WXML Final Report: Number Theory and Noise

M. Conroy, G. Dorfsman-Hopkins, Xinwei Fan, Elana Lessing, Christine Wolf

Spring 2016

1 Introduction

Students investigated properties of integer sequences via sound files created for each sequence. For an increasing sequence x_n of positive integers, a sound file is created by setting the x_n-th sample to a constant, non-zero value, and all other samples to zero. The CD-standard sampling rate of 44100 samples per second was used.

2 Quadratic residues - Christine Wolf

The Quadratic Residues subproject of the Number Theory and Noise project for the WXML attempted to explore the large-scale behavior of quadratic residue sequences for various moduli, as well as how this behavior related to particular characteristics of the chosen modulus. Because of the rapid information absorption enabled by the use of sound files, characteristic patterns and features were able to be detected even with relatively large moduli. Additional data collection programs were also used to supplement the data from the sound files, such as a program that read and recorded the size and frequency of occurence of gaps between residues.

The initial investigation revolved largely around prime moduli. It was discovered that the sound files generated from the quadratic residues of prime moduli tended to have distinctive features – a noisy foreground sound with a distinct pulsing "ping" in the background. This was consistent across the tested primes but contrasted with numbers that were numerically close to the primes, suggesting that characteristics of quadratic residue sequences are more determined by the factorization of the modulus than by its numerical value. This idea was further supported by the simple to prove lemma that if $m = kn$ for some integers k, m, and n, then all quadratic residues of n are also residues of m. That is, adding additional prime factors to a number can only remove quadratic residues.

This property was further explored through moduli that were pure powers of primes. The powers of any given prime tended to have a pure sound that decreased in pitch as the power increased, which made sense given the previous lemma. These prime powers also tended to have the same sizes of gaps across powers, as well as approximately the same ratios – for example, powers of 7 seem to have gaps of size 1, 2, 3, and 4 only, with the ratio between them fairly consistent no matter what the power.

Adding even one additional prime factor to these pure prime sequences usually quickly eliminates the pure quality of the sound. The natural conclusion that it does this by upsetting the formerly predictable size and ratio of gaps seems to hold true. Through looking at the spectrograms of sound files created from prime powers and prime powers with an additional distinct prime factor, we can see that the prime powers tend to have one dominant white line in their spectrogram at a particular frequency. The lines of both primes involved in the modulus with two distinct prime factors appear in its spectrogram, but as the power of the prime factors increase the lines become more and more blurred until they are barely distinguishable. This makes sense with the increasing "noisiness" of the sound as power increases.

2.1 Future directions

Current data suggests that there may be a pattern in the number, size, and distribution of gaps between quadratic residues of a given modulus, particularly when the modulus is prime or a prime power. Because the noise of the sound file for a sequence is determined by the gaps in the sequence, examination of the sound files coupled with other methods of data collection might yield more interesting discoveries.

3 Digit-based sequences - Elana Lessing

I was researching sound files produced by digits based sequences. In specific, my research was on two different families of digit based sequences. The first family is composed of sequences generated by integers whose digit's sum is dividable by a certain natural number and the second family is composed of sequences generated by integers whose product of their digits are under a certain natural number.

However, before I begin discussing my results it is also important to note that when sounds are generated from integer sequences, the sound is not the values of the sequences themselves. Instead, because of the way our sound files are generated, the sound represents the gaps (or lack of) between different values of the sequence. In addition, I would like to clarify some terminology. I will be discussing repeating patterns in the sound files and when I do so I am referring to points where the sound produced is identical or the value of the gaps between terms of the sequence is repeated.

3.1 Family One

As I mentioned before, this family of sequences are defined by all integers whose digit's sum is divisible by a certain integer m. However, I would like to construct a more rigorous definition. To do so, let us define $\{x_n\}$ the sequence defined using the natural number m and n as a arbitrary natural number.
Note that if n has i digits and we can rewrite n as the following:

$$n = 10^i n_i + 10^{i-1} n_{i-1} + \cdots + 10 n_2 + n_0$$

Using this form of n, we can say that

$$n \in \{x_n\} \text{ if m} \left| \sum_{k=1}^{k=i} a_k \right.$$

In this analysis I will be discussing sequences for all natural numbers m such that $0 < \text{m} < 10$. One of the first features I noticed was that all of the sound files I have generated have repeating patterns. Some of these patterns are longer or shorter, but the sound files will always repeat themselves eventually, which means that the pattern of gaps between different values of the sequence will repeat itself.

While there are patterns of various lengths, in the sequences I have generated there have only been two sequences for which the pattern has an length of 1; when $\text{m} = 3$ and $\text{m} = 9$. These two sequences sound very distinct from the others and I will be dividing my analysis into three parts accordingly (the first part when $\text{m} = 1$, the second when $\text{m} = 3$ or $\text{m} = 9$ and the third part for all the other cases).

3.1.1 m = 1

I have including $\text{m} = 1$ more for the sake of completeness than because there is anything interesting to say about the sound files for particular sequence. Every natural number is disable by 1 so $n \in \{x_n\}$ for all $n \in \mathbb{N}$, which means that this sequence has a density of 1 in $\mathbb{N}$. Since every natural number is in this sequence, the sound generated from this sequence is silence.

3.1.2 m = 3, 9

The first terms of the sequence $\{a_n\}$ generated using $\text{m} = 3$ are $\{3, 6, 9, 12, 15, 18, 21, 24, 27, 30, ...\}$[1]
The first terms of the sequence $\{b_n\}$ generated using $\text{m} = 9$ are $\{9, 18, 27, 36, 45, 54, 63, 72, 81, 90, ...\}$[2]

When playing these sound files, it is very apparent that both of these sequences generate a high pitched sound and that both of these sounds are very pure. There is no inconsistencies or any clicking sounds in these sound files. To discuss why this is, we can start by observing the terms of these sequence that I have shown. When $\text{m} = 3$ the sequence appears to be equivalent to the sequence of all multiples of 3 and when $\text{m} = 9$ the sequence appears to be equivalent to the sequence of all multiples of 9.

To show that this is true for the entire sequence we can use the well known result that that a number is divisible by 3 or 9 if and only if the digits add up to a number that is divisible by 3 or 9 respectively. From this result, it follows that $\{a_n\}$ can be rewritten as $\{a_n\} = 3n$ and is the sequence of every third number. In addition, it follows that $\{b_n\}$ can be rewritten as $\{b_n\} = 9n$ and is the sequence of every ninth number.

We can now use this result to discuss the density of these sequences. Since a_n is every third number and b_n is every ninth number, we know that a_n is less dense that b_n. This is reflected in

[1]Sequence A008585 in OEIS
[2]Sequence A008591 in OEIS

the sound files because a_n generates a much higher sound than b_n does. In fact, the sound file generated by a_n is so high pitched that it can be difficult to hear for some people.

The question then becomes when is the exact density of these sequences. Since we know a_n is composed of every third integer and b_n is composed of every ninth integer, is would make sense to say that the density of a_n in $\mathbb{N}$ is $\frac{1}{3}$ and the density of b_n in $\mathbb{N}$ is $\frac{1}{9}$. Note that this is also saying that, when m = 2 and m = 3, the density of the sequences generated in $\mathbb{N}$ are $\frac{1}{\mathrm{m}}$. However, this is only an informal argument and not a formal proof, an issue I will be addressing later.

3.1.3 m = 2, 4, 5, 6, 7, 8

The first terms of the sequence $\{c_n\}$ generated using m = 2 are $\{2, 4, 6, 8, 11, 13, 15, 17, 19, 20, ...\}$[3]
The first terms of the sequence $\{d_n\}$ generated using m = 4 are $\{4, 8, 13, 17, 22, 26, 31, 35, 39, 40, ...\}$[4]
The first terms of the sequence $\{e_n\}$ generated using m = 5 are $\{5, 14, 19, 23, 28, 32, 37, 41, 46, 50, ...\}$[5]
The first terms of the sequence $\{f_n\}$ generated using m = 6 are $\{6, 15, 24, 33, 39, 42, 48, 51, 57, 60, ...\}$[6]
The first terms of the sequence $\{g_n\}$ generated using m = 7 are $\{7, 16, 25, 34, 43, 52, 59, 61, 68, 70, ...\}$[7]
The first terms of the sequence $\{h_n\}$ generated using m = 8 are $\{8, 17, 26, 35, 44, 53, 62, 71, 79, 80, ...\}$[8]

While the cases where m = 3 and m = 9 were very pure sounds, it is immediately apparent that this group of sequences are different at first listen. In contrast to the pure sounds discussed before, all of these sounds have many clicks in them. While you could play the m = 3 and m = 9 sound files at any point and you would hear the same sound, this is no longer true as the sound changes from moment to moments (though it does still repeat patterns). However, we can still talk about some very interesting patterns in these sequences.

For example, there is the sequence $\{c_n\}$, which is generated with m = 2. If you observe the terms I have shown above, you can notice that this is approximately every other term of the sequence. However, there are some exceptions and notably all of these exceptions fall on a value disable by 10.[9] To talk about why this occurs, we need to discuss the behavior of this sequence in a bit more depth. In specific, we know that when the sum of all the digits besides the 1s digit has a certain parity, 1s digits of the same parity will correspond to terms in the sequence.[10] This means that while the parity of the sum of the digits besides the 1s digit stays constant, every other term will be in this sequence. So it follows that only times the pattern can be broken is when the parity of the sum of the digits besides the 1s digits changes, which is only possible around values divisible by 10.

---

[3]Sequence A054683 in OEIS

[4]Sequence A268620 in OEIS

[5]Sequence A227793 in OEIS

[6]I have submitted this sequence to OEIS and it is currently pending review

[7]Sequence A273159 in OEIS

[8]Sequence A273188 in OEIS

[9]Note that there are values disable by 10 that are not exceptions. For example; 98, 100 and 102 are all terms of this sequence.

[10]If the sum of all the digits besides the 1s digit has an even parity than adding an even 1s digit to this will result in an even number, which is divisible by 2 and in the sequence. However, adding an odd 1s digit will result in an odd number, which is not divisible by 2 and not in the sequence.
If the sum of all the digits besides the 1s digit has an odd parity than adding an odd 1s digit to this will result in an even, which is divisible by 2 and in the sequence. However, adding an even 1s digit will result in an odd number, which is not divisible by 2 and not in the sequence.

This pattern becomes very noticeable in the the sound file. The sound files generated by $\{c_n\}$ is a series of fairly regular clicks with some moments of irregularity. These moments of irregularity correspond to the exceptions that I was discussing above. While the sound files is moving though values of the sequence too quickly for your ear to catch the individual points, all these exceptions combined disrupt the sound enough for you to hear them.

While this pattern can not be articulated as eloquently for larger values of m, for all values of m you can use a similar idea to construct a regular pattern with some exceptions present which generates irregularities in the sound files.

Next, I would like to discuss the some other elements of these sounds files. As the m values grow larger, the sound produced becomes lower. While the amount of clicks and variation within a single sound file make it a bit difficult to discuss the pitch variation quantitatively, I can say that the difference in the pitch very is noticeable when the sound files are played after one another. In addition, it is very noticeable that the clicks are spaced further apart from each other as the m values increases. Knowing this information, it would make scene that the density of the sequences in $\mathbb{N}$ decreases as m grows larger. This is also backed up by the terms of the series I have included. Averaging the gaps between the first 10 terms (these are the terms that I have included above) we get the following chart

m	Average gap
2	2
4	4
5	5
6	6
7	7
8	8

While this is certainly not a proof, we can see and clear pattern where m (when m = 2, 4, 5, 6, 7, 8) and the average gaps of the first 10 terms of the sequence are equivalent. Using this information, we can begin to discuss the density in $\mathbb{N}$ of these sequences. We know that roughly every mth natural number gives us a terms of the sequence so we can say that the density is approximately $\frac{1}{\text{m}}$ in $\mathbb{N}$. However, I would like to again clarify that this is a informal argument and not a proof.

3.1.4 Next Steps

So far, I have made informal arguments about the density of these sequences in $\mathbb{N}$ when $0 < \text{m} < 10$. You can see what I have claimed the densities are for varies m values below:

m	Density claim
1	1/1
2	1/2
3	1/3
4	1/4
5	1/5
6	1/6
7	1/7
8	1/8
9	1/9

Using this table, I can generalize this claim to be that the density in $\mathbb{N}$ of these sequences is $\frac{1}{\text{m}}$.

Proving this claim is my overall next step. In order to do this, I have started proving it for specific values of m. I believe I have proved the density in $\mathbb{N}$ is $\frac{1}{3}$ for when m $= 3$ and I have begun to prove the density of additional m values. I am hoping to do this for all the sequences discussed here (when $0 < \text{m} < 10$). After I complete this, I would then move on to generalizing proof for all values of m $\in \mathbb{N}$.

3.2 Family Two

Definition

I was looking at sequences which is defined as all the integers for which the product of their digits is less then a certain integer m. However, there is one additional rule that is needed in the definition. Most numbers contain a zero digit in them so a majority of numbers were being included in the sequence regardless m I choose which was not creating very interesting sound files. In order to remedy this I have been counting 0 terms as 1 terms in my products.

To construct a more rigorous definition of the sequence, let us say that $\{x_n\}$ is the sequence constructed using the natural number m and let us say that n is a arbitrary natural number with i digits.
Note that if n has i digits and we can rewrite n as the following:

$$n = 10^i n_i + 10^{i-1} n_{i-1} + \cdots + 10 n_2 + n_0$$

Using this form of n, we can say that $n \in \{x_n\}$ if

$$\text{m} < \prod_{k=0}^{k=i} n'_k$$

when

$$n'_k = \begin{cases} 1 & \text{if } n_k = 0 \\ n_k & \text{if } n_k \neq 0 \end{cases}$$

3.2.1 Analysis

All of these sounds have a similar sounding files. However, these sound files do not have a repetitive pattern as the previous family did. Instead, these sound files are composed of a chucks of sound following by a chuck of silence. As the sound files plays, the chucks of sound become progressively shorter and the chunks of silence become larger. Eventually, there is silence with an occasional sound being played.

The question then becomes why does this pattern occur. Let us use the case of
m $= 40$ as an example. We know that $3 * 4^2 = 48$ which greater than 40, so 344 will not be in out sequence. However, we can also say that 334 through 339 will also be excluded from our sequence because we are increasing the value of our product. In addition, since we are discounting digits with the value of 0, we can say that 3340 through 3390 is excluded from out sequence. In fact, we can say that $334 * 10^n$ through $339 * 10^n$ will be excluded from our sequence for any natural number

n. Note not only that this interval will get larger as n increases, but that this interval is not the only interval that will do this. Any value that is not in the sequence can be used to construct an interval similar to this one, all of which will grow larger as n increases and exclude more values from this sequence.[11] However, there will never be a point where there are no remaining values in these sequence. Using a similar technique, we can say that 1 will be in this sequence so $1 * 10^n$ will be in this sequence for any natural number n.

While all the sound files generated using this technique follow a similar pattern, there is differences depending on the m value chosen. Most notably a smaller m value generates a sequence that becomes largely silence much faster simply because there are more combinations of numbers that result in a product larger that the m value.

3.2.2 Next Steps

While I have some observations on this family of sequences, I have yet to quantify them. While I have some ideas for how to do so, perhaps by comparing the density of values in this sequence under a certain natural number for different values of m, I have not yet applied this technique. A long term goal would be finding a formula for the density under a certain natural number in terms of m and that natural number.

4 Long sounds and approximations - Xinwei Fan

For the first half of the quarter, I was concentrating on making long sound files of integer sequences like primes and abundant numbers. Sound files made using integer sequences with an upper bound of 10^6 produces a time of 23 seconds. While this is long enough for a quick sense of what the noise is like, sequences may not reveal audible properties unless over long periods of time. In this case, for some sequences, I tried to create and listen to long sound files to find out what we can hear in them. The biggest challenge I met is to generate the number sequences used for producing sounds. Because long sound files, as long as 30 minutes, require huge amount of numbers. This made my progress really slow, since I spent several days just to generate the number sequences in text files.

For the second half of my research, I changed the way I looked into the sequences: instead of trying to hear long sound files, I wanted to figure out how much density matters to sounds. To do this, I simulated sound of abundant number sequence through density. The natural density of the set of abundant numbers is between 0.2474 and 0.2480. The way I do this is by generating random sequence with the same density and see how close can the sounds be. For random sequence, I included all multiple of 6, all multiples of 20, and randomly generated the rest. The result is cheerful: the audio with multiples of 6 and 20 sounds pretty similar to abundants than audio of sequence only contains multiples of 6. From here I start to feel like I am producing something others never tried, and how intriguing can mathematics be. If I am going to continue this project, I think I will try to discover the relationship between density and audios further in other number sequences.

[11] I have chosen this particular interval as an example, but the same is true for any interval of numbers which are not in the sequence