WXML Final Report: Number Theory and Noise

Sherry Chen Bobby Pedersen Nile Wynar

Winter 2018

MENTORS: DR. MATTHEW CONROY, GABRIEL DORFSMAN-HOPKINS, NIKOLAS EPTAMINITAKIS

1 Introduction

1.1 The initial problem

Our project investigates the possibilities arising from representing sequences of positive integers as sound. We wish to describe the properties of the sounds using the properties of the sequences with number theory. A digital audio file is created from a given set A of positive integers by setting sample number i to a non-zero constant c for all i in the set. All other samples are set to zero. We use the standard CD-audio sampling rate of 44100 samples per second, $\Delta t = \frac{1}{44100} = 0.0000226757$ seconds. Take the set of prime numbers as an example: whenever the program encounters a prime number that number is assigned the number 1 and all composite numbers are given the number 0. This can be seen in the waveform (Figure 1).

2 Progress

2.1 Computational

Sherry's Focuses I started my work with sequences involving prime numbers, and later focused on sequences involving digital roots. One interesting

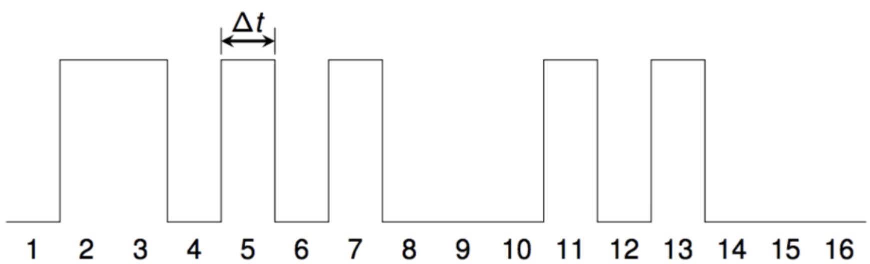

Figure 1: Example of the waveform derived from the sequence of prime numbers

sequence I have discovered is A034048: numbers with multiplicative digital root value 0.

The definition of multiplicative digital root of an integer is that multiplying the digits of this integer, then multiplying the digits of numbers derived from it, until the remaining number has only one digit. The remaining number is the multiplicative digital root of this integer.

One interesting property of this sequence is that unlike many other sequences I have encountered, of which the natural density is often zero, this sequence actually has natural density 1.

Bobby's Focuses I explored variances in Beatty sequences of different irrational numbers, as well as studying the reasons why Beatty sequences sound so similar regardless of which irrational they are based off of. I used approximations of the irrational via continued fractions as well as Fourier transforms to make determinations about why Beatty sequences sound the way they do.

Some possilbe explanations for why Beatty sequences have the particular sound they do is seen in the Fourier plots that display many frequencies within the Beatty sequences that are nearly overlapping but not quite. I assume this creates a phenomenon similar to beat frequencies and accounts for why Beatty sequences are generally unpleasant sounding.

Other reasoning for why Beatty sequences sound as they do can be seen in the fact that the difference sequences are near periodic, describing why they have such steady sounds and don't tend to vary to much from beginning to end.

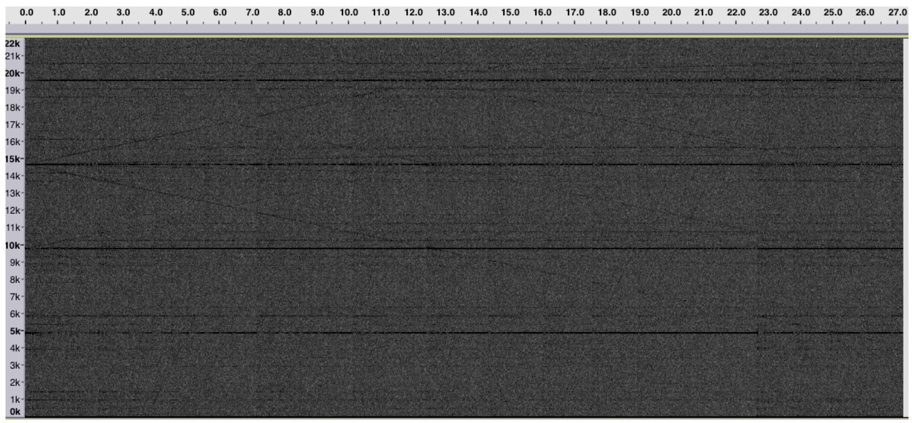

Figure 2: Spectrogram of A061910: Numbers n such that the digit sum of n^2 is a square

Nile's Focuses My goal with the sequence A061910, sum of digits of n^2 is a square, from last quarter, was to explain the diagonal lines in the sound's spectrogram (Figure 2). Since these diagonal lines persisted when created the spectrogram for A237525, sum of digits of n^3 is a cube, I assumed that the diagonal line phenomenon in the spectrogram was a property of taking the sum of digits of a number raised to a power and finding the digit sums that were integers raised to that same power. However, I found that this was not the case. I calculated that the most common digit sum of n^2 up to 10^6 is 54. Then I created the sound that corresponded to numbers n such that the sum of digits of n^2 is 54. I expected the diagonal lines to disappear, instead they were more prevalent (Figure 3). This shows that the diagonal line property in A061910's spectrogram does not come from this sequence's very specific definition, but rather something that probably arises from taking the digit sum of integers raised to a power. However, I am still not sure about this and am unable to explain this phenomenon. Hopefully, next year I can continue this exploration.

After running in circles around A061910, I decided to move on but I continued to work with the theme of digit sums as I did last quarter. I discovered an interesting sequence that I wanted to focused on, A004207. This sequence is the sum of digits of previous terms in base 10 starting with $a_0 = 1$. The n^{th} term is represented by

$$a_n = a_{n-1} + s(a_{n-1}, 10),$$

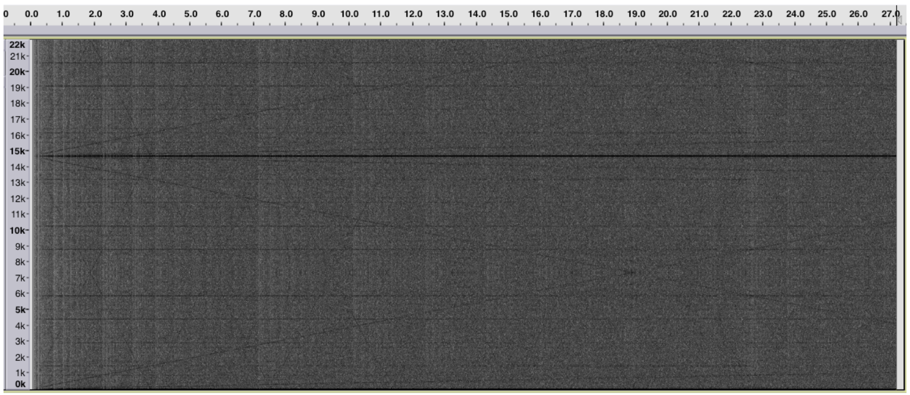

Figure 3: Spectrogram of A061910 variation: Numbers n such that the digit sum of n^2 is 54

where $s(x, 10)$ is the sum of digits of x in base 10. Judging by the spectrogram (Figure 4), one would think the sound is periodic. It looks as though the sound repeats itself every two seconds or so. However, I used Pari/GP to calculate terms in the sequence farther to see if this sound was periodic. 6666706 is in the sequence, so I created the sound starting with $a_0 = 6666706$. Although the spectrogram looks very similar, in that it has a repetitive two second pattern, the sound itself sounds different. This led me to believe that the sound is not periodic and I prove this in the section below.

There is a dark spectral line at 14701 Hz, and 44100/14701 = about 3. This arises from the fact that there are no multiples of three in the sequence A004207. This sparked my curiosity to look at this sequence in other bases. In the next section I will prove a few facts about the sequences formed by the sum of digits of previous terms in other bases.

2.2 Theoretical

Regarding A034048 (Numbers with multiplicative digital root 0) has natural density 1:

Lemma 1. *If the natural density of a set of positive integers is 0, then the natural density of its complement with respect to positive integers is 1.*

Proof. Let S be a set of positive integers. Let x $\in \mathbb{Z}_{>0}$. Then, let S(x) be the number of elements of S that are less than x. Suppose the natural density of

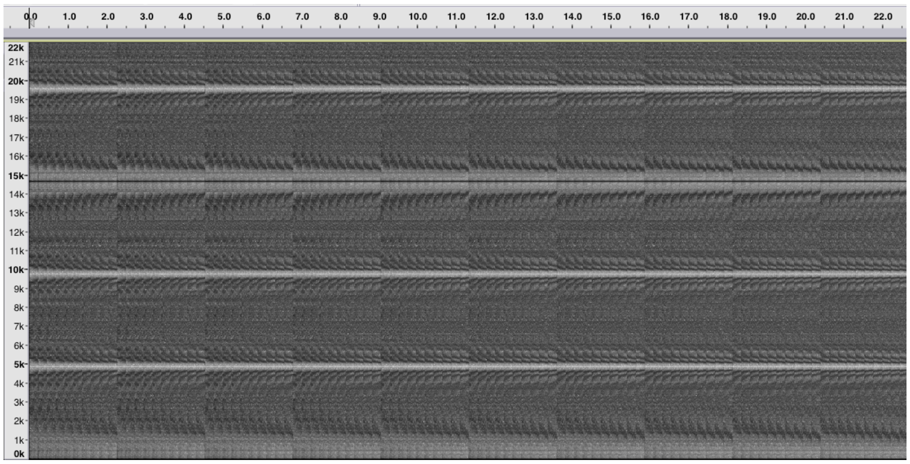

Figure 4: Spectrogram of A004207: Sum of digits of previous terms with $a_0 = 1$

S is 0; that is, suppose

$$\lim_{x\to\infty} \frac{S(x)}{x} = 0.$$

Let S^C be the complement of S with respect of $\mathbb{Z}_{>0}$. Then, let $S^C(x)$ be the number of elements of S^C that are less than x. Since S and S^C are complements of each other with respect to $\mathbb{Z}_{>0}$,

$$S(x) + S^C(x) = x.$$

Therefore,

$$\begin{aligned}\lim_{x\to\infty} \frac{S^C(x)}{x} &= \lim_{x\to\infty} \frac{x - S(x)}{x} \\ &= \lim_{x\to\infty} \left(1 - \frac{S(x)}{x}\right) \\ &= 1 - \lim_{x\to\infty} \frac{S(x)}{x} = 1.\end{aligned}$$

where the nature density of S^C is $\lim\limits_{x\to\infty} \frac{S^C(x)}{x}$. Hence, the natural density of S^C is 1. $\square$

Lemma 2. *Suppose S is a set of positive integers, and the natural density of S is 0. Then the natural density of any subset of S is 0.*

Proof. Let S be a set of positive integers. Suppose S has natural density 0. Let x $\in \mathbb{Z}_{>0}$. Then, let S(x) be the number of elements of S that are less than x. Hence,

$$\lim_{x\to\infty} \frac{S(x)}{x} = 0.$$

Let $T \subseteq S$. Let T(x) be the number of elements of T that are less than x,

$$0 \leq T(x) \leq S(x).$$

Therefore,

$$0 \leq \frac{T(x)}{x} \leq \frac{S(x)}{x}.$$

Since $\lim\limits_{x\to\infty} \frac{S(x)}{x} = 0$, according to the Squeeze Theorem,

$$\lim_{x\to\infty} \frac{T(n)}{n} = 0.$$

Therefore, the natural density of T is 0. $\square$

Lemma 3. *Let a, n be positive integers. The partial sum of a n-term power series $\sum\limits_{i=1}^{n} a^i$ is $\frac{a^{n+1}-a}{a-1}$.*

Proof. Let a, n be positive integers. Let $L = \sum\limits_{i=1}^{n} a^i$. Then

$$aL = \sum_{i=2}^{n+1} a^i,$$

and so

$$\begin{aligned} L &= \frac{1}{a-1}(a-1)L \\ &= \frac{1}{a-1}(aL - L) \\ &= \frac{1}{a-1}(\sum_{i=2}^{n+1} a^i - \sum_{i=1}^{n} a^i) \\ &= \frac{a^{n+1}-a}{a-1}. \end{aligned}$$

$\square$

Theorem. *A sequence of positive integers A034048 is defined as numbers with multiplicative digit root value 0. The natural density of A034048 is 1.*

Proof. Let x $\in \mathbb{Z}_{>0}$. Let $A(x) = |\{a \in \mathbb{Z}_{>0} : a \notin A034048 \wedge a \leq x\}|$. To prove that the natural density of A0304048 is 1, we want to show that $\lim_{x\to\infty} \frac{A(x)}{x} = 0$.

Let F(x) $= |\{f \in \mathbb{Z}_{>0}$: f has no zero digits $\wedge\ f \leq x\}|$. For every integer with multiplicative digit root value $\neq 0$, it must has no zero digits. Therefore, if $\lim_{x\to\infty} \frac{F(x)}{x} = 0$, then $\lim_{x\to\infty} \frac{A(x)}{x} = 0$, and so the natural density of A034048 is 1.

Assume x has n digits. Then

$$x \geq 10^{n-1},$$

and

$$F(x) \leq \sum_{i=1}^{n} 9^i = \frac{9^{n+1} - 9}{8},$$

where 9^i is the number of i-digit integers which have no zero digits.

Therefore,

$$\frac{F(x)}{x} \leq \frac{\frac{9^{n+1}-9}{8}}{10^{n-1}},$$

where we obtain $\lim_{n\to\infty} \frac{\frac{9^{n+1}-9}{8}}{10^{n-1}}$ as following:

$$\begin{aligned}\lim_{n\to\infty} \frac{\frac{9^{n+1}-9}{8}}{10^{n-1}} &= \lim_{n\to\infty} \left(\frac{9^{n+1}}{8 \times 10^{n-1}} - \frac{9}{8 \times 10^{n-1}} \right) \\ &= \lim_{n\to\infty} \left(\frac{81}{8} \left(\frac{9}{10} \right)^{n-1} - \frac{9}{8} \left(\frac{1}{10} \right)^{n-1} \right) = 0.\end{aligned}$$

Since $\lim_{n\to\infty} \frac{\frac{9^{n+1}-9}{8}}{10^{n-1}} = 0$, $\lim_{x\to\infty} \frac{F(x)}{x} = 0$. Hence, $\lim_{x\to\infty} \frac{A(x)}{x} = 0$, and the natural density of A034048 is 1. □

Regarding $a_n = a_{n-1} + s(a_{n-1}, b)$**:**

Definitions

- A010062, A010063, A010065, A010066, A010068, A010069, A010071, A010072, A004207: sum of digits of all previous terms in base 2, 3, 4, 5, 6, 7, 8, 9, and 10 respectively with $a_0 = 1$

- $\mathrm{s}(n, b)$ = sum of digits of n in base b

- $a_n = a_{n-1} + s(a_{n-1}, b)$, represents the terms of a sequence usually defined as the sum of digits in base b of all previous terms

- d_n, the difference sequence of a_n in base b, $d_n = a_n - a_{n-1} = s(a_{n-1}, b)$ for $n > 0$

- $N(n, b)$ = number of digits of n in base b

Theorem. *The waveform created from a sequence of the form* $a_n = a_{n-1} + s(a_{n-1}, b)$ *is not periodic.*

Proof. Assume that the waveform of a_n is periodic. Define the following characteristic equation that is used to define the waveform,

$$\chi(n) = \begin{cases} 1 & \text{if } \exists i : n = a_i \\ 0 & \text{if } \forall i : n \neq a_i \end{cases}.$$

Since $\chi(n) = 1$ corresponds to the top of the sound waves, the difference sequence represents the distance between the peaks of the waves, so it is necessary for d_n to be periodic in order for the waveform of a_n to be periodic. Since d_n is positive and finite for all n, this implies that d_n is bounded. So there exists a real number $M > 0$ such that $d_n < M$ for all natural numbers n. We will produce M consecutive integers, each of which will have a sum of digits greater than M. Let

$$x_0 = \sum_{j=0}^{M-1} b^{N(M,b)+j} = (\underbrace{1\ldots1}_{\text{M ones}}\ \overbrace{0\ldots0}^{\text{N(M, b) zeros}})_b$$

So define $x_n = x_0 + n$ for $1 \leq n < M$. x_n is defined with $N(M, b)$ many zeros. Let

$$L = (1\ \underbrace{0\ldots0}_{\text{N(M, b) zeros}}\ )_b > M$$

so adding M to L, does not change the leading 1 of L. So $s(x_n, b) > M$ for all x_n. If $a_i < x_j$ and $a_{i+1} > x_j$ for all i and all x_j between x_0 and x_M, then $a_{i+1} - a_i > M$ since $x_M - x_0 = M$. But $a_{i+1} - a_i = d_{i+1} < M$. So since x_n consists of M many consecutive integers and since $d_n < M$ for all n, for at least one i and j, $a_i = x_j$. So then $s(a_i, b) > M$, but $d_{i+1} = s(a_i, b) > M$ so this contradicts that the difference sequence is bounded. So the waveform of a_n is not periodic. □

Theorem. *For $n > 0$, $a_n \equiv 2a_{n-1}$ (mod $b-1$).*

Proof. Let

$$N = (n_m \ldots n_0)_b = \sum_{i=0}^{m} n_i \cdot b^i$$

represent N in base b. So $N \equiv \sum_{i=0}^{m} n_i = s(N, b) \pmod{b-1}$. Since $a_n = a_{n-1} + s(a_{n-1}, b)$, then $a_n \equiv a_{n-1} + a_{n-1} \pmod{b-1}$ so $a_n \equiv 2 \cdot a_{n-1} \pmod{b-1}$. □

Corollary 1. *The sequence A004207 mod 9, is (1 2 4 8 7 5) repeating and none of its terms are multiples of three.*

Proof. $b = 10$, so $b - 1 = 9$. By the theorem above, $a_n \equiv 2a_{n-1} \pmod 9$. $a_0 = 1$ so $a_1 = 1$. Hence, $a_1 \equiv 1 \pmod 9$ which implies that $a_2 \equiv 2$ (mod 9), so $a_3 \equiv 4 \pmod 9$, $a_4 \equiv 8 \pmod 9$, $a_5 \equiv 16 \equiv 7 \pmod 9$, $a_6 \equiv 14 \equiv 5$ (mod 9), $a_7 \equiv 10 \equiv 1 \pmod 9$, and so the cycle continues. If $n = 3m$ for some integer m, then $n \equiv 0$ or 3 or 6 (mod 9). Since (1 2 4 8 7 5) does not contain 0, 3, or 6, then there are no multiples of 3 in the sequence. □

Corollary 2. *For $n > 2$, all terms of the sequence A010066 are multiples of four.*

Proof. $b = 5$, so $b-1 = 4$. $a_0 = 1$ by definition, so $a_1 = 1$ and by the theorem above, $a_2 \equiv 2 \pmod 4$, so $a_3 \equiv 4 \equiv 0 \pmod 4$. So for $n \geq 3, a_n \equiv 0$ (mod 4), so all future terms of A010066 are multiples of four. □

Nile's Contribution to the OEIS The Pari code for A004207 and A010066 on the OEIS was impossible to understand so I added simpler code so that people in the future can more easily access these sequences. Below are my contributions to A004207 and A010066 respectively.

```
(PARI) a = 1; print1(a, ", "); for(i = 1, 50, print1(a, ", ");
a = a + sumdigits(a)); \\ Nile Nepenthe Wynar, Feb 10 2018
```

```
(PARI) a = 1; print1(a, ", "); for(i = 1, 40, a = a + sumdigits(a, 5);
print1(a, ", ")); \\ Nile Nepenthe Wynar, Feb 10 2018
```