Distribution of Gaps in PLRS and Phase Transitions

Amanda Bower, Ginny Hogan, Rachel Insoft, Shiyu Li, Philip Tosteson, and Kevin Vissuet;
Advisor: Steven J. Miller (sjm1@williams.edu)
Number Theory and Probability Group - SMALL 2012 - Williams College

1. Distribution of Gaps in Generalized Zeckendorf Decompositions

1.1 Introduction

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: $2012 = 1597 + 377 + 34 + 3 + 1 = F_{16} + F_{13} + F_8 + F_3 + F_1$.

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

We can generalize these theorems to all Positive Linear Recurrence Sequences

Positive Linear Recurrence Sequences: $H_{n+1} = c_1 H_n + c_2 H_{n-1} + \cdots + c_L H_{n-L+1},\ n \geq L$ 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 $L = 1$.

1.2 Our Results

Kangaroo Recurrence: A Kangaroo recurrence of ℓ hops of length g is defined as $K_{n+1} = K_n + K_{n-g} + K_{n-2g} + \cdots + K_{n-\ell g}$.

Theorem: In a Kangaroo Recurrence, the probability of obtaining a gap of length

- $j \geq\ g+1$ is $P(j) = (\lambda_{g,\ell} - 1)^2 \left(\frac{a_1}{C_{\text{Lek}}}\right) \lambda_{g,\ell}^{-j}$.
- $j = g$ is $P(j) = \left(\frac{a_1}{C_{\text{Lek}}}\right) \lambda_{g,\ell}^{-2g}$.

Proof Idea. Let $X_{i,i+j}(n) = \#\{m \in [K_n, K_{n+1})$: decomposition of m includes K_i, K_{i+j}, but not K_q for $i < q < i+j\}$.

Let $Y(n) =$ total number of gaps in decompositions for integers in $[K_n, K_{n+1})$.

$$P(j) = \lim_{n \to \infty} \frac{1}{Y(n)} \sum_{i=1}^{n-j} X_{i,i+j}(n).$$

Generalized Lekkerkerker $\Rightarrow\ Y(n) \sim (C_{\text{Lek}} n + d)(K_{n+1} - K_n)$.

We can calculate $X_{i,i+j}(n) = \text{Left}*\text{Right} = (K_{i+1} - K_i)(K_{n-i-j+2} - K_{n-i-j+1} - (K_{n-i-j+1} - K_{n-i-j}))$.

1.3 Generalized Results

Theorem: Let $H_{n+1} = c_1 H_n + c_2 H_{n-1} + \cdots + c_L H_{n+1-L}$ be a positive linear recurrence of length L where $c_i \geq 1$ for all $1 \leq i \leq L$. Then

$$P(j) = \begin{cases} 1 - (\frac{a_1}{C_{\text{Lek}}})(\lambda_1^{-n+2} - \lambda_1^{-n+1} + 2\lambda_1^{-1} + a_1^{-1} - 3) & \text{for } j = 0 \\ \lambda_1^{-1}(\frac{1}{C_{\text{Lek}}})(\lambda_1(1 - 2a_1) + a_1) & \text{for } j = 1 \\ (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{\text{Lek}}}\right) \lambda_1^{-j} & \text{for } j \geq 2. \end{cases}$$

2. Longest Gaps in Zeckendorf Decompostions

2.1 Introduction

Gaps: If $x \in [F_n, F_{n+1})$ has Zeckendorf decomposition $x = F_n + F_{n-g_1} + F_{n-g_2} + \cdots + F_{n-g_k}$, we define the *gaps* in its decomposition to be $\{g_1, g_1 - g_2, \cdots, g_{k-1} - g_k\}$. The longest gap of a Zeckendorf decomposition is the gap that is greatest in terms of the measure of length.

2.2 Results

Cumulative Distribution Function in Fibbonaccis Pick x randomly from the interval $[F_n, F_{n+1})$. We prove explicitly the cumulative distribution of x's longest gap.

Theorem: Let $r = \phi^2/(\phi^2 + 1)$ (ϕ the golden mean). Define f as $f(n) = \log rn / \log \phi\ + u$ for some fixed $u \in \mathbb{R}$. Then, as $n \to \infty$, the probability that $x \in [F_n, F_{n+1})$ has longest gap less than or equal to $f(n)$ converges to

$$\mathbb{P}(L(x) \leq f(n,u)) = e^{e^{(1-u)\log\phi + \{f(n)\}}}$$

Corollary: If $f(n,u)$ grows any **slower** or ***faster*** than $\log n / \log \phi$, then $\mathbb{P}(L(x) \leq f(n))$ goes to **0** or ***1*** respectively.

Mean and Variance
We can use the **CDF** to determine the regular distribution function, and particularly the mean and variance. Let

$$P(u) = \mathbb{P}\left(L(x) \leq \frac{\log(\frac{\phi^2}{\phi^2+1} n)}{\log \phi} + u\right),$$

then the distribution of the longest gap is **approximately** $\frac{\mathrm{d}}{\mathrm{d}u} P(u)$.

The mean is given by

$$\mu = \int_{-\infty}^{\infty} u \frac{\mathrm{d}}{\mathrm{d}u} P(u) \mathrm{d}u.$$

In the continuous approximation, the mean is (γ is the Euler- Mascheroni constant)

$$\frac{\log\left(\frac{\phi^2}{\phi^2+1} n\right)}{\log \phi} - \gamma.$$

2.3 General PLRS

There is a critical root , $z_f \to 1/\lambda_1$ exponentially as $f \to \infty$.

PLRS Cumulative Distribution: Let λ_i be the eigenvalues of the recurrence, and p_i their coefficients. Define

$$\begin{aligned} \mathcal{G}(x) &= \prod_{i=2}^{L} \left(x - \frac{1}{\lambda_i}\right) \\ \mathcal{P}(x) &= (c_1 - 1)x^{t_1} + c_2 x^{t_2} + \cdots + c_L x^{t_L} \\ \mathcal{R}(x) &= c_1 x^{t_1} + c_2 x^{t_2} + \cdots + (c_L - 1)x^{t_L} \\ \mathcal{M}(x) &= 1 - c_1 x - c_2 x^{t_2+1} - \cdots - c_L x^{t_L+1}. \end{aligned}$$

The cumulative distribution of the longest gap in $[H_n, H_{n+1})$ is:

$$\mathbb{P}(L(x) < f) = \frac{-\mathcal{P}(z_f)\ /\ (p_1\lambda_1 - p_1)}{z_f \mathcal{M}'(z_f) + f\ z_f^f\ \mathcal{R}(z_f) + z_f^{f+1} \mathcal{R}'(z_f)} \left(\frac{1}{z_f \lambda_1}\right)^n + H(n,f)$$

where there exists ϵ with $1/\lambda_1 < \epsilon < 1$, such that $H(n,f) \ll f\epsilon^n$.

3. Phase Transitions

3.1 Abelian and Non-Abelian Cases

Since we are now looking at groups we need an analogous definition. So the sumset becomes $S \cdot S = \{xy : x, y \in S\}$., while the sum-difference becomes $S \cdot S^{-1} = \{xy^{-1} : x, y \in S\}$.

Lemma for Sumsets of Cyclic Groups: If $S, T \subseteq \mathbb{Z}/n\mathbb{Z}$ and if $k \in \mathbb{Z}/n\mathbb{Z}$ then

$$\mathbb{P}(k \notin S \cdot T) = O((3/4)^n)$$

Lemma for Sumdiff of Cyclic Groups: If $S, T \subseteq \mathbb{Z}/n\mathbb{Z}$ and if $k \in \mathbb{Z}/n\mathbb{Z}$ then $\mathbb{P}(k \notin S \cdot S^{-1W}) = \frac{f(n/d)^d}{2^n} \leq (\varphi/2)^n$ where $\gcd(k,n) = d$ and $f(n) = F(n+1) + F(n-1)$ where $F(n)$ is the n^{th} Fibonacci number.

Dihedral Groups: If S is a random subset of D_{2n} (if $\alpha \in D_{2n}$ then $\mathbb{P}(\alpha \in S) = 1/2$) then

$$\lim_{n \to \infty} \mathbb{P}(|S \cdot S| = |S \cdot S^{-1}|) = 1.$$

[Semi-Direct Products: For the group $\mathbb{Z}/n\mathbb{Z} \rtimes \mathbb{Z}/m\mathbb{Z}$, if either n or m go to infinity then, $\mathbb{P}(|S \cdot S| = |S \cdot S^{-1}|) = 1$.

Optimal Refinement of Keeler's Theorem: Let $P = C_1 \cdots C_r$ be a product of r disjoint in S_N, where n is the number of entries in P. Then P can be undone by a product of γ of $n+r+2$ dishing transpositions in S_{n+2} each containing at least one of the outside entries $x = n+1$, $y = n+2$. Moreover, this result is best possible in the sense that $n+r+2$ cannot be replaced by a smaller number.

Abelian Groups: As the size of an abelian group approaches infinity, then $\mathbb{P}(|S \cdot S| = |S \cdot S^{-1}|) = 1$.

3.2 Probability Decaying in N

Martin and O'Bryant, 2006: Positive percentage of sets are MSTD when sets chosen with uniform probability. **Hegarty and Miller, 2008:** When elements chosen with probability $p(N) \to 0$ as $N \to \infty$, then $|A - A| > |A + A|$ almost surely. For $s > d$, consider the **Generalized Sumset** $A_{s,d} = A + \cdots + A - A - \cdots - A$ where we have s plus signs and d minus signs. Let $h = s + d$. We want to study the size of this set as a function of s,d, and δ for probability $p(N) = cN^{-\delta}$. We call the critical value the phase transition because it is the value at which the order of the number of repeated elements is as large as the number of distinct elements.

Our goal: Extend the results of Hegarty-Miller to the case of Generalized Sumsets and determine where the phase transition occurs for $h > 2$.

Three different cases for δ: Fast Decay: $\delta > \frac{h-1}{h}$; Critical Decay: $\delta = \frac{h-1}{h}$; Slow Decay: $\delta < \frac{h-1}{h}$.
Fast Decay: For $\delta > \frac{h-1}{h}$, the set with more differences is larger 100% of the time. **Critical Decay:** In the two-case, for $g(x) = 2\sum \frac{(-1)^{k-1} x^k}{(k+1)!}$, $S \sim g\left(\frac{c^2}{2}\right) N$ and $D \sim g\left(c^2\right) N$. For $A + A + A$,
$g(x) = \sum (-1)^{k-1} \left(\frac{1}{k12^k} + \frac{c_k}{(-8)^k}\right) x^k$ with $c_k = \int_{-\frac{1}{\sqrt{3}}}^{\frac{1}{\sqrt{3}}} (x^2 - 1)^k dx$. For $A + A - A$, $g(x) = \sum_{k=1}^{m} (-1)^{k-1} \frac{1}{k!} ((-\frac{3}{8})^k c_k + \frac{1}{k}) x^k$.