Number Theory and Probability Group SMALL 2012

L. Alpoge, A. Bower, V. Hogan, R. Insoft, S. Li, V. Luo, S. J. Miller, N. Triantafillou, P. Tosteson, K. Vissuet

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

SMALL, August 7, 2012

Random Matrix Theory (Luo and Triantafillou)

Goals

- We study distributions of structured ensembles. More structure and less averaging leads to new behavior.
- Generalize results on weighted structured ensembles.
- Look at transitions as the amount of structure changes.

Weighted Toeplitz Matrices

For fixed n, we consider $N \times N$ weighted Toeplitz matrices, whose entries are iidrv from a p with mean 0, variance 1 and finite higher moments and randomly chosen $\epsilon_{ij} \in \{-1, 1\}$ with Prob$(\epsilon_{ij} = 1) = p$. A weighted Toeplitz matrix is of the form

$$\begin{pmatrix} \epsilon_{11}b_0 & \epsilon_{12}b_1 & \cdots & \epsilon_{1(N-1)}b_{N-2} & \epsilon_{1N}b_{N-1} \\ \epsilon_{21}b_1 & \epsilon_{22}b_0 & \cdots & \epsilon_{2(N-1)}b_{N-3} & \epsilon_{2N}b_{N-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \epsilon_{(N-1)1}b_{N-2} & \epsilon_{(N-1)2}b_{N-3} & \cdots & \epsilon_{(N-1)(N-1)}b_0 & \epsilon_{(N-1)N}b_1 \\ \epsilon_{N1}b_{N-1} & \epsilon_{N2}b_{N-2} & \cdots & \epsilon_{N(N-1)}b_1 & \epsilon_{NN}b_0 \end{pmatrix}$$

Configurations

By eigenvalue trace lemma, k^{th} uncentered moment is

$$\frac{1}{N^{1+k/2}}\mathbb{E}\left(\sum_{i_1,\dots,i_k} a_{i_1i_2}a_{i_2i_3}\cdots a_{i_{k-1}i_k}a_{i_ki_1}\right).$$

In Toeplitz ensembles, all terms on a diagonal are the same, so we relabel $a_{i_ji_{j+1}}$ as $b_{|i_j-i_{j+1}|}$.

Lemma

The only terms that contribute to the $2k^{\text{th}}$ moment of the limiting spectral measure are terms where the b's are matched in exactly pairs.

Configurations

Lemma

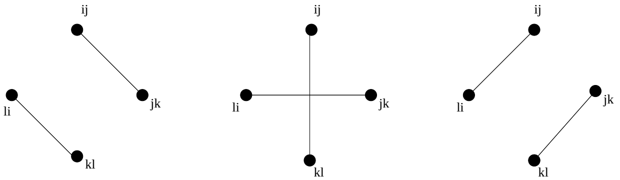
The only terms that contribute to the $2k^{\text{th}}$ moment of the limiting spectral measure are terms where the *b*'s are matched in exactly pairs.

Sketch of Proof: degree of freedom argument.

Idea: Count contribution from each pairing.

If $a_{i_j i_{j+1}} = a_{i_k i_{k+1}}$, then $|i_j - i_{j+1}| = -|i_k - i_{k+1}|$.

Examples of Configurations



In top left configuration, $a_{ij} = a_{jk}$, $a_{li} = a_{kl}$
$\Rightarrow |i-j| = -|j-k|, |l-i| = -|k-l|$.

Past Weighted Toeplitz results

Theorem: Beckwith-Miller-Shen (SMALL 2011)

Consider the weighted ensemble where the (i, j)th and (j, i)th entries of these matrices are multiplied by a randomly chosen $\epsilon_{ij} \in \{1, -1\}$, with $\text{Prob}(\epsilon_{ij} = 1) = p$.

For $p = 1/2$, the limiting spectral measure is the semi-circle. For all other p, the limiting measure has unbounded support, converges to original ensemble's limiting measure as $p \to 1$ (weakly convergent, surely if density is even).

New Results

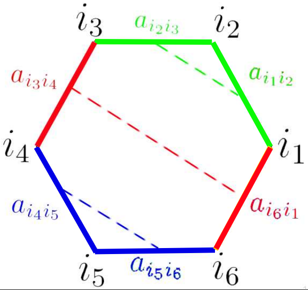
Theorem: LMT '12

1. For palindromic Toeplitz matrices, the depression of the contribution depends only on the crossing number.
2. Given any matrix ensemble, when $p = 1/2$, the limiting spectral measure is the semicircle distribution (special dependicies allowed between matrix elements).
3. Any distribution that had unbounded or bounded support before weighting still has unbounded or bounded support after weighting.

1. Depression Dependency on Number of Crossings

Theorem: Dependency on Number of Crossings

- For palindromic Toeplitz matrices, depression of contribution depends only on crossing number.
- Dependency does not hold for doubly palindromic Toeplitz, consider 6th moment.



2. Depression of at least $(2p-1)^2$

Theorem: Depression at least $(2p-1)^2$

- $x(c)$ is original contribution from the specified configuration, $2k^{\text{th}}$ moment, and $e(c)$ number of vertices in crossing pairs.
- Consider “nice" ensembles, i.e., highly palindromic Toeplitz.
- Noncrossing: contrib. at most $(x(c)-1)(2p-1)^4+1$ and at least $(x(c)-1)(2p-1)^{2k}+1$.
- Crossing: contribution at most $x(c)(2p-1)^{e(c)}$ and at least $x(c)(2p-1)^{2k}$.

3. Interpolation

Theorem: Interpolation

- Consider general real symmetric matrix ensemble.
- Noncrossing: contribution to the $2k^{\text{th}}$ moment reduced from $x(c)$ to at most $(2p-1)^2(x(c)-1)+1$.
- Crossing: contribution to $2k^{\text{th}}$ moment reduced from $x(c)$ to at most $(2p-1)^2x(c)$.
- When $p=\frac{1}{2}$, obtain semicircle distribution.

Low-lying zeroes of Maass form L-functions
(Alpoge)

Katz-Sarnak

- Birch-Swinnerton-Dyer: values of L-functions near a "central point" are ridiculously interesting.

Katz-Sarnak

- Birch-Swinnerton-Dyer: values of L-functions near a "central point" are ridiculously interesting.
- Every analytic number theorist ever: zeroes of L-functions are ridiculously interesting.

Katz-Sarnak

- Birch-Swinnerton-Dyer: values of L-functions near a "central point" are ridiculously interesting.
- Every analytic number theorist ever: zeroes of L-functions are ridiculously interesting.

Katz-Sarnak

- Birch-Swinnerton-Dyer: values of L-functions near a "central point" are ridiculously interesting.
- Every analytic number theorist ever: zeroes of L-functions are ridiculously interesting.
- Let's combine the two.

Katz-Sarnak

- Birch-Swinnerton-Dyer: values of L-functions near a "central point" are ridiculously interesting.
- Every analytic number theorist ever: zeroes of L-functions are ridiculously interesting.
- Let's combine the two. We'll study zeroes of L-functions near some central point.

Katz-Sarnak

- Birch-Swinnerton-Dyer: values of L-functions near a "central point" are ridiculously interesting.
- Every analytic number theorist ever: zeroes of L-functions are ridiculously interesting.
- Let's combine the two. We'll study zeroes of L-functions near some central point.
- If we replace $\mathbb{Q}$ by $\mathbb{F}_q(t)$, we know the Riemann hypothesis (and more: Deligne). Deligne's proof uses the action of a "monodromy group."

Katz-Sarnak

- Birch-Swinnerton-Dyer: values of L-functions near a "central point" are ridiculously interesting.
- Every analytic number theorist ever: zeroes of L-functions are ridiculously interesting.
- Let's combine the two. We'll study zeroes of L-functions near some central point.
- If we replace $\mathbb{Q}$ by $\mathbb{F}_q(t)$, we know the Riemann hypothesis (and more: Deligne). Deligne's proof uses the action of a "monodromy group." Who is this guy in "real life" (over $\mathbb{Q}$)?

Katz-Sarnak

- Birch-Swinnerton-Dyer: values of L-functions near a "central point" are ridiculously interesting.
- Every analytic number theorist ever: zeroes of L-functions are ridiculously interesting.
- Let's combine the two. We'll study zeroes of L-functions near some central point.
- If we replace $\mathbb{Q}$ by $\mathbb{F}_q(t)$, we know the Riemann hypothesis (and more: Deligne). Deligne's proof uses the action of a "monodromy group." Who is this guy in "real life" (over $\mathbb{Q}$)?
- Katz-Sarnak: study the distribution of zeroes (of a family of L-functions) near this central point (via hitting them with neutrons), and you shall find out...

- Iwaniec-Luo-Sarnak figured it out for modular forms: it's $\mathrm{O} := \lim_{n\to\infty} \mathrm{O}(n)$, whatever that means.

- Iwaniec-Luo-Sarnak figured it out for modular forms: it's $\mathrm{O} := \lim_{n\to\infty} \mathrm{O}(n)$, whatever that means.
- Modular forms "are" two-dimensional representations (hom.s $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\text{something})$).

- Iwaniec-Luo-Sarnak figured it out for modular forms: it's $\mathrm{O} := \lim_{n\to\infty} \mathrm{O}(n)$, whatever that means.
- Modular forms "are" two-dimensional representations (hom.s $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2$(something)).
- But they're not all two-dimensional representations.

- Iwaniec-Luo-Sarnak figured it out for modular forms: it's $\mathrm{O} := \lim_{n\to\infty} \mathrm{O}(n)$, whatever that means.
- Modular forms "are" two-dimensional representations (hom.s $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2$(something)).
- But they're not all two-dimensional representations. What's left are the *Maass forms*.

- Iwaniec-Luo-Sarnak figured it out for modular forms: it's $\mathrm{O} := \lim_{n\to\infty} \mathrm{O}(n)$, whatever that means.
- Modular forms "are" two-dimensional representations (hom.s $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2$(something)).
- But they're not all two-dimensional representations. What's left are the *Maass forms*.
- Last year's SMALL group figured out what the monodromy group *should* be: O.

- Iwaniec-Luo-Sarnak figured it out for modular forms: it's $\mathrm{O} := \lim_{n\to\infty} \mathrm{O}(n)$, whatever that means.
- Modular forms "are" two-dimensional representations (hom.s $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2$(something)).
- But they're not all two-dimensional representations. What's left are the *Maass forms*.
- Last year's SMALL group figured out what the monodromy group *should* be: O. I proved it.

- Iwaniec-Luo-Sarnak figured it out for modular forms: it's $\mathrm{O} := \lim_{n\to\infty} \mathrm{O}(n)$, whatever that means.
- Modular forms "are" two-dimensional representations (hom.s $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2$(something)).
- But they're not all two-dimensional representations. What's left are the *Maass forms*.
- Last year's SMALL group figured out what the monodromy group *should* be: O. I proved it.
- Let's find out how!

- But first, a fun aside. (Something I found during YMC.)

- But first, a fun aside. (Something I found during YMC.)
- Goldfeld-Kontorovich (famous number theorist and his former student):

- But first, a fun aside. (Something I found during YMC.)
- Goldfeld-Kontorovich (famous number theorist and his former student): “we obtain the low-lying zero densities for… $\mathrm{GL}(3)$ Maass forms.”

- But first, a fun aside. (Something I found during YMC.)
- Goldfeld-Kontorovich (famous number theorist and his former student): "we obtain the low-lying zero densities for... $\mathrm{GL}(3)$ Maass forms."
- "The methods presented here are capable of wide generalization... it should be possible to determine the symmetry types of families associated to... $\mathrm{GL}(n)$ for any $n \geq 2$. We hope to return to this topic in a future publication."

Theorem

Research is not for the faint of heart.

□

My proof

- Main method: a trace formula (a huge common generalization of Schur's orthogonality of characters and Poisson's summation formula).

My proof

- Main method: a trace formula (a huge common generalization of Schur's orthogonality of characters and Poisson's summation formula).
- Main insights:
- First: throw out everything from before and start from scratch.

My proof

- Main method: a trace formula (a huge common generalization of Schur's orthogonality of characters and Poisson's summation formula).
- Main insights:
- First: throw out everything from before and start from scratch.
- Second: use the only real tool of analytic number theory: contour integration. (Push a contour "off to ∞".)

My proof

- Main method: a trace formula (a huge common generalization of Schur's orthogonality of characters and Poisson's summation formula).
- Main insights:
- First: throw out everything from before and start from scratch.
- Second: use the only real tool of analytic number theory: contour integration. (Push a contour "off to ∞".)
- Third: use Poisson summation.

My proof

- Main method: a trace formula (a huge common generalization of Schur's orthogonality of characters and Poisson's summation formula).
- Main insights:
- First: throw out everything from before and start from scratch.
- Second: use the only real tool of analytic number theory: contour integration. (Push a contour "off to ∞".)
- Third: use Poisson summation.
- Fourth: use Fourier inversion (and Taylor expand copiously).

My proof (continued)

- Fifth (and most important): if $A = B$, then $B = A$.

My proof (continued)

- Fifth (and most important): if $A = B$, then $B = A$. With this leap of imagination, apply Poisson summation again!

My proof (continued)

- Fifth (and most important): if $A = B$, then $B = A$. With this leap of imagination, apply Poisson summation again!
- Sixth: Integrals with a very quickly oscillating factor (that doesn't stop) are tiny.

My proof (continued)

- Fifth (and most important): if $A = B$, then $B = A$. With this leap of imagination, apply Poisson summation again!
- Sixth: Integrals with a very quickly oscillating factor (that doesn't stop) are tiny.
- Seventh: Sums are ugly. Integrals are nice. So smooth out sums into integrals!

My proof (continued)

- Fifth (and most important): if $A = B$, then $B = A$. With this leap of imagination, apply Poisson summation again!
- Sixth: Integrals with a very quickly oscillating factor (that doesn't stop) are tiny.
- Seventh: Sums are ugly. Integrals are nice. So smooth out sums into integrals!
- Eighth: Don't be afraid to assume GRH if you are trying to say something about the zeroes of generalized L-functions...

My proof (continued)

- Fifth (and most important): if $A = B$, then $B = A$. With this leap of imagination, apply Poisson summation again!
- Sixth: Integrals with a very quickly oscillating factor (that doesn't stop) are tiny.
- Seventh: Sums are ugly. Integrals are nice. So smooth out sums into integrals!
- Eighth: Don't be afraid to assume GRH if you are trying to say something about the zeroes of generalized L-functions…
- Ninth: Never give up!!!

My proof (continued)

- Fifth (and most important): if $A = B$, then $B = A$. With this leap of imagination, apply Poisson summation again!
- Sixth: Integrals with a very quickly oscillating factor (that doesn't stop) are tiny.
- Seventh: Sums are ugly. Integrals are nice. So smooth out sums into integrals!
- Eighth: Don't be afraid to assume GRH if you are trying to say something about the zeroes of generalized L-functions...
- Ninth: Never give up!!! Oh, and apply Poisson summation four more times.

My proof (continued)

- Fifth (and most important): if $A = B$, then $B = A$. With this leap of imagination, apply Poisson summation again!
- Sixth: Integrals with a very quickly oscillating factor (that doesn't stop) are tiny.
- Seventh: Sums are ugly. Integrals are nice. So smooth out sums into integrals!
- Eighth: Don't be afraid to assume GRH if you are trying to say something about the zeroes of generalized L-functions...
- Ninth: Never give up!!! Oh, and apply Poisson summation four more times.
  $A = B = A = B = A = B = \cdots$

My proof (continued)

- And, most importantly, ask your computer if things you’re trying to prove are actually true.

My proof (continued)

- And, most importantly, ask your computer if things you're trying to prove are actually true.
- Hopefully most of that made sense to everybody.

My proof (continued)

- And, most importantly, ask your computer if things you're trying to prove are actually true.
- Hopefully most of that made sense to everybody.
- If you're itching to know what has kept me locked away for all these weeks, the above encodes a writeup with 13 parameters, final error terms that take up entire pages, and a (fully detailed) writeup that will take about 100 pages, if not more.

My proof (continued)

- And, most importantly, ask your computer if things you're trying to prove are actually true.
- Hopefully most of that made sense to everybody.
- If you're itching to know what has kept me locked away for all these weeks, the above encodes a writeup with 13 parameters, final error terms that take up entire pages, and a (fully detailed) writeup that will take about 100 pages, if not more.
- Yuck!!!

My proof (continued)

- And, most importantly, ask your computer if things you're trying to prove are actually true.
- Hopefully most of that made sense to everybody.
- If you're itching to know what has kept me locked away for all these weeks, the above encodes a writeup with 13 parameters, final error terms that take up entire pages, and a (fully detailed) writeup that will take about 100 pages, if not more.
- Yuck!!! The summary I gave is more fun, I think!

My proof (continued)

- And, most importantly, ask your computer if things you're trying to prove are actually true.
- Hopefully most of that made sense to everybody.
- If you're itching to know what has kept me locked away for all these weeks, the above encodes a writeup with 13 parameters, final error terms that take up entire pages, and a (fully detailed) writeup that will take about 100 pages, if not more.
- Yuck!!! The summary I gave is more fun, I think!

Thanks!!!!

Gaps in Generalized Zeckendorf Decompositions (Bower, Insoft, Li and Tosteson)

Previous Results

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

Previous Results

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

Zeckendorf's Theorem

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

Previous Results

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

Zeckendorf's Theorem

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

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

Previous Results

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

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 (1952)

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.

Previous Generalizations

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

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

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 $L = 1$.

Previous Generalizations

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

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

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

Previous Generalizations

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

$$H_{n+1} = c_1 H_n + c_2 H_{n-1} + \cdots + c_L H_{n-L+1}, \; n \geq L$$

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
- Lekkerkerker: Average number summands is $C_{\mathrm{Lek}} n + d$.

Previous Generalizations

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

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

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
- Lekkerkerker: Average number summands is $C_{\mathrm{Lek}} n + d$.
- Central Limit Type Theorem

Distribution of Gaps

For $H_{i_1} + H_{i_2} + \cdots + H_{i_n}$, 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 $H_{i_1} + H_{i_2} + \cdots + H_{i_n}$, the gaps are the differences:

$$i_n - i_{n-1},\ i_{n-1} - i_{n-2},\ \ldots,\ i_2 - i_1.$$

Example: For $H_1 + H_8 + H_{18}$, the gaps are 7 and 10.

Distribution of Gaps

For $H_{i_1} + H_{i_2} + \cdots + H_{i_n}$, the gaps are the differences:

$$i_n - i_{n-1}, \; i_{n-1} - i_{n-2}, \; \dots, \; i_2 - i_1.$$

Example: For $H_1 + H_8 + H_{18}$, the gaps are 7 and 10.

Definition

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

Distribution of Gaps

For $H_{i_1} + H_{i_2} + \cdots + H_{i_n}$, the gaps are the differences:

$$i_n - i_{n-1},\ i_{n-1} - i_{n-2},\ \dots,\ i_2 - i_1.$$

Example: For $H_1 + H_8 + H_{18}$, the gaps are 7 and 10.

Definition

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

Big Question: What is $P(m) = \lim_{n\to\infty} P_n(m)$?
Big Question: What is the distribution of the longest gap?

Positive Linear Recurrences of Any Length

Theorem

Let $H_{n+1} = c_1 H_n + \cdots + c_L H_{n+1-L}$ be a Positive Linear Recurrence Sequence, then, if $j \geq L$,

$$P(j) \;=\; (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j},$$

where λ_1 is the largest root of the characteristic polynomial of the recurrence.

Positive Linear Recurrences of Any Length

Theorem

Let $H_{n+1} = c_1 H_n + \cdots + c_L H_{n+1-L}$ be a Positive Linear Recurrence Sequence, then, if $j \geq L$,

$$P(j) \;=\; (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j},$$

where λ_1 is the largest root of the characteristic polynomial of the recurrence.

What can we say about the distribution of gaps $< L$ for any PLRS?

Positive Linear Recurrences of Any Length

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* $P(j) =$

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

Proof Set Up of case $j \geq 2$

Theorem

If $j \geq 2$*, then* $P(j) = (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j}$.

Let $X_{i,i+j}(n) = \#\{m \in [H_n, H_{n+1})$: decomposition of m includes H_i, H_{i+j}, but not H_q for $i < q < i+j\}$.

Proof Set Up of case $j \geq 2$

Theorem

If $j \geq 2$, then $P(j) = (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j}$.

Let $X_{i,i+j}(n) = \#\{m \in [H_n, H_{n+1})$: decomposition of m includes H_i, H_{i+j}, but not H_q for $i < q < i+j\}$.

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

Proof Set Up of case $j \geq 2$

Theorem

If $j \geq 2$, then $P(j) = (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j}$.

Let $X_{i,i+j}(n) = \#\{m \in [H_n, H_{n+1})$: decomposition of m includes H_i, H_{i+j}, but not H_q for $i < q < i+j\}$.

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

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

Proof Set Up of case $j \geq 2$

Theorem

If $j \geq 2$, then $P(j) = (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j}$.

Let $X_{i,i+j}(n) = \#\{m \in [H_n, H_{n+1})$: decomposition of m includes H_i, H_{i+j}, but not H_q for $i < q < i+j\}$.

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

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

Generalized Lekkerkerker:
$\Rightarrow\ Y(n) \sim (C_{Lek} n + d)(H_{n+1} - H_n)$.

A Quick Counting Lesson: How do we count $X_{i,i+j}$?

We need to see the number of legal decompositions with a gap of length j.

Can count how many legal decompositions exist to the left and right of the gap.

Lemma

Let $H_{n+1} = c_1 H_n + \cdots + c_L H_{n+1-L}$ be a Positive Linear Recurrence Sequence, then the number of legal decompositions which contain H_m as the largest summand is $H_{m+1} - H_m$.

Calculating $X_{i,i+j}$

Theorem

If $j \geq 2$, then $P(j) = (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j}$.

In the interval $[H_n, H_{n+1})$:
How many decompositions contain a gap from H_i to H_{i+j}?

Calculating $X_{i,i+j}$

Theorem

If $j \geq 2$, then $P(j) = (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j}$.

In the interval $[H_n, H_{n+1})$:
How many decompositions contain a gap from H_i to H_{i+j}?

Left: For the indices less than i: $H_{i+1} - H_i$ choices.

Calculating $X_{i,i+j}$

Theorem

If $j \geq 2$, then $P(j) = (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j}$.

In the interval $[H_n, H_{n+1})$:
How many decompositions contain a gap from H_i to H_{i+j}?

Left: For the indices less than i: $H_{i+1} - H_i$ choices.

Right: For the indices greater than $i+j$:
$H_{n-i-j+2} - H_{n-i-j+1} - (H_{n-i-j+1} - H_{n-i-j})$ choices.

Calculating $X_{i,i+j}$

Theorem

If $j \geq 2$, then $P(j) = (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j}$.

In the interval $[H_n, H_{n+1})$:
How many decompositions contain a gap from H_i to H_{i+j}?

Left: For the indices less than i: $H_{i+1} - H_i$ choices.

Right: For the indices greater than $i + j$:
$H_{n-i-j+2} - H_{n-i-j+1} - (H_{n-i-j+1} - H_{n-i-j})$ choices.

So $X_{i,i+j}(n) = \text{Left} * \text{Right} =$
$(H_{i+1} - H_i)(H_{n-i-j+2} - H_{n-i-j+1} - (H_{n-i-j+1} - H_{n-i-j}))$.

Final Steps of the Proof

For sufficiently large n, $H_n \approx a_1\lambda_1^n$.

Final Steps of the Proof

For sufficiently large n, $H_n \approx a_1\lambda_1^n$.

Then with some algebra...

Final Steps of the Proof

For sufficiently large n, $H_n \approx a_1 \lambda_1^n$.

Then with some algebra...

$$P(j) = (\lambda_1 - 1)^2 \left(\frac{a_1}{C_{Lek}}\right) \lambda_1^{-j}.$$

Gaps in Generalized Zeckendorf Decompositions (Bower, Insoft, Li and Tosteson)

Longest Gap:

Big question: Given a random number x in the interval $[F_n, F_{n+1})$, what is the probability that x has longest gap equal to r?

Our Method

What we do:

Our Method

What we do:

- Recast the problem through combinatorics.

Our Method

What we do:

- Recast the problem through combinatorics.
- Obtain generating functions!

Our Method

What we do:

- Recast the problem through combinatorics.
- Obtain generating functions!
- Get the important relationships.

Our Method

What we do:

- Recast the problem through combinatorics.
- Obtain generating functions!
- Get the important relationships.
- Analyze limiting behavior.

Cumulative Distribution Function

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

Cumulative Distribution Function

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

Theorem

Let $r = \phi^2/(\phi^2+1)$. Set $f(n) = \log rn/\log\phi + u$ for some fixed $u \in \mathbb{Z}$. As $n \to \infty$, the probability that $x \in [F_n, F_{n+1})$ has longest gap at most $f(n)$ converges to

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

Cumulative Distribution Function

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

Theorem

Let $r = \phi^2/(\phi^2+1)$. Set $f(n) = \log rn/\log\phi + u$ for some fixed $u \in \mathbb{Z}$. As $n \to \infty$, the probability that $x \in [F_n, F_{n+1})$ has longest gap at most $f(n)$ converges to

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

Immediate Corollary: If $f(n,u)$ grows any **slower** or ***faster*** than $\log n/\log\phi$, then $\mathbb{P}(L(x) \le f(n))$ 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**.

Mean and Variance

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

Let

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

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** $\frac{\mathrm{d}}{\mathrm{du}} P(u)$.

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** $\frac{\mathrm{d}}{\mathrm{du}} P(u)$.

The mean is given by

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

The variance follows similarly.

Mean and Variance

So the mean is about

$$\mu = \frac{\log\left(\frac{\phi^2}{\phi^2+1}\right)}{\log\phi} + \int_{-\infty}^{\infty} e^{-e^{(1-u)\log\phi}} e^{(1-u)\log\phi} \log\phi \, du.$$

Mean and Variance

So the mean is about

$$\mu = \frac{\log\left(\frac{\phi^2}{\phi^2+1}\right)}{\log\phi} + \int_{-\infty}^{\infty} e^{-e^{(1-u)\log\phi}} e^{(1-u)\log\phi} \log\phi \, du.$$

$$= \frac{1}{\log\phi}\left(\log\left(\frac{\phi^2}{\phi^2+1}\right) - \int_0^{\infty} \log(w)\cdot e^{-w}\, dw\right).$$

Mean and Variance

So the mean is about

$$\mu = \frac{\log\left(\frac{\phi^2}{\phi^2+1}\right)}{\log\phi} + \int_{-\infty}^{\infty} e^{-e^{(1-u)\log\phi}} e^{(1-u)\log\phi} \log\phi \, du.$$

$$= \frac{1}{\log\phi}\left(\log\left(\frac{\phi^2}{\phi^2+1}\right) - \int_0^{\infty} \log(w)\cdot e^{-w}\, dw\right).$$

Theorem

In the continuous approximation, the mean is

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

Positive Linear Recurrence Sequences

This method can be greatly generalized to **Positive Linear Recurrence Sequences** (linear recurrences with non-negative coefficients). WLOG:

$$H_{n+1} = c_1 H_{n-(t_1=0)} + c_2 H_{n-t_2} + \cdots + c_L H_{n-t_L}.$$

Positive Linear Recurrence Sequences

This method can be greatly generalized to **Positive Linear Recurrence Sequences** (linear recurrences with non-negative coefficients). WLOG:

$$H_{n+1} = c_1 H_{n-(t_1=0)} + c_2 H_{n-t_2} + \cdots + c_L H_{n-t_L}.$$

Theorem (Zeckendorf's Theorem for *PLRS* recurrences)

*Any $b \in \mathbb{N}$ has a unique **legal** decomposition into sums of H_n, $b = a_1 H_{i_1} + \cdots + a_{i_k} H_{i_k}$.*

Here **legal** reduces to non-adjacency of summands in the Fibonacci case.

Generating Function for PLRS

The **number** of $b \in [H_n, H_{n+1})$, with longest gap $< f$ is the coefficient of x^{n-s} in the generating function:

$$\sum_{k\geq 0}\Bigg[\left((c_1-1)x^{t_1}+\cdots+(c_L-1)x^{t_L}\right)\left(\frac{x^{s+1}-x^f}{1-x}\right)+$$

$$x^{t_1}\left(\frac{x^{s+t_2-t_1+1}-x^f}{1-x}\right)+\cdots+x^{t_{L-1}}\left(\frac{x^{s+t_L-t_{L-1}}+1-x^f}{1-x}\right)\Bigg]^k\times$$

$$\frac{1}{1-x}\left(c_1-1+c_2x^{t_2}+\cdots+c_Lx^{t_L}\right)$$

A geometric series!

Phase Transitions
(Hogan)

Past Results

- Martin and O'Bryant, 2006: Positive percentage of sets are MSTD (more sum than difference) when sets chosen with uniform probability. Surprising: $x + y = y + x$ but $x - y$ usually not $y - x$.

Past Results

- Martin and O'Bryant, 2006: Positive percentage of sets are MSTD (more sum than difference) when sets chosen with uniform probability. Surprising: $x + y = y + x$ but $x - y$ usually not $y - x$.

- Iyer, Lazarev, Miller, Zhang, 2011: Generalized results above to an arbitrary number of summands.

Phase Transition

Theorem (Hegarty-Miller): $\mathscr{S} = |A+A|$, $\mathscr{D} = |A-A|$, $g(x) := 2\left(\frac{e^{-x}-(1-x)}{x}\right)$. Take $k \in \{0,\dots,N-1\}$ with probability $p(N) \to 0$, then if

- $p(N) = o(N^{-1/2})$:

$$\mathscr{S} \sim \frac{(N\cdot p(N))^2}{2} \quad \text{and} \quad \mathscr{D} \sim 2\mathscr{S} \sim (N\cdot p(N))^2.$$

- $p(N) = c\cdot N^{-1/2}$:

$$\mathscr{S} \sim g\left(\frac{c^2}{2}\right)N \quad \text{and} \quad \mathscr{D} \sim g(c^2)N.$$

- $N^{-1/2} = o(p(N))$: Let $\mathscr{S}^c := (2N+1) - \mathscr{S}$, $\mathscr{D}^c := (2N+1) - \mathscr{D}$. Then

$$\mathscr{S}^c \sim 2\cdot\mathscr{D}^c \sim 4/p(N)^2.$$

Generalized Sumsets

Definition

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 $h = s + d$.

Generalized Sumsets

Definition

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 $p(N) = cN^{-\delta}$.

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

Cases for δ

To answer, we must consider three different cases for δ.

Cases for δ

To answer, we must consider three different cases for δ.

- Fast Decay: $\delta > \frac{h-1}{h}$.
- Critical Decay: $\delta = \frac{h-1}{h}$.
- Slow Decay: $\delta < \frac{h-1}{h}$.

Cases for δ

To answer, we must consider three different cases for δ.

- Fast Decay: $\delta > \frac{h-1}{h}$.
- Critical Decay: $\delta = \frac{h-1}{h}$.
- Slow Decay: $\delta < \frac{h-1}{h}$.

These three cases correspond to the speed at which the probability of choosing elements decays to 0.

Fast Decay

- For $\delta > \frac{h-1}{h}$, the set with more differences is larger 100% of the time.

Fast Decay

- For $\delta > \frac{h-1}{h}$, the set with more differences is larger 100% of the time.

- Compute the number of distinct h-tuples.

Fast Decay

- For $\delta > \frac{h-1}{h}$, the set with more differences is larger 100% of the time.
- Compute the number of distinct h-tuples.
- For h-tuples $a = (a_1, \cdots, a_h)$, $b = (b_1, \cdots, b_h)$, define indicator variable $Y_{a,b}$ to be 1 when a and b generate the same element.

Fast Decay

- For $\delta > \frac{h-1}{h}$, the set with more differences is larger 100% of the time.
- Compute the number of distinct h-tuples.
- For h-tuples $a = (a_1, \cdots, a_h)$, $b = (b_1, \cdots, b_h)$, define indicator variable $Y_{a,b}$ to be 1 when a and b generate the same element.
- Bound the expected value and variance of the sum of these indicator variables.

Fast Decay

- For $\delta > \frac{h-1}{h}$, the set with more differences is larger 100% of the time.

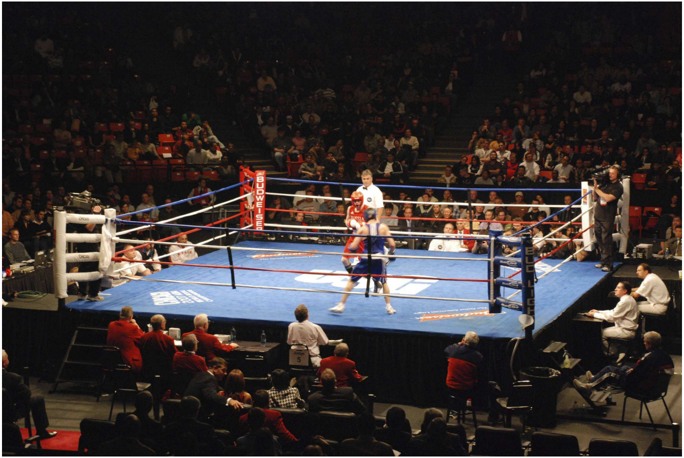
- Compute the number of distinct h-tuples.

- For h-tuples $a = (a_1, \cdots, a_h)$, $b = (b_1, \cdots, b_h)$, define indicator variable $Y_{a,b}$ to be 1 when a and b generate the same element.

- Bound the expected value and variance of the sum of these indicator variables.

- Chebyshev's Inequality: $\mathrm{Prob}(|X - \mu_X| \geq k\sigma_X) \leq 1/k^2$.

Sumsets vs Sumdifferences (Vissuet)



In the Red Corner:

- Sumset $:= S + S = \{xy : x, y \in S\}$

In the Red Corner:

- Sumset $:= S + S = \{xy : x, y \in S\}$
- Underdog

In the Red Corner:

- Sumset $:= S + S = \{xy : x, y \in S\}$
- Underdog
- Weakness: For abelian groups we have that $xy = yx$

In the Blue Corner:

- Sumdifference $:= S - S = \{xy^{-1} : x, y \in S\}$

In the Blue Corner:

- Sumdifference $:= S - S = \{xy^{-1} : x, y \in S\}$
- Reigning Champion

In the Blue Corner:

- Sumdifference $:= S - S = \{xy^{-1} : x, y \in S\}$
- Reigning Champion
- Weakness: $x \cdot x^{-1}$ is the identity $\forall x \in S$.

Rules of the match

- The winner of each bout is the one who usually has the bigger cardinality.

Rules of the match

- The winner of each bout is the one who usually has the bigger cardinality.
- The match will consist of 3 different venues and will be best 2 out of 3.

Rules of the match

- The winner of each bout is the one who usually has the bigger cardinality.
- The match will consist of 3 different venues and will be best 2 out of 3.
- Question that needs to be asked before we start:

Rules of the match

- The winner of each bout is the one who usually has the bigger cardinality.
- The match will consist of 3 different venues and will be best 2 out of 3.
- Question that needs to be asked before we start:
- Are You Ready To Rumbbllllee?

First Venue

The Match

- Sumsets terribly loses the first 13 of $\aleph_0$ rounds because there does not exist a subset of $[0, 14]$ such that $|S+S| > |S-S|$.

The Match

- Sumsets terribly loses the first 13 of $\aleph_0$ rounds because there does not exist a subset of $[0, 14]$ such that $|S + S| > |S - S|$.

- However, In Round 14, with the help of Conway, Sumset gets a jab in with the set: $\{0, 2, 3, 4, 7, 11, 12, 14\}$.

The Match

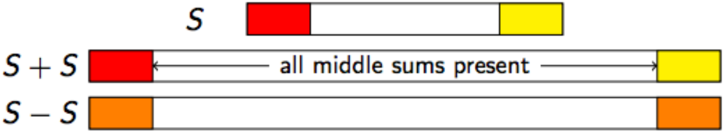
- Sumsets terribly loses the first 13 of $\aleph_0$ rounds because there does not exist a subset of $[0, 14]$ such that $|S + S| > |S - S|$.

- However, In Round 14, with the help of Conway, Sumset gets a jab in with the set: $\{0, 2, 3, 4, 7, 11, 12, 14\}$.

- With the help of Coaches Martin and O'Bryant, Sumsets realizes that if he wants to win it has to concentrate on having a better "fringe."

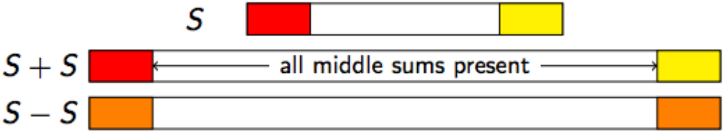
Sumset's tactics

- Key Idea: In the $\mathbb{Z}$ case, fringe matters most, middle sums and differences are present with high probability



Sumset's tactics

- Key Idea: In the $\mathbb{Z}$ case, fringe matters most, middle sums and differences are present with high probability



- *If we choose the "fringe" of S cleverly, the middle of S will become largely irrelevant. - Martin and O'Bryant's inspiring words*

First results

- So with the help of Martin and O'Bryant, Sumsets learns that there exists a positive percentage of subsets that are sum dominated.

First results

- So with the help of Martin and O'Bryant, Sumsets learns that there exists a positive percentage of subsets that are sum dominated.

- Sadly the percentage of sum dominated sets is estimated to be .00045%.

First results

- So with the help of Martin and O'Bryant, Sumsets learns that there exists a positive percentage of subsets that are sum dominated.

- Sadly the percentage of sum dominated sets is estimated to be .00045%.

- Sumset loses the first bout.

The Second Venue

The Second Venue

The Match Round 1

- Sumset faces some difficulties in the $\mathbb{Z}/n\mathbb{Z}$ venue because there is no fringe.

The Match Round 1

- Sumset faces some difficulties in the $\mathbb{Z}/n\mathbb{Z}$ venue because there is no fringe.

- Luckily for Sumset, because carousel go round and round there are know many ways to write each element.

The Match Round 1

- Sumset faces some difficulties in the $\mathbb{Z}/n\mathbb{Z}$ venue because there is no fringe.

- Luckily for Sumset, because carousel go round and round there are know many ways to write each element.

Theorem

If we let S be a random subset of $\mathbb{Z}/n\mathbb{Z}$ (if $\alpha \in D_{2n}$ then $\mathbb{P}(\alpha \in S) = 1/2$) then

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

The Match Round 2

Theorem

Similar results hold for Abelian Groups, Dihedral Groups, and Semi-direct Products of cyclic groups.

The Match Round 2

Theorem

Similar results hold for Abelian Groups, Dihedral Groups, and Semi-direct Products of cyclic groups.

Although for any finite n, there are more subsets S of D_{2n} such that $|S+S| > |S-S|$, the judges still decided to call the boat a draw due to limiting behavior.

The Third Venue

The Third Venue

The Match

- The free group was Sumset's strength, it is no longer in an abelian group.

The Match

- The free group was Sumset's strength, it is no longer in an abelian group.

- Not only that, but Sumdifference's weakness is still there ($x \cdot x^{-1}$ is the identity for all $x \in S$).

The Match

- The free group was Sumset's strength, it is no longer in an abelian group.
- Not only that, but Sumdifference's weakness is still there ($x \cdot x^{-1}$ is the identity for all $x \in S$).
- The match was very one sided.

Ping Pong

Theorem (Free Group)

If we let $\langle a, b \rangle_l$ be all words up to length l and $S \subseteq \langle a, b \rangle_l$ then as l goes to infinity we have that:

$$\mathbb{P}(|S \cdot S| \geq |S \cdot S^{-1}|) = 1.$$


TAKE THAT
SUM DIFFERENCE!
Stanford
ATHLETICS
JOOLA