A statistical mechanics approach to music

Lauren H. Li
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
(Dated: May 19, 2023)

Typically regarded as an artistic practice, music is a common medium for creativity and free expression. However, music also possesses a rich underlying structure, and music theory seeks to identify rules that govern the human perception of sound. Here, we build upon previous works to utilize the toolbox of statistical mechanics to explore the structure of music. We leverage a dissonance model introduced by [1] and use a mean field approximation to recreate the 12-tone structure of classical Western music. Additionally, we conduct Monte Carlo simulations to investigate the interactions between musical tones.

INTRODUCTION

Music is ubiquitous across all cultures and is commonly used for human artistic expression. While musical practices vary across the globe, many characteristics of music, such as restriction to a finite set of pitches, is almost universal. People have long studied music theory using a top-down approach; for example, the theory of harmony and counterpoint is a core subject in Western conservatories. Recently physicists have began studying music using a bottom-up approach, leveraging the tools of statistical mechanics [1, 2]. In fact, certain aspects of music lend itself very well to statistical mechanical models.

In Western music, there are 12 pitch classes: tones with fundamental frequencies that differ by a factor of two (an octave interval) are perceived to be in the same pitch class. The 12 pitch classes are commonly arranged around a circle, known as the *Circle of Fifths* (Fig. 1a). Notes that neighbor each other around the circle are generally perceived to be better sounding when played in unison. In the more recent neo-Riemannian theory [3], notes are arranged on a triangular lattice, known as a Tonnetz (Fig. 1b). Connected notes on the grid can be reached via simple musical transformations.

These circular patterns and lattice arrangements are reminiscent of many concepts in statistical mechanics [4]; here, we explore some of these ideas in relation to classical music theory. First, we outline a general model that quantifies dissonance between different pitches. Then, we reproduce a key result of [1] that uses a mean field approach to recover the 12 pitch classes of tonal music. Finally, we apply models of dissonance to the clock model and perform Monte Carlo simulations in an attempt to recover an ordered lattice of music pitches.

DISSONANCE MODEL

Here, we outline the dissonance model presented in [1]. As in the classical study of sound waves, a *tone* is characterized by its fundamental frequency f (also known as the pitch), but is also composed of a set of overtones with amplitudes $\alpha_n A$ and frequencies $\phi_n f$, where $\phi_n = n \in \mathbb{N}$

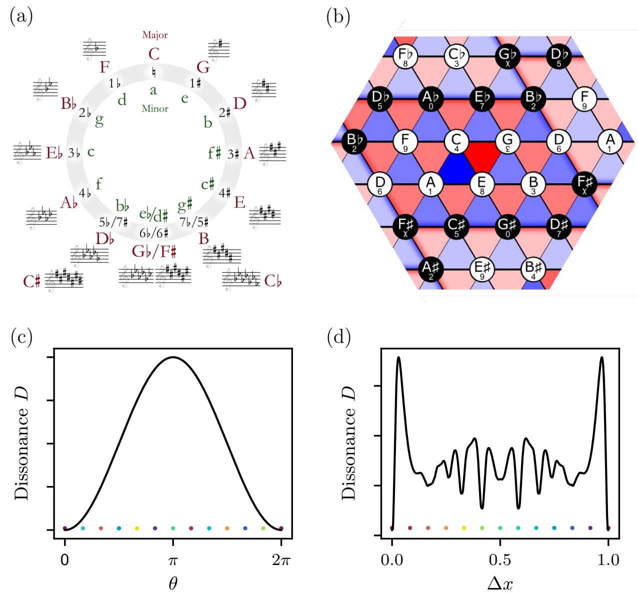

FIG. 1. Models of musical pitch. Classical structures for musical pitches: (a) the Circle of Fifths and (b) the Tonnetz from neo-Riemannian theory.[a] Quantitative models of dissonance: (c) cosine function where notes around the Circle of Fifths map to $\theta \in [0, 2\pi]$, (d) general dissonance function that describes the difference in frequency $\Delta x = \log_2(f_a/f_b)$ between two pitches. Color dots represent the 12 different pitch classes in Western music.

[a] Retrieved from `wikipedia.org/wiki/Circle_of_fifths` and `wikipedia.org/wiki/Tonnetz`, respectively.

and $\alpha_n = 1/n$. Given a set of tones $\{t_i\}$, the total dissonance perceived by an observer is given by

$$D_{\text{tot}} = \sum_{ij} D_{ij}, \tag{1}$$

where $D_{ij} = D(f_i, f_j)$ is the dissonance between two tones with pitches f_i and f_j. We assume that all tones overlap over the same amount of time. To quantify the dissonance of two tones, which could contain overtones, we first introduce a function $d(f_a, f_b)$ that characterizes the dissonance between two pure tones, with consist of only the fundamental frequency f_a and f_b, respectively. According to [1], the form of this dissonance function can be found to be

$$d(f_a, f_b) = e^{-[\ln(|\Delta x|/w_c)]^2/w_c}, \tag{2}$$

where $\Delta = \log_2(f_a/f_b)$ and w_c is the Δx at which the greatest dissonance is attained. Throughout this paper, we set $w_c = 0.03$, in accordance to [1]; this yields the dissonance function in Fig. 1d). Note that $\Delta x = 1$ corresponds to pitches that are an octave apart.

MEAN FIELD APPROACH

Here, we use a mean field approximation presented in [1] to reproduce the 12 tones found in classical music. We consider the thermodynamic limit of tones, which comprise pitches distributed according to the probability distribution $P(x)$, where the pitch $x = \log_2(f/f_{\text{ref}}$ for some reference frequency f_{ref}. For the purpose of simulation, we set the reference to Middle C, which has frequency $f_{\text{ref}} = 261.6$ Hz. Finally, we assume periodicity $P(x) = P(x+1)$, as the pitch distribution should be the same in every octave; in this sense, the total dissonance is given by

$$D_p(x) = \sum_{n=-\infty}^{\infty} D(x+n), \tag{3}$$

which sums over all possible octaves. By considering only a single octave, the total dissonance is then given by

$$D_{\text{tot}} = \frac{1}{2}\int_0^1 \int_0^1 P(x) D_p(x-y) P(y)\, dy dx. \tag{4}$$

Likewise, the entropy is

$$S = -\int_0^1 P(x) \ln P(x)\, dx. \tag{5}$$

To identify the equilibrium behavior, we seek to find the distribution $P(x)$ that minimizes the free energy

$$F = D_{\text{tot}} - TS, \tag{6}$$

given the constraint that $P(x)$ is normalized. Using the method of Lagrange multipliers and the variational principle [1], stable solutions of $P(x)$ can be found. We show $P(x)$ for different values of T in Fig. 2. Indeed, we recover the 12 pitches of music for $T_c \approx 16.2$. For $T < T_c$, peaks in $P(x)$ exist, but there are generally fewer than 12. On the other hand, for $T > T_c$, $P(x)$ flattens to a uniform distribution. These patterns can be interpreted as a phase transition from an ordered to disordered phase, where the most commonly used musical structures emerge at the at the critical temperature T_c.

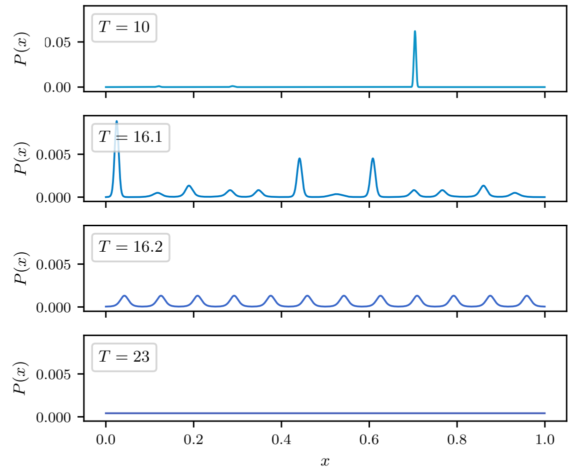

FIG. 2. Distribution of pitches in an octave for different temperatures T. We identify the critical temperature as $T_c = 16.2$, which gives rise to the familiar 12-tone structure of Western music.

MONTE CARLO SIMULATIONS

Inspired by neo-Riemannian theory, we also conduct Monte Carlo simulations of tones on a 2D square lattice. We implement the Metropolis-Hastings algorithm, where points on the lattice are initialized to to some semitone $s_i = 1, 2, \ldots, 12$, positing the 12-tone structure of music. We then construct a clock model, where nearest- and next nearest-neighbor sites have interaction

$$\beta H_C = -\sum_{\langle i,j \rangle} J(s_i, s_j), \tag{7}$$

and $J(s_i, s_j)$ is either a cosine function (Fig. 1c) or a frequency dissonance function (Fig. 1d). Performing our simulations on a 50×50 lattice, we identify the low and high temperature behaviors the tone lattice (Fig. 3).

In the case where interactions are modeled by a cosine function, tones are evenly distributed around the circle at angles $\theta = 2\pi s_i/12$. We see that tones on the lattice separate into distinct regions at low T (Fig. 3a). Furthermore, since our model is a discretized version of the XY model, we also observe vortices. By our construction, tones around the vortices cycle around the Circle of Fifths.

In the case where interactions are given by the calculated dissonance function, we see that tone also separate into distinct regions at low T (Fig. 3b). While some structure is present, more research is needed to make definitive conclusions. Nevertheless, we observe some instances where tones that are close on the Circle of Fifths also appear close on the tone lattice.

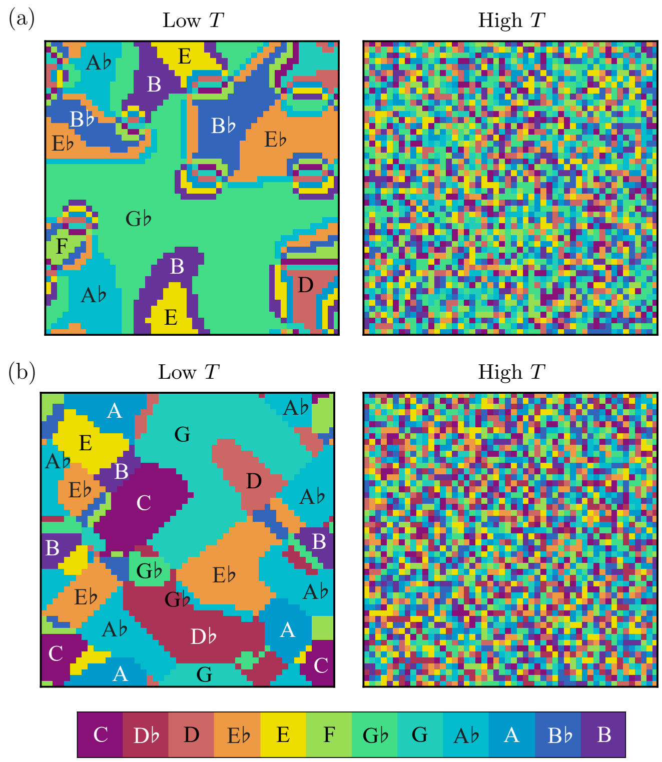

FIG. 3. Lattice models of pitch. Clock model with interactions according to (a) a cosine function or (b) the calculated dissonance function. At low temperature, tones separate into regions of the same pitch. Some musical structure emerges by observing neighboring regions.

DISCUSSION

In this work, we leverage methods from statistical mechanics to study music from first principles. We reproduced key results from [1] and conducted new simulations that could lead to fundamental insights on the theory of music and human perception. We hope our work will inspire other such interdisciplinary studies that join science and the arts.

Acknowledgements

The author is thanks Prof. Mehran Kardar and Changnan Peng for their outstanding teaching of Statistical Mechanics II (8.334, Spring 2023).

[1] Jesse Berezovsky, "The structure of musical harmony as an ordered phase of sound: A statistical mechanics approach to music theory," Science advances **5**, eaav8490 (2019).

[2] Huay Din and Jesse Berezovsky, "Critical behavior and the kibble-zurek mechanism in a musical phase transition," Plos one **18**, e0280227 (2023).

[3] Richard Cohn, "Introduction to neo-riemannian theory: a survey and a historical perspective," Journal of Music Theory , 167–180 (1998).

[4] Mehran Kardar, Statistical physics of fields (Cambridge University Press, 2007).