WXML Final Report: Number Theory and Noise

Nile Wynar　　Lisa Yan

Fall 2017

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

1 Introduction

1.1 The initial problem

Our project investigates the possibilities arising from representing sets 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). Because we created our sounds from a variety of integer sequences, there was no reason to expect the results to be melodious, and they weren't.

1.2 New directions

This quarter, we decided to explore more with the program PARI/GP, especially using it for large primes. Initially, we tried to program our own

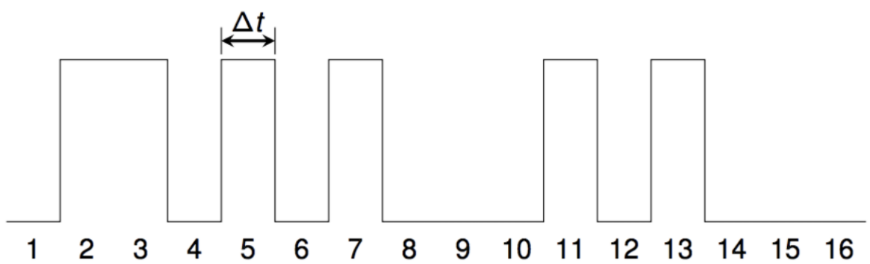

Figure 1: Example of the waveform derived from the sequence of prime numbers

primality test in Python; however, it wasn't efficient and would take too long. PARI/GP implements an efficient primality method called "isprime". One minor problem with PARI/GP is that it takes longer to process for more numbers and in order for our noises to be longer enough for us to analyze, we need over a million numbers in a sequences. In addition, we decided to start looking at the spectrogram and analyzing the anomaly in the spectrograms we dont understand. For example, the diagonal lines in A061910s spectrogram.

2 Progress

2.1 Computational

Nile's Focuses I was most interested in sequences involving the digit sum. They are relatively periodic because of properties of the digit sum. Since $digitsum(n) = m \Rightarrow digitsum(n + 10^k - 1) = m$ most of the time – with the exception of numbers n that end in zeros – we can see why the sound would be repetitive and basically periodic. Also, sequences such as these can be decomposed into sequences of the form $A_n = \{n \in \mathbb{N} \mid digitsum(n) = \{a\}\}$. For example, the sequence A028839 which is the numbers n such that the digit sum of n is a square can be decomposed into three separate sequences who add up to the whole sequence. This is because the maximum possible digit sum for integers up to $1.2 * 10^6$ is 55 so there are only six possible squares $\{1, 4, 9, 16, 25, 36, 49\}$. So we can deduce that the total wave form can be written as a sum of six different waveform. However, since a digit sum of 1, 4, 9 and 49 are so unlikely, I approximated A028839 with the three

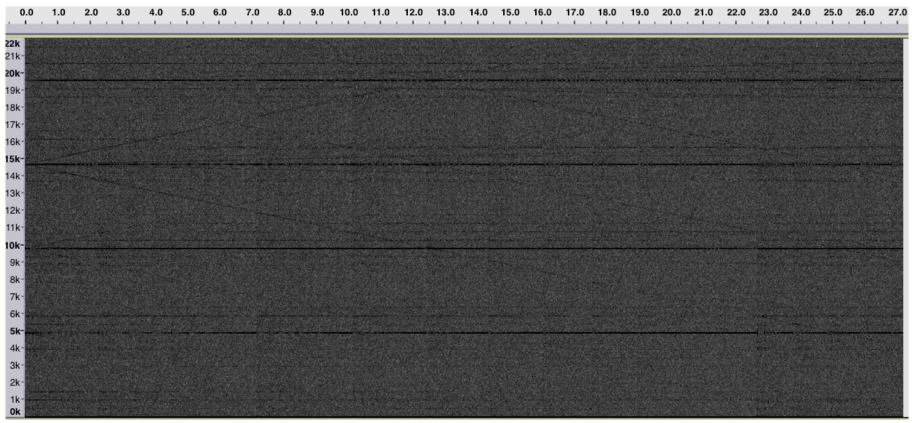

Figure 2: Spectrogram of A061910: Numbers n such that the digit sum of n^2 is a square

sequences that represent the $digitsum(n) = \{4^2\}$, $digitsum(n) = \{5^2\}$, and $digitsum(n) = \{6^2\}$. This process can be used on any sequence of this form and so we can always determine how many separate wave forms are needed for the decomposition of the original sequence.

Another sequence I was interested in was A061910 (Figure 2) which is the sum of digits of n^2 is a square. The spectrogram of this sound has diagonal lines that we have not been able to explain yet. I created A061910 with the base two digit sum and noticed that these diagonal lines start at the bottom left corner, which is different than where they start with the base ten digit sum. To try to understand this phenomenon, I created the sequence A237525, which is numbers n such that the sum of digits of n^3 is a cube. This sound's spectrogram also had these diagonal lines, however they are curved. I plan to vary the set that determines which digit sums we take in our sequence and the power we raise the integer to to see if either of these variables change the diagonal lines in the sound's respective spectrogram. Hopefully, this will help us understand what property of this sequence creates these diagonal lines.

Lisa's Focuses I was interested in the Beatty sequences, especially the sequence of $\left\lfloor \sqrt(2)n \right\rfloor$ taking the floor of the positive multiples of a positive irrational number. First, I generated various Beatty sequences to see if I can find a pattern, for example, I generated A022844, A286428, and A001951. It seems that all the differences of the sequences seem to be 1 or 2. Later,

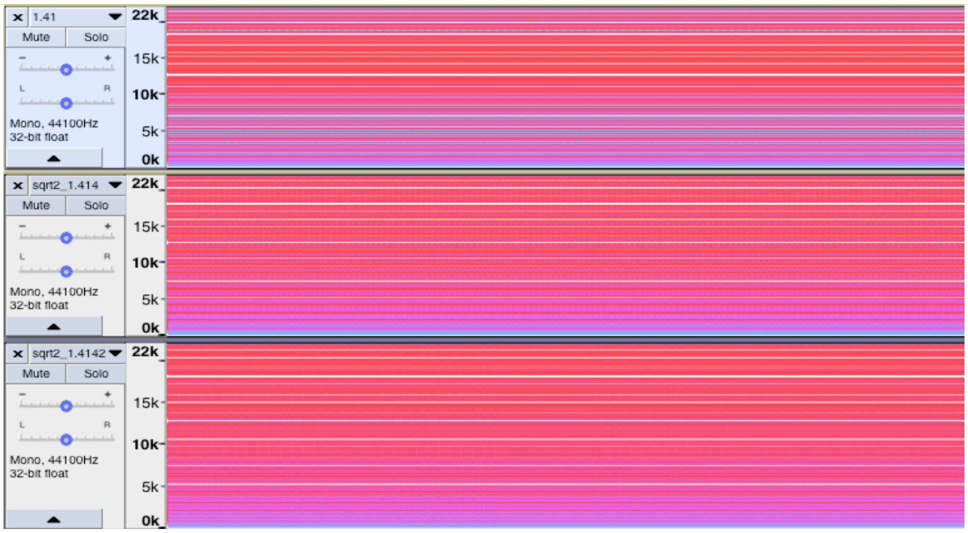

Figure 3: Spectrograms of three sequences generated using decimal approximation of $\left\lfloor\sqrt(2)n\right\rfloor$

I started just focusing on A001951, which is the sequence of $\left\lfloor\sqrt(2)n\right\rfloor$. I was interested in how the programs interpret irrational numbers, so I used both decimal approximation and continued fraction approximation to estimate the sequence. I made a series of sound with $\lfloor 1.4n\rfloor$, $\lfloor 1.41n\rfloor$, $\lfloor 1.414n\rfloor$, and $\lfloor 1.4142n\rfloor$ (Figure **??**). When I first made the sequence of $\lfloor 1.4n\rfloor$ with Python, the spectrogram shows clicking sounds even though there shouldnt be any. We quickly realized that the clicking sounds were due to a rounding error in Python. We generated the sequence again in PARI/GP and the clicking sounds disappeared. Because of the rounding errors with decimal places, I decided to use fractions and approximate $\sqrt{2}$ using continued fractions. I made another set of sounds with $\left\lfloor\frac{7}{5}n\right\rfloor$, $\left\lfloor\frac{17}{12}n\right\rfloor$, $\left\lfloor\frac{41}{29}n\right\rfloor$, $\left\lfloor\frac{99}{70}n\right\rfloor$, $\left\lfloor\frac{239}{169}n\right\rfloor$, and $\left\lfloor\frac{577}{408}n\right\rfloor$. One interesting observation about the spectrograms is that as the approximation becomes more accurate, there are more horizontal lines in the spectrogram (Figure 4). This is because as the fraction gets closer to $\sqrt{2}$, the denominator of the fraction gets larger. As the denominator gets larger, the fundamental frequency is lower, and because our range only goes up to $22,000$, we can see more horizontal lines on the spectrogram.

ss

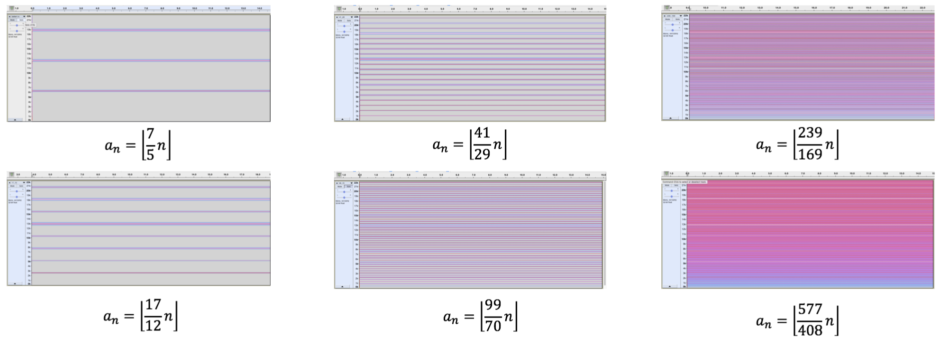

Figure 4: Spectrograms of all the sequences generated using continued fraction of $\left\lfloor \sqrt(2) n \right\rfloor$

2.2 Theoretical

Regarding $\lfloor n \log n \rfloor$: We attempted to prove that the differences of this sequence's difference sequence is only in the set $\{-1, 0, 1\}$. We were able to get upper and lower bounds for the difference sequence $f(n+1) - f(n)$, where $f(n) = \lfloor n \log n \rfloor$, using the approximation for the natural log of n since $\log(n+1) - \log(n) < \frac{1}{n}$.

$$1 + \lfloor \log n \rfloor \leq f(n+1) - f(n) \leq 2 + \lfloor \log n \rfloor$$

Which yields that the difference sequence, $d(n) \leq 2$. At least this doesn't contradict our original goal, however we still have work to do.

Regarding $\left\lfloor \frac{p}{q} n \right\rfloor$: We observed some interesting patterns in sequences in the form $\left\lfloor \frac{p}{q} n \right\rfloor$, in which p and q are relatively prime, and started to prove them:

- The sequence of differences of $a_n = \left\lfloor \frac{p}{q} n \right\rfloor$, which is $a_n = \left\lfloor \frac{p}{q} n - \frac{p}{q}(n-1) \right\rfloor$, seems to have a period of q.
- The sequence of $a_n = \left\lfloor \frac{p}{q} n \right\rfloor \; mod \; q$ seems to have a period of p.

We have proved that the sequence of difference of $\left\lfloor \frac{p}{q} n \right\rfloor$, which is $\left\lfloor \frac{p}{q} n - \frac{p}{q}(n-1) \right\rfloor$, does have a period of q. However, we still need to prove that q is the minimal number required for a period to happen through the fact that p and q are relatively prime.

Regarding A295389: We contributed a sequence to the Online Encyclopedia of Integer Sequences! The sequence is the numbers n such that the sum of digits of n is squarefree.

3 Future directions

Next quarter, we will continue exploring sequences involving the base ten digit sum as well as other bases. In addition, we wish to understand the spectrograms of some of our sequences, such as, the spectrogram of integers n such that the sum of digits of n^2 is a square, $a_n = \lfloor n \log n \rfloor$, and the spectrograms of the approximation sequences of the Beatty sequence $a_n = \left\lfloor \sqrt(2) n \right\rfloor$. Since we've looked at the sequences of differences in various sequences, we wish to compute the density of sequences within certain intervals and see how these intervals affect the respective densities. Regarding the Beatty sequences, we wish to understand Beatty sequences with different approximations of the irrational part, and prove our observations: 1.The period of the sequence of differences of $a_n = \left\lfloor \frac{p}{q} n \right\rfloor$ is q. 2.Prove that the period of the sequence $a_n = \left\lfloor \frac{p}{q} n \right\rfloor$ *mod* q is p. Finally, we want to continue discovering more novel sequences and contribute them to the OEIS.