Cookie Monster Meets the Fibonacci Numbers, II. Mmmmmmm – Theorems!

Murat Koloğlu, Gene Kopp, Steven J. Miller and Yinghui Wang
http://www.williams.edu/Mathematics/sjmiller/public_html

CANT 2011
Introduction
Goals of the Talk

- Chat about ‘fun’ properties of Fibonacci numbers.
- Ask the right questions.
- Notice the right perspective.
- Techniques: generating functions, partial fractions.

Thanks to colleagues from the Williams College 2010 SMALL REU program (especially Ed Burger, David Clyde, Cory Colbert, Carlos Dominguez, Gea Shin and Nancy Wang).
Previous Results

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

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

Zecckendorf’s Theorem

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

Example: \( 2011 = 1597 + 377 + 34 + 3 = F_{16} + F_{13} + F_8 + F_3 \).
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:** $2011 = 1597 + 377 + 34 + 3 = F_{16} + F_{13} + F_8 + F_3.$

**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.
New 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 $[F_{2010}, F_{2011})$; $F_{2010} \approx 10^{420}$. 
Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing $C$ identical cookies among $P$ distinct people is $\binom{C + P - 1}{P - 1}$. 
Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing $C$ identical cookies among $P$ distinct people is \( \binom{C+P-1}{P-1} \).

**Proof:** Consider $C + P - 1$ cookies in a line.
### Preliminaries: The Cookie Problem

#### The Cookie Problem

The number of ways of dividing $C$ identical cookies among $P$ distinct people is $\binom{C + P - 1}{P - 1}$.

**Proof:** Consider $C + P - 1$ cookies in a line. **Cookie Monster** eats $P - 1$ cookies:
Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing $C$ identical cookies among $P$ distinct people is $\binom{C+P-1}{P-1}$.

Proof: Consider $C + P - 1$ cookies in a line. Cookie Monster eats $P - 1$ cookies: $\binom{C+P-1}{P-1}$ ways to do.
Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing \( C \) identical cookies among \( P \) distinct people is \( \binom{C+P-1}{P-1} \).

*Proof:* Consider \( C + P - 1 \) cookies in a line. **Cookie Monster** eats \( P - 1 \) cookies: \( \binom{C+P-1}{P-1} \) ways to do. Divides the cookies into \( P \) sets.
Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing \( C \) identical cookies among \( P \) distinct people is \( \binom{C+P-1}{P-1} \).

Proof: Consider \( C + P - 1 \) cookies in a line. **Cookie Monster** eats \( P - 1 \) cookies: \( \binom{C+P-1}{P-1} \) ways to do. Divides the cookies into \( P \) sets.

Example: 8 cookies and 5 people (\( C = 8, P = 5 \)):
Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing $C$ identical cookies among $P$ distinct people is $\binom{C + P - 1}{P - 1}$.

Proof: Consider $C + P - 1$ cookies in a line. **Cookie Monster** eats $P - 1$ cookies: $\binom{C + P - 1}{P - 1}$ ways to do. Divides the cookies into $P$ sets.

Example: 8 cookies and 5 people ($C = 8, P = 5$):
Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing \( C \) identical cookies among \( P \) distinct people is \( \binom{C+P-1}{P-1} \).

Proof: Consider \( C + P - 1 \) cookies in a line. Cookie Monster eats \( P - 1 \) cookies: \( \binom{C+P-1}{P-1} \) ways to do. Divides the cookies into \( P \) sets.

Example: 8 cookies and 5 people (\( C = 8, P = 5 \)): 

- 17
Preliminaries: The Cookie Problem

The Cookie Problem

The number of ways of dividing $C$ identical cookies among $P$ distinct people is $\binom{C + P - 1}{P - 1}$.

Proof: Consider $C + P - 1$ cookies in a line. 
Cookie Monster eats $P - 1$ cookies: $\binom{C + P - 1}{P - 1}$ ways to do. Divides the cookies into $P$ sets.
Example: 8 cookies and 5 people ($C = 8$, $P = 5$):
Preliminaries: The Cookie Problem: Reinterpretation

Reinterpreting the Cookie Problem

The number of solutions to $x_1 + \cdots + x_P = C$ with $x_i \geq 0$ is
\[
\binom{C+P-1}{P-1}.
\]
Preliminaries: The Cookie Problem: Reinterpretation

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

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

$$1 \leq i_1 < i_2 < \cdots < i_{k-1} < i_k = n, \ i_j - i_{j-1} \geq 2.$$
Reinterpreting the Cookie 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 \( F_n \).

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

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

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

\[
d_1 + d_2 + \cdots + d_k = n - 2k + 1, \ d_j \geq 0.
\]
Preliminaries: The Cookie Problem: Reinterpretation

Reinterpreting the Cookie 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 $F_n$.

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

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

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

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

Cookie counting $\Rightarrow p_{n,k} = \binom{n-2k+1-k-1}{k-1} = \binom{n-k}{k-1}$. 
An Erdos-Kac Type Theorem
(note slightly different notation)
Generalizing Lekkerkerker

**Theorem (KKMW 2010)**

As $n \rightarrow \infty$, the distribution of the number of summands in Zeckendorf’s Theorem is a Gaussian.

**Figure:** Number of summands in $[F_{2010}, F_{2011}]$
Generalizing Lekkerkerkerker: Erdos-Kac type result

**Theorem (KKMW 2010)**

As $n \to \infty$, the distribution of the number of summands in Zeckendorf’s Theorem is a Gaussian.

**Sketch of proof:** Use Stirling’s formula,

$$n! \approx n^n e^{-n} \sqrt{2\pi n}$$

to approximate binomial coefficients, after a few pages of algebra find the probabilities are approximately Gaussian.
(Sketch of the) Proof of Gaussianity

The probability density for the number of Fibonacci numbers that add up to an integer in \([F_n, F_{n+1}]\) is

\[ f_n(k) = \left( \frac{n^{-1} - k}{k} \right) \frac{1}{F_n} \]

Consider the density for the \(n + 1\) case. Then we have, by Stirling

\[
f_{n+1}(k) = \binom{n - k}{k} \frac{1}{F_n} = \frac{(n - k)!}{(n - 2k)!k!} \frac{1}{F_n} = \frac{1}{\sqrt{2\pi}} \frac{(n - k)^{n-k+\frac{1}{2}}}{k^{k+\frac{1}{2}}(n - 2k)^{n-2k+\frac{1}{2}}} \frac{1}{F_n}
\]

plus a lower order correction term.

Also we can write \(F_n = \frac{1}{\sqrt{5}} \phi^{n+1} = \frac{\phi}{\sqrt{5}} \phi^n\) for large \(n\), where \(\phi\) is the golden ratio (we are using relabeled Fibonacci numbers where \(1 = F_1\) occurs once to help dealing with uniqueness and \(F_2 = 2\)). We can now split the terms that exponentially depend on \(n\).

\[
f_{n+1}(k) = \left( \frac{1}{\sqrt{2\pi}} \sqrt{\frac{(n - k)}{k(n - 2k)}} \frac{\sqrt{5}}{\phi} \right) \left( \phi^{-n} \frac{(n - k)^{n-k}}{k^k(n - 2k)^{n-2k}} \right).
\]

Define

\[
N_n = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{(n - k)}{k(n - 2k)}} \frac{\sqrt{5}}{\phi}, \quad S_n = \phi^{-n} \frac{(n - k)^{n-k}}{k^k(n - 2k)^{n-2k}}.
\]

Thus, write the density function as

\[ f_{n+1}(k) = N_n S_n \]

where \(N_n\) is the first term that is of order \(n^{-1/2}\) and \(S_n\) is the second term with exponential dependence on \(n\).
(Sketch of the) Proof of Gaussianity (cont)

Model the distribution as centered around the mean by the change of variable \( k = \mu + \chi \sigma \) where \( \mu \) and \( \sigma \) are the mean and the standard deviation, and depend on \( n \). The discrete weights of \( f_n(k) \) will become continuous. This requires us to use the change of variable formula to compensate for the change of scales:

\[
f_n(k)dk = f_n(\mu + \chi \sigma)\sigma dx.
\]

Using the change of variable, we can write \( N_n \) as

\[
N_n = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n-k}{k(n-2k)}} \sqrt{\frac{\phi}{\sqrt{5}}}
\]

\[
= \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1 - k/n}{(k/n)(1 - 2k/n)}} \sqrt{\frac{\phi}{\sqrt{5}}}
\]

\[
= \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1 - \frac{\mu + \sigma x}{n}}{\left(\frac{\mu + \sigma x}{n}\right)(1 - 2\frac{\mu + \sigma x}{n})}} \sqrt{\frac{\phi}{\sqrt{5}}}
\]

\[
= \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1 - C - y}{(C + y)(1 - 2C - 2y)}} \sqrt{\frac{\phi}{\sqrt{5}}}
\]

where \( C = \frac{\mu}{n} \approx \frac{1}{(\phi + 2)} \) (note that \( \phi^2 = \phi + 1 \)) and \( y = \frac{\sigma x}{n} \). But for large \( n \), the \( y \) term vanishes since \( \sigma \sim \sqrt{n} \) and thus \( y \sim n^{-1/2} \). Thus

\[
N_n \approx \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{1 - C}{C(1 - 2C)}} \sqrt{\frac{\phi}{\phi}} = \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{(\phi + 1)(\phi + 2)}{\phi}} \sqrt{\frac{\phi}{\phi}} = \frac{1}{\sqrt{2\pi n}} \sqrt{\frac{5(\phi + 2)}{\phi}} = \frac{1}{\sqrt{2\pi \sigma^2}}
\]

since \( \sigma^2 = n\frac{\phi}{5(\phi + 2)} \).
For the second term $S_n$, take the logarithm and once again change variables by $k = \mu + x\sigma$,

\[
\log(S_n) = \log \left( \phi^{-n} \frac{(n - k)^{(n-k)}}{k^k(n - 2k)^{(n-2k)}} \right) \\
= -n \log(\phi) + (n - k) \log(n - k) - (k) \log(k) \\
- (n - 2k) \log(n - 2k) \\
= -n \log(\phi) + (n - (\mu + x\sigma)) \log(n - (\mu + x\sigma)) \\
- (\mu + x\sigma) \log(\mu + x\sigma) \\
- (n - 2(\mu + x\sigma)) \log(n - 2(\mu + x\sigma)) \\
= -n \log(\phi) \\
+(n - (\mu + x\sigma)) \left( \log(n - \mu) + \log \left( 1 - \frac{x\sigma}{n - \mu} \right) \right) \\
- (\mu + x\sigma) \left( \log(\mu) + \log \left( 1 + \frac{x\sigma}{\mu} \right) \right) \\
- (n - 2(\mu + x\sigma)) \left( \log(n - 2\mu) + \log \left( 1 - \frac{x\sigma}{n - 2\mu} \right) \right) \\
= -n \log(\phi) \\
+(n - (\mu + x\sigma)) \left( \log \left( \frac{n}{\mu} - 1 \right) + \log \left( 1 - \frac{x\sigma}{n - \mu} \right) \right) \\
- (\mu + x\sigma) \log \left( 1 + \frac{x\sigma}{\mu} \right) \\
- (n - 2(\mu + x\sigma)) \left( \log \left( \frac{n}{\mu} - 2 \right) + \log \left( 1 - \frac{x\sigma}{n - 2\mu} \right) \right).
\]
(Sketch of the) Proof of Gaussianity (cont)

Note that, since \( n/\mu = \phi + 2 \) for large \( n \), the constant terms vanish. We have \( \log(S_n) \)

\[
\begin{align*}
&= -n \log(\phi) + (n - k) \log \left( \frac{n}{\mu} - 1 \right) - (n - 2k) \log \left( \frac{n}{\mu} - 2 \right) + (n - (\mu + x\sigma)) \log \left( 1 - \frac{x\sigma}{n - \mu} \right) \\
&\quad - (\mu + x\sigma) \log \left( 1 + \frac{x\sigma}{\mu} \right) - (n - 2(\mu + x\sigma)) \log \left( 1 - \frac{x\sigma}{n - 2\mu} \right) \\
&= -n \log(\phi) + (n - k) \log(\phi + 1) - (n - 2k) \log(\phi) + (n - (\mu + x\sigma)) \log \left( 1 - \frac{x\sigma}{n - \mu} \right) \\
&\quad - (\mu + x\sigma) \log \left( 1 + \frac{x\sigma}{\mu} \right) - (n - 2(\mu + x\sigma)) \log \left( 1 - \frac{x\sigma}{n - 2\mu} \right) \\
&= n(- \log(\phi) + \log(\phi^2) - \log(\phi)) + k(\log(\phi^2) + 2 \log(\phi)) + (n - (\mu + x\sigma)) \log \left( 1 - \frac{x\sigma}{n - \mu} \right) \\
&\quad - (\mu + x\sigma) \log \left( 1 + \frac{x\sigma}{\mu} \right) - (n - 2(\mu + x\sigma)) \log \left( 1 - 2 \frac{x\sigma}{n - 2\mu} \right) \\
&= (n - (\mu + x\sigma)) \log \left( 1 - \frac{x\sigma}{n - \mu} \right) - (\mu + x\sigma) \log \left( 1 + \frac{x\sigma}{\mu} \right) \\
&\quad - (n - 2(\mu + x\sigma)) \log \left( 1 - 2 \frac{x\sigma}{n - 2\mu} \right).
\end{align*}
\]
Finally, we expand the logarithms and collect powers of \( x\sigma / n \).

\[
\log(S_n) = (n - (\mu + x\sigma)) \left( -\frac{x\sigma}{n - \mu} - \frac{1}{2} \left( \frac{x\sigma}{n - \mu} \right)^2 + \ldots \right) \\
- (\mu + x\sigma) \left( \frac{x\sigma}{\mu} - \frac{1}{2} \left( \frac{x\sigma}{\mu} \right)^2 + \ldots \right) \\
- (n - 2(\mu + x\sigma)) \left( -2 \frac{x\sigma}{n - 2\mu} - \frac{1}{2} \left( \frac{2x\sigma}{n - 2\mu} \right)^2 + \ldots \right)
\]

\[
= (n - (\mu + x\sigma)) \left( -\frac{x\sigma}{n(\phi + 1)} - \frac{1}{2} \left( \frac{x\sigma}{n(\phi + 1)} \right)^2 + \ldots \right) \\
- (\mu + x\sigma) \left( \frac{x\sigma}{n(\phi + 2)} - \frac{1}{2} \left( \frac{x\sigma}{n(\phi + 2)} \right)^2 + \ldots \right) \\
- (n - 2(\mu + x\sigma)) \left( -\frac{2x\sigma}{n(\phi + 2)} - \frac{1}{2} \left( \frac{2x\sigma}{n(\phi + 2)} \right)^2 + \ldots \right)
\]

\[
= \frac{x\sigma}{n} \left( -\left(1 - \frac{1}{\phi + 2}\right) \left(\frac{\phi + 2}{\phi + 1}\right) - 1 + 2 \left(1 - \frac{2}{\phi + 2}\right) \frac{\phi + 2}{\phi} \right) \\
- \frac{1}{2} \left( \frac{x\sigma}{n} \right)^2 \left( -2 \frac{\phi + 2}{\phi + 1} + \frac{\phi + 2}{\phi + 1} + 2(\phi + 2) - (\phi + 2) + 4 \frac{\phi + 2}{\phi} \right)
\]

\[+ O \left( n(x\sigma/n)^3 \right) \]
(Sketch of the) Proof of Gaussianity (cont)

\[
\log(S_n) = \frac{x\sigma}{n} \left( -\frac{\phi + 1}{\phi + 2} \frac{\phi + 2}{\phi + 1} - 1 + 2 \frac{\phi}{\phi + 2} \frac{\phi + 2}{\phi} \right) \\
- \frac{1}{2} \left( \frac{x\sigma}{n} \right)^2 n(\phi + 2) \left( - \frac{1}{\phi + 1} + 1 + \frac{4}{\phi} \right) \\
+ O \left( n \left( \frac{x\sigma}{n} \right)^3 \right) \\
= -\frac{1}{2} \left( \frac{x\sigma}{n} \right)^2 (\phi + 2) \left( \frac{3\phi + 4}{\phi(\phi + 1)} + 1 \right) + O \left( n \left( \frac{x\sigma}{n} \right)^3 \right) \\
= -\frac{1}{2} \left( \frac{x\sigma}{n} \right)^2 (\phi + 2) \left( \frac{3\phi + 4 + 2\phi + 1}{\phi(\phi + 1)} \right) + O \left( n \left( \frac{x\sigma}{n} \right)^3 \right) \\
= -\frac{1}{2} x^2 \sigma^2 \left( \frac{5(\phi + 2)}{\phi n} \right) + O \left( n (x\sigma/n)^3 \right).
\]
(Sketch of the) Proof of Gaussianity (cont)

But recall that

$$\sigma^2 = \frac{\phi n}{5(\phi + 2)}.$$ 

Also, since $\sigma \sim n^{-1/2}$, $n \left(\frac{x \sigma}{n}\right)^3 \sim n^{-1/2}$. So for large $n$, the $O\left(n \left(\frac{x \sigma}{n}\right)^3\right)$ term vanishes. Thus we are left with

$$\log S_n = -\frac{1}{2} x^2$$

$$S_n = e^{-\frac{1}{2} x^2}.$$ 

Hence, as $n$ gets large, the density converges to the normal distribution:

$$f_n(k) dk = N_n S_n dk$$

$$= \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{1}{2} x^2} \sigma dx$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2} dx.$$ 

$\square$
| Generalizations
|---|

**Erdos-Kac**

**Method**

**Open Questions**

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

- **Lekkerkerker**

- **Central Limit Type Theorem**
Generalizing Lekkerkerkerker

Generalized Lekkerkerkerker’s Theorem

The average number of summands in the generalized Zeckendorf decomposition for integers in \([H_n, H_{n+1})\) tends to \(Cn + d\) as \(n \to \infty\), where \(C > 0\) and \(d\) are computable constants determined by the \(c_i\)'s.

\[
C = -\frac{y'(1)}{y(1)} = \frac{\sum_{m=0}^{L-1} (s_m + s_{m+1} - 1)(s_{m+1} - s_m)y^m(1)}{2 \sum_{m=0}^{L-1} (m + 1)(s_{m+1} - s_m)y^m(1)}.
\]

\[
s_0 = 0, s_m = c_1 + c_2 + \cdots + c_m.
\]

\(y(x)\) is the root of \(1 - \sum_{m=0}^{L-1} \sum_{j=s_m}^{s_{m+1}-1} x^j y^{m+1}\).

\(y(1)\) is the root of \(1 - c_1 y - c_2 y^2 - \cdots - c_L y^L\).
Central Limit Type Theorem

As \( n \to \infty \), the distribution of the number of summands, i.e., \( a_1 + a_2 + \cdots + a_m \) in the generalized Zeckendorf decomposition \( \sum_{i=1}^{m} a_i H_i \) for integers in \([H_n, H_{n+1})\) is Gaussian.
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 ≤ $i < m$), $a_m \in \{1, \ldots, 9\}$.
- 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 ≤ $i < n$) 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.
Method of General Proof
Generating Function (Example: Binet’s Formula)

**Binet’s Formula**

\[
F_1 = F_2 = 1; \quad F_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^n - \left( \frac{-1+\sqrt{5}}{2} \right)^n \right].
\]
Generating Function (Example: Binet’s Formula)

**Binet’s Formula**

\[ F_1 = F_2 = 1; \quad F_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{-1 + \sqrt{5}}{2} \right)^n \right]. \]

- **Recurrence relation:** \( F_{n+1} = F_n + F_{n-1} \)  
  (1)

- **Generating function:** \( g(x) = \sum_{n>0} F_n x^n. \)

(1) \[ \sum_{n \geq 2} F_{n+1} x^{n+1} = \sum_{n \geq 2} F_n x^{n+1} + \sum_{n \geq 2} F_{n-1} x^{n+1} \]

\[ \Rightarrow \sum_{n \geq 3} F_n x^n = \sum_{n \geq 2} F_n x^{n+1} + \sum_{n \geq 1} F_n x^{n+2} \]

\[ \Rightarrow \sum_{n \geq 3} F_n x^n = x \sum_{n \geq 2} F_n x^n + x^2 \sum_{n \geq 1} F_n x^n \]

\[ \Rightarrow g(x) - F_1 x - F_2 x^2 = x(g(x) - F_1 x) + x^2 g(x) \]

\[ \Rightarrow g(x) = x/(1 - x - x^2). \]
Partial Fraction Expansion (Example: Binet’s Formula)

- **Generating function:** \( g(x) = \sum_{n>0} F_n x^n = \frac{x}{1-x-x^2} \).

- **Partial fraction expansion:**

\[
\Rightarrow g(x) = \frac{x}{1-x-x^2} = \frac{1}{\sqrt{5}} \left( \frac{\frac{1+\sqrt{5}}{2} x}{1 - \frac{1+\sqrt{5}}{2} x} - \frac{-\frac{1+\sqrt{5}}{2} x}{1 - \frac{-1+\sqrt{5}}{2} x} \right).
\]

**Coefficient of** \( x^n \) **(power series expansion):**

\[
F_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1+\sqrt{5}}{2} \right)^n - \left( \frac{-1+\sqrt{5}}{2} \right)^n \right] \quad \text{- Binet’s Formula!}
\]

(Using geometric series: \( \frac{1}{1-r} = 1 + r + r^2 + r^3 + \cdots \).
Differentiating Identities and Method of Moments

- Differentiating identities
  Example: Given a random variable $X$ such that
  $\text{Prob}(X = 1) = \frac{1}{2}$, $\text{Prob}(X = 2) = \frac{1}{4}$, $\text{Prob}(X = 3) = \frac{1}{8}$, ... ,
  then what’s the mean of $X$ (i.e., $E[X]$)?
  
  Solution: Let $f(x) = \frac{1}{2}x + \frac{1}{4}x^2 + \frac{1}{8}x^3 + \cdots = \frac{1}{1-x/2} - 1$.

  \[ f'(x) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4}x + 3 \cdot \frac{1}{8}x^2 + \cdots. \]

  \[ f'(1) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + \cdots = E[X]. \]

- Method of moments: Random variables $X_1$, $X_2$, ... .
  If the $\ell^{\text{th}}$ moment $E[X_\ell^n]$ converges to that of the standard normal distribution ($\forall \ell$), then $X_n$ converges to a Gaussian.

  Standard normal distribution:
  $2m^{\text{th}}$ moment: $(2m - 1)!! = (2m - 1)(2m - 3) \cdots 1$, 
  $(2m - 1)^{\text{th}}$ moment: 0.
New Approach: Case of Fibonacci Numbers

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

- **Recurrence relation:**
  \[ N \in [F_{n+1}, F_{n+2}) : N = F_{n+1} + F_t + \cdots, t \leq n - 1. \]
  \[ p_{n+1,k+1} = p_{n-1,k} + p_{n-2,k} + \cdots \]
New Approach: Case of Fibonacci Numbers

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

- **Recurrence relation:**
  
  $$N \in [F_{n+1}, F_{n+2}): \quad N = F_{n+1} + F_t + \cdots, \ t \leq n - 1.$$ 
  
  $$p_{n+1,k+1} = p_{n-1,k} + p_{n-2,k} + \cdots$$
  
  $$p_{n,k+1} = p_{n-2,k} + p_{n-3,k} + \cdots$$
New Approach: Case of Fibonacci Numbers

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

- **Recurrence relation:**
  \[ N \in [F_{n+1}, F_{n+2}) : N = F_{n+1} + F_t + \cdots, t \leq n - 1. \]
  \[ p_{n+1,k+1} = p_{n-1,k} + p_{n-2,k} + \cdots \]
  \[ p_{n,k+1} = p_{n-2,k} + p_{n-3,k} + \cdots \]
  \[ \Rightarrow p_{n+1,k+1} = p_{n,k+1} + p_{n-1,k}. \]
New Approach: Case of Fibonacci Numbers

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

- **Recurrence relation:**

  \[ N \in [F_{n+1}, F_{n+2}) : N = F_{n+1} + F_t + \cdots, t \leq n - 1. \]

  \[ p_{n+1,k+1} = p_{n-1,k} + p_{n-2,k} + \cdots \]

  \[ p_{n,k+1} = p_{n-2,k} + p_{n-3,k} + \cdots \]

  \[ \Rightarrow p_{n+1,k+1} = p_{n,k+1} + p_{n-1,k}. \]

- **Generating function:**

  \[ \sum_{n,k>0} p_{n,k} x^k y^n = \frac{y}{1 - y - xy^2}. \]

- **Partial fraction expansion:**

  \[ \frac{y}{1 - y - xy^2} = -\frac{y}{y_1(x) - y_2(x)} \left( \frac{1}{y - y_1(x)} - \frac{1}{y - y_2(x)} \right) \]

  where \( y_1(x) \) and \( y_2(x) \) are the roots of \( 1 - y - xy^2 = 0 \).

**Coefficient of \( y^n \):**

\[ g(x) = \sum_{n,k>0} p_{n,k} x^k. \]
New Approach: Case of Fibonacci Numbers (Continued)

$K_n$: the corresponding random variable associated with $k$.

\[ g(x) = \sum_{n,k>0} p_{n,k} x^k. \]

- **Differentiating identities:**
  \[ g(1) = \sum_{n,k>0} p_{n,k} = F_{n+1} - F_n, \]
  \[ g'(x) = \sum_{n,k>0} kp_{n,k} x^{k-1}, \]
  \[ g'(1) = g(1) E[K_n], \]
  \[ (xg'(x))' = \sum_{n,k>0} k^2 p_{n,k} x^{k-1}, \]
  \[ (xg'(x))'|_{x=1} = g(1) E[K_n^2], \]
  \[ (x(xg'(x))')'|_{x=1} = g(1) E[K_n^3], \]

Similar results hold for the centralized $K_n$: $K'_n = K_n - E[K_n]$.

- **Method of moments** (for normalized $K'_n$):
  \[ E[(K'_n)^{2m}] / (SD(K'_n))^{2m} \to (2m - 1)!!, \]
  \[ E[(K'_n)^{2m-1}] / (SD(K'_n))^{2m-1} \to 0. \]

\[ \Rightarrow K_n \to \text{Gaussian}. \]
New Approach: General Case

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

- **Recurrence relation:**
  - **Fibonacci:** \( p_{n+1,k+1} = p_{n,k+1} + p_{n,k} \).
  - **General:** \( p_{n+1,k} = \sum_{m=0}^{L-1} \sum_{j=s_m}^{s_{m+1}-1} p_{n-m,k-j} \).
    where \( s_0 = 0, s_m = c_1 + c_2 + \ldots + c_m \).

- **Generating function:**
  - **Fibonacci:** \( \frac{y}{1-xy-xy^2} \).
  - **General:**
    \[
    \sum_{n \leq L} p_{n,k} x^k y^n - \sum_{m=0}^{L-1} \sum_{j=s_m}^{s_{m+1}-1} x^j y^{m+1} \sum_{n<L-m} p_{n,k} x^k y^n \over \left(1 - \sum_{m=0}^{L-1} \sum_{j=s_m}^{s_{m+1}-1} x^j y^{m+1}\right). \]
New Approach: General Case (Continued)

- Partial fraction expansion:
  Fibonacci: \[- \frac{y}{y_1(x) - y_2(x)} \left( \frac{1}{y - y_1(x)} - \frac{1}{y - y_2(x)} \right) \cdot \]
  General:
  \[- \frac{1}{\sum_{j=s_{L-1}}^{s_L-1} x^j} \sum_{i=1}^{L} \frac{B(x, y)}{(y - y_i(x)) \prod_{j \neq i} (y_j(x) - y_i(x))} \cdot \]

  \[B(x, y) = \sum_{n \leq L} \rho_{n,k} x^k y^n - \sum_{m=0}^{L-1} \sum_{j=s_m}^{L-1, s_{m+1} \leq L} x^j y^{m+1} \sum_{n < L - m} \rho_{n,k} x^k y^n,\]
  \[y_i(x): \text{root of } 1 - \sum_{m=0}^{L-1} \sum_{j=s_m}^{s_{m+1} \leq L} x^j y^{m+1} = 0.\]

Coefficient of \(y^n\): \[g(x) = \sum_{n,k>0} \rho_{n,k} x^k.\]

- Differentiating identities
- Method of moments \(\Rightarrow K_n \rightarrow \text{Gaussian}\)
Future Research
Further Research (hopefully SMALL 2011)

1. Are there similar results for linearly recursive sequences with arbitrary integer coefficients (i.e. negative coefficients are allowed in the defining relation)?

2. Lekkerkerkerkerker's theorem, and the Gaussian extension, are for the behavior in intervals $[F_n, F_{n+1})$. Do the limits exist if we consider other intervals, say $[F_n + g_1(F_n), F_n + g_2(F_n))$ for some functions $g_1$ and $g_2$?

3. For the generalized recurrence relations, what happens if instead of looking at $\sum_{i=1}^n a_i$ we study $\sum_{i=1}^n \min(1, a_i)$? In other words, we only care about how many distinct $H_i$'s occur in the decomposition.

4. What can we say about the distribution of the gaps / largest gap between summands in the Zeckendorf decomposition? Appropriately normalized, how do they behave?
Appendix:
Combinatorial Identities and Lekkerkerker’s Theorem
Needed Binomial Identity

**Binomial identity involving Fibonacci Numbers**

Let $F_m$ denote the $m^{th}$ Fibonacci number, with $F_1 = 1$, $F_2 = 2$, $F_3 = 3$, $F_4 = 5$ and so on. Then

$$\sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} \binom{n-1-k}{k} = F_{n-1}.$$

**Proof by induction:** The base case is trivially verified. Assume our claim holds for $n$ and show that it holds for $n + 1$. We may extend the sum to $n - 1$, as $\binom{n-1-k}{k} = 0$ whenever $k > \left\lfloor \frac{n-1}{2} \right\rfloor$. Using the standard identity that

$$\binom{m}{\ell} + \binom{m}{\ell+1} = \binom{m+1}{\ell+1},$$

and the convention that $\binom{m}{\ell} = 0$ if $\ell$ is a negative integer, we find

$$\sum_{k=0}^{n} \binom{n-k}{k} = \sum_{k=0}^{n} \left[ \binom{n-1-k}{k-1} + \binom{n-1-k}{k} \right]$$

$$= \sum_{k=1}^{n} \binom{n-1-k}{k-1} + \sum_{k=0}^{n} \binom{n-1-k}{k}$$

$$= \sum_{k=1}^{n} \binom{n-2-(k-1)}{k-1} + \sum_{k=0}^{n} \binom{n-1-k}{k} = F_{n-2} + F_{n-1}$$

by the inductive assumption; noting $F_{n-2} + F_{n-1} = F_n$ completes the proof.
Preliminaries for Lekkerkerker’s Theorem

\[
\mathcal{E}(n) := \sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} k \binom{n-1-k}{k}.
\]

Average number of summands in \([F_n, F_{n+1})\) is

\[
\frac{\mathcal{E}(n)}{F_{n-1}} + 1.
\]

Recurrence Relation for \(\mathcal{E}(n)\)

\[
\mathcal{E}(n) + \mathcal{E}(n - 2) = (n - 2)F_{n-3}.
\]
Recurrence Relation

Recurrence Relation for $\mathcal{E}(n)$

$$\mathcal{E}(n) + \mathcal{E}(n - 2) = (n - 2)F_{n-3}.$$ 

Proof by algebra (details later):

$$\begin{align*}
\mathcal{E}(n) &= \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} k \binom{n-1-k}{k} \\
&= (n - 2) \sum_{\ell=0}^{\lfloor \frac{n-3}{2} \rfloor} \binom{n-3-\ell}{\ell} - \sum_{\ell=0}^{\lfloor \frac{n-3}{2} \rfloor} \ell \binom{n-3-\ell}{\ell} \\
&= (n - 2)F_{n-3} - \mathcal{E}(n - 2).
\end{align*}$$
Solving Recurrence Relation

Formula for $\mathcal{E}(n)$ (i.e., Lekkerkerker’s Theorem)

$$\mathcal{E}(n) = \frac{nF_{n-1}}{\phi^2 + 1} + O(F_{n-2}).$$

$$\sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-1)^\ell (\mathcal{E}(n - 2\ell) + \mathcal{E}(n - 2(\ell + 1)))$$

$$= \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} (-1)^\ell (n - 2 - 2\ell)F_{n-3-2\ell}.$$

Result follows from Binet’s formula, the geometric series formula, and differentiating identities: $\sum_{j=0}^{m} jx^j = x^{m+1}x^m(x-1)-(x^{m+1}-1)\over(x-1)^2$. Details later in the appendix.
Derivation of Recurrence Relation for $\mathcal{E}(n)$

\[
\mathcal{E}(n) = \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} k \binom{n-1-k}{k}
\]

\[
= \sum_{k=1}^{\lfloor \frac{n-1}{2} \rfloor} k \frac{(n-1-k)!}{k!(n-1-2k)!}
\]

\[
= \sum_{k=1}^{\lfloor \frac{n-1}{2} \rfloor} (n-1-k) \frac{(n-2-k)!}{(k-1)!(n-1-2k)!}
\]

\[
= \sum_{k=1}^{\lfloor \frac{n-1}{2} \rfloor} (n-2-(k-1)) \frac{(n-3-(k-1))}{(k-1)!(n-3-2(k-1))!}
\]

\[
= \sum_{\ell=0}^{\lfloor \frac{n-3}{2} \rfloor} (n-2-\ell) \binom{n-3-\ell}{\ell}
\]

\[
= (n-2) \sum_{\ell=0}^{\lfloor \frac{n-3}{2} \rfloor} \binom{n-3-\ell}{\ell} - \sum_{\ell=0}^{\lfloor \frac{n-3}{2} \rfloor} \ell \binom{n-3-\ell}{\ell}
\]

\[
= (n-2)F_{n-3} - \mathcal{E}(n-2),
\]

which proves the claim (note we used the binomial identity to replace the sum of binomial coefficients with a Fibonacci number).
Formula for $\mathcal{E}(n)$

\[ \mathcal{E}(n) = \frac{n F_{n-1}}{\varphi^2 + 1} + O(F_{n-2}). \]

**Proof:** The proof follows from using telescoping sums to get an expression for $\mathcal{E}(n)$, which is then evaluated by inputting Binet’s formula and differentiating identities. Explicitly, consider

\[
\begin{align*}
\left\lfloor \frac{n-3}{2} \right\rfloor \sum_{\ell=0} \left( -1 \right)^\ell (\mathcal{E}(n - 2\ell) + \mathcal{E}(n - 2(\ell + 1))) &= \sum_{\ell=0} \left( -1 \right)^\ell (n - 2 - 2\ell) F_{n-3-2\ell} \\
\left\lfloor \frac{n-3}{2} \right\rfloor \sum_{\ell=0} \left( -1 \right)^\ell (n - 3 - 2\ell) F_{n-3-2\ell} + \sum_{\ell=0} \left( -1 \right)^\ell (2\ell) F_{n-3-2\ell} \\
&= \sum_{\ell=0} \left( -1 \right)^\ell (n - 3 - 2\ell) F_{n-3-2\ell} + O(F_{n-2});
\end{align*}
\]

while we could evaluate the last sum exactly, trivially estimating it suffices to obtain the main term (as we have a sum of every other Fibonacci number, the sum is at most the next Fibonacci number after the largest one in our sum).
Formula for $\mathcal{E}(n)$ (continued)

We now use Binet's formula to convert the sum into a geometric series. Letting $\varphi = \frac{1+\sqrt{5}}{2}$ be the golden mean, we have

$$F_n = \frac{\varphi}{\sqrt{5}} \cdot \varphi^n - \frac{1 - \varphi}{\sqrt{5}} \cdot (1 - \varphi)^n$$

(our constants are because our counting has $F_1 = 1$, $F_2 = 2$ and so on). As $|1 - \varphi| < 1$, the error from dropping the $(1 - \varphi)^n$ term is $O(\sum_{\ell \leq n} n) = O(n^2) = o(F_{n-2})$, and may thus safely be absorbed in our error term. We thus find

$$\mathcal{E}(n) = \frac{\varphi}{\sqrt{5}} \sum_{\ell=0}^{\lfloor n/2 \rfloor} (n - 3 - 2\ell)(-1)^\ell \varphi^{n-3-2\ell} + O(F_{n-2})$$

$$= \frac{\varphi^{n-2}}{\sqrt{5}} \left[ (n - 3) \sum_{\ell=0}^{\lfloor n/2 \rfloor} (-\varphi^{-2})^\ell - 2 \sum_{\ell=0}^{\lfloor n/2 \rfloor} \ell(-\varphi^{-2})^\ell \right] + O(F_{n-2}).$$
Formula for $\mathcal{E}(n)$ (continued)

We use the geometric series formula to evaluate the first term. We drop the upper boundary term of $(-\varphi^{-1})^\left\lfloor \frac{n-3}{2} \right\rfloor$, as this term is negligible since $\varphi > 1$. We may also move the 3 from the $n - 3$ into the error term, and are left with

$$
\mathcal{E}(n) = \frac{\varphi^{n-2}}{\sqrt{5}} \left[ \frac{n}{1 + \varphi^{-2}} - 2 \sum_{\ell=0}^{\left\lfloor \frac{n-3}{2} \right\rfloor} \ell (-\varphi^{-2})^\ell \right] + O(F_{n-2})
$$

$$
= \frac{\varphi^{n-2}}{\sqrt{5}} \left[ \frac{n}{1 + \varphi^{-2}} - 2S\left( \left\lfloor \frac{n-3}{2} \right\rfloor, -\varphi^{-2} \right) \right] + O(F_{n-2}),
$$

where

$$
S(m, x) = \sum_{j=0}^{m} j x^j.
$$

There is a simple formula for $S(m, x)$. As

$$
\sum_{j=0}^{m} x^j = \frac{x^{m+1} - 1}{x - 1},
$$

applying the operator $x \frac{d}{dx}$ gives

$$
S(m, x) = \sum_{j=0}^{m} j x^j = x \frac{(m+1)x^m(x-1) - (x^{m+1} - 1)}{(x-1)^2} = \frac{mx^{m+2} - (m+1)x^{m+1} + x}{(x-1)^2}.
$$
Taking $x = -\varphi^{-2}$, we see that the contribution from this piece may safely be absorbed into the error term $O(F_{n-2})$, leaving us with

$$
\mathcal{E}(n) = \frac{n\varphi^{n-2}}{\sqrt{5}(1 + \varphi^{-2})} + O(F_{n-2}) = \frac{n\varphi^n}{\sqrt{5}(\varphi^2 + 1)} + O(F_{n-2}).
$$

Noting that for large $n$ we have $F_{n-1} = \frac{\varphi^n}{\sqrt{5}} + O(1)$, we finally obtain

$$
\mathcal{E}(n) = \frac{nF_{n-1}}{\varphi^2 + 1} + O(F_{n-2}). \quad \square
$$