Random Numerical Semigroups

by **Santiago Morales Duarte**

Thesis submitted in fulfilment of the requirements for the degree of
Bachelor of Science
under the supervision of Tristram Bogart

Department of Mathematics
Faculty of Science
Universidad de los Andes
October 28, 2023

Abstract

This is the text of the abstract, providing a brief summary of the context, research problem, main contributions and conclusions of this work.

Dedication

For my mother. (name of person - this is optional!)

Acknowledgements

I would like to thank (supervisor, family, research collaborators, anyone else who significantly helped you with this work).

Santiago Morales Duarte
October 28, 2023
Bogotá, Colombia

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The start of the introduction provides some context and brief background. This is a test for github.

1.1 Research Objectives and Overview

The research question which this Thesis aims to answer is...

The specific research objectives of this Thesis are:

1. Objective 1
2. Objective 2

Chapter 2 provides a comprehensive review of literature which is relevant to the overall aim. This includes ...

Chapter 4 aims to ...

This chapter resulted in the following publications:

- D. R. Franklin and K. J. Wilson, "A LaTeX Thesis Template for the School of Electrical and Data Engineering," *IEEE Transactions on LaTeX Thesis Templates*, vol. 1, no. 1, Oct. 2021
- D. R. Franklin and K. J. Wilson, "A LaTeX Thesis Template for the School of Electrical and Data Engineering," *IEEE Transactions on LaTeX Thesis Templates*, vol. 1, no. 1, Oct. 2021

Chapter 5 aims to ...

This chapter resulted in the following publications:

- D. R. Franklin and K. J. Wilson, "A LaTeX Thesis Template for the School of Electrical and Data Engineering," *IEEE Transactions on LaTeX Thesis Templates*, vol. 1, no. 1, Oct. 2021

Chapter 6 aims to ...

This chapter resulted in the following publications:

- D. R. Franklin and K. J. Wilson, “A LaTeX Thesis Template for the School of Electrical and Data Engineering,” *IEEE Transactions on LaTeX Thesis Templates*, vol. 1, no. 1, Oct. 2021

Finally, Chapter 7 summarises the results and implications of this work, and provides recommended directions for continuation of this work in the future.

1.1.1 Additional Research Contributions

A number of additional research publications and presentations are listed below:

- xxx

Test [1]

Chapter 2

The Probabilistic Method

2.1 Introduction

The probabilistic method is a powerful tool, with applications in Combinatorics, Graph Theory, Number Theory and Computer Science. It is a nonconstructive method that proves the existence of an object with a certain property by showing that the probability that a randomly chosen object has that property is greater than zero. This requires an appropiate sample space.

The method is best illustrated by an example:

Definition 2.1.1. A *tournament* is a directed graph T on n vertices such that for every pair of vertices $i, j \in V(T)$, exactly one of the edges (i, j) or (j, i) is in $E(T)$. [2]

Section 2.2 discusses Theme 1. Section 2.3 discusses Theme 2....

2.2 Union Bound

2.2.1 Example

2.2.2 Subtopic B

2.2.3 Subtopic C

2.3 Linearity of Expectation

Etc. etc. 2.1.1

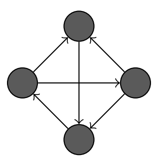


Figure 2.1: A tournament on 4 vertices.

2.4 Second Moment Method

2.5 Threshold Functions

Let $n \in \mathbb{N}$ and $0 \leq p \leq 1$. The random graph $G(n,p)$ is a probability space over the set of graphs on n labeled vertices determined by

$$\Pr[\{i,j\} \in G] = p$$

with these events mutually independent [2]. Given a graph theoretic property A, there is a probability that $G(n,p)$ satisfies A, which we write as $\Pr[G(n,p) \vDash A]$.

Definition 2.5.1. $r(n)$ is a threshold function for a graph theoretic property A if

1. When $p(n) \in o(r(n))$, $\lim_{n\to\infty} \Pr[G(n,p(n)) \vDash A] = 0$,
2. When $r(n) \in o(p(n))$, $\lim_{n\to\infty} \Pr[G(n,p(n)) \vDash A] = 1$,

or vice versa. [2]

We give an example of a threshold function which illustrates a common method for proving that a function is a threshold.

Threshold function for having isolated vertices

Let G be a graph on n labeled vertices. An isolated vertex of G is a vertex which does not belong to any of the edges of G. Let A be the property that G contains an isolated vertex. We will prove that $r(n) = \dfrac{\ln n}{n}$ is a threshold for A.

For each vertex i in G define the variable

$$X_i = \begin{cases} 1 & \text{if } i \text{ is an isolated vertex,} \\ 0 & \text{if } i \text{ is not an isolated vertex.} \end{cases}$$

Now, the probability that a vertex i is isolated is $(1-p)^{n-1}$ since it is the probability that none of the other $n-1$ vertices is connected to i. Let $X = \sum_{i=1}^{n} X_i$, then the expected number of isolated vertices is

$$E[X] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \Pr[X_i] = n(1-p)^{n-1}.$$

Let $p = k\dfrac{\ln n}{n}$ for $k \in \mathbb{R}_{>0}$. Then

$$\begin{aligned} \lim_{n\to\infty} E[X] &= \lim_{n\to\infty} n\left(1 - k\frac{\ln n}{n}\right)^{n-1} \\ &= ne^{-k\ln n} = n^{1-k}. \end{aligned}$$

Therefore, $\lim_{n\to\infty} E[X] = 0$ if $k > 1$. Since $E[X] \geq \Pr[X > 0]$, we conclude that

$$\lim_{n\to\infty} \Pr[G(n,p) \vDash A] = \lim_{n\to\infty} \Pr[X > 0] = 0.$$

Now, for $k < 1$, the fact that $\lim_{n\to\infty} E[X] = \infty$ is not enough to conclude that $\lim_{n\to\infty} \Pr[G(n,p) \vDash A] = 1$. We have to use the second moment method.

Theorem. *If $E[X] \to \infty$ and $Var[X] = o(E[X]^2)$, then $\lim_{n\to\infty} Pr[X > 0] = 1$. [2]*

Proof. We will prove that, in this case, $Var[X] = o(E[X]^2)$. First,

$$\begin{aligned}\sum_{i\neq j} E[X_iX_j] &= \sum_{i\neq j} \Pr[X_i = X_j = 1] \\ &= n(n-1)(1-p)^{n-1}(1-p)^{n-2} \\ &= n(n-1)(1-p)^{2n-3},\end{aligned}$$

for if i is an isolated vertex, then there is no edge between i and j so we only have to account for the remaining $n-2$ edges that contain j.

Thus, since $\sum_{i=1}^n E[X_i^2] = \sum_{i=1}^n E[X_i] = E[X]$ and $\lim_{n\to\infty} p = 0$,

$$\begin{aligned}\lim_{n\to\infty} \frac{\mathrm{Var}[X]}{E[X]^2} = \lim_{n\to\infty} \frac{E(X^2) - E[X]^2}{E[X]^2} &= \lim_{n\to\infty} \frac{\sum_{i=1}^n E[X_i^2] + \sum_{i\neq j} E[X_iX_j]}{E[X]^2} - 1 \\ = 0 + \lim_{n\to\infty} \frac{n(n-1)(1-p)^{2n-3}}{n^2(1-p)^{2n-2}} - 1 &= \lim_{n\to\infty} \frac{1}{1-p} - 1 = 0.\end{aligned}$$

We conclude that $\mathrm{Var}[X] \in o(E[X]^2)$ and so, if $k < 1$,

$$\lim_{n\to\infty} \Pr[G(n,p) \vDash A] = \lim_{n\to\infty} \Pr[X > 0] = 1.$$

Therefore, $r(n) = \frac{\ln n}{n}$ is a threshold function for property A.

Chapter 3

Numerical Semigroups

3.1 Introduction

Some contextural overview of what we are going to discuss. Include the first example in the book.

Section 2.2 discusses Theme 1. Section 2.3 discusses Theme 2....

A *numerical semigroup* is a subset $S \subset \mathbb{N}_0$ which is closed under addition, i.e. $a, b \in S$ implies $a + b \in S$. For instance, $\mathbb{N}_0$, $\mathbb{N}_0 \setminus \{0\}$, $2\mathbb{N}_0$ are all numerical semigroups, but $\mathbb{N}_0 \setminus \{2\}$ is not. Some literature requires that a semigroup has a finite complement in $\mathbb{Z}_{\geq 0}$ [3], but we prefer the more general definition.

3.2 Invariants

3.2.1 Subtopic A

3.2.2 Subtopic B

3.2.3 Subtopic C

3.3 Wilf's Conjecture

Etc. etc.

Chapter 4

Random Numerical Semigroups

4.1 Introduction

In this Chapter, XXX is presented. Note: to cross-reference to other parts of the document you do so like this - see Section 5.2.2.

Section 4.2 discusses Theme 1. Section 4.4 discusses Theme 2....

4.2 Box Model

Figure 4.1 illustrates the key elements of Thesis synthesis. LaTeX will tend to place figures where it wants (they 'float' - generally they should be at the top or bottom of a page); you can override the default behaviour if you want, but you probably don't want to bother doing that until after your content is pretty much done. Instead, keep the figures as close as possible to the text; you can tweak this afterwards if you want by adding an option:

```
\begin{figure}[htb]
```

This says put it here, if you can, otherwise at the top, or otherwise at the bottom. BUT I strongly suggest not using **h** as it looks terrible.

Maybe you need an inline URL at this point: Here's one! `https://en.wikibooks.org/wiki/LaTeX`.

Mathematics can be inline, for example $x = \int_0^\infty y^2 dy$, or can be in display mode, as shown in (4.1):

$$F(x, y, z) = \sqrt{x^{y^z}} \tag{4.1}$$

You probably want to show some of your results in a table, such as Table 4.1.

The basic creative process is shown in Figure 4.2. Specifically, Figure 4.2a shows the first step, while Figure 4.2b and 4.2c show the remaining key steps of the procedure.

4.2.1 Subtopic A

Algorithm 1 shows a classical algorithm typeset in LaTeX. This is also a float.

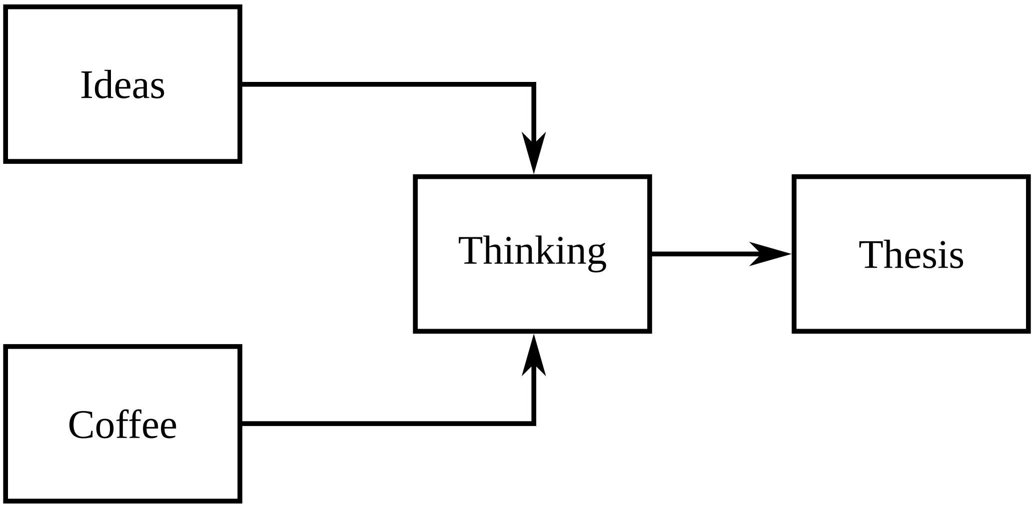
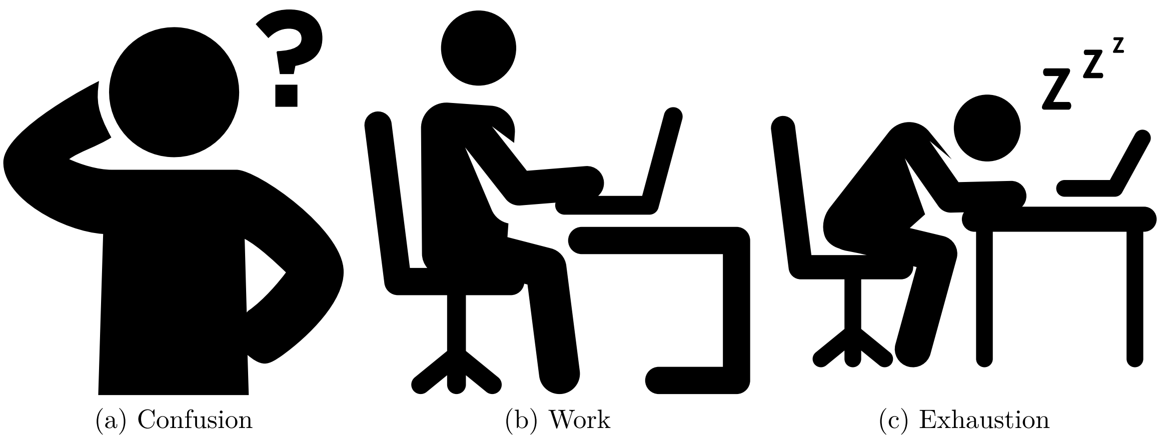


Figure 4.1: This the approximate process for producing a Thesis. It typically takes 3-4 years.

Table 4.1: Table captions normally go at the top.

Left column	μ	σ
Item 1	0.3	0.5
Item 2	0.9	0.4



(a) Confusion (b) Work (c) Exhaustion

Figure 4.2: The stages of the creative process.

Algorithm 1 An algorithm with caption

```
Require: n ≥ 0
Ensure: y = x^n
  y ← 1
  X ← x
  N ← n
  while N ≠ 0 do
    if N is even then
      X ← X × X
      N ← N/2                                  ▷ This is a comment
    else if N is odd then
      y ← y × X
      N ← N − 1
    end if
  end while
```

4.2.2 Results

4.2.3 Subtopic C

4.3 ER-type model

We generate a random numerical semigroup with a model similar to the Ërdos-Renyi model for random graphs.

Definition 4.3.1. For $p \in [0,1]$ and $M \in \mathbb{N}$, a random numerical semigroup $S(M,p)$ is a probability space over the set of semigroups $S = \langle \mathcal{A} \rangle$ with $\mathcal{A} \subset \{1, ..., M\}$, determined by

$$\Pr[n \in \mathcal{A}] = p,$$

with these events mutually independent.

4.4 Downward model

Etc. etc.

4.5 Conclusion

Chapter 5

Experiments

5.1 ER-type model experiments

In this Chapter, XXX is presented. Include pseudocode.

Section 5.2 discusses Theme 1. Section 5.3 discusses Theme 2....

5.2 Downward model experiments

5.2.1 Subtopic A

5.2.2 Subtopic B

5.2.3 Subtopic C

5.3 Theme 2

Etc. etc.

Chapter 6

Results

6.1 Introduction

In this Chapter, XXX is presented.

Section 6.2 discusses Theme 1. Section 6.4 discusses Theme 2....

6.2 Lower Bound

6.3 Expected Frobenius Number

We prove a Theorem found in [4] without the use of the simplicial complex.

Theorem 6.3.1. *Let $S \sim S(M, p)$, where $p = p(M)$ is a monotone decreasing function of M. If $\frac{1}{M} \ll p \ll 1$, then S is cofinite, i.e., the set of gaps is finite, a.a.s and*

$$\lim_{M\to\infty} E[e(S)] = \lim_{M\to\infty} E[g(S)] = \lim_{M\to\infty} E[F(S)] = \infty.$$

Proof. Let $X := \min(S \setminus \{0\})$ be a random variable. Then, for $0 < n \leq M$,

$$\Pr[X = n] = p(1-p)^{n-1},$$

and so

$$\begin{aligned} E[X] &= \sum_{n=0}^{\infty} n\Pr[X = n] = \sum_{n=0}^{M} np(1-p)^{n-1} = p\frac{d}{dp}\left[-\sum_{n=0}^{M}(1-p)^n\right] \\ &= p\frac{d}{dp}\frac{(1-p)^{M+1}-1}{p} = p\frac{1-(1-p)^{M+1}-(M+1)(1-p)^M p}{p^2} \\ &= \frac{1-(1-p)^M - M(1-p)^M p}{p} \geq \frac{1-e^{-Mp}-Mpe^{-Mp}}{p}. \end{aligned}$$

Thus, since $\lim_{M\to\infty} Mp = \infty$, then $\lim_{M\to\infty} Mpe^{-Mp} = \lim_{M\to\infty} e^{-Mp} = 0$, which implies that

$$\lim_{M\to\infty} E[X] = \lim_{M\to\infty} \frac{1-e^{-Mp}-Mpe^{-Mp}}{p} = \infty.$$

Also, note that if $p = \frac{c}{M}$, $c \in \mathbb{R}_+$ $(0 < e^{-c} + ce^{-c} < 1)$,

$$\lim_{M\to\infty} E[X] = \lim_{M\to\infty} \frac{1 - e^{-c} - ce^{-c}}{p} = \infty.$$

Proof. Fix $a \in \mathbb{N}$ such that $a > 11$ and let $A = \{1, \ldots, \lfloor \frac{a}{p} \rfloor\}$. Since $\frac{1}{M} \ll p$, we have that $\lfloor \frac{a}{p} \rfloor \leq M$ for large enough M. Consider the following events:

- E_1: No generator selected is less than $\frac{1}{ap}$.

  Let X_1 be the number of generators selected from $\{1, \ldots, \lfloor \frac{1}{ap} \rfloor\}$. Then

$$\Pr[\overline{E_1}] = \Pr[X_1 > 0] \leq E[X_1] \leq p \cdot \frac{1}{ap} = \frac{1}{a}.$$

- E_2 : At most $\frac{3a}{2}$ generators are selected from A.

  Let X_2 be the number of generators selected in A, then X_2 is a binomial random variable with $n = \frac{a}{p}$ and we can use the bound (Feller I can add this to the appendix)

$$\Pr[\overline{E_2}] = \Pr\left[X_2 > \frac{3a}{2}\right] \leq \frac{\frac{3a}{2}(1-p)}{(\frac{3a}{2} - a)^2} \leq \frac{6}{a}.$$

  Also, note that by the union bound

$$\Pr[E_1 \wedge E_2] \leq 1 - \frac{1}{a} - \frac{6}{a} = 1 - \frac{7}{a}.$$

- E_3 : At least $\frac{a}{2}$ generators are selected from A.

  Similarly, we can use the bound for the other tail of the distribution so that

$$\Pr[\overline{E_3}] = \Pr\left[X_2 < \frac{a}{2}\right] \leq \frac{(n - \frac{a}{2})p}{(np - \frac{a}{2})^2} = \frac{a - (\frac{a}{2})p}{(\frac{a}{2})^2} \leq \frac{4}{a}.$$

- E_4 : The generators selected from A are minimal.

  Let $Y_{(1)}, Y_{(2)}, \ldots, Y_{(k)}$ denote the first k generators selected in A. Assume E_1 and E_2. We have that E_1 implies $Y_{(1)} \geq \frac{1}{ap}$ and E_2 implies $k \leq \frac{3a}{2}$.

  First we bound for the probability that, given E_1 and E_2, $b \in A$ is selected as a generator. By conditional probability

$$\begin{aligned}\Pr[b \text{ is selected}] = {} & \Pr[b \text{ is selected}|E_1 \wedge E_2]\Pr[E_1 \wedge E_2] \\ & + \Pr[b \text{ is selected}|\overline{E_1 \wedge E_2}]\Pr[\overline{E_1 \wedge E_2}],\end{aligned}$$

  and so

$$\Pr[b \text{ is selected}|E_1 \wedge E_2] \leq \frac{\Pr[b \text{ is selected}]}{\Pr[E_1 \wedge E_2]} \leq \frac{p}{1 - \frac{7}{a}}.$$

  Now, note that $Y_{(2)}$ is not minimal if a multiple of $Y_{(1)}$ is selected in A. Thus, if we fix $Y_{(1)} = y_1 \geq \frac{1}{ap}$, $Y_{(1)}$ is not minimal if $b \in \{2y_1, 3y_1, \ldots, c_1 y_1\}$ is selected, where $c_1 y_1$ is the largest multiple of y_1 which does not exceed $\frac{a}{p}$. Since $y_1 \geq \frac{1}{ap}$, we have that $c_1 \leq a^2$. Then, using the union bound,

$$\Pr[Y_{(2)} \text{ is not minimal}|E_1 \wedge E_2 \wedge Y_{(1)} = y_1] \leq \frac{pa^2}{1-\frac{7}{a}}.$$

If we sum over all possible y_1, we get that

$$\Pr[Y_{(2)} \text{ is not minimal}|E_1 \wedge E_2] \leq \frac{pa^2}{1-\frac{7}{a}}.$$

Similarly, for $2 \leq t \leq k$ and fixed $Y_{(1)} = y_1, \ldots, Y_{(t-1)} = y_{t-1}$, $Y_{(t)}$ is not minimal if the first $t-1$ numbers selected from A can generate $Y_{(t)}$. For the possible numbers generated by the first t numbers selected, there are at most a^2 choices for each coefficient, so there are at most a^{2t} such linear combinations. Then

$$\Pr[Y_{(t)} \text{ is not minimal}|E_1 \wedge E_2] \leq \frac{pa^{2t}}{1-\frac{7}{a}}.$$

Therefore, since $Y_{(1)}$ is always minimal, we can use the union bound and $k \leq \frac{3a}{2}$ to conclude that

$$\Pr[E_4|E_1 \wedge E_2] \geq 1 - \frac{p}{1-\frac{7}{a}} \sum_{t=1}^{\frac{3a}{2}-1} a^{2t} = 1 - o(1).$$

Thus,

$$\Pr[E_4] = \Pr[E_4|E_1 \wedge E_2]\Pr[E_1 \wedge E_2] \geq 1 - \frac{7}{a} - o(1).$$

Finally, note that by union bound,

$$\Pr[E_4 \wedge E_3] \geq 1 - \frac{11}{a} - o(1).$$

Therefore, for every $N \in \mathbb{N}$ and $\varepsilon > 0$, there exists K such that $M \geq K$ implies

$$\Pr[f(S) > N],\ \Pr[g(S) > N],\ \Pr[e(S) > N] > 1 - \varepsilon.$$

.

6.3.1 Subtopic A

6.3.2 Subtopic B

6.3.3 Subtopic C

6.4 Lower Bound result

I conjecture that the hypothesis that q is prime can be relaxed.

6.4.1 Lemma

- Let q be a prime number and S be a random subset of $\mathbb{Z}_q$ of size $4\lfloor 3\log_2 q\rfloor$. As q tends to infinity, $2\lfloor 3\log_2 q\rfloor S$ covers $\mathbb{Z}_q$ almost always.

Let q be a prime number and let $s \in \mathbb{N}$ such that $s \leq q$. Let S be a uniformly random subset of $\mathbb{Z}_q$ of size s, that is,

$$\Pr(S) = \frac{1}{\binom{q}{s}}.$$

For a given $z \in \mathbb{Z}_q$ and $k \in \mathbb{N}$ for which $k \leq s/2$, let

$$N_z^k := \left\{ K \subset \mathbb{Z}_q : |K| = k, \sum_{t\in K} t = z \right\}.$$

Note that $|N_z^k| = \frac{1}{q}\binom{q}{k}$, since $K \in N_0^k$ if and only if $K + k^{-1}z \in N_z^k$ for every $z \in \mathbb{Z}_q$.

For $K \in N_z^k$, let A_K be the event that $K \subset S$. Let X_K be the indicator variable of A_K. We define the random variable

$$X_z = \sum_{K\in N_z^k} X_K.$$

Note that X_z counts the number of sets of size k which add up to z. We provide two ways of finding $E[X_z]$. The first one uses that, for any $K \subset N_z^k$,

$$E[X_K] = \Pr[A_K] = \frac{\binom{q-k}{s-k}}{\binom{q}{s}},$$

and so we get that

$$E[X_z] = \sum_{K\in N_z^k} E[X_K] = |N_z^k| E[X_K] = \frac{1}{q}\binom{q}{k}\frac{\binom{q-k}{s-k}}{\binom{q}{s}} = \frac{1}{q}\binom{s}{k}.$$

This motivates the second way, for we know that

$$\sum_{z\in Z_q} X_z = \binom{s}{k} = \sum_{z\in Z_q} E[X_z].$$

As in the argument for finding $|N_z^k|$, for every $z \in \mathbb{Z}_q$,

$$E[X_0] = \sum_{K\in N_0^k} E[X_K] = \sum_{K\in N_0^k} E[X_{K+k^{-1}z}] = \sum_{K\in N_z^k} E[X_K] = E[X_z].$$

Therefore, we also find that

$$E[X_z] = \frac{1}{p}\binom{s}{k}. \tag{6.1}$$

Now, for $K, L \in N_z^k$, let $j \in \mathbb{N}$ and define

$$\Delta_j := \sum_{|K\cap L|=j} \Pr[A_K \wedge A_L].$$

Fix $j \leq k$, then

$$\Pr[A_K \wedge A_L] = \frac{\binom{q-2k+j}{s-2k+j}}{\binom{q}{s}}.$$

We can bound the number of events for which $|K \cap L| = j$. First we choose K as any set in N_z^k and then we choose the remaining $k-j$ elements as any subset of $\mathbb{Z}_q \setminus K$ with size $k-j$. Thus,

$$\Delta_j \leq \frac{\binom{p}{k}\binom{q-k}{k-j}\binom{q-2k+j}{s-2k+j}}{q\binom{q}{s}}.$$

This implies that, using (1),

$$\begin{aligned}
\frac{\Delta_j}{E[X_z]^2} &\leq \frac{\binom{q}{k}\binom{q-k}{k-j}\binom{q-2k+j}{s-2k+j}}{\frac{1}{q}\binom{s}{k}\frac{1}{q}\binom{s}{k}q\binom{q}{s}} \\
&= \frac{\frac{q!}{(q-k)!k!}\frac{(p-k)!}{(k-j)!(q-2k+k)!}\frac{(q-2k+j)!}{(s-2k+j)!(q-s)!}}{\frac{1}{q}\binom{s}{k}\frac{s!}{(s-k)!k!}\frac{q!}{(q-s)!s!}} \\
&= \frac{q\binom{s-k}{k-j}}{\binom{s}{k}}.
\end{aligned}$$

Let $s = 4\lfloor 3\log_2 q\rfloor$ and $k = 2\lfloor 3\log_2 q\rfloor$, where $\alpha \in (0,1)$. Using that $\binom{s-k}{t}$ is maximized at $t = \lfloor (s-k)/2 \rfloor$,

$$\frac{\Delta_j}{E[X_z]^2} \leq \frac{q\binom{2\lfloor 3\log_2 q\rfloor}{\lfloor 3\log_2 q\rfloor}}{\binom{4\lfloor 3\log_2 q\rfloor}{2\lfloor 3\log_2 q\rfloor}} \leq \frac{q}{\binom{2\lfloor 3\log_2 q\rfloor}{\lfloor 3\log_2 q\rfloor}} \leq \frac{q}{2^{\lfloor 3\log_2 q\rfloor}} \sim \frac{1}{q^2},$$

since $\binom{2\lfloor q^\alpha\rfloor}{\lfloor 3\log_2 q\rfloor}^2 \leq \binom{4\lfloor 3\log_2 q\rfloor}{2\lfloor 3\log_2 q\rfloor}$ (I can prove this in a lemma or in the appendix).

This proves that

$$\Pr[X_z = 0] \leq \frac{\Delta}{E[X_z]^2} = \sum_{j=0}^{k}\frac{\Delta_j}{E[X_z]^2} \leq \frac{(k+1)}{q^2}.$$

Therefore, by the union bound,

$$\Pr[\bigvee_{z\in\mathbb{Z}_q} X_z = 0] \leq \frac{(k+1)}{q}.$$

6.5 Theorem

- Let $g(x)$ be a function for which $x(\log x)^2 \in o(g(x))$ .Then

$$\lim_{p\to 0}\Pr\left[F(S) \leq g\left(\frac{1}{p}\right)\right] = 1.$$

The proof of this Theorem consists of several parts. The strategy is to prove that the Ápery set of a subsemigroup of S is completed before step $g\left(\frac{1}{p}\right)$ with high probability, since $F(S)$ is less than the maximum element of this Ápery set. The proof has the following structure:

1. First, we will find a step for which a prime q is chosen with high probability (E1).
2. Then, in the spirit of the Lemma, we will find a step such that s elements, which are different modulo q, are chosen with high probability (E2).
3. Finally, we will apply the Lemma to the Ápery set of a subsemigroup of S generated by the subset in part 2.

Proof.

Part 1

Let $h(x)$ be a function such that $h(x) \in o(x(\log x)^2)$ and $x \log x \in o(h(x))$. Let $t(x) = 20x \log x$. Consider the event E_1 that there exists a prime $q \in S$, such that

$$t\left(\frac{1}{p}\right) \leq q \leq h\left(\frac{1}{p}\right).$$

Let q_n be the n-th prime number and let k_x be the number of primes between $20x \log x$ and $h(x)$. For $n \geq 6$, by the Prime Number Theorem,

$$n(\log n + \log\log n - 1) < q_n < n(\log n + \log\log n) = o(h(n)).$$

Thus, $n = o(k_n)$ (I can prove this if it is not clear) and, for every $c > 0$,

$$\lim_{p\to 0} \Pr[\neg E_1] \geq \lim_{p\to 0}(1-p)^{k_{\frac{1}{p}}} \geq \lim_{p\to 0}(1-p)^{\frac{c}{p}} = e^{-c}.$$

Therefore,

$$\lim_{p\to 0} \Pr[E_1] = 1.$$

Part 2

Now, assume E_1. Then S contains a prime number q for which

$$t\left(\frac{1}{p}\right) \leq q \leq h\left(\frac{1}{p}\right).$$

Let $s = 4\lfloor 3\log_2 q\rfloor$, as in the Lemma.

Let $A := \{q+1, q+2, \ldots, 2q\}$. Consider the event **E2** that at least s generators are selected in A. Let X_1 be the number of generators selected in A, then X_1 is a binomial random variable with parameters $n = q$ and p. Then, in a similar way to $E2$ in Theorem 1, we use the Binomial Distribution Tail Bound to show that, assuming that p is small enough so that $qp > s$ for all possible q,

$$\Pr[\overline{E_2}|E_1] = \Pr\left[X_1 < s\right] \leq \Pr\left[X_2 < s\right] \leq \frac{(n-s)p}{(np-s)^2} = \frac{(q-s)p}{(qp-s)^2}.$$

Thus, bounding by the worst case asymptotically, (needs to be explained better)

$$\lim_{p\to 0} P[\overline{E_2}|E_1] = \lim_{p\to 0} \frac{\left(h\left(\frac{1}{p}\right) - 4\left\lfloor 3\log_2 h\left(\frac{1}{p}\right)\right\rfloor\right)p}{\left(20\log\frac{1}{p} - 4\lfloor 3\log_2 t\left(\frac{1}{p}\right)\rfloor\right)^2} = 0.$$

We conclude that

$$\lim_{p\to 0} \Pr[E_2|E_1] = 1,$$

and so

$$\lim_{p\to 0} \Pr[E_1 \wedge E_2] = \lim_{p\to 0} \Pr[E_2|E_1]\Pr[E_1] = 1.$$

Part 3

Finally, assume E_1 and E_2. Let $B = \{Y_1, \dots, Y_s\}$ be a randomly selected subset of size s of the generators selected in E_2. Since the generators are chosen randomly and $|A| = q$, we can apply the Lemma to the Ápery set of the subsemigroup generated by B, denoted by $G(B)$, and conclude that the Ápery set of $G(B)$ will be completed before step $h\left(\frac{1}{p}\right) 2 \left\lfloor 3 \log_2 h\left(\frac{1}{p}\right)\right\rfloor$ with high probability as $p \to 0$.

Thus, if $g(x)$ be a function for which $x(\log x)^2 \in o(g(x))$ (Probably needs to be explained better),

$$\lim_{p\to 0} \Pr\left[F(G(B)) \le g\left(\frac{1}{p}\right)\right] = 1.$$

Since $F(S) \le F(G(B))$, we conclude that

$$\lim_{p\to 0} \Pr\left[F(S) \le g\left(\frac{1}{p}\right)\right] = 1.$$

6.6 Conclusion

Chapter 7

Conclusions and Future Work

7.1 Summary of Outcomes

7.2 Recommendations & Future Work

7.3 Concluding Remarks

In summary, ...

Bibliography

[1] D. R. Franklin and K. J. Wilson, "A LaTeX Thesis Template for the School of Electrical and Data Engineering," *IEEE Transactions on LaTeX Thesis Templates*, vol. 1, no. 1, Oct. 2021.

[2] N. Alon and J. H. Spencer, *The probabilistic method.* John Wiley & Sons, 2016.

[3] S. Chapman, R. Garcia, and C. O'Neill, "Beyond coins, stamps, and chicken mcnuggets: An invitation to numerical semigroups," *A Project-Based Guide to Undergraduate Research in Mathematics: Starting and Sustaining Accessible Undergraduate Research*, pp. 177–202, 2020.

[4] J. De Loera, C. O'Neill, and D. Wilburne, "Random numerical semigroups and a simplicial complex of irreducible semigroups," *arXiv preprint arXiv:1710.00979*, 2017.

[5] K. Frankston, J. Kahn, B. Narayanan, and J. Park, "Thresholds versus fractional expectation-thresholds," *Annals of Mathematics*, vol. 194, no. 2, pp. 475–495, 2021.

[6] J. Park and H. T. Pham, "A proof of the kahn-kalai conjecture," *arXiv e-prints*, arXiv–2203, 2022.

Appendix A

Example Appendix

Here you might present some additional results, derivations, proofs etc. that were not included in the main text.

Appendix B

Software Documentation

Here's an example source code listing, where the code is read in from an external file:

```
% Function to create a nice rotating animated GIF of 3D volumetric data V

function animation (V)

h = volshow (V, 'BackgroundColor', [0 0 0], 'Renderer', 'MaximumIntensityProjection', 'CameraPosition', [2 2 0], 'CameraUpVector', [1 0 0], 'ColorMap', jet);

camproj ('perspective');

N = 500;

filename = 'animation.gif';
vec = linspace(0, 4 * pi(), N)';
myPosition = 2 * [zeros(size(vec)) cos(vec) sin(vec)];

for idx = 1:N
% Update current view.
    h.CameraPosition = myPosition(idx, :);
% Use getframe to capture image.
    I = getframe(gcf);

    [indI, cm] = rgb2ind (I.cdata,256);
% Write frame to the GIF File.
    if idx == 1
        imwrite(indI, cm, filename, 'gif', 'Loopcount', inf, 'DelayTime', 0.05);
    else
        imwrite(indI, cm, filename, 'gif', 'WriteMode', 'append', 'DelayTime', 0.05);
    end
end
```

B.1 Code Availability

All scripts and source code used for simulation and analysis of the ... are available here:

`https://bitbucket.org/username/gitrepo.git`

B.2 Software Requirements

- MATLAB code is confirmed working with version XXXX;
- Simulations require the use of gcc version XXX or llvm/clang version YYYY

B.3 Simulation Code - How to Run