\varphi + 2} + O(1) \), where \( \varphi = \frac{1 + \sqrt{5}}{2} \) is the golden mean.

Let \( f_n \) and \( g_n \) be the cumulative density functions for \( X_n' \) and \( Y_n' \), respectively. By Lemma 2.3, we have

\[
g_n \left( x - \frac{q(n)}{\sigma_y(n)} \right) \leq f_n(x) \leq g_n \left( x + \frac{q(n)}{\sigma_y(n)} \right).
\]

Since \( \sigma_y(n) \to \infty \), we may add the restriction to \( q(n) \) that

\[
q(n) = o(\sigma_y(n)) = o\left( \sqrt{n} \right).
\]

\(^4\)Many of the references give proofs both for the case of the Fibonacci numbers as well as for more general recurrences; see [20] for a simple proof using just Stirling’s formula, which yields that the mean grows on the order of \( \alpha(n) \) and the standard deviation grows on the order of \( \sqrt{\alpha(n)} \).
Since \( \{g_n\}_n \) converges pointwise to the cumulative distribution function for the standard normal, say \( g(x) \), and since

\[
\lim_{n \to \infty} g_n \left( x - \frac{q(n)}{\sigma_y(n)} \right) = \lim_{n \to \infty} g_n \left( x + \frac{q(n)}{\sigma_y(n)} \right) = g(x),
\]

it follows that \( \{f_n\}_n \) also converges pointwise to \( g(x) \). \( \square \)

4. Conclusion and Future Work

We were able to handle the behavior of the number of Zeckendorf summands of numbers drawn from small intervals by finding a correspondence between Zeckendorf decompositions in the interval \( [m, m + F_{\alpha(n)}] \) and in the interval \( [0, F_{\alpha(n)}] \) when a certain gentle condition is placed on the integers \( m \) we consider. The key step was to show that almost surely an integer \( m \) chosen uniformly at random from \( [F_n, F_{n+1}] \) will permit the construction of a bijection onto the interval \( [0, F_{\alpha(n)}] \). Our results follow from previous results on the Gaussian behavior of the number of Zeckendorf summands in this interval.

Our arguments hold for more general recurrence relations (see [4]), though the arguments become more technical. There are two approaches to proving an analogue of the key step, specifically showing that for almost all \( m \) we have a sufficiently large gap in the middle section. One approach is to appeal to some high powered machinery that shows the distribution of the longest gap between summands for \( m \in [F_n, F_{n+1}] \) is strongly concentrated about \( C \log \log n \), where \( C \) is some constant depending on the recurrence. Results along these lines are known for many recurrences; see [3, 5]. Of course, these results contain far more than we need; we do not need to know there is a gap as large as \( C \log \log n \), but rather just that there is a gap a little longer than the length of the recurrence.

References


GAUSSIAN BEHAVIOR OF ZECKENDORF DECOMPOSITIONS


E-mail address: ajb5@williams.edu

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

E-mail address: pdynes@clemson.edu

DEPARTMENT OF MATHEMATICAL SCIENCES, CLEMSON UNIVERSITY, CLEMSON, SC 29634

E-mail address: xe1@williams.edu

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

E-mail address: bmcdon11@u.rochester.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ROCHESTER, ROCHESTER, NY 14627

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: ktor.student@manhattan.edu

DEPARTMENT OF MATHEMATICS, MANHATTAN COLLEGE, RIVERDALE, NY 10471

E-mail address: ctturnagebutterbaugh@gmail.com

DEPARTMENT OF MATHEMATICS, NORTH DAKOTA STATE UNIVERSITY, FARGO, ND 58102

E-mail address: mweinstein@g.hmc.edu

DEPARTMENT OF MATHEMATICS, HARVEY MUDD COLLEGE, CLAREMONT, CA 91711

DECEMBER 2014 53