

# Exploring Characteristics of Rhythm Structure Using Statistical Mechanics

Robert St.Clair<sup>1</sup> and Jesse Berezovsky<sup>1</sup>

<sup>1</sup>Department of Physics, Case Western Reserve University

May 3, 2023

## Abstract

We apply methods from statistical mechanics to explore a model of musical rhythm. We have created a model that yields the probability of a beat occurring in each discrete time bin. Using these probabilities, we can calculate note length distributions. Our research has shown that there theoretically exist ordered and disordered rhythmic phases based on quantities known as temperature and chemical potential. We have seen our theoretical models reflect note distributions of known musical compositions and believe that these derived models, although simplified, support intuitive beliefs about rhythm that we have known.

Code, sheet music, and MIDI files used in this paper can be found [here](#).

# 1 Motivation & Introduction

New and exciting studies have emerged recently that model harmonies using statistical mechanics. In his 2019 paper, Jesse Berezovsky has shown that harmonies can be modeled using statistical mechanics, and these models reflect traditional patterns that we see in the construction of musical pieces [1]. This is important because modeling music in this specific way is a new way of looking at music theory that may help us learn more about why we perceive things as pleasant sounding.

We seek to develop a similar model that can be used to better understand rhythmic structure. We hope that our research will help push the boundaries of music composition as well as lead to a deeper understanding of how we perceive rhythm.

In *Rhythm and Transforms*, William Sethares defines rhythm as a periodic phenomenon, which is characterized by period, phase, and ordered elements [2]. It is also proposed that we can represent a single period of rhythm by defining an underlying time grid where sites on this grid can take on either a value of “0” or “1” depending on if a note is absent or present. This time grid is a repeating vector of time points, so no matter how the start point pattern is shifted, it can be viewed as the same rhythm [2].

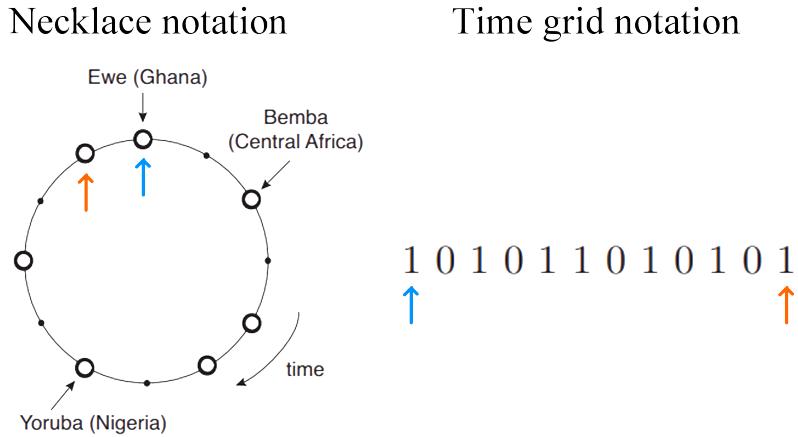


Figure 1: “Necklace notation” (circular rhythmic notation) of three African rhythms: the Ewe of Ghana, the Bemba of Central Africa, and the Yoruba of Nigeria. Note that these are all the same periodic rhythm despite different starting points (phases). Below is the time grid representation of this rhythm with the start in phase with the Ewe rhythm [2].

Using these simplifying assumptions, as well as some we developed, our core model assumes the following:

- Three beats are perceived as rhythmic if they are spaced equally in time [3].
- Humans can only perceive sounds so fast (around 5-10 ms), so we can represent time as a set of discrete time bins when modelling a rhythm.
- Each individual time bin has an independent probability of a beat occurring.
- Composers and songwriters must balance order and complexity of a rhythm. Too complex and the rhythm is too random and not recognizable enough. Too ordered and it becomes too generic and unexciting.

Under these assumptions, we can assume that there is some parameter that governs the defined trade-off between rhythmic order and complexity. In a physical system, this quantity is temperature [4]. Assuming that each time bin has a probability of holding an on beat and that the probability of an on beat does not affect the probabilities of the other time site, we can use a Fermi-Dirac distribution to model the probability of these spaces being occupied, and then use these probabilities in our time grid to model rhythm.

## 2 Model

We can define rhythmic structure as a one-dimensional chain of discrete moments in time where a beat or note can occur. Each segment of this chain represents a probability of there being a “1” in that space of the time grid. For simplicity, we divide these discrete time bins such that the smallest note in a rhythm occupies exactly one time bin rather than making the smallest division 10 milliseconds as mentioned above. This is due to the fact that not all rhythm operates at such a speed. A chain of occupied time states (1-1-1-1) denotes a continuous beat of this smallest note. For example, let us say this note is an eighth note. A chain of every other state (1-0-1-0) would then represent a continuous beat of quarter notes.

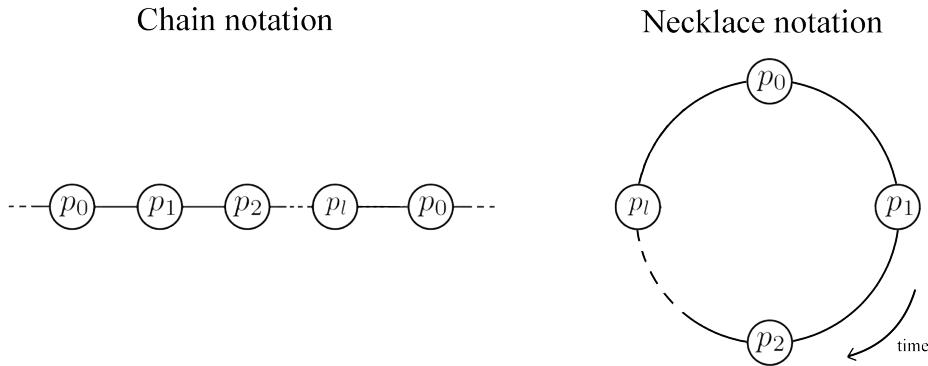


Figure 2: The chain of probabilities for each time bin, shown in a linear chain notation as well as the necklace notation. With  $l$  time bins, there are  $l$  probabilities within this chain that each represent an individual time bin. These probabilities repeat in a similar way to the time grid. If any of these probabilities lie between 0 and 1, then variations in the time grid will occur as the probabilities repeat.

Under our assumptions, it is reasonable for us to utilize a model derived from Fermi-Dirac statistics. We use Fermi-Dirac statistics to model rhythm because it uses the assumption of non-interacting particles to describe the macroscopic properties of a system [4], and allows for the introduction of a temperature term  $T$  as well as a chemical potential term  $\mu$ . Like in physical systems, we expect temperature to govern the structure of rhythm with lower temperatures resulting in more rigid structured rhythm and higher temperatures to result in more disordered but also more complex rhythms. The quantity of  $\mu$  relates to particle density in physical systems. In this model, it is a quantity that is related to the density of beats. We expect to see phase transitions occur between more ordered and less ordered phases at boundaries, another observation from when the same models are applied to physical systems.

We modify the model slightly by replacing the quantity of energy with a term we call rhythmicity. Rhythmicity is a quantity that describes how rhythmic a configuration of probabilities is based on how likely it is that a note is present in each time site. We calculate the rhythmicity by looking at a note and seeing when other notes are likely to occur as well as how spaced they are from each other.

At each possible site, rhythm is “perceived” based on the notes that precede and follow this point in time. If there is a note present in the current time bin, then the rhythm perceived depends on if there are notes in equally distant time bins [3]. For some distance  $s$  away from the current time bin and some maximum range of interaction  $N$ , we can say that the total perception of rhythm is equal to

$$R_{tot} = \frac{1}{N\bar{B}^2} \sum_{j=-\infty}^{\infty} B_j \sum_{s=1}^N B_{j-s} B_{j+s} \quad (1)$$

Where  $\bar{B}$  is the average occupancy of all time bins, and  $B_j$  denotes the on or off value of a note in a time bin. A nonzero term is added to the outer sum if and only if there is a note present in the current bin as well as notes present a distance  $s$  away in both directions. We can additionally calculate the *change* in rhythmicity for a given site  $n$ , removing the need for an infinite sum. Since  $B_n$  can be present in Equation 1 as  $B_j$ ,  $B_{j-s}$ , or  $B_{j+s}$ , we can substitute in the values  $n = j$ ,  $n = j - s$ , and  $n = j + s$  to this equation and solve for  $\Delta R_n$  for these three cases:

$$\Delta R_n = \frac{1}{N\bar{B}^2} \sum_{s=1}^N (B_{n-s} B_{n+s} + B_{n+s} B_{n+2s} + B_{n-s} B_{n-2s}) \quad (2)$$

If we assume a looping chain of  $L$  independent probabilities  $p_l$  as discussed in the previous section, then we can make the assumption that the average occupancy of a single repetition of this chain is  $\bar{B} = \frac{1}{L} \sum_{l=0}^{L-1} p_l$ . Utilizing a mean field approximation [4] such that the interaction with each time site is given by its average value, we obtain the following equation for the change in rhythmicity:

$$\Delta R_n = \frac{L}{(\sum p_l)^2} \sum_{l=0}^{L-1} (p_{-l} + 2p_l + p_{2l}) \quad (3)$$

Where  $p_{\pm} = p_m$  with  $m = n \pm l \pmod{L}$  and  $p_{2l} = p_k$  with  $k = n + 2l \pmod{L}$ . Making the adjustment for rhythmicity as the energy term in the Fermi-Dirac distribution, the resulting probability for a beat to be present at each site  $l$  in time is:

$$p_l = \frac{1}{1 + \exp(-(\Delta R_l + \mu)/T)} \quad (4)$$

This choice of model works well because it encapsulates this trade-off between order and complexity that is present in countless pieces of music. As mentioned, we expect there exist different rhythmic phases that depend on rhythmic temperature just like how there are different phases of matter that occur at different temperatures. As water cools, it turns from disorderly gaseous vapor to somewhat ordered water to a structured solid ice. Similarly, we should expect to see rhythmic structure transition from disordered patterns to ones with more structure as our rhythmic system cools.

The grand potential of a configuration [4] can be described as

$$\Omega = U - TS - \mu N \quad (5)$$

Where  $U$  is the energy of a configuration,  $S$  is the entropy, and  $N$  is the number of particles in a state. If we substitute  $U$  for the quantity  $\bar{R}_{tot}$ , the total rhythmicity, and  $N$  for the average probability across all states  $\bar{p}$ , then this equation becomes

$$\Omega = -\bar{R}_{tot} - T\bar{S} - \mu\bar{p} \quad (6)$$

where the average total entropy  $\bar{S}$  across  $L$  time bins can be defined as

$$\bar{S} = \frac{1}{L} \sum_l p_l \log p_l \quad (7)$$

When the grand potential is minimized, we then have an equilibrium solution. This is reflected in the Landau theory of phase transitions, where a phase is defined by the equilibrium solutions that are minimums of a free energy function [5]. In our model, the grand potential serves as such a function. Phase transitions are characterized by a change in the amount of stable equilibrium solutions, or when a stable solution changes from a minimum to a maximum and becomes unstable.

### 3 Method

We seek to analyze how our proposed model stands up to musical compositions present in the real world. To do this, we cover the steps taken to solve for the probabilities that a note will occur at each independent time site as well as how we use this information to create a phase diagram. We will then cover how we select musical compositions to analyze and the steps we take to analyze them.

#### 3.1 Theoretical Distributions

Using the model that we have created, we are able to find the probabilities that a note will occur in each time bin,  $p_l$ . Our first method of solving for these probabilities was to solve Equation 4 for the each probability  $p_l$ . For a probability vector of two beat sites, this is not that difficult to solve. We solved it analytically and graphically and obtained two equations relating the probability of the first site (site A) to the probability of the second (site B). Below are plots that are used to determine the probabilities of two given temperature and chemical potential pairs by finding the intersection between these two equations.

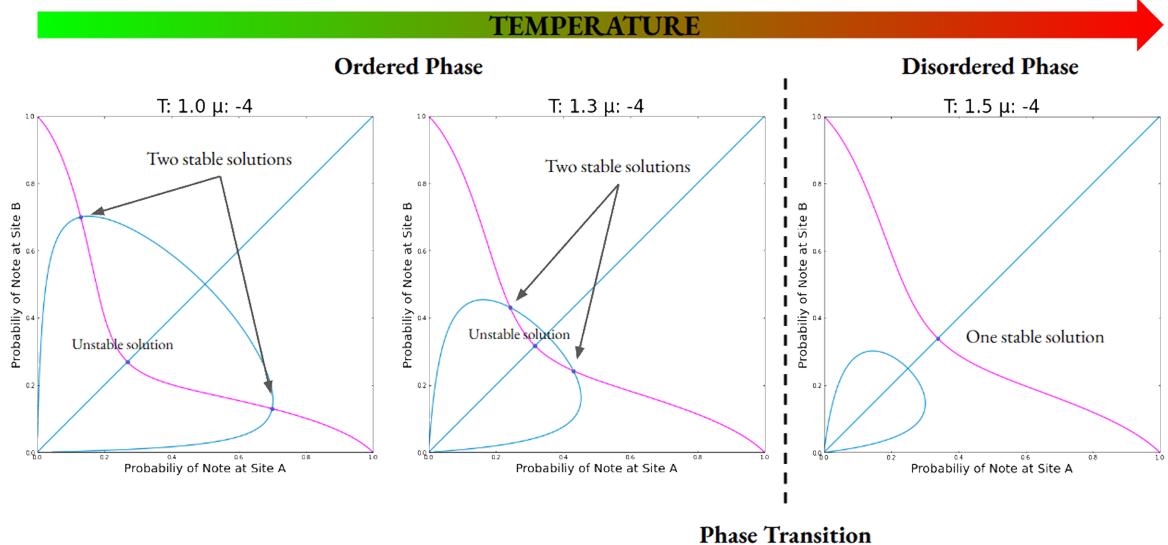


Figure 3: The graphical solutions for the two time bin model at increasing temperatures, with temperatures of 1.0, 1.3, and 1.5 for the left, middle, and right plots, respectively. As the temperature increases, the solutions get closer together and we notice a shift from three solutions (two stable) to one singular stable solution, characteristic of a phase change.

The probabilities of each site can be shown by where the curves intersect. These solutions for the probabilities are shown in Figure 3. For a temperature of 1.0, we found that there are two stable solutions which are symmetric across the line  $y = x$ . These two solutions can be described by the following relations:  $f_{A1} = f_{B2}$  and  $f_{A2} = f_{B1}$ , where  $f_A$  is the probability of a note present in site A and  $f_B$  corresponds to the probability of a note in site B. The third solution, where  $f_{A3} = f_{B3}$  is an unstable solution. This is due to the fact that when we calculate the grand potential (free energy)

of our rhythmic system using Equation 6, this particular solution is at a local maximum of the grand potential.

As temperature increases, we see a shift in the number of solutions. As we shift from three solutions to one, the solution where  $f_A = f_B$  becomes stable as the free energy at that point switches from a local maximum to a minimum. This shift is evidence of a phase transition as temperature increases. Indeed, if the probabilities of both sites are equal (and not 0 or 1), then we can treat every time bin as being random. Any rhythm generated by this distribution will be unstructured as there is no variation in the probabilities.

This same solution techniques can be used with more probability sites, but it becomes both analytically difficult to do. (Moving up from two-dimensional to four or eight-dimensional space is not necessarily the easiest to visualize!) We then solve for these probabilities using iteration since the rhythmicity in Equation 4 depends on the probabilities,  $p_l$ , so these probabilities occur on both sides of the equation. We iteratively solve for each  $p_l$  simultaneously then, and stop iterating once the values have converged.

After obtaining solutions for a range of  $T$  and  $\mu$ , we can look at the site probabilities and obtain information describing the structure of our underlying probability chain. We can then use this information to create a phase diagram for our model.

### 3.2 Use of Known Compositions

If we believe that our model can represent known music, then we must compare the distributions obtained by our model to note length distributions in known compositions. For the scope this paper, we limit our data of known compositions to pieces that are able to be represented well by MIDI files. We choose MIDI files because notes within this data structure are encoded with discrete time values. This way, we can easily measure when notes fall into certain time bins. Since MIDI files can have up to 16 different tracks, all with different instruments, compositions that are represented by this file type often separate melodies and accompaniments as well as instruments. This allows us to isolate a particular part of the song (generally the dominant melody) to analyze with our model.

We limit our choice of MIDI files to those that also have corresponding sheet music. This way, we are able to check that the note length distributions that we obtain from these pieces actually reflect the written music. It also allows for all pieces that we look at to have notes at standard time divisions. We use the website MuseScore to obtain both MIDI files and sheet music for the compositions that we look at [6]. The website has a wide selection of both classical and contemporary music for us to choose from. We did not want to limit our selection to any certain composer or era of music.

To parse this information and obtain note length distributions, we utilize the Python library Mido, which can decode MIDI files into their tracks and notes while keeping speed, timing, tone, track, and instrument information [7]. Each note in a MIDI file is stored as a “message”, and each one of these messages has a time associated with it, which denotes when the note occurs. Unlike traditional sheet music, MIDI files do not break notes up into measures. These note messages contain  $\Delta t$  information instead, telling us the time between notes. Using this, we can determine how far apart each note is and come up with a length for it. When notes occur at the same time, such as a chord, we do not weigh this any differently than one note when determining the note length distribution. We also standardize the note length information by dividing all lengths by the minimum note length found. This way, all notes can be expressed as a multiple of the smallest note in the composition.

We also make the simplifying assumption that rests are rhythmically the same as a held note. This is because the probabilities calculated with our model are for the likelihood of a rhythmic hit occurring at that time bin, not a sustained tone. A scheme of how we parse MIDI files is shown in Figure 4.

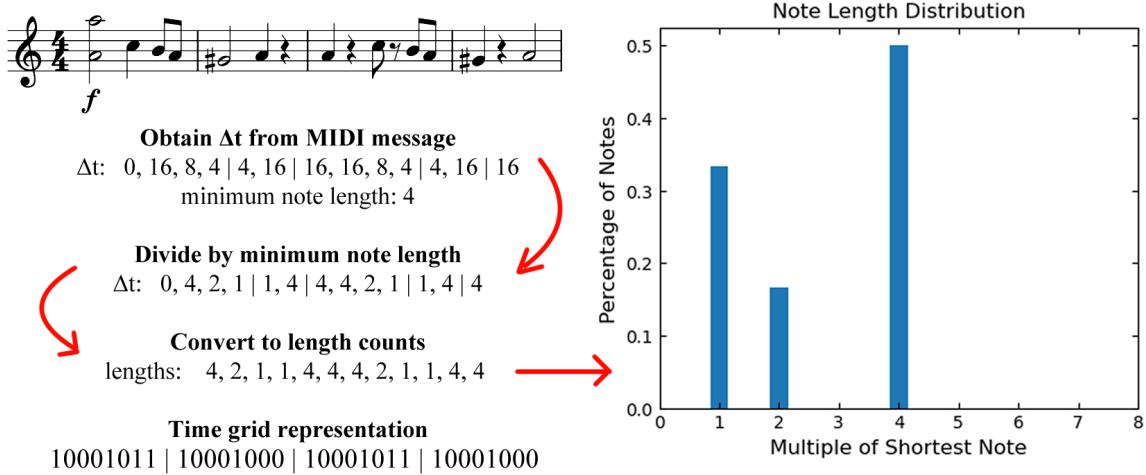


Figure 4: An example rhythm and an outline of how our code parses a MIDI file. We first store the  $\Delta t$  values provided by the MIDI messages in an array and then divide by the smallest nonzero  $\Delta t$  found. We then omit any lengths of zero, as these represent chords or the start of a piece. We then have a vector with counts of each note length that can be plotted as a histogram. When we plot the histogram, we normalize the distribution as to compare with theoretical ones from our model. A time grid representation can also be obtained as a result of our parsing.

Our last choice in selecting music is that we only chose to select pieces written in 2,4, or 8 as the current model that we are using has eight sites. Keeping the beats per measure in our selected compositions to these numbers allows us to see if our model does properly reflect rhythm on this scale. If we were to look at our model with 3 or 9 independent sites, then we would want to use music written in 3 rather than 4. Changing the number of time bins would require the recalculation of all the probabilities at each  $T$  and  $\mu$  pair and the creation of a new phase diagram. We seek to investigate this and discuss it more in the Future Plans section.

## 4 Results

### 4.1 Phase Diagram

Having conceptually shown how probabilities are calculated and that there is evidence of different rhythmic phases, we move on to a larger site model to analyze rhythm. Using a eight-site model, we calculated the probabilities  $p_l$  at a range of temperature  $T$  and chemical potential  $\mu$ . The changes in a probability chain with a constant  $\mu = -6.0$  over a temperature sweep from 0.1 to 2.0 are shown in Figure 5.

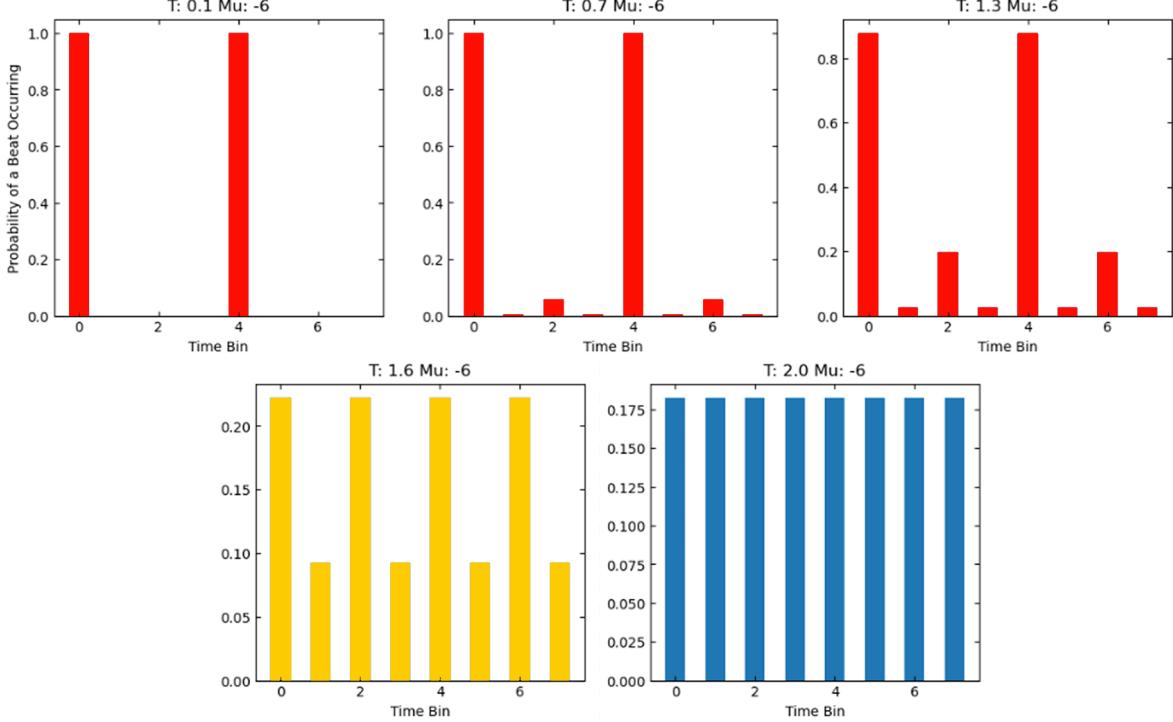


Figure 5: Plots of  $p_l$  at a constant  $\mu = -6.0$  as  $T$  is increased from 0.1 to 2.0. The plots are colored as to indicate the system at different phases, which can be seen by the number of periods contained within these  $L = 8$  sites: two in the red plots, four in the yellow plots, and eight in the blue.

Doing this over a range of  $\mu$ , we can obtain similar probability distributions. We can then compare the heights of key peaks and use the ratio between them to gain knowledge about the structure of these probabilities, such as the frequency of repetition of the same probabilities. In order to visualize these differences between probability sites, we assign RBG values to these specific ratios:

- Red: The ratio between the highest peak and the peak halfway (peaks 0 and 4). If the peaks are the same, then there is at least a repetition of the same pattern every 4 sites.
- Green: Ratio between the highest peak and peak 2. This color tells us about the repetition every two sites, or if the same probabilities are repeated four times within a configuration.
- Blue: Ratio of the first two peaks, which tells us if the same probability is repeated between two adjacent time bins.

When combined, these three colors allow us to see if the rhythm structure repeats every 2 sites, 4 sites, or 8 sites. The resulting phase diagram is shown below in Figure 6.

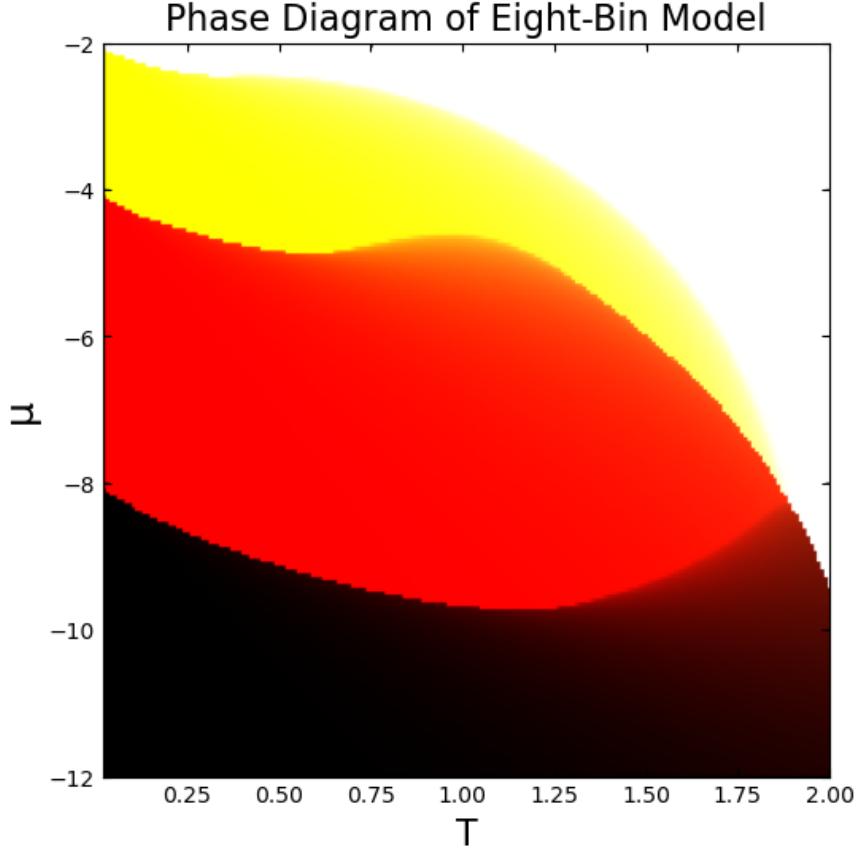


Figure 6: The phase diagram for  $L = 8$  time bins. There are four distinct phases that appear: one for probabilities repeating every eight sites (black), one for every four (red), one for every two (yellow), and a phase where all sites have the same probability (white). This last phase can be considered as a “disordered” phase, since there is no observed structure in the probabilities.

## 4.2 Rhythm Structure in Different Phases

Using our phase diagram, we explored the rhythmic structure in each of these phases as well as how they relate to each other. Using the  $p_l$  values from the model, we randomly generated a large time grid where each  $l$ th time bin was generated by the corresponding  $p_l$ . With a large enough generated time grid, we can calculate the note length distribution by counting the space between note occurrences. This is similar to reversing the last steps described in Figure 4. At a large enough grid, the normalized note length counts converge on their true values. These generated note length distributions are how we compare our model to known music.

In our phase diagram investigation, we first explored some example distributions in each phase. I will refer to the white phase as the disordered phase, the yellow phase as the four period phase, the red as the two period phase, and the black as the single period phase. Example distributions from various points throughout the phase diagram are shown in Figure 7.

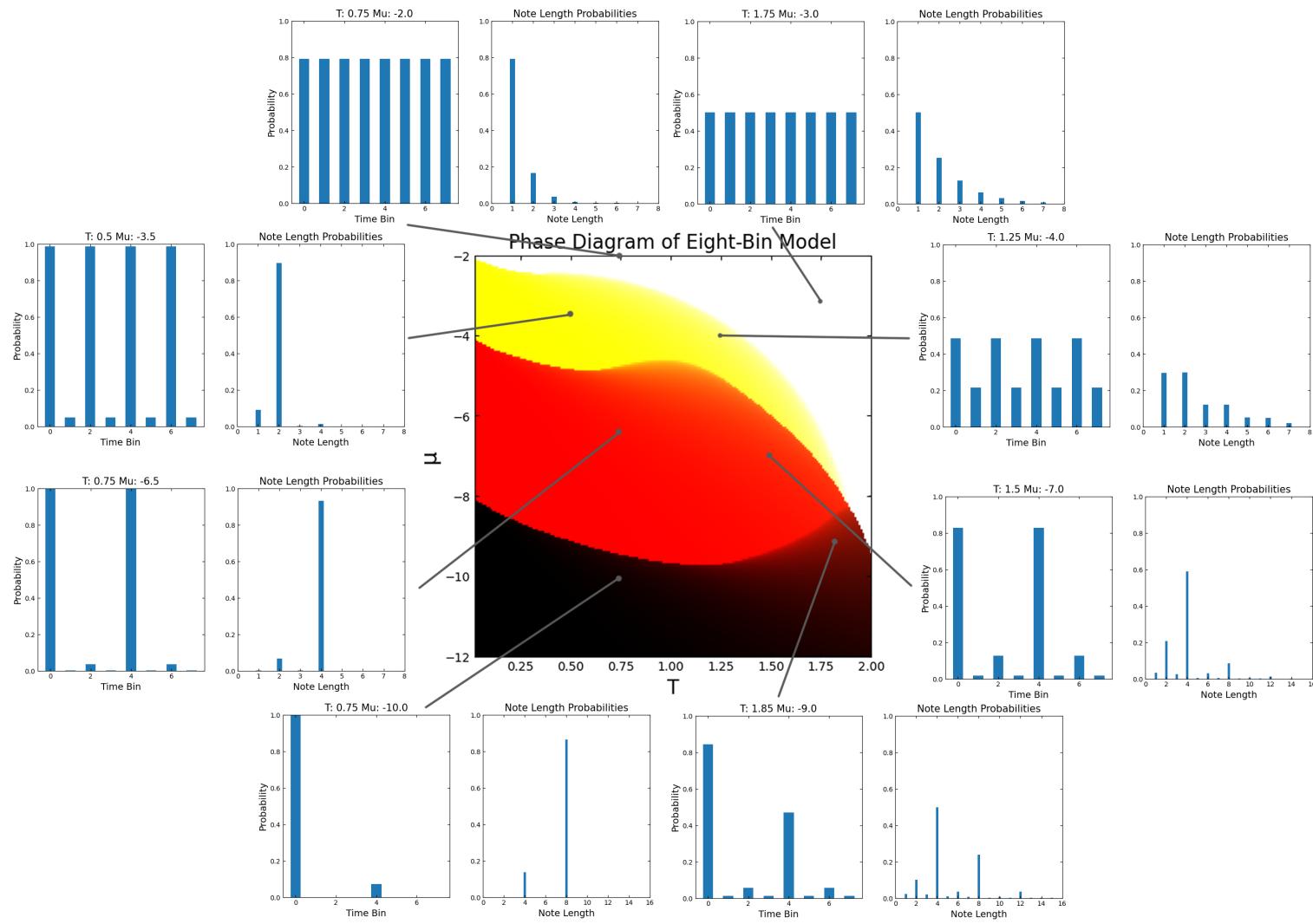


Figure 7: The probability and note length distributions for  $L = 8$  time bins at various points on the phase diagram. Similarities between phases are observed.

Notable observations include that the disordered phase has note length distributions that appear very Poisson-like. Such a distribution is characteristic of random independent events in a span of time, which adds to the belief that this phase is rhythmically disordered. We also notice the general trend of increased variation as  $T$  increases, which is indicated by the presence of more “unusual” note lengths at higher temperatures of the same phase. For example, at  $T = 0.75$  and  $\mu = -6.5$ , the only possible notes have a length of 2 and 4, which maps well to eighth and quarter notes. However, at  $T = 1.5$  and  $\mu = -7.0$ , we see the emergence of half notes, and small amounts of sixteenth, dotted eighth, and dotted quarter notes as possibilities. We also see incredibly low probabilities for certain note lengths that have no conventional name. One might assume that such notes do not have conventional names because they are not used often in common rhythmic structure.

We observed that our model can produce the same distributions at different points on our phase diagram. Remarkably similar rhythm patterns are found within different phases, as shown by Figure 8. Although the probability configuration gets more spread out as we reduce  $\mu$ , at these specific points there is virtually no probability at the sites between the present peaks. This results in note length distributions with notes of double the value in each lower phase. If we play a rhythm generated by the solution at  $T = 1.9, \mu = -11.5$  at a speed four times that of a rhythm generated from  $T = 0.5, \mu = -3.0$ , we should expect to see the same patterns emerge. We could easily write all of these states as a two-bin system with probabilities of 0.98 and 0.13 in each time bin as long as they are synced up in tempo. Due to this, it then seems that our model is able to produce degenerate rhythmic states.

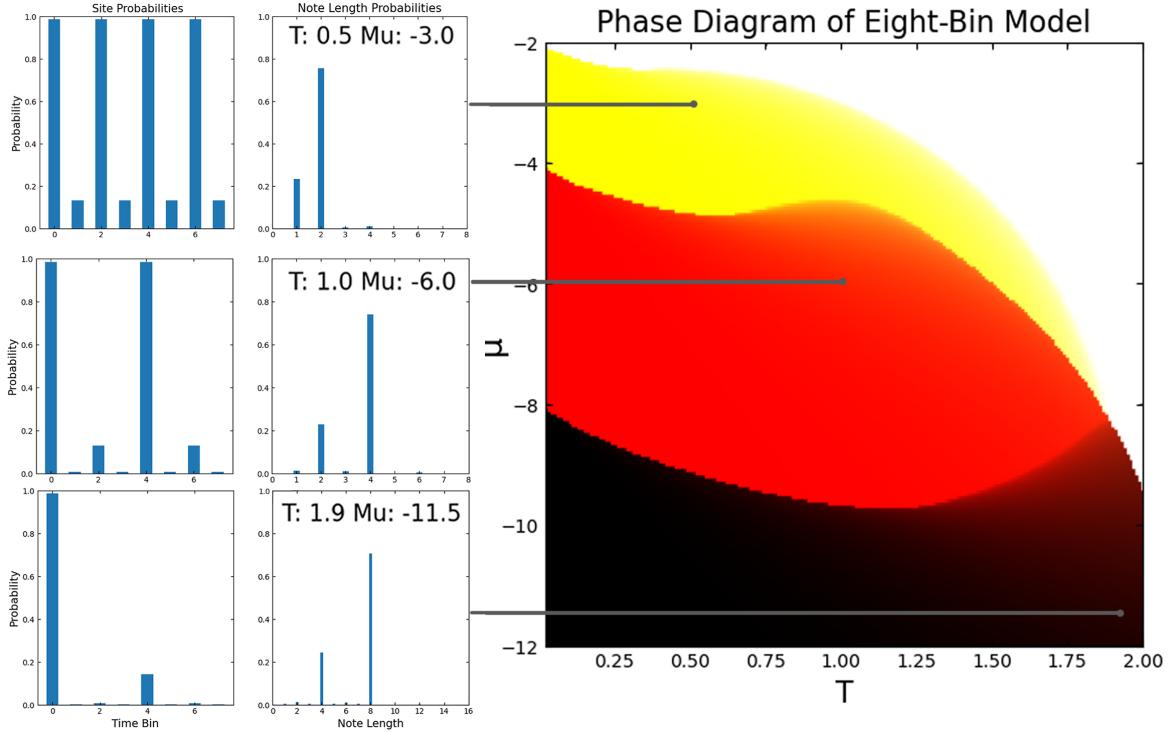


Figure 8: Evidence of the same distributions being obtained at different points in different phases on our diagram. If the rhythms generated by the points in the two period and single period phases are sped up by a factor of two and four respectively, then they will play at the same speed as the rhythm generated by the point in the four period phase and be qualitatively identical. This can be seen as evidence of degenerate states within our model.

### 4.3 Comparison to Known Compositions

In this section, we analyze a selection of known compositions and observe if our model is able to reflect the note length distributions obtained from these pieces of music. The sheet music for these pieces can be found in the GitHub repository mentioned in the Introduction section. We attempted to select a wide variety of musical genres in order to see how our model performs under a “more full” spectrum of music. A wide selection of music may also result in the observation of rhythmically similar compositions that may not intuitively seem alike at first glance. We might expect songs belonging to similar genres to be clustered because rhythmic characteristics may be similar.

The note distributions of various pieces along with close representations of these distributions generated by our model are shown in the figures below.

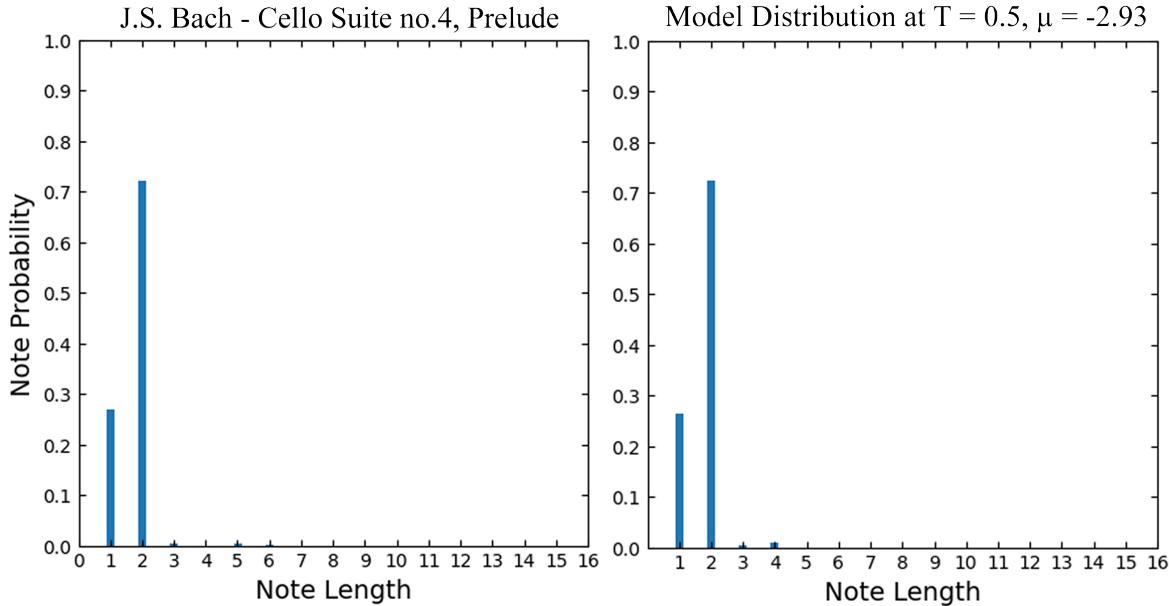


Figure 9: A Bach composition which corresponds to the distribution generated at  $T = 0.5, \mu = -2.93$ . The Prelude movement was isolated due to the composition having significant rhythmic changes between movements.

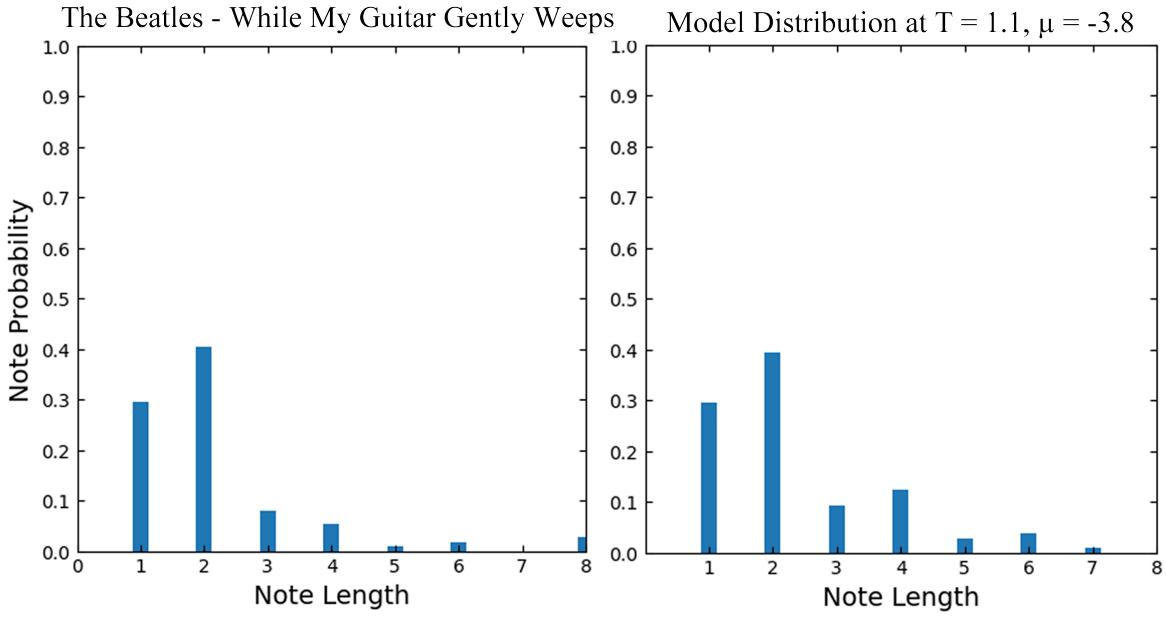


Figure 10: A classic Beatles song approximately represented by the values  $T = 1.1, \mu = -3.8$ . Only the vocals of this song were used.

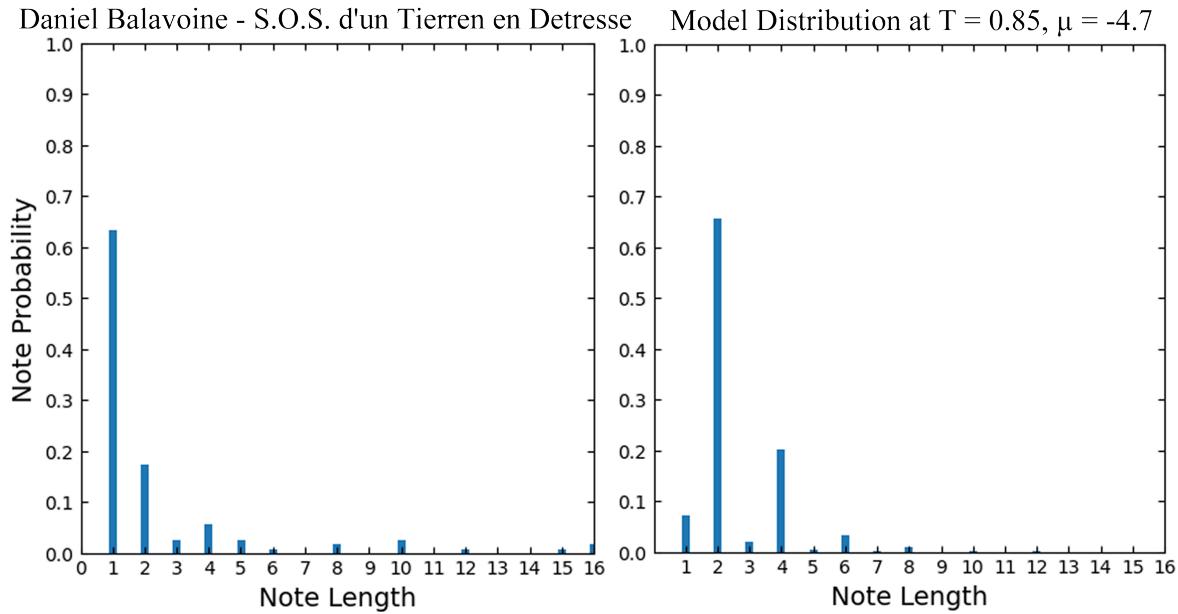


Figure 11: A vocally challenging song from Canadian-French rock opera *Starmania*. Represented by  $T = 0.85, \mu = -4.7$ . If we look at the 2, 4, and 6 note lengths in the model distribution, we can compare them to the 1, 2, and 3 peaks of the data. Perhaps the probability of the small peaks in the model is low enough for no notes to appear in the short time range of this composition.

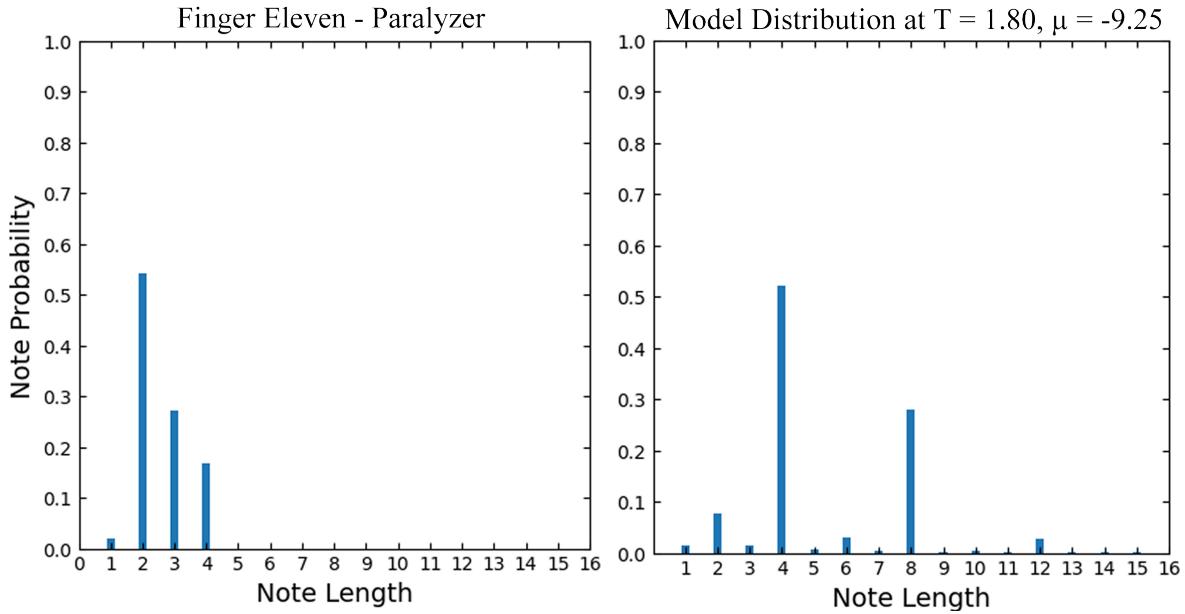


Figure 12: The chorus of a hard rock song is analyzed here. Similar to the Balavoine song, perhaps the short range of the chorus (just a few measures) leads to some notes from our model of  $T = 1.80, \mu = -9.25$  not occurring.

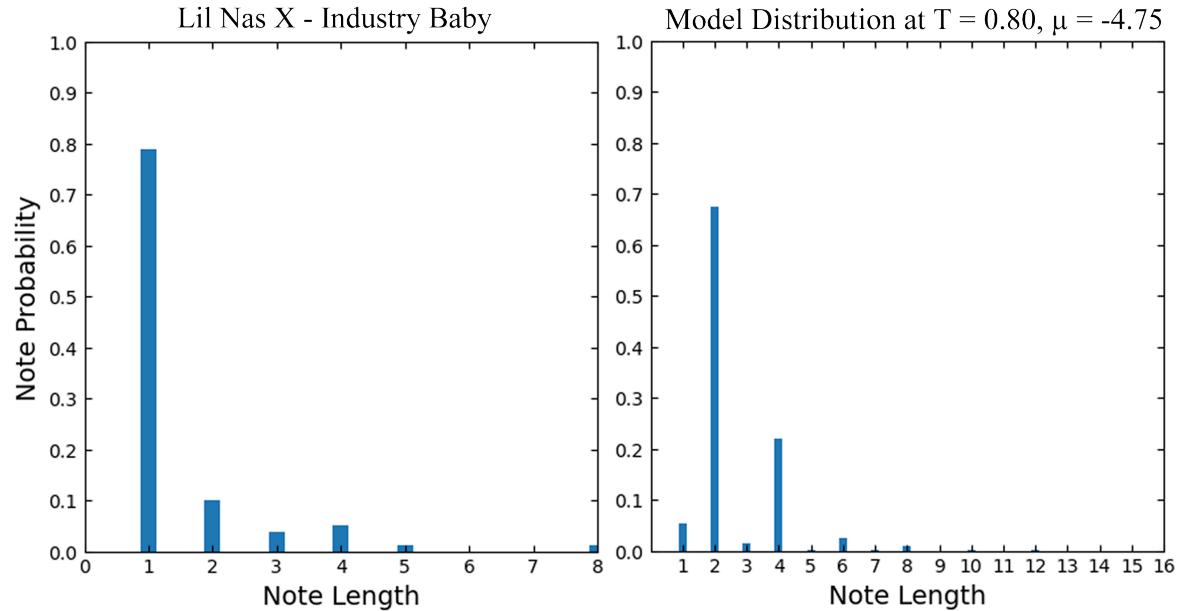


Figure 13: A recently popular modern pop song approximated by  $T = 0.80, \mu = -4.75$ .

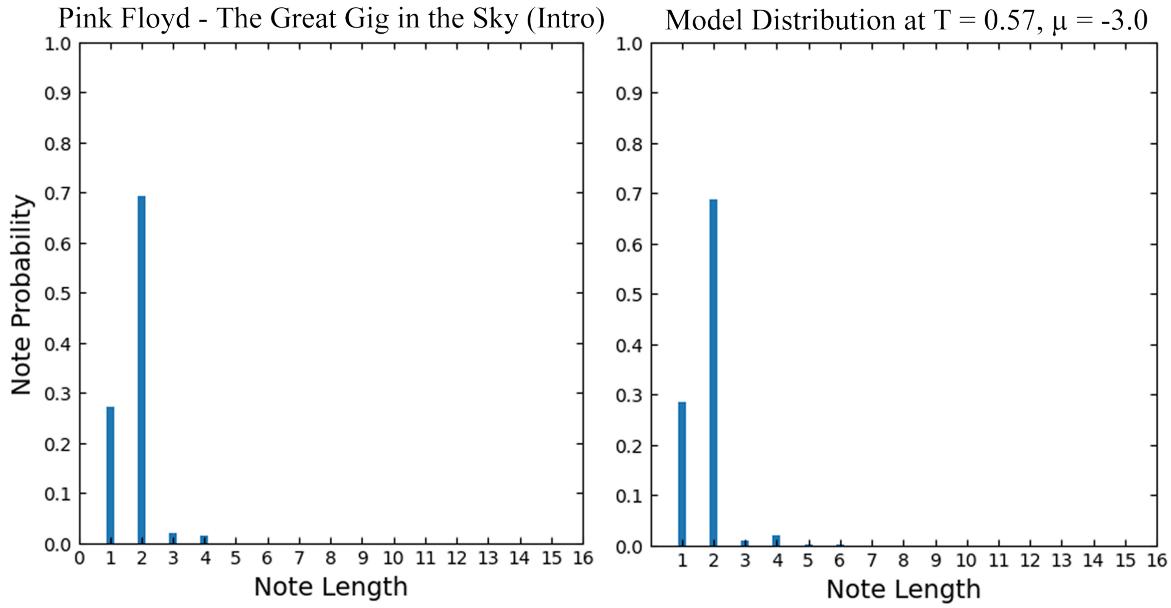


Figure 14: Experimental song by Pink Floyd, the introduction melody played on the piano is shown and represented well by the distribution at  $T = 0.57, \mu = -3.0$ .

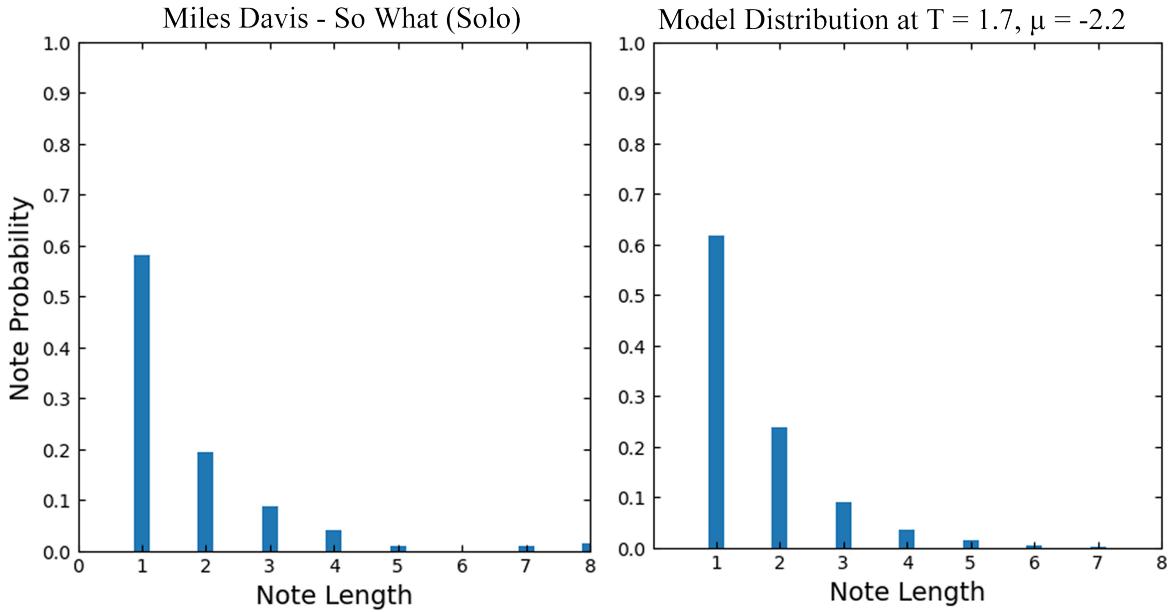


Figure 15: Trumpet solo from a Jazz song. This is the only piece that we looked at which existed in the disordered phase at  $T = 1.7, \mu = -2.2$ . This makes sense for a composition of this nature. Jazz is known for its less rigid musical and rhythmic structure, and a solo pushes those limits as they are often improvised.

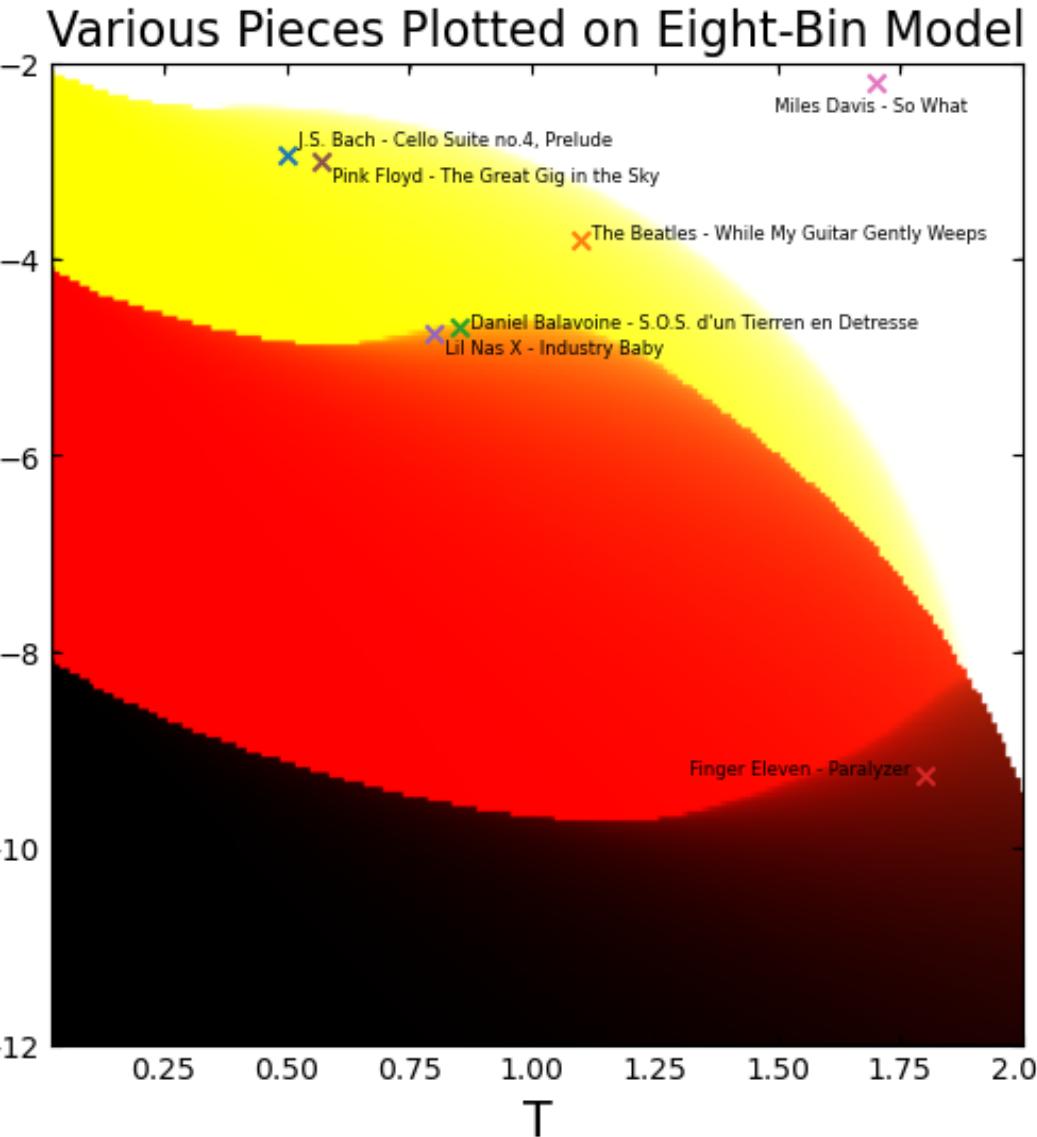


Figure 16: Another plot of the phase diagram with the above songs plotted based on their approximate corresponding  $T$  and  $\mu$  values. According to our model, the Prelude of Bach's Cello suite no. 4 is rhythmically similar to Pink Floyd's "The Great Gig in the Sky". Additionally, the forlorn rock opera S.O.S. appears to be rhythmically similar to hit Lil Nas X song "Industry Baby". The other compositions lay in their own respective regions. It may be reasonable to assume that a hard rock song like Finger Eleven's "Paralyzer" rests in the single period phase, as such a song may have more structural variation compared to a Beatles song, which occupies the same phase as a classical composition and a Pink Floyd song.

## 5 Discussion & Conclusion

We proposed a means of modeling rhythm through statistical mechanics and the Fermi-Dirac distribution by assuming that rhythm is generated by a repeating chain of length  $L$  of independent probabilities and that there are underlying parameters  $T$  and  $\mu$  that govern rhythm similar to how these quantities govern structure within physical systems.

We showed through a simple case that stable solutions exist for this model, and the solutions change as  $T$  changes, indicative of the presence of different distinct structural phases.

We then used our model to construct a full phase diagram for a configuration of  $L = 8$  independent time bin probabilities, where distinct rhythmic phases emerged. We observed that rhythmic structure generally is less ordered at high  $T$  and more ordered at low  $T$ . This decrease of order signified by the increased likelihood of notes that have no classical notation as temperature rises until the disorder phase is reached and the distribution becomes Poisson-like.

We then proceeded to look at known compositions to see if our proposed model could approximate rhythmic distributions seen in such musical pieces. The results are shown in the section above in Figures 9 through 16. It can be seen that our model sometimes does very well in approximating human compositions in some cases, and does not quite represent the composition in other cases. This shows that we are at least heading in the right direction, but with approximations such as non-interacting particle sites and rhythmic perception at large time scales, we may be simplifying rhythm too much and may need to develop a more complex model.

We do not seek to reinvent music theory with this paper, but we do seek to ground our intuitive sense of rhythm that humans have known for ages in a mathematical model to explain this intuitive sense of rhythm. At a glance, this model may be a step in the right direction as we are able to both describe rhythm as an ordered phases as well as use such a description to visualize the underlying structure of a specific rhythm.

## 6 Future Plans

Currently, we have only investigated distributions with values of  $L$  that have 2 as a sole factor. We would like to investigate our model when  $L$  is 3 or 9. We also expect interesting results when looking at a  $L$  that has more than just one factor, such as 6 or 12. We have started some initial work on a  $L = 6$  model and have found that the system can tend to a ordering of every 2 sites as well as every 3. We would like to see when the system has equilibrium states in each respective pattern, as this may tell us about the uniqueness and stability of rhythmic *meter* in addition to learning about configurations.

We have also seen that the distributions obtained from the model so not always correctly reflect smaller variations in note occurrences. This may be because in the calculation of  $R_{tot}$ , we assume rhythmic perception on an infinite time scale. This assumption inherently does not capture human perception, as we do not necessarily keep track of notes and rhythm that occurred in the distant past. Due to this, we seek to develop a short range correlation for our model [3]. This may allow for more small fluctuations in note distributions.

The assumption that the time bin sites do not interact with each other may be an oversimplification. A composer likely takes into account the placement of notes when creating a composition, and the presence of “on” beats nearby may influence the presence of a beat in the current time bin more than our model assumes. A possible remedy to this would be to use a interacting fermionic model to model this complexity.

Finally, we are limited in our analysis by our use of MIDI files. Although sufficient for the sake of this paper, we do not generally see discrete sound recordings in everyday life. Most of these MIDI files are simplified adaptations of the corresponding compositions. Such simplifications may leave out nuance in the playing of rhythm or differences in swing or waltz type rhythms. A scheme for rhythm and beat detection is outlined in [2], which may be worthwhile in exploring if we want to expand our analysis of known music to the most common medium of .mp3 recordings.

## 7 References

- [1] J. Berezovsky, “The structure of musical harmony as an ordered phase of sound: A statistical mechanics approach to music theory.” *Science advances* vol. 5,5 eaav8490. 17 May 2019, doi:10.1126/sciadv.aav8490
- [2] W. A. Sethares, *Rhythm and Transforms*. London: Springer, 2007.
- [3] J. Berezovsky, “Musical Rhythm as an Ordered Phase,” unpublished.
- [4] C. Kittel, *Elementary Statistical Physics*. Mineola, New York: John Wiley Sons, Inc, 1958.
- [5] C. Kittel and H. Kroemer, *Thermal Physics*, 2nd ed. New York, New York: W.H. Freeman and Company, 1980.
- [6] W. Schweer, Sheet Music MIDI Files. <https://musescore.com>. MuseScore Online Sheet Music.
- [7] O. M. Bjørndalen, “Mido - MIDI Objects for Python,” <https://mido.readthedocs.io>. Read the Docs, 05-Oct-2021.