frontiers in Applied Mathematics and Statistics

ORIGINAL RESEARCH
published: 09 June 2020
doi: 10.3389/fams.2020.00018

A Dynamical Systems Approach to Spectral Music: Modeling the Role of Roughness and Inharmonicity in Perception of Musical Tension

Michal Hadrava[1,2,3]* *and Jaroslav Hlinka*[2,3]

[1] Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czechia, [2] Department of Complex Systems, Institute of Computer Science of the Czech Academy of Sciences, Prague, Czechia, [3] RP3 Applied Neuroscience and Neuroimaging, National Institute of Mental Health, Klecany, Czechia

Tension-resolution patterns seem to play a dominant role in shaping our emotional experience of music. In traditional Western music, these patterns are mainly expressed through harmony and melody. However, many contemporary musical compositions employ sound materials lacking any perceivable pitch structure, rendering the two compositional devices useless. Still, composers like Tristan Murail or Gérard Grisey manage to implement the patterns by manipulating spectral attributes like roughness and inharmonicity. However, in order to understand the music of theirs and the other proponents of the so-called "spectral music," one has to eschew traditional categories like pitch, harmony, and tonality in favor of a lower-level, more general representation of sound—which, unfortunately, music-psychological research has been reluctant to do. In the present study, motivated by recent advances in music-theoretical and neuroscientific research into a the highly related phenomenon of dissonance, we propose a neurodynamical model of musical tension based on a spectral representation of sound which reproduces existing empirical results on spectral correlates of tension. By virtue of being neurodynamical, the proposed model is generative in the sense that it can simulate responses to arbitrary sounds.

Keywords: music, neurodynamics, timbre, tension, dissonance, roughness, inharmonicity, periodicity

OPEN ACCESS

Edited by:
Plamen Ch. Ivanov,
Boston University, United States

Reviewed by:
George Datseris,
Max-Planck-Institute for Dynamics and Self-Organisation (MPG), Germany
Axel Hutt,
Inria Nancy—Grand-Est Research Centre, France

***Correspondence:**
Michal Hadrava
mihadra@gmail.com

Specialty section:
This article was submitted to Dynamical Systems, a section of the journal Frontiers in Applied Mathematics and Statistics

Received: *18 November 2019*
Accepted: *04 May 2020*
Published: *09 June 2020*

Citation:
Hadrava M and Hlinka J (2020) A Dynamical Systems Approach to Spectral Music: Modeling the Role of Roughness and Inharmonicity in Perception of Musical Tension. Front. Appl. Math. Stat. 6:18. doi: 10.3389/fams.2020.00018

1. INTRODUCTION

Music gives rise to some of the strongest emotional experiences in our lives. Even though the first surviving theoretical treatments of the power of music to move the soul were written in the fifth century B.C. [1], the origin of this power still largely remains a mystery. However, both musicological and music-psychological evidence seems to converge on the theory that music arouses emotions by a sophisticated play of tension-resolution patterns [2]. For instance, many authors describe the musical language of Richard Wagner (1813–1883) as "the language of longing" [2]; it may not be merely a coincidence that Wagner's common practice was to introduce a dissonant chord, making the listener "long for" a more consonant chord to "resolve" the dissonance, and then keep the listener in tension by delaying the resolution or slap him right away with another dissonance [[2], p. 334–339].

Creating and resolving tension is an easy task for composers who follow the nineteenth century Western tradition (as, e.g., most "mainstream" composers of film music do); any standard textbook on harmony and voice leading provides them with plenty of recipes [e.g., [3]] tested by centuries of musical practice and decades of psychological research [e.g., [4] and references therein]. However, over the course of the twentieth century, many composers enriched their palette with sounds possessing neither definite pitch (precluding melody) nor perceivable voice structure (precluding harmony), thus venturing out into territories about which traditional theory has nothing to say but *hic sunt leones* [5]. Still, while tantalizing their audience with a sound palette ranging from pure tones to the most atrocious noises, they seek control over how their music is experienced by the listener as much as their more conservative colleagues do [6].

Devoid of any perceivable pitch structure, the ferocious sound materials contemporary art music is at times so fond of can only be conceived in terms of loudness and timbre. This forces any composer seeking a full control over these "beasts" to dive from the lofty heights of venerable musical abstractions like pitch, harmony, and tonality to the cold depths of spectral representations of sound. However, beauty emerges even from such depths; by careful manipulation of roughness and inharmonicity composers like Tristan Murail or Gérard Grisey "tense" their audience no less than Richard Wagner by his mastery of tonal harmony; indeed, the term "spectral music," used when referring to the music style pioneered by the former two composers [7], does not tell the whole story.

As usual, music-psychological research somewhat lagged behind compositional practice; loud music has been shown to be perceived as more tense than soft music [8]; likewise, roughness (that is perception of rapid beating due to interference of close frequencies) seems to be positively correlated with tension [9, 10]. A recent study assessed the effect of specific timbre attributes on the perception of tension [11], confirming particularly the role of roughness, inharmonicity (deviation of the constituent frequencies from integer multiples of a fundamental tone) and spectral flatness. However, the functional forms and mechanisms through which such stimuli aspects combine to give rise to perceived tension are still unclear.

For standard Western musical intervals, roughness is a principal source of perceived dissonance in musical material, which thus gives rationale to mathematical models of musical dissonance [12]. In Stolzenburg [13], a mathematical model of dissonance has been proposed and shown to correlate highly with empirical psychophysical data. The core idea of the model can be illustrated on a simple example: Consider an interval of perfect fifth, the most consonant interval after the octave, which, in the standard Western tuning, corresponds to the distance of seven semitones. Hence, denoting f_1 and f_2 the fundamental frequencies of the tones spanning the interval, $[f_1, f_2] = [f_1, f_1 2^{\frac{7}{12}}]$ ($2^{\frac{1}{12}}$ is the frequency ratio corresponding to a semitone distance in the standard Western tuning). First, approximate $[f_1, f_2]$ with a pair of coprime natural numbers, $\Omega' = [i, j]$, such that $\frac{i}{j} \approx \frac{f_1}{f_2} = 2^{\frac{7}{12}}$, e.g., $\Omega' = [2, 3]$. Then, the dissonance of the interval will be 2, the minimum of Ω'. Likewise, for an interval of major sixth, considered less consonant than P5, with frequencies $[f_1, f_2] = [f_1, f_1 2^{\frac{9}{12}}]$, we have $\Omega' = [3, 5]$, the dissonance being 3 in this case. Finally, for a dissonant interval of minor seventh with frequencies $[f_1, f_2] = [f_1, f_1 2^{\frac{9}{12}}]$, $\Omega' = [4, 7]$ and the dissonance will be 4. In general, the dissonance of any vector of frequencies, f, approximated as $\Omega' \in \mathbb{Z}^n_{\geq 1}$, is assumed to be proportional to the minimum element of Ω'. Note that the dissonance of an interval estimated this way does not change if we include *harmonics* (integer multiples) of the constituent frequencies. However, it does change if one uses a different rational approximation of the frequency ratio; incidentally, all standard Western intervals except for the octave are characterized by an irrational frequency ratio. In Stolzenburg [13], this inconvenience is dealt with by averaging over several alternative approximations.

The quantity above, called *relative periodicity* [see Definition 6 in [13], p. 17], is equivalent to obtaining the period of the fastest oscillation having the frequencies in question as its harmonics, in particular the period assessed in cycles of the lowest frequency in question. Interestingly, this oscillation has been experimentally observed to be represented in the auditory brainstem response to the intervals listed above [14], with the representation being particularly faithful for relatively more consonant intervals.

Motivated by the latter observation, we put forward a neurodynamical model of tension which is in line with the basic concepts of pitch perception of complex sounds and reproduces the results concerning the effect of roughness and inharmonicity reported in Farbood and Price [11] and, at the same time, provides a dynamical interpretation of relative periodicity [13]. In this regard we follow suit of existing studies which apply the dynamical systems theory to composition [15] and analysis [16, 17] of music.

2. METHODS

Everyone interested in neurodynamical modeling faces the same basic dilemma: which model to use? For modeling perception of music, the most common choices are the leaky integrate-and-fire (LIF) model [18–21] and a canonical model for gradient-frequency networks of Wilson-Cowan-type neural oscillators [22–25]. Still, neuroimaging methods are far from giving us an assurance that among the myriad possible models one of these is the "correct" one. Hence, to improve our chances, instead of adhering to a particular model right from the beginning, we take a whole class of models as our point of departure in the hope that the class is wide enough to include a good approximation to the actual biological system. More precisely, we proceed by derivation of a normal form to which any of the class members can be transformed through a continuous near-identity change of variables and parameters and (possibly) a time scaling. Then, analysis of the entire class effectively reduces to analysis of the normal form [26].

The pioneering work of the latter approach in our field is Large et al. [22]; the model proposed therein can even be fit to auditory brainstem responses to musical intervals [24, 27]. Consequently, one could argue that we already have a neurodynamical model of

tension at hand. However, to prove analytically that the latter is indeed a decent model of tension, we would first have to simplify the model substantially for the specific purpose, adopting thus a similar strategy as in previous applications of the model [23, 28, 29]. Moreover, the class of systems of which the latter is a canonical model [22, 27] is not the class of systems we are interested in, as shown below.

Our choice of model class is motivated by the observation that the auditory system is sensitive to the periodicity of the signal (see section 1). A possible explanation for the observation is that the system comprises an array of oscillatory "detectors" with external auditory signal input; these can be viewed (and indeed physically seem to be) ordered tonotopically with respect to their eigenfrequency. Within this framework, eliciting a sustained oscillation in one of the oscillators represents detection of the corresponding period in the input signal. Or, on a continuous scale, the more sustained an oscillation is, the more "'confident" is the auditory system that the input signal exhibits the corresponding period. In order for a stimulus to elicit an oscillation in a model belonging to our class of choice, it needs to destabilize the (originally stable) quiescent state; if this shift in stability is relatively small or intermittent, the oscillation will have small or fluctuating amplitude.

To avoid introducing unnecessary complexity, we start building our model class of interest by considering the simplest possible model of an oscillator:

$$\begin{aligned}\dot{x}_1 &= -y_1\\ \dot{y}_1 &= x_1\,,\end{aligned}$$

where $x_1, y_1 \in \mathbb{R}$ and $\dot{x}_1, \dot{y}_1$ are time derivatives. In matrix form:

$$\begin{bmatrix}\dot{x}_1\\ \dot{y}_1\end{bmatrix} = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ y_1\end{bmatrix}.$$

Here, x_1 and y_1 could be interpreted as the amount of local inhibitory and excitatory synaptic activity, respectively, but the particular physiological interpretation of the variables is not important for our discussion. In line with the scenario outlined above, we want the oscillator to transit from a quiescent state (say, $[x_1, y_1] = [0, 0]$) to an oscillation when subject to an input having the oscillator's period (1 in this case). By definition, the spectrum of such an input consists of frequencies which are a subset of $1, 2, \ldots, M$. As we later aim to introduce the possibility of modeling the effect of inputs with frequencies deviating from perfect multiples of the base frequency, we allow already in the basic model for repeated frequency components of the input, yielding the input frequency vector $[\omega_1, \omega_2, \ldots, \omega_n] = \Omega = [1, 1, \ldots, 1, 2, 2, \ldots, 2, \ldots, M, M, \ldots, M]$, where M is the highest frequency contained in the input. Representing the i-th frequency component of the input with frequency ω_i as $[x_{i+1}(t), y_{i+1}(t)] \in \mathbb{R}^2$, we extend the oscillator with forcing by the input:

$$\begin{bmatrix}\dot{x}_1\\ \dot{y}_1\end{bmatrix} = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ y_1\end{bmatrix} + \begin{bmatrix}f(x_1, y_1, x_2(t), y_2(t), x_3(t), y_3(t), \ldots, x_{n+1}(t), y_{n+1}(t))\\ g(x_1, y_1, x_2(t), y_2(t), x_3(t), y_3(t), \ldots, x_{n+1}(t), y_{n+1}(t))\end{bmatrix},$$

where $f(\cdot), g(\cdot)$ are smooth real functions.

Two more steps are needed in order to make the model class amenable to derivation of a normal form. First, we need to rewrite this non-autonomous system as an autonomous one. This is straightforward, since $[x_{i+1}(t), y_{i+1}(t)]$, being frequency components of a signal, are harmonic oscillations. Let $x = [x_1, y_1, x_2, y_2, \ldots, x_{n+1}, y_{n+1}]$:

$$\dot{x} = \begin{bmatrix}0 & -1 & & & & & & & & \\ 1 & 0 & & & & & & & & \\ & & 0 & -\omega_1 & & & & & & \\ & & \omega_1 & 0 & & & & & & \\ & & & & 0 & -\omega_2 & & & & \\ & & & & \omega_2 & 0 & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & 0 & -\omega_n & \\ & & & & & & & \omega_n & 0 & \end{bmatrix} x + \begin{bmatrix}f(x)\\ g(x)\\ 0\\ 0\\ 0\\ 0\\ \vdots\\ 0\\ 0\end{bmatrix}.$$

Second, we expand the $f(\cdot)$ and $g(\cdot)$ functions around the origin:

$$\dot{x} = \begin{bmatrix}0 & -1 & & & & & & & & \\ 1 & 0 & & & & & & & & \\ & & 0 & -\omega_1 & & & & & & \\ & & \omega_1 & 0 & & & & & & \\ & & & & 0 & -\omega_2 & & & & \\ & & & & \omega_2 & 0 & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & 0 & -\omega_n & \\ & & & & & & & \omega_n & 0 & \end{bmatrix} x + \begin{bmatrix}f_2(x)\\ g_2(x)\\ 0\\ 0\\ 0\\ 0\\ \vdots\\ 0\\ 0\end{bmatrix} + \begin{bmatrix}f_3(x)\\ g_3(x)\\ 0\\ 0\\ 0\\ 0\\ \vdots\\ 0\\ 0\end{bmatrix} + \begin{bmatrix}\mathcal{O}(x^4)\\ \mathcal{O}(x^4)\\ 0\\ 0\\ 0\\ 0\\ \vdots\\ 0\\ 0\end{bmatrix}, \tag{1}$$

where $f_d(x)$, $g_d(x)$ are homogeneous polynomials of degree d and $\mathcal{O}(x^4)$ is a polynomial of degree 4 or more.

Before delving into the actual derivation, a few remarks are in place. First, the idea of modeling the auditory system as a tonotopically arranged array (or rather a series of arrays) of oscillators is in fact not new [23, 24, 27]. Further, since the lowest frequency of signal with harmonic spectrum corresponds to its perceived pitch, Equation (1) is, by design, a pitch detector. Then, instantiating Equation (1) with a range of eigenfrequencies can be thought of as matching a template to the signal; consequently, our model can be considered a neurodynamical implementation of template matching postulated as a possible mechanism underlying pitch perception [30].

As the first step of the derivation, we diagonalize the linear part of Equation (1) using the following two matrices:

$$T = \frac{1}{2}\begin{bmatrix} 1 & 1 & & & & & & \\ -\imath & \imath & & & & & & \\ & & 1 & 1 & & & & \\ & & -\imath & \imath & & & & \\ & & & & 1 & 1 & & \\ & & & & -\imath & \imath & & \\ & & & & & & \ddots & \\ & & & & & & 1 & 1 \\ & & & & & & \imath & -\imath \end{bmatrix},$$

$$T^{-1} = \begin{bmatrix} 1 & \imath & & & & & & \\ 1 & -\imath & & & & & & \\ & & 1 & \imath & & & & \\ & & 1 & -\imath & & & & \\ & & & & 1 & \imath & & \\ & & & & 1 & -\imath & & \\ & & & & & & \ddots & \\ & & & & & & 1 & \imath \\ & & & & & & 1 & -\imath \end{bmatrix},$$

where $\imath$ is the imaginary unit. Denoting the linear part of Equation (1) as A, the diagonalized matrix reads:

$$T^{-1}AT = \begin{bmatrix} \imath & & & & & & & & \\ & -\imath & & & & & & & \\ & & \imath\omega_1 & & & & & & \\ & & & -\imath\omega_1 & & & & & \\ & & & & \imath\omega_2 & & & & \\ & & & & & -\imath\omega_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \imath\omega_n & \\ & & & & & & & & -\imath\omega_n \end{bmatrix}. \quad (2)$$

The diagonalization defines a change of variables:

$$\begin{bmatrix} z_1 \\ w_1 \\ z_2 \\ w_2 \\ z_3 \\ w_3 \\ \vdots \\ z_{n+1} \\ w_{n+1} \end{bmatrix} = \zeta = T^{-1}x = \begin{bmatrix} 1 & \imath & & & & & & \\ 1 & -\imath & & & & & & \\ & & 1 & \imath & & & & \\ & & 1 & -\imath & & & & \\ & & & & 1 & \imath & & \\ & & & & 1 & -\imath & & \\ & & & & & & \ddots & \\ & & & & & & 1 & \imath \\ & & & & & & 1 & -\imath \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ x_3 \\ y_3 \\ \vdots \\ x_{n+1} \\ y_{n+1} \end{bmatrix}$$

$$= \begin{bmatrix} x_1 + \imath y_1 \\ x_1 - \imath y_1 \\ x_2 + \imath y_2 \\ x_2 - \imath y_2 \\ x_3 + \imath y_3 \\ x_3 - \imath y_3 \\ \vdots \\ x_{n+1} + \imath y_{n+1} \\ x_{n+1} - \imath y_{n+1} \end{bmatrix}.$$

After this change, Equation (1) reads:

$$\dot{z_k} = \dot{x_k} + \imath\dot{y_k} = \quad (3)$$
$$= \begin{cases} -y_k + \imath x_k + f(x) = \imath^2 y_k + \imath x_k + f(x) & k = 1 \\ = \imath z_k + f_{zeta}(\zeta), & \\ \imath\omega_{k-1} z_k, & 1 < k \le n \end{cases}$$
$$\dot{w_k} = \dot{x_k} - \imath\dot{y_k} = \quad (4)$$
$$= \begin{cases} -y_k - \imath x_k + g(x) = \imath^2 y_k - \imath x_k + g(x) & k = 1 \\ = -\imath w_k + g_\zeta(\zeta), & \\ -\imath\omega_{k-1} w_k, & 1 < k \le n, \end{cases}$$

with $g_\zeta(\zeta) = \overline{f_\zeta}(\zeta)$, where and $\bar{\cdot}$ is complex conjugate. From here, with a slight abuse of notation, we drop for simplicity the subscript of $g_\zeta(\zeta)$ which is a function in the complex vector space corresponding to $g(x)$ from section 2, that is obtained through the change of variables applied, i.e., $g_\zeta(\zeta) = g_\zeta(T^{-1}x) = g(x)$; likewise for $f(\zeta)$. Our subject of study will be the normal form of the system comprising Equations (3) and (4).

When in normal form, $f(\zeta)$ in Equation (3) will only contain monomials of the form

$$c_{rspq} z^{[r,p]} w^{[s,q]} = c_{rspq} z_1^r z_2^{p_1} z_3^{p_2} \cdots z_{n+1}^{p_n} w_1^s w_2^{q_1} w_3^{q_2} \cdots w_{n+1}^{q_n}, \quad (5)$$

where $p = [p_1, p_2, \ldots, p_n]$, $q = [q_1, q_2, \ldots, q_n]$, and $[r, s, p, q]$ is a nonnegative integer solution to the equation

$$r - s + p\Omega - q\Omega = 1. \quad (6)$$

Analogously, the exponents in $g(\zeta)$ solve

$$r - s + p\Omega - q\Omega = -1 \quad (7)$$

(see Equation 2). Since $f(\zeta)$ and $g(\zeta)$ contain neither constant nor linear terms (see section 2),

$$r+s+|p|+|q| = r+s+(p_1+p_2+\cdots+p_n)+(q_1+q_2+\cdots+q_n) \ge 2.$$

For conciseness, from now on, whenever $\nu = [r, s, p, q]$ satisfies the above inequality, we write $\nu \in S$ when it solves Equation (6) and $\nu \in \widetilde{S}$ when it solves Equation (7). Note that $[r, s, p, q] \in S$ iff $[s, r, q, p] \in \widetilde{S}$. Consequently, in the normal form,

$$g(\zeta) = \sum_{[r,s,p,q]\in\widetilde{S}} c_{rspq} z^{[r,p]} w^{[s,q]} = \sum_{[r,s,p,q]\in S} c_{srqp} z^{[s,q]} w^{[r,p]} =$$
$$\overline{f(\zeta)} = \sum_{[r,s,p,q]\in S} \overline{c_{rspq} z^{[r,p]}}\, \overline{w}^{[s,q]} = \sum_{[r,s,p,q]\in S} \overline{c_{rspq}} w^{[r,p]} z^{[s,q]}$$

and, by equality of polynomials,

$$c_{srqp} = \overline{c_{rspq}}$$

for all $[r, s, p, q] \in S$.

To broaden the class of systems covered by our normal form, we unfold Equation (3) using small parameters $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}^n$:

$$\dot{z_k} = \begin{cases} (\alpha + \imath)z_k + f(\zeta), & k = 1 \\ (\imath + \imath\beta_{k-1})\omega_{k-1}z_k, & k \geq 2 \end{cases}$$
$$\dot{w_k} = \begin{cases} (\alpha - \imath)w_k + g(\zeta), & k = 1 \\ (-\imath - \imath\beta_{k-1})\omega_{k-1}w_k, & k \geq 2 \end{cases}.$$

This way, in addition to the models with Taylor expansion around the origin of the form (Equation 1), our class now includes models whose Taylor expansion around the origin has the form

$$\dot{x} = \begin{bmatrix} \alpha & -1 & & & & & & & \\ 1 & \alpha & & & & & & & \\ & & \imath\omega_1\beta_1 & -\omega_1 & & & & & \\ & & \omega_1 & -\imath\omega_1\beta_1 & & & & & \\ & & & & \imath\omega_2\beta_2 & -\omega_2 & & & \\ & & & & \omega_2 & -\imath\omega_2\beta_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \imath\omega_n\beta_n & -\omega_n \\ & & & & & & & \omega_n & -\imath\omega_n\beta_n \end{bmatrix} x + \begin{bmatrix} f_2(x) \\ g_2(x) \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} f_3(x) \\ g_3(x) \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \mathcal{O}(x^4) \\ \mathcal{O}(x^4) \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}. \quad (8)$$

The corresponding normal form then reads

$$\begin{aligned} \dot{z_1} &= (\alpha + \imath)z_1 + \sum_{[r,s,p,q]\in S} c_{rspq} z^{[r,p]} w^{[s,q]} \\ \dot{w_1} &= (\alpha - \imath)w_1 + \sum_{[r,s,p,q]\in S} \overline{c_{rspq}} z^{[s,q]} w^{[r,p]} \\ \dot{z_2} &= \imath(1 + \beta_1)\omega_1 z_2 \\ \dot{w_2} &= -\imath(1 + \beta_1)\omega_1 w_2 \\ \dot{z_3} &= \imath(1 + \beta_2)\omega_2 z_3 \\ \dot{w_3} &= -\imath(1 + \beta_2)\omega_2 w_3 \\ &\vdots \\ \dot{z_{n+1}} &= \imath(1 + \beta_n)\omega_n z_{n+1} \\ \dot{w_{n+1}} &= -\imath(1 + \beta_n)\omega_n w_{n+1}. \end{aligned} \quad (9)$$

Intuitively, the α parameter makes it possible to change the stability of the origin (see section 2.1) whereas the β_k parameters allow the model to resonate when the inputs are not exactly its harmonics. In fact, it might happen that two different inputs approximate the same harmonic. That is, their frequencies equal $(1 + \beta_i)\omega_i$ and $(1 + \beta_j)\omega_j$, respectively, with $\omega_i = \omega_j$. This is the reason why we allowed for duplicate frequencies in the input frequency vector Ω (see the beginning of section 2).

It might appear that Equation (9) only models period detection in a full harmonic spectrum (up to the n-th harmonic). However, it can be easily adapted to a more general stimulus with frequencies $H^{\text{harm}} \subset \Omega$ by removing the dependence on those z_j and w_j for which $\Omega_j \notin H^{\text{harm}}$ from the equation for $\dot{z}_1$ and $\dot{w}_1$ (assuming the dimension of the system is large enough to accommodate any stimulus of practical interest). Since the right-hand sides of the equations are polynomials, it suffices to zero-out the coefficients of those terms containing nonzero powers of the offending z_j or w_j. This will turn out to be useful when extending the model to an array by making scaled copies of Equation (9); while changing the eigenfrequency by scaling the left-hand side, we can zero-out coefficients as needed to reflect the changing relation between the eigenfrequency and the inputs.

It might be interesting to compare Equation (9) to [[22], Equation 15] reproduced below in a form that facilitates the comparison [see also [27], Equation A.1], which is also a normal form for an oscillator coupled to a set of sinusoidal inputs:

$$\dot{z}_1 = (\alpha + \imath)z_1 + \sum_{[r,s,p,q]} c_{rspq} z^{[r,p]} w^{[s,q]}, \quad (10)$$

where

$$c_{rspq} = \begin{cases} d_{rspq} & p = q = 0 \\ \sqrt{\epsilon}^{s+|p|+|q|-1} & \text{otherwise.} \end{cases} \quad (11)$$

In Equation (10), the sum runs over all such vectors $[r, s, p, q]$ for which either $r = s + 1$, $s \in \mathbb{Z}_{>0}$, $p = q = 0$, or $r = 0$, $s \in \mathbb{Z}_{\geq 0}$, $p, q \in \mathbb{Z}^n_{\geq 0}$, $r + s + |p| + |q| > 0$. In contrast, the sum in Equation (9) only runs over the nonnegative solutions to Equation (6), whose coefficients are determined by input frequencies. Hence, whereas Equation (10) applies to inputs of arbitrary frequencies, Equation (9) requires that the frequencies are close to the harmonics of the oscillator's eigenfrequency. On the other hand, in contrast to Equation (9), Equation (10) presupposes a particular form of the coefficients (see Equation 11). Consequently, Equation (9) covers neither a subset, nor a superset of the models covered by Equation (10).

2.1. Model Analysis

In this subsection, we analyse the normal form (Equation 9) derived in section 2. More precisely, we study the stability of the origin ($z_1 = w_1 = 0$). The choice of the origin as the focus of this section is motivated by our previous (arbitrary) choice of the origin as a "quiescent state" of the oscillator (see the beginning of section 2). The reason why we treat its *stability* in such a detail here is that it determines whether, e.g., the oscillator stays quiet (its period was not detected in the input signal) or oscillates (its period was detected in the signal)—see the beginning of section 2.

As the first step of the analysis, note that all solutions to Equation (9) are of the form

$$\begin{aligned}&[z_1, z_2, z_3, \ldots, z_{n+1}]\\&= [z_1(t), \rho_1 e^{\imath(1+\beta_1)\omega_1 t}, \rho_2 e^{\imath(\varphi_2+(1+\beta_2)\omega_2 t)}, \ldots, \rho_n e^{\imath(\varphi_n+(1+\beta_n)\omega_n t)}]\\&[w_1, w_2, w_3, \ldots, w_{n+1}]\\&= [w_1, \rho_1 e^{-\imath(1+\beta_1)\omega_1 t}, \rho_2 e^{-\imath(\varphi_2+(1+\beta_2)\omega_2 t)}, \ldots, \rho_n e^{-\imath(\varphi_n+(1+\beta_n)\omega_n t)}]\,.\end{aligned} \tag{12}$$

Consequently, using the simplified notation $\Omega = [1, 1, \ldots, 1, 2, 2, \ldots, 2, \ldots, N, N, \ldots, N]$ and $\beta = [\beta_1, \beta_2, \ldots, \beta_n]$ as above, and introducing $\rho = [\rho_1, \rho_2, \ldots, \rho_n]$, $\varphi = [0, \varphi_2, \ldots, \varphi_n]$, and $\Omega_\beta = \Omega \circ \beta = [\omega_1\beta_1, \ldots, \omega_n\beta_n]$, we can drop equations for $\dot{z}_2, \dot{z}_3, \ldots, \dot{z}_{n+1}$ and $\dot{w}_2, \dot{w}_3, \ldots, \dot{w}_{n+1}$ from Equation (9) and write

$$\begin{aligned}\dot{z}_1 &= (\alpha + \imath)z_1\\&\quad + \sum_{[r,s,p,q]\in S} c_{rspq} z_1^r w_1^s \rho^{p+q} e^{\imath(p\varphi - q\varphi)} e^{\imath(p(\Omega+\Omega_\beta) - q(\Omega+\Omega_\beta))t}\\\dot{w}_1 &= (\alpha - \imath)w_1\\&\quad + \sum_{[r,s,p,q]\in S} \overline{c_{rspq}} z_1^s w_1^r \rho^{p+q} e^{\imath(q\varphi - p\varphi)} e^{\imath(q(\Omega+\Omega_\beta) - p(\Omega+\Omega_\beta))t}\,.\end{aligned} \tag{13}$$

Introducing new coordinates relative to a rotating frame of reference $e^{\imath t}$,

$$\begin{aligned}u &= z_1 e^{-\imath t}\\v &= w_1 e^{\imath t}\,,\end{aligned} \tag{14}$$

and new parameters,

$$\Delta_{pq} = p\Omega_\beta - q\Omega_\beta\,, \tag{15}$$

Equation (13) reduces to:

$$\begin{aligned}\dot{u} = \dot{z}_1 e^{-\imath t} - \imath z_1 e^{-\imath t} &= e^{-\imath t}(\dot{z}_1 - \imath z_1)\\&= \alpha u + \sum_{[r,s,p,q]\in S} c_{rspq} u^r v^s \rho^{p+q} e^{\imath(p\varphi - q\varphi)} e^{\imath\Delta_{pq}t}\end{aligned} \tag{16}$$

$$\begin{aligned}\dot{v} = \dot{w}_1 e^{\imath t} + \imath w_1 e^{\imath t} &= e^{\imath t}(\dot{w}_1 + \imath w_1)\\&= \alpha v + \sum_{[r,s,p,q]\in S} \overline{c_{rspq}} u^s v^r \rho^{p+q} e^{\imath(q\varphi - p\varphi)} e^{-\imath\Delta_{pq}t}\,.\end{aligned} \tag{17}$$

As we will see in section 3, under a rather generic restriction on Equations (16) and (17), the stability of $u = v = 0$ crucially depends on the relative periodicity and the inharmonicity of the input, as formalized in previous studies, and hence its perceived tension (see section 1). Namely, we require that Equation (16) contains no linear terms in v and, symmetrically, Equation (17) has no linear terms in u. For the remaining linear terms, Equation (6) reduces to

$$p\Omega - q\Omega = 0\,. \tag{18}$$

Assuming the origin is a fixed point of the system of Equations (16) and (17), its stability is determined by the Jacobian of the system evaluated at the origin:

$$\begin{aligned}(J)_{[0,0]} &= \begin{bmatrix}\left(\frac{\partial \dot{u}}{\partial u}\right)_{0,0} & \left(\frac{\partial \dot{u}}{\partial v}\right)_{0,0}\\ \left(\frac{\partial \dot{v}}{\partial u}\right)_{0,0} & \left(\frac{\partial \dot{v}}{\partial v}\right)_{0,0}\end{bmatrix}\\&= \begin{bmatrix}\alpha + \sum_{[1,0,p,q]\in S} & 0\\ c_{10pq}\rho^{p+q}e^{\imath(p\varphi - q\varphi)}e^{\imath\Delta_{pq}t} & \\ 0 & \alpha + \sum_{[1,0,p,q]\in S}\\ & \overline{c_{10pq}}\rho^{p+q}e^{\imath(q\varphi - p\varphi)}e^{-\imath\Delta_{pq}t}\end{bmatrix}\\&= \begin{bmatrix}\alpha & 0\\ 0 & \alpha\end{bmatrix}\\&\quad + \sum_{[1,0,p,q]\in S} \rho^{p+q}\begin{bmatrix}c_{10pq}e^{\imath(p\varphi - q\varphi)} & 0\\ 0 & \overline{c_{10pq}}e^{\imath(q\varphi - p\varphi)}\end{bmatrix}\begin{bmatrix}e^{\imath\Delta_{pq}t} & 0\\ 0 & e^{-\imath\Delta_{pq}t}\end{bmatrix}\\&= A + \sum_{[1,0,p,q]\in S} \rho^{p+q} B_{pq}\begin{bmatrix}e^{\imath\Delta_{pq}t} & 0\\ 0 & e^{-\imath\Delta_{pq}t}\end{bmatrix},\end{aligned} \tag{19}$$

where

$$\begin{aligned}A &= \begin{bmatrix}\alpha & 0\\ 0 & \alpha\end{bmatrix}\\B_{pq} &= \begin{bmatrix}c_{10pq}e^{\imath(p\varphi - q\varphi)} & 0\\ 0 & \overline{c_{10pq}}e^{\imath(q\varphi - p\varphi)}\end{bmatrix}.\end{aligned}$$

In particular, the fixed point solution at the origin is stable, if all the eigenvalues of the Jacobian have negative real parts; while it is unstable if at least one eigenvalue of the Jacobian has positive real part. Apparently, without input ($\rho = 0$), the stability is solely determined by the matrix A, particularly the unfolding parameter α. For positive α, the fixed point at origin is unstable, while for negative alpha, the fixed point at origin is stable. Note that whereas the original class of models (section 2) only covers systems in which the origin is marginally stable ($\alpha = 0$), the "unfolded" class (Equation 8) encompasses the entire spectrum of stability of the origin.

With input, one can view the Jacobian as the matrix A perturbed by a time-dependent term consisting of a sum of oscillators with amplitudes depending exponentially on $p + q$ (with base ρ) and frequencies Δ_{pq}. Thus, if we consider the neural auditory system as spontaneously possessing stable fixed point for a given pitch-detector, i.e., its $\alpha < 0$, only inputs with high amplitude ρ and/or spectral content giving rise to suitable solutions $[1, 0, p, q] \in S$ with small value of $(p + q)$ can perturb the matrix A sufficiently for the fixed point to lose stability at least transiently (note the complicated periodic behavior of the perturbation on the right-hand side), and the pitch-detector show input-modulated oscillatory behavior. An example of such scenario is presented later in section 3.

Let us now in detail assess which monomials appear on the right-hand side of the reduced equations. Note that all solutions to Equation (18) correspond to non-negative integer linear combinations of a finite set of minimal solutions, i.e., they have the structure:

$$p\Omega - q\Omega = 0 \iff [p,q] = kM,\ M \in \mathbb{Z}_{\geq 0}^{l\times 2n},\ k \in \mathbb{Z}_{\geq 0}^{l}, \quad (20)$$

where M is a matrix with the i-th row, denoted $[m_i, n_i]$, equal to the i-th minimum solution to Equation (18) [31]. Consequently (see Equation 15),

$$\begin{aligned}\Delta_{pq} &= p\Omega_\beta - q\Omega_\beta = [p,q][\Omega_\beta, -\Omega_\beta] = kM[\Omega_\beta, -\Omega_\beta] \\ &= \sum_j k_j [m_j, n_j][\Omega_\beta, -\Omega_\beta] = \sum_j k_j (m_j - n_j)\Omega_\beta\,,\end{aligned}$$

$$e^{\imath\,\Delta_{pq}t} = e^{\imath\left(\sum_j k_j (m_j-n_j)\Omega_\beta\right)t} = \prod_j \left(e^{\imath\,(m_j-n_j)\Omega_\beta t}\right)^{k_j},$$

and

$$(J)_{[0,0]} = A + \sum_{k\in\mathbb{Z}_{>0}^{l}} B_{kM} \prod_j \left(\rho^{m_j+n_j} \begin{bmatrix} e^{\imath\,(m_j-n_j)\Omega_\beta t} & 0 \\ 0 & e^{-\imath\,(m_j-n_j)\Omega_\beta t}\end{bmatrix}\right)^{k_j}. \quad (21)$$

As noted above, we model stimulation with $H^{\mathrm{harm}} \subset \Omega$ by zeroing-out those monomials containing nonzero powers of z_j or w_j for which $\Omega_j \notin H^{\mathrm{harm}}$. Consequently, B_{kM} will be nonzero if and only if

$$kM = kM'\,,$$

where $M' \in \mathbb{Z}_{\geq 0}^{l\times 2|H^{\mathrm{harm}}|}$ is a submatrix of M whose rows, $[m'_i, n'_i]$, satisfy

$$\begin{aligned} m'_{ij} > 0 &\implies \Omega_j \in H^{\mathrm{harm}}\,, \\ n'_{ij} > 0 &\implies \Omega_j \in H^{\mathrm{harm}}\end{aligned}$$

(see Equations 18 and 20).

3. RESULTS

In this section, we show how the system of Equations (16) and (17) reacts to stimulation with complex tones varying in relative periodicity and inharmonicity. For the specific case of complex tones consisting of two harmonics, analytical treatment is feasible as we are basically dealing with an interval comprising two pure tones. Let the frequency ratio of the two harmonics be approximated as $i{:}j$, $H^{\mathrm{harm}} = [i,j]$, where $i \leq j$. **Table 1** summarizes the rows of M' together with the corresponding frequencies of the complex exponentials in Equation (21) (i.e., $\pm(m_j - n_j)\Omega_\beta$) and the exponents of ρ (i.e., $m_j + n_j$) for this case. Note that both the frequencies (in absolute value) and the exponents grow monotonically with i, which is precisely the relative periodicity of the interval (see section 1). Hence, as long as B_{kM} does not grow superexponentially with i, the amplitudes of the complex exponentials (i.e., $\rho^{m_j+n_j}$ in Equation 21) increase with decreasing relative periodicity of the interval.

TABLE 1 | Context-derived heterogeneous functions of monocyte subsets.

m'_k	n'_k	$m'_k + n'_k$	$(m'_k - n'_k)\,[\beta_i, \beta_j]\,[i,j]$
[1, 0]	[1, 0]	[2, 0]	0
[0, 1]	[0, 1]	[0, 2]	0
[j, 0]	[0, i]	[j, i]	$ij(\beta_i - \beta_j)$
[0, i]	[j, 0]	[j, i]	$-ij(\beta_i - \beta_j)$

Quantities appearing in Equation (21) for the special case of stimulation with an interval (that is, $H^{\mathrm{harm}} = [i,j]$, $i \leq j$).

Further, it can be shown that the frequencies above also grow (in absolute value) with the inharmonicity of the interval, the other factor in perception of musical tension considered here. Let f_1 and f_2 denote the lower and the higher frequency of the interval, respectively, that is,

$$\begin{aligned} f_1 &= (1+\beta_i)i\,, \\ f_2 &= (1+\beta_j)j = \big((1+\beta_i) + (\beta_j - \beta_i)\big)\,j \end{aligned} \quad (22)$$

(see Equation 12). Additionally, let

$$f_0 = 1 + \beta_i\,.$$

The inharmonicity of the interval $[f_1, f_2]$ with respect to the fundamental frequency f_0 is defined as its weighted Manhattan distance to the interval $[if_0, jf_0]$, comprising the i-th and the j-th harmonic of f_0. The distance is weighted by the squared signal amplitudes and normalized by f_0 and the sum of the squared signal amplitudes [11]. In our particular case, the inharmonicity $I^{ij}_{f_1 f_2}$ is equal to

$$\begin{aligned} I^{ij}_{f_1 f_2} &= \frac{2}{f_0}\frac{|f_0 i - f_0 i|\rho_i^2 + |\big(f_0 + (\beta_j - \beta_i)\big)\,j - f_0 j|\rho_j^2}{\rho_i^2 + \rho_j^2} \\ &= \frac{2}{f_0}\frac{j|\beta_j - \beta_i|\rho_j^2}{\rho_i^2 + \rho_j^2} \\ \frac{f_0\left(\rho_i^2 + \rho_j^2\right)}{2\rho_j^2} I &= j|\beta_j - \beta_i|\,. \end{aligned}$$

Indeed, the frequencies grow (in absolute value) with the inharmonicity of the interval. Consequently, noting that A governs the stability of the fixed point solution at origin without input ($\rho = 0$), we conclude that, under the above assumption on B_{kM}, pure-tone intervals with lower relative periodicity and lower inharmonicity (i.e., those perceived as less tense) cause a higher-amplitude and slower fluctuation of the driven system eigenvalues around those of A than those with higher relative periodicity and inharmonicity (perceived as more tense).

Note that there is an ambiguity of approximation represented by a choice of i and j in Equation (22). Further, Equations (16) and

(17) are far from being a global model of perception of tension even in complex tones consisting of just two harmonics; they are local in the sense that they only model perception of a particular tone with respect to a particular approximation. Last but not least, we still have to demonstrate that the above fluctuations of stability translate to features of the oscillatory dynamics in a meaningful way. We address these issues now when considering the general case of stimulation with a complex tone consisting of more than two harmonics. To this end, we construct an array of models like Equation (16) differing in eigenfrequency. Here, each eigenfrequency represents a choice of f_0 in Equation (22) so that the entire array essentially works as a pitch detector. Time traces from simulations of such an array are depicted in **Figures 1**, **2**.

The equations for the array were derived by applying the above restriction on linear terms to Equation (9), writing-out the inputs (Equations 12, 16, 17) and using Equations (6) and (18),

$$\begin{aligned}\dot{z}_1 = &\left(\alpha + \imath + \sum_{\substack{p\Omega - q\Omega = 0 \\ |p|+|q|\geq 1}} c_{10pq}\rho^{p+q}e^{\imath(p-q)\varphi}e^{\imath(p-q)\Omega_\beta t}\right) z_1 \\ &+ z_1|z_1|^2 \sum_{k=0}^{\infty}|z_1|^{2k} \sum_{p\Omega - q\Omega = 0} c_{(s+1)spq}\rho^{p+q}e^{\imath(p-q)\varphi}e^{\imath(p-q)\Omega_\beta t} \\ &+ \sum_{\substack{[r,s,p,q]\in S \\ p\Omega - q\Omega \neq 0 \\ r+s\geq 2}} c_{rspq} z_1^r \overline{z_1}^s \rho^{p+q} e^{-\imath(r-s-1)t} e^{\imath(p-q)\varphi} e^{\imath(p-q)\Omega_\beta t}\,, \quad (23)\end{aligned}$$

then truncating the higher-order terms,

$$\begin{aligned}\dot{z}_1 = &\left(\alpha + \imath + \sum_{\substack{p\Omega - q\Omega = 0 \\ p=(\omega_j)_i,\, q=(\omega_i)_j}} c_{10pq}\rho^{p+q}e^{\imath(p-q)\varphi}e^{\imath(p-q)\Omega_\beta t}\right) z_1 \\ &+ c_{2100} z_1 |z_1|^2\,,\end{aligned}$$

and, finally, setting

$$c_{2100} = -1\,, c_{10pq} = 1\,, \varphi = 0\,, z_1 = z_{1k}\,, \rho = \rho_k\,,$$

and scaling the time for convenience by the eigenfrequency, f_k, which yields a parametrically-forced normal form for supercritical Andronov-Hopf bifurcation:

$$\dot{z}_{1k} = f_k \left(\left(\alpha + \imath + \sum_{\substack{p\Omega - q\Omega = 0 \\ p=(\omega_j)_i,\, q=(\omega_i)_j}} \rho_k^{p+q} e^{\imath(p-q)\Omega_\beta t} \right) z_{1k} - z_{1k}|z_{1k}|^2 \right). \quad (24)$$

Here, $(\omega_j)_i$ signifies a vector with ω_j at the i-th position and zero otherwise. Ω and β are set as

$$\Omega = [\omega_1, \omega_2, \ldots, \omega_{54}] = [1, 1, \ldots, 1, 2, 2, \ldots, 2, \ldots, 6, 6, \ldots, 6]$$

$$\beta = \Omega_{24TET}\frac{1}{\Omega} - 1$$

(see Equations 1, 8), where each element of Ω_{24TET} approximates the corresponding element of Ω as a power of $2^{\frac{1}{24}}$ (the 24-tone equal-tempered tuning); Ω and Ω_{24TET} are aligned in such a way that $\omega_5 = \omega_{24TET,5} = 1$ and hence $\beta_5 = 0$. The oscillators (Equation 24) receive connections from a bank of input units—linear oscillators with eigenfrequencies spanning from $B\flat_0$ to B_4 in quarter-tone steps. In accordance with Equation (8), each oscillator (Equation 24) is only connected to input units with frequencies ($f_k\omega_i\beta_i$ in Equation 8, after scaling by f_k) approximating its harmonics (in the above tuning) and, additionally, to frequencies up to 4 quarter-tones below and above these. In other words, it does not have fixed homogeneous connectivity input strength from all input units, but rather receives (weighted) input only from input units with frequencies close to its first six approximate harmonics; the connectivity of each oscillator is thus effectively defined by a connectivity pattern or kernel consisting of six unimodal elementary Gaussian kernels ($k(l) = e^{-0.5l^2}$; $l \in \{-4, \ldots, 4\}$ and $k(l) = 0$ otherwise) centered at the harmonics. See visualization of the connectivity kernel in **Figure 3**. Moreover, only the connections emanating from the input units whose frequencies are included in the stimulus are set to have nonzero amplitude in the respective simulation. Note that by fixing a set of eigenfrequencies (corresponding to different choices of f_0 in Equation 22) and input units and restricting the connectivity to (near-)harmonics, there remains no ambiguity in approximation of the input; each oscillator, as long as the input falls within the reach of its connectivity kernel, approximates the input in its own, unique, way.

All simulations were run from initial conditions

$$z_{1k}(t_0 = 0) = 0.001\,,$$

with

$$\alpha = -0.001\,,$$

a parameter setting corresponding to the (almost loss of) stability (without input) of the fixed point $z_{1k} = 0$. Three alternative inputs were applied, whose spectra can be seen in **Figure 4**. The first corresponds to harmonic input with the C tone at its base plus its first five harmonics (tones with integer multiple frequencies of the base tone). The second input results from a transformation of the first which increases inharmonicity while the third is a result of a transformation which increases roughness.

As can be seen from **Figures 1**, **2**, both transformations seem to increase fluctuation of stability of the origin, as predicted by our analysis pertaining to two-frequency stimulation. This results in an increase in amplitude modulation across the oscillator array and a corresponding decrease in peak amplitude (see **Figure 5**). In other words, an increase in perceived musical tension seems to be related to an increase in fluctuation of stability of the origin which manifests itself as an absence of a stable dominant amplitude peak. These preliminary observations are largely confirmed by computing the minimum and the maximum of each oscillator's amplitude trace (see **Figure 6**). Consequently, we put forward the absence of a stable dominant amplitude peak as a hallmark of perceived musical tension in our model.

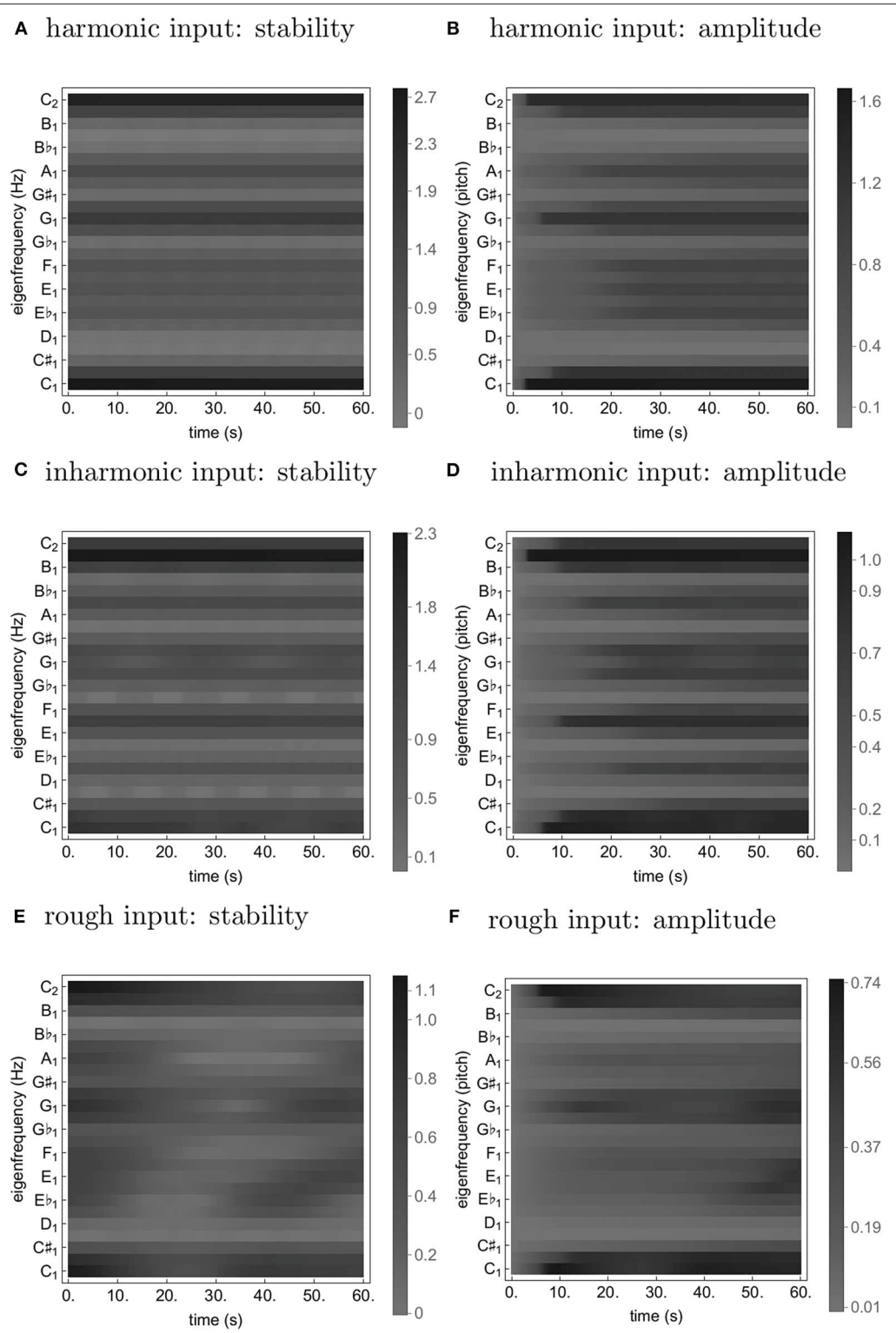

FIGURE 1 | Time traces from simulations of Equation (24) with a soft harmonic **(A,B)**, inharmonic **(C,D)**, and rough **(E,F)** input. The stability **(A,C,E)** is quantified as $\Re\left((J)_{[0,0]}\right)$ (see Equation 2); the amplitude **(B,D,F)** is simply $|z_{1k}|$.

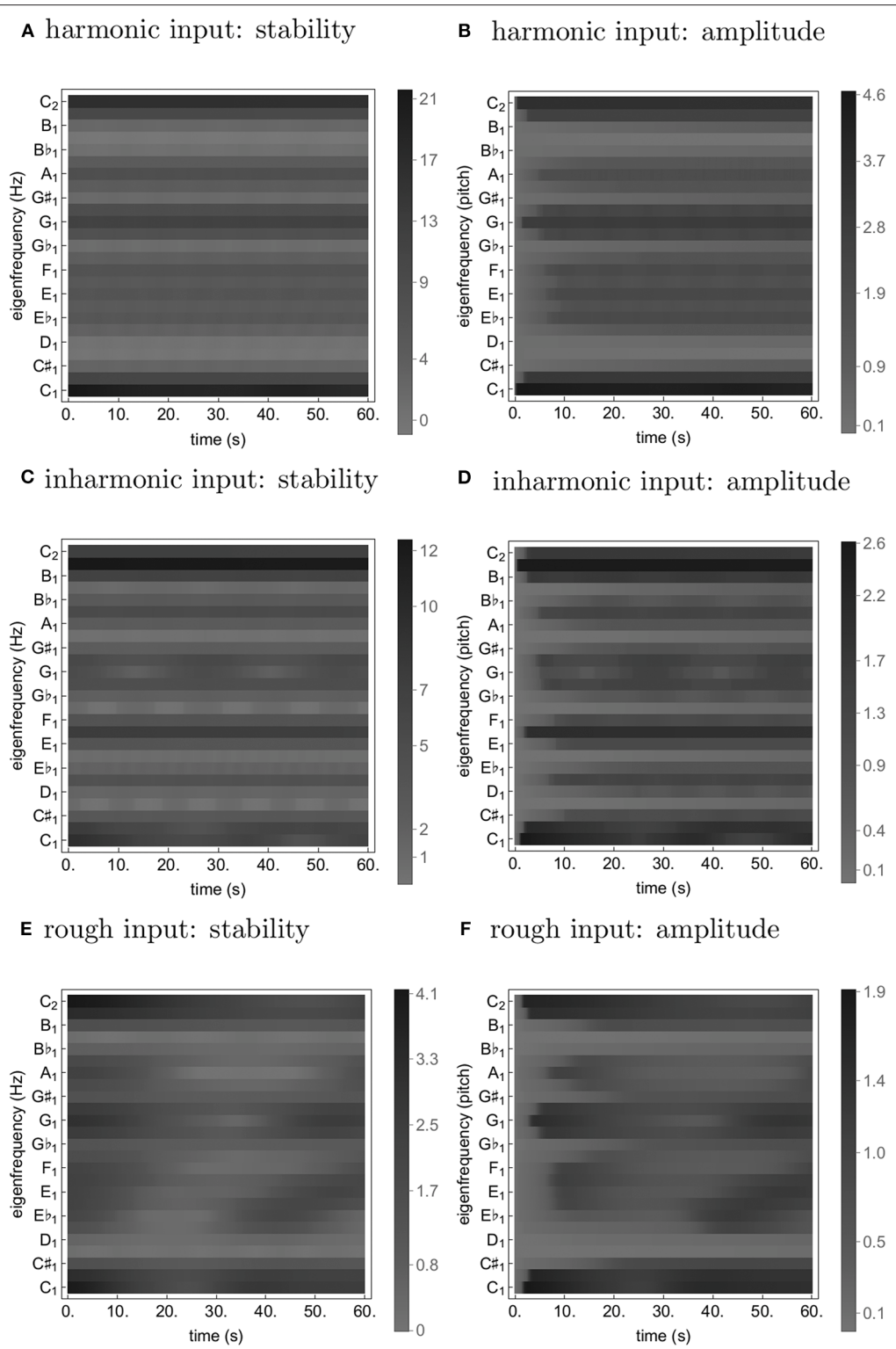

FIGURE 2 | Time traces from simulations of Equation (24) with a loud harmonic **(A,B)**, inharmonic **(C,D)**, and rough **(E,F)** input. The stability **(A,C,E)** is quantified as $\Re\left((J)_{[0,0]}\right)$ (see Equation 2); the amplitude **(B,D,F)** is simply $|z_{1k}|$.

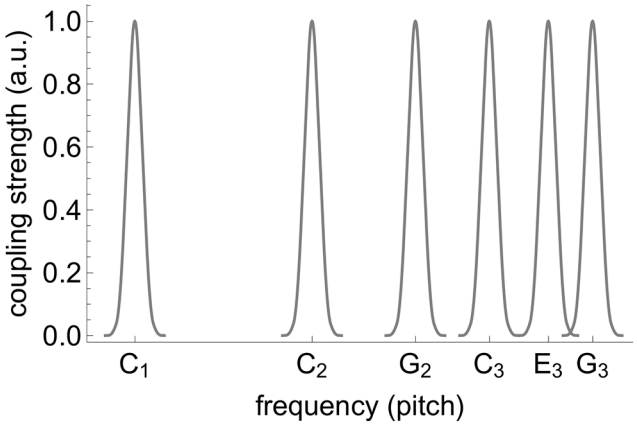

FIGURE 3 | Connectivity of the oscillator with eigenfrequency C_1 ($k = 1$, see Equation 24). The first "blob" represents connections which stem from input units with eigenfrequencies ranging from $B\flat_0$ to D_1 and target the oscillator's "slots" corresponding to $\rho_{1,1}, \rho_{1,2}, \ldots, \rho_{1,9}$ in Equation (24); likewise, the second blob connects $B\flat_1, \ldots, D_2$ to $\rho_{1,10}, \rho_{1,11}, \ldots, \rho_{1,18}$ etc. The connectivity of any other oscillator is obtained by shifting this kernel so that the center of the first blob is aligned with the oscillator's eigenfrequency.

4. DISCUSSION

We propose the absence of a stable unambiguous pitch detection modeled as the absence of a pronounced amplitude peak in an array of oscillators to be a correlate of timbre-induced musical tension. In the class of oscillators we chose for populating the array, the amplitude of the limit cycle is determined by the stability of the origin; if the stability switches between a stable and an unstable regime fast enough, the amplitude doesn't have enough time to grow. We show that the frequency and magnitude of this switching depends on inharmonicity and roughness of the input to the oscillator. Imagine such an oscillator is actually present in the brain; when subject to a tense (inharmonic and/or rough) stimulus, it will remain almost silent, leading to an "unclear," "unstable," "difficult to memorize" etc. percept (see **Figures 1**, **2D,F**). In contrast, a less tense stimulus would result in a "clear", "stable", "easy to memorize" etc. percept (see **Figures 1**, **2B**). Of course, neurophysiological and neuroimaging evidence shows the what we have in our head is not a single oscillator, but rather an entire bank of them; we show in our simulations that the results of our analysis of single oscillator generalize to an array of them in the sense that the average oscillation amplitude across the array is lower for tense than less tense stimuli (see **Figure 6**).

Of course, tension is clearly not a one-dimensional phenomenon and different aspects of it could be related to different aspects of the underlying neurodynamics. For instance, in a nonlinear model like the one proposed here, loudness of the input is going to affect both the general amplitude of the oscillations and their temporal fluctuations—in a frequency-dependent manner, as our example simulations for two loudness levels suggest. We consider disentangling these not necessarily orthogonal dimensions of tension as a natural extension of the currently proposed modeling framework.

We have proposed a neurodynamical model of musical tension (see Equation 24) which reproduces existing empirical results on timbral correlates of tension, is consistent with neuroimaging findings [14] in that consonant stimuli compared to dissonant stimuli elicit more sustained periodic neuronal activity of higher amplitude, and due to its generative nature can provide prediction of perceived tension of an arbitrary sound input. More precisely, we have demonstrated that both inharmonicity and roughness make the spectrum of the simulated signal flatter and more variable (wider range over time) (see **Figures 1**, **2**, **6**). Note that while [14] quantified the periodicity by the amplitude of the autocorrelation peak of the signal spectra, we rather proposed the absence of a temporally persistent, pronounced amplitude peak in the spectrum of the elicited neural activity as a possible correlate of tension—a related indicator that is also present in the results presented in [14]. One might even speculate, based on the similarity of the spectrum to the major key profile [32], that the same principles underlie perception of tonality.

Considering the simulation results reported above in more detail, we note that the overall increase in fluctuation of stability of the origin for the inharmonic and the rough input as compared to the harmonic one can be explained based on the analytical insights into the dynamics of a single oscillator obtained earlier. More precisely, the nearly-harmonic relations in the inharmonic and the rough input introduce oscillating terms into most of the oscillators' coupling functions; the increase of amplitude modulation is, in turn, accounted for by the fact that the amplitude of the stable limit cycle of Equation (24) is determined by the stability of the origin. The decrease of the peak amplitude is, for the inharmonic input, probably due to the connectivity; there are no exact harmonic relations in the input and hence no oscillator can align its connectivity kernel optimally with the input (see **Figure 3**). For the rough input, it might be a consequence of scaling down the amplitudes of its frequency components to keep the overall loudness equal to that of the other inputs which consist of fewer harmonics (see **Figure 4**).

As for our general approach, a few comments are in order. First, for the sake of simplicity, we chose a subclass of *multiple centers* [a generalization of *double centers*; see [26]] as our family of models. It might be an interesting avenue for future research to determine whether there are other families of models in which relative periodicity and inharmonicity of the input plays such an important role.

Also for the sake of simplicity, we only considered relative periodicity and inharmonicity of pure-tone dyads. For general sounds, we would be dealing with the set of nonnegative solutions to a general linear Diophantine equation (Equation 18). To the best of our knowledge, the structure of the set (its minimum generators) can only be determined algorithmically [e.g., [31]]. This makes analytical insights virtually impossible in the general case.

Further, concerning the phenomenon wherein loud music is perceived as more tense than soft music [8], we argue that, replacing the bank of input units with a model of cochlea, the effective input generated by a loud harmonic spectrum would be very similar to the rough input used in the simulations reported

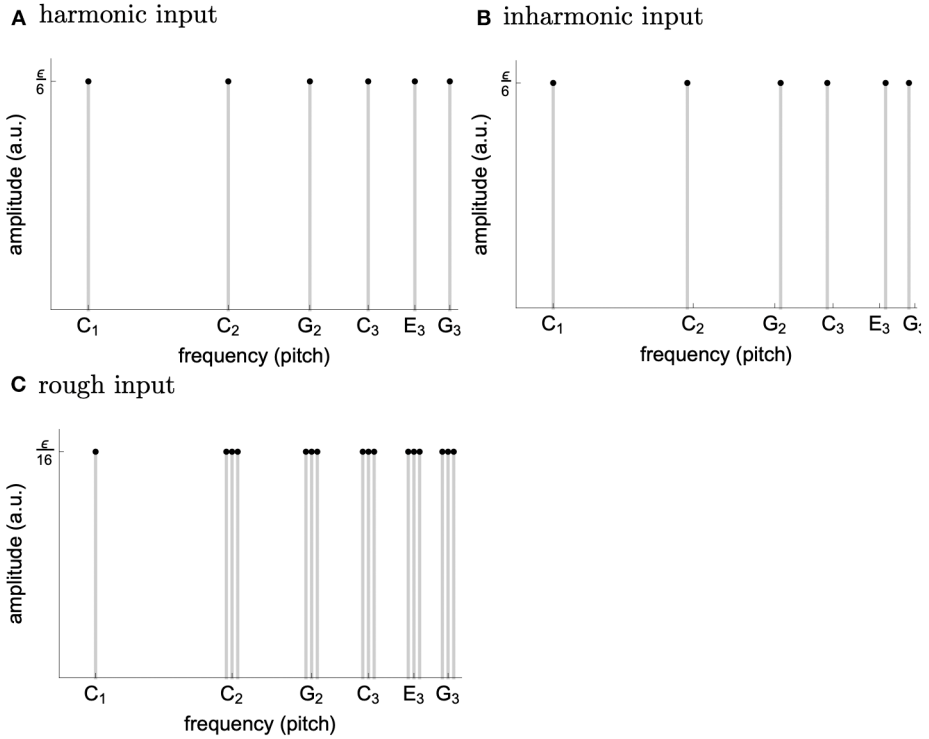

FIGURE 4 | Spectrum of the harmonic **(A)**, inharmonic **(B)**, and rough **(C)** input used in the simulations. The construction loosely follows the procedure from [11]. We chose $\epsilon = 3$ for the soft and $\epsilon = 5.4$ for the loud inputs.

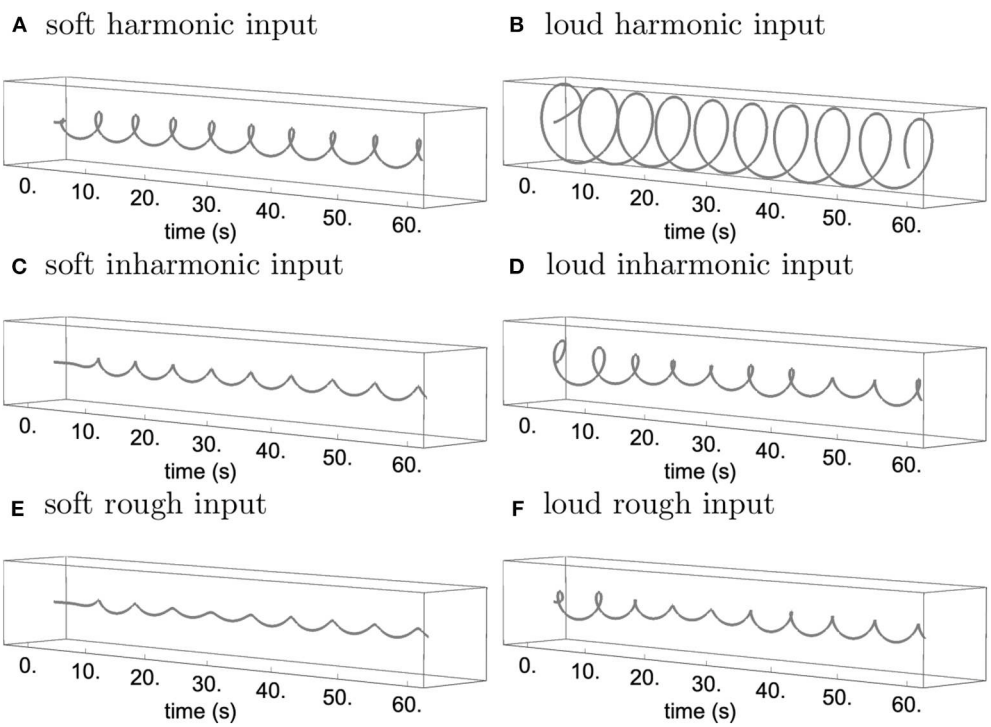

FIGURE 5 | Trajectory of the C_1 oscillator with soft/low-amplitude input corresponding to **Figure 1 (A,C,E)** and loud/high-amplitude input corresponding to **Figure 2 (B,D,F)** for the harmonic **(A,B)**, inharmonic **(C,D)**, and rough input **(E,F)**, rendered in the complex plane.

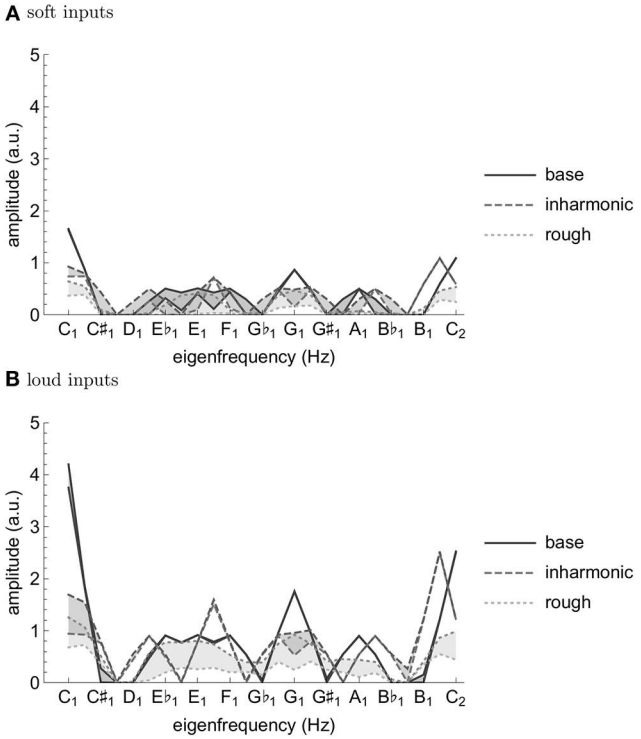

FIGURE 6 | Minima and maxima of the oscillator amplitude traces from simulations with soft/low-amplitude input corresponding to **Figure 1 (A)** and loud/high-amplitude input corresponding to **Figure 2 (B)** with the first 20 s dropped to attenuate the effect of the same initial conditions. Incidentally, note the similarities between **Figure 6** and the major key profile from Krumhansl and Kessler [32].

here. More precisely, we expect the loud harmonic spectrum to displace not only those segments of the basilar membrane whose eigenfrequencies match the harmonics, but also the adjacent segments (see **Figure 4**). This way, the effect of loudness would be accounted for by a combination of cochlear physiology and sensitivity of our model to roughness.

Finally, even though the choice of spectral representation was motivated by our interest in contemporary art music, especially the so-called "spectral music," the model presented here is applicable to any kind of music; indeed, even music composed with traditional categories in mind ends up being rendered as sound which can be fed into our model.

To conclude, mapping perception to neurodynamics is hard. However, from time to time, a favorable constellation of research sheds light on the underlying physiology. The fruitful concept of *relative periodicity* [13] suggests that roughness, as one of the perceptual "dimensions" of timbre contributing to tension, might originate in (neural) resonance. Indeed, in this study, we have shown that the dynamics of stability of the origin in a wide class of periodically forced nonlinear oscillators crucially depends on the relative periodicity of the input and, additionally, on its inharmonicity. Since roughness and inharmonicity are principal constituents of perceived tension, we have effectively put forward a possible neurodynamical explanation of musical tension. Moreover, for a particular model belonging to the above class, we have demonstrated by simulations that tense inputs result in an absence of a persistent dominant peak in the spectrum of the time series generated by the model.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

JH and MH contributed to conception, theoretical analysis, design of the study, manuscript revision, read and approved the submitted version. MH implemented the simulations and visualizations and wrote the first draft of the manuscript.

FUNDING

This study was funded by the project Nr. LO1611 with a financial support from the MEYS under the NPU I program and with institutional support RVO:67985807.

ACKNOWLEDGMENTS

We thank Pavel Sanda and Hana Markova for reading the manuscript and providing helpful comments.

REFERENCES

1. Anderson W, Mathiesen TJ. *Ethos*. Oxford University Press (2001). Available online at: http://www.oxfordmusiconline.com/subscriber/article/grove/music/09055
2. Huron D. *Sweet Anticipation*. Cambridge, MA: The MIT Press (2006).
3. Aldwell E, Schachter C, Cadwallader A. *Harmony & Voice Leading*. Exeter: Schirmer Cengage Learning (2011).
4. Lerdahl F, Krumhansl CL. Modeling tonal tension. *Music Percept*. (2007) **24**:329–66. doi: 10.1525/mp.2007.24.4.329
5. Murail T. The revolution of complex sounds. *Contemp Music Rev*. (2005) **24**:121–35. doi: 10.1080/07494460500154780
6. Murail T. Target practice. *Contemp Music Rev*. (2005) **24**:149–71. doi: 10.1080/07494460500154814
7. Fineberg J. Spectral music. *Contemp Music Rev*. (2000) **19**:1–5. doi: 10.1080/07494460000640221
8. Ilie G, Thompson WF. A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Percept*. (2006) **23**:319–30. doi: 10.1525/mp.2006.23.4.319
9. Bigand E, Parncutt R, Lerdahl F. Perception of musical tension in short chord sequences: the influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Percept Psychophys*. (1996) **58**:125–41. doi: 10.3758/BF03205482
10. Pressnitzer D, McAdams S, Winsberg S, Fineberg J. Perception of musical tension for nontonal orchestral timbres and its relation to psychoacoustic roughness. *Percept Psychophys*. (2000) **62**:66–80. doi: 10.3758/BF03212061
11. Farbood MM, Price KC. The contribution of timbre attributes to musical tension. *J Acoust Soc Am*. (2017) **141**:419–27. doi: 10.1121/1.4973568

12. Hutchinson W, Knopoff L. The acoustic component of Western consonance. *J New Music Res*. (1978) **7**:1–29. doi: 10.1080/09298217808570246
13. Stolzenburg F. Harmony perception by periodicity detection. *J Math Music*. (2015) **9**:215–38. doi: 10.1080/17459737.2015.1033024
14. Lee KM, Skoe E, Kraus N, Ashley R. Neural transformation of dissonant intervals in the auditory brainstem. *Music Percept*. (2015) **32**:445–59. doi: 10.1525/mp.2015.32.5.445
15. Bidlack RA. *Music From Chaos: Nonlinear Dynamical Systems as Generators of Musical Materials*. San Diego, CA: University of California (1990).
16. Boon JP, Decroly O. Dynamical systems theory for music dynamics. *Chaos*. (1995) **5**:501–8. doi: 10.1063/1.166145
17. Hennig H, Fleischmann R, Fredebohm A, Hagmayer Y, Nagler J, Witt A, et al. The nature and perception of fluctuations in human musical rhythms. *PLoS ONE*. (2011) **6**:e26457. doi: 10.1371/journal.pone.0026457
18. Coombes S, Lord GJ. Intrinsic modulation of pulse-coupled integrate-and-fire neurons. *Phys Rev E*. (1997) **56**:5809. doi: 10.1103/PhysRevE.56.5809
19. Lots IS, Stone L. Perception of musical consonance and dissonance. *J R Soc Interface*. (2008) **5**:1429–34. doi: 10.1098/rsif.2008.0143
20. Heffernan B, Longtin A. Pulse-coupled neuron models as investigative tools for musical consonance. *J Neurosci Methods*. (2009) **183**:95–106. doi: 10.1016/j.jneumeth.2009.06.041
21. Ushakov YV, Dubkov AA, Spagnolo B. Spike train statistics for consonant and dissonant musical accords in a simple auditory sensory model. *Phys Rev E*. (2010) **81**:041911. doi: 10.1103/PhysRevE.81.041911
22. Large EW, Almonte FV, Velasco MJ. A canonical model for gradient frequency neural networks. *Phys D*. (2010) **239**:905–11. doi: 10.1016/j.physd.2009.11.015
23. Large EW. A Dynamical systems approach to musical tonality. In: Huys R, Jirsa VK, editors. *Nonlinear Dynamics in Human Behavior. Vol. 328 of Studies in Computational Intelligence*. Berlin; Heidelberg: Springer (2011). p. 193–211. doi: 10.1007/978-3-642-16262-6_9
24. Large EW, Almonte FV. Neurodynamics, tonality, and the auditory brainstem response. *Ann NY Acad Sci*. (2012) **1252**:E1–7. doi: 10.1111/j.1749-6632.2012.06594.x
25. Large EW, Kim JC, Flaig NK, Bharucha JJ, Krumhansl CL. A neurodynamic account of musical tonality. *Music Percept*. (2016) **33**:319–31. doi: 10.1525/mp.2016.33.3.319
26. Murdock J. *Normal Forms and Unfoldings for Local Dynamical Systems*. New York, NY: Springer-Verlag (2003). doi: 10.1007/b97515
27. Lerud KD, Almonte FV, Kim JC, Large EW. Mode-locking neurodynamics predict human auditory brainstem responses to musical intervals. *Hear Res*. (2014) **308**:41–9. doi: 10.1016/j.heares.2013.09.010
28. Kim JC, Large EW. Signal processing in periodically forced gradient frequency neural networks. *Front Comput Neurosci*. (2015) **9**:152. doi: 10.3389/fncom.2015.00152
29. Kim JC, Large EW. Mode locking in periodically forced gradient frequency neural networks. *Phys Rev E*. (2019) **99**:022421. doi: 10.1103/PhysRevE.99.022421
30. Parncutt R. Template-matching models of musical pitch and rhythm perception. *J New Music Res*. (1994) **23**:145–67. doi: 10.1080/09298219408570653
31. Clausen M, Fortenbacher A. Efficient solution of linear diophantine equations. *J Symbol Comput*. (1989) **8**:201–16. doi: 10.1016/S0747-7171(89)80025-2
32. Krumhansl CL, Kessler EJ. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychol Rev*. (1982) **89**:334. doi: 10.1037/0033-295X.89.4.334

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.