Distribution of summands in generalized Zeckendorf decompositions

Amanda Bower (UM-Dearborn), Louis Gaudet (Yale University), Rachel Insoft (Wellesley College), Shiyu Li (Berkeley), Steven J. Miller and Philip Tosteson (Williams College)

http://www.williams.edu/Mathematics/sjmiller

AMS Session on Number Theory, I
Room 12, Mezzanine Level, San Diego
Wednesday January 9, 2013, 3:30p.m.
Introduction
Goals of the Talk

- Explain consequences of combinatorial perspective.
- Perspective important: misleading proofs.
- Highlight techniques.
- Some open problems.

Joint work at SMALL (Undergraduate REU Program) at Williams College in 2010, 2011 and 2012.
Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$;

$F_1 = 1$, $F_2 = 2$, $F_3 = 3$, $F_4 = 5$, ...
Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1};$
$F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5, \ldots.$

Zeckendorf’s Theorem
Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.
Previous Results

Fibonacci Numbers: \( F_{n+1} = F_n + F_{n-1} \);

\( F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5, \ldots \).

Zeckendorf’s Theorem

Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.

Example: \( 2013 = 1597 + 377 + 34 + 5 = F_{16} + F_{13} + F_8 + F_4 \).
Previous Results

Fibonacci Numbers: $F_{n+1} = F_n + F_{n-1}$; $F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5, \ldots$.

Zeckendorf’s Theorem

Every positive integer can be written uniquely as a sum of non-consecutive Fibonacci numbers.

Example: $2013 = 1597 + 377 + 34 + 5 = F_{16} + F_{13} + F_8 + F_4$.

Lekkerkerkerker’s Theorem (1952)

The average number of summands in the Zeckendorf decomposition for integers in $[F_n, F_{n+1})$ tends to $\frac{n}{\phi^2 + 1} \approx .276n$, where $\phi = \frac{1+\sqrt{5}}{2}$ is the golden mean.
Old Results

Central Limit Type Theorem

As $n \to \infty$, the distribution of the number of summands in the Zeckendorf decomposition for integers in $[F_n, F_{n+1})$ is Gaussian (normal).

Figure: Number of summands in $[F_{2010}, F_{2011})$; $F_{2010} \approx 10^{420}$. 
New Results: Bulk Gaps: $m \in [F_n, F_{n+1})$ and $\phi = \frac{1 + \sqrt{5}}{2}$

$$m = \sum_{j=1}^{k(m)=n} F_j, \quad \nu_{m;n}(x) = \frac{1}{k(m)-1} \sum_{j=2}^{k(m)} \delta \left( x - (i_j - i_{j-1}) \right).$$

**Theorem (Zeckendorff Gap Distribution)**

Gap measures $\nu_{m;n}$ converge almost surely to average gap measure where $P(k) = \frac{1}{\phi^k}$ for $k \geq 2$.

**Figure:** Distribution of gaps in $[F_{1000}, F_{1001}); F_{2010} \approx 10^{208}$. 
New Results: Longest Gap

**Theorem (Longest Gap)**

As \( n \to \infty \), the probability that \( m \in [F_n, F_{n+1}) \) has longest gap less than or equal to \( f(n) \) converges to

\[
\text{Prob} (L_n(m) \leq f(n)) \approx e^{-e^{\log n - f(n) / \log \phi}}
\]

**Immediate Corollary:** If \( f(n) \) grows slower or faster than \( \log \phi \cdot \log n \), then \( \text{Prob}(L_n(m) \leq f(n)) \) goes to 0 or 1, respectively.
Gaussian Behavior
Reinterpreting the Cookie (or Stars and Bars) Problem

The number of solutions to $x_1 + \cdots + x_P = C$ with $x_i \geq 0$ is $\binom{C+P-1}{P-1}$.
Reinterpreting the Cookie (or Stars and Bars) Problem

The number of solutions to $x_1 + \cdots + x_P = C$ with $x_i \geq 0$ is $\binom{C+P-1}{P-1}$.

Let $p_{n,k} = \# \{ N \in [F_n, F_{n+1}) : \text{the Zeckendorf decomposition of } N \text{ has exactly } k \text{ summands} \}$. 
From The Cookie Problem to Gaussian Behavior

Reinterpreting the Cookie (or Stars and Bars) Problem

The number of solutions to $x_1 + \cdots + x_P = C$ with $x_i \geq 0$ is $inom{C + P - 1}{P - 1}$.

Let $p_{n,k} = \# \{ N \in [F_n, F_{n+1}) : \text{the Zeckendorf decomposition of } N \text{ has exactly } k \text{ summands} \}$.

For $N \in [F_n, F_{n+1})$, the largest summand is $F_n$.

$$N = F_{i_1} + F_{i_2} + \cdots + F_{i_{k-1}} + F_n,$$

$$1 \leq i_1 < i_2 < \cdots < i_{k-1} < i_k = n, \quad i_j - i_{j-1} \geq 2.$$
From The Cookie Problem to Gaussian Behavior

Reinterpreting the Cookie (or Stars and Bars) Problem

The number of solutions to \( x_1 + \cdots + x_P = C \) with \( x_i \geq 0 \) is \( \binom{C + P - 1}{P - 1} \).

Let \( p_{n,k} = \# \{ N \in [F_n, F_{n+1}): \text{the Zeckendorf decomposition of } N \text{ has exactly } k \text{ summands} \} \).

For \( N \in [F_n, F_{n+1}) \), the largest summand is \( F_n \).

\[
N = F_{i_1} + F_{i_2} + \cdots + F_{i_{k-1}} + F_n, \\
1 \leq i_1 < i_2 < \cdots < i_{k-1} < i_k = n, \ i_j - i_{j-1} \geq 2.
\]

\[
d_1 := i_1 - 1, \ d_j := i_j - i_{j-1} - 2 \ (j > 1), \\
d_1 + d_2 + \cdots + d_k = n - 2k + 1, \ d_j \geq 0.
\]
From The Cookie Problem to Gaussian Behavior

Reinterpreting the Cookie (or Stars and Bars) Problem

The number of solutions to $x_1 + \cdots + x_P = C$ with $x_i \geq 0$ is
$${C + P - 1 \choose P - 1}.
$$

Let $p_{n,k} = \# \{ N \in [F_n, F_{n+1}) : \text{the Zeckendorf decomposition of } N \text{ has exactly } k \text{ summands} \}$.

For $N \in [F_n, F_{n+1})$, the largest summand is $F_n$.

$$N = F_{i_1} + F_{i_2} + \cdots + F_{i_{k-1}} + F_n,$$

$$1 \leq i_1 < i_2 < \cdots < i_{k-1} < i_k = n, \ i_j - i_{j-1} \geq 2.$$

$$d_1 := i_1 - 1, \ d_j := i_j - i_{j-1} - 2 \ (j > 1).$$

$$d_1 + d_2 + \cdots + d_k = n - 2k + 1, \ d_j \geq 0.$$

Cookie counting $\Rightarrow p_{n,k} = {n-2k+1 + k-1 \choose k-1} = {n-k \choose k-1}$. 
Generalizations

Generalizing from Fibonacci numbers to linearly recursive sequences with arbitrary nonnegative coefficients.

\[ H_{n+1} = c_1 H_n + c_2 H_{n-1} + \cdots + c_L H_{n-L+1}, \quad n \geq L \]

with \( H_1 = 1, \) \( H_{n+1} = c_1 H_n + c_2 H_{n-1} + \cdots + c_n H_1 + 1, \) \( n < L, \)

coefficients \( c_i \geq 0; \) \( c_1, c_L > 0 \) if \( L \geq 2; \) \( c_1 > 1 \) if \( L = 1. \)

- **Zeckendorf**: Every positive integer can be written uniquely as \( \sum a_i H_i \) with natural constraints on the \( a_i \)'s (e.g. cannot use the recurrence relation to remove any summand).
- **Lekkerkerkerker**
- **Central Limit Type Theorem**
Example: the Special Case of $L = 1$, $c_1 = 10$

\[ H_{n+1} = 10H_n, \ H_1 = 1, \ H_n = 10^{n-1}. \]

- Legal decomposition is decimal expansion: \( \sum_{i=1}^{m} a_i H_i: \)
  \[
  a_i \in \{0, 1, \ldots, 9\} \ (1 \leq i < m), \ a_m \in \{1, \ldots, 9\}.\]

- For \( N \in [H_n, H_{n+1}) \), \( m = n \), i.e., first term is \( a_n H_n = a_n 10^{n-1} \).

- \( A_i \): the corresponding random variable of \( a_i \).
  The \( A_i \)'s are independent.

- For large \( n \), the contribution of \( A_n \) is immaterial.
  \( A_i \ (1 \leq i < n) \) are identically distributed random variables with mean 4.5 and variance 8.25.

- Central Limit Theorem: \( A_2 + A_3 + \cdots + A_n \to \text{Gaussian} \)
  with mean \( 4.5n + O(1) \)
  and variance \( 8.25n + O(1) \).
Far-difference Representation

**Theorem (Alpert, 2009) (Analogue to Zeckendorf)**

Every integer can be written uniquely as a sum of the $\pm F_n$'s, such that every two terms of the same (opposite) sign differ in index by at least 4 (3).

**Example:** $1900 = F_{17} - F_{14} - F_{10} + F_6 + F_2$.

$K$: # of positive terms, $L$: # of negative terms.

**Generalized Lekkerkerkerker’s Theorem**

As $n \to \infty$, $E[K]$ and $E[L] \to n/10$. $E[K] - E[L] = \varphi/2 \approx .809$.

**Central Limit Type Theorem**

As $n \to \infty$, $K$ and $L$ converges to a bivariate Gaussian.

- $\text{corr}(K, L) = -(21 - 2\varphi)/(29 + 2\varphi) \approx -.551$, $\varphi = \frac{\sqrt{5}+1}{2}$.
- $K + L$ and $K - L$ are independent.
Gaps in the Bulk
Distribution of Gaps

For $F_{i_1} + F_{i_2} + \cdots + F_{i_n}$, the gaps are the differences $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \ldots, i_2 - i_1$. 
Distribution of Gaps

For $F_{i_1} + F_{i_2} + \cdots + F_{i_n}$, the gaps are the differences $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \ldots, i_2 - i_1$.

Example: For $F_1 + F_8 + F_{18}$, the gaps are 7 and 10.
Distribution of Gaps

For $F_{i_1} + F_{i_2} + \cdots + F_{i_n}$, the gaps are the differences $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \ldots, i_2 - i_1$.

Example: For $F_1 + F_8 + F_{18}$, the gaps are 7 and 10.

Let $P_n(k)$ be the probability that a gap for a decomposition in $[F_n, F_{n+1})$ is of length $k$. 
For $F_{i_1} + F_{i_2} + \cdots + F_{i_n}$, the gaps are the differences $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \ldots, i_2 - i_1$.

Example: For $F_1 + F_8 + F_{18}$, the gaps are 7 and 10.

Let $P_n(k)$ be the probability that a gap for a decomposition in $[F_n, F_{n+1})$ is of length $k$.

What is $P(k) = \lim_{n \to \infty} P_n(k)$?
Distribution of Gaps

For $F_{i_1} + F_{i_2} + \cdots + F_{i_n}$, the gaps are the differences $i_n - i_{n-1}, i_{n-1} - i_{n-2}, \ldots, i_2 - i_1$.

Example: For $F_1 + F_8 + F_{18}$, the gaps are 7 and 10.

Let $P_n(k)$ be the probability that a gap for a decomposition in $[F_n, F_{n+1})$ is of length $k$.

What is $P(k) = \lim_{n \to \infty} P_n(k)$?

Can ask similar questions about binary or other expansions: $2012 = 2^{10} + 2^9 + 2^8 + 2^7 + 2^6 + 2^4 + 2^3 + 2^2$. 
Main Results

**Theorem (Base $B$ Gap Distribution)**

For base $B$ decompositions, $P(0) = \frac{(B-1)(B-2)}{B^2}$, and for $k \geq 1$, $P(k) = c_B B^{-k}$, with $c_B = \frac{(B-1)(3B-2)}{B^2}$.

**Theorem (Zeckendorf Gap Distribution)**

For Zeckendorf decompositions, $P(k) = \frac{\phi(\phi-1)}{\phi^k}$ for $k \geq 2$, with $\phi = \frac{1+\sqrt{5}}{2}$ the golden mean.
Main Results

• \( H_n: H_{n+1} = c_1 H_n + c_2 H_{n-1} + \cdots + c_L H_{n+1-L} \) a positive linear recurrence of length \( L \) where \( c_i \geq 1 \) for all \( 1 \leq i \leq L \).

• \( \lambda_1 > 1 \): largest root (in absolute value) of characteristic polynomial of \( H_n \).

• Generalized Binet: \( H_n = a_1 \lambda_1^n + \cdots \).

Theorem

Notation as above, probability of a gap of length \( j \) is

\[
\begin{cases} 
1 - \left( \frac{a_1}{C_{Lek}} \right) (\lambda_1^{-n+2} - \lambda_1^{-n+1} + 2\lambda_1^{-1} + a_1^{-1} - 3) & j = 0 \\
\lambda_1^{-1} \left( \frac{1}{C_{Lek}} \right) (\lambda_1 (1 - 2a_1) + a_1) & j = 1 \\
(\lambda_1 - 1)^2 \left( \frac{a_1}{C_{Lek}} \right) \lambda_1^{-j} & j \geq 2
\end{cases}
\]
Proof of Fibonacci Result

Lekkerkerker $\Rightarrow$ total number of gaps $\sim F_{n-1} \frac{n}{\phi^2 + 1}$.

Let $X_{i,j} = \# \{ m \in [F_n, F_{n+1}) : \text{decomposition of } m \text{ includes } F_i, F_j, \text{ but not } F_q \text{ for } i < q < j \}$. 

$$P(k) = \lim_{n \to \infty} \frac{\sum_{i=1}^{n-k} X_{i,i+k}}{F_{n-1} \frac{n}{\phi^2 + 1}}.$$
Proof sketch of almost sure convergence

\[ m = \sum_{j=1}^{k(m)} F_{ij}, \]
\[ \nu_{m,n}(x) = \frac{1}{k(m)-1} \sum_{j=2}^{k(m)} \delta(x - (i_j - i_{j-1})) . \]

\[ \mu_{m,n}(t) = \int x^t d\nu_{m,n}(x). \]

Show \( \mathbb{E}_m[\mu_{m,n}(t)] \) equals average gap moments, \( \mu(t) \).

Show \( \mathbb{E}_m[(\mu_{m,n}(t) - \mu(t))^2] \) and \( \mathbb{E}_m[(\mu_{m,n}(t) - \mu(t))^4] \) tend to zero.

Key ideas: (1) Replace \( k(m) \) with average (Gaussianity); (2) use \( X_{i, i+g_1, j, j+g_2} \).
References
References


