VPS VIBRATIONS IN PHYSICAL SYSTEMS

DOI: 10.21008/j.0860-6897.2023.2.09
Vibrations in Physical Systems, 2023, 34(2), 2023209

Impact on the sensation of consonance by controlling spectral components of simultaneous pitches consisting a chord in sound synthesis using the additive method

Marek PLUTA

Department of Mechanics and Vibroacoustics, AGH University of Krakow, al. Mickiewicza 30, 30-059 Krakow, Poland

Corresponding author: Marek PLUTA, email: pluta@agh.edu.pl

Abstract In pitched sounds the ratios between frequencies of spectral components remain close to natural numbers. In acoustic instruments these ratios deviate from exact values due to damping or boundary conditions, complicating use of musical systems defining frequency ratios associated with musical intervals. If such system is imposed on the fundamental frequencies of pitches consisting a chord, higher spectral components may deviate noticeably, strengthening or weakening the effects of beating and roughness caused by their proximity, and changing chord characteristics in terms of consonance-dissonance gradation. Unlike acoustic instruments, sound synthesizers can precisely adjust components of generated signal, and use this phenomenon as a controllable timbre effect. For this purpose spectral components of a chord within the range of beating and roughness are modified to gradually strengthen or weaken both phenomena. The study presents the assumptions of the effect and its implementation in a sound synthesizer based on the additive method.

Keywords: consonance, dissonance, sound timbre, additive synthesis.

1. Introduction

One of the qualities attributed to vertical, i.e. simultaneously sounding structures of pitches in music, such as harmonic intervals or chords, is their consonance. In the most basic approach all such multi-pitch structures can be categorised either as consonant, or dissonant, with profound musical implications regarding their function in harmony, and place in harmonic progressions within a musical work.

More detailed study of the phenomenon, based on psychoacoustics rather than simple arithmetic ratios of component frequencies, reveals that a gradation of consonance is better approach than a simple binary categorisation [1–4]. Still, with regards to particular chords, their assessment within the consonance-dissonance continuum is generally regarded as fixed. However, the problem becomes more complex when various, and not necessary time-constant musical tuning systems defining interval frequency ratios are applied, which is usually the case with most acoustic, and also with some of more advanced electronic instruments. Furthermore, in such instruments component frequencies often deviate from natural ratios. Therefore, fundamental frequency ratios between pitches within a chord, defined by musical interval systems, are combined with non-natural frequency ratios of components consisting individual pitched sounds. It leads to a complex scenario, where the actual degree of consonance is difficult to predict.

While it is a possible problem for acoustic instruments, leading to difficulties in obtaining acceptable pitch intonation in consonant structures, or contrary – sufficiently harsh dissonance, the phenomenon may be exploited in selected sound synthesis methods, where component frequencies of individual pitches can be controlled and adjusted in several ways, thus changing chord characteristics in terms of consonance-dissonance gradation. This would provide a new, controllable timbre effect for sound synthesizers. The effect could be applied to morph a single chord between dissonant and consonant without changing its pitch structure, though it would be stronger in case of originally dissonant structures, with more spectral components to modify.

This study discusses the basis of the phenomenon, assumptions of the effect, and presents an example of its implementation in a sound synthesizer based on the additive method.

2. Consonance and dissonance

A common simplification assumes that pitched sounds are multi-tone compounds with frequency ratios between spectral components remaining very close to natural numbers. Their pitch is equivalent to pitch of a sine tone of the same frequency as fundamental tone of a multi-tone. However, in acoustic instruments these frequency ratios deviate from exact values due to e.g., damping or boundary conditions [5]. In case of a single grand piano string representing note C4 (f_0 = 261.63 Hz) it has been measured that the third component is already detuned towards higher frequencies by 2.2 ct (0.13 %), which can be perceptible [6, 7], and the sixth component is detuned by as much as 9.8 ct (or 0.57 %) [8]. In case of a guitar string representing note E2 (f_0 = 82.41 Hz) component detuning can exceed 20 ct [9]. The values may vary significantly even in the same instrument, due to string wear [9] and other factors, such as material defects or initial irregularity of string linear density, but also string-constraints related frequency-dependent boundary conditions [5, 8].

Chords are usually composed of three or more sounds characterised by different pitches. Their function and musical meaning is dictated by structure of internal pitch distances, measured with musical intervals [10]. An interval represents a certain ratio of fundamental frequencies between two pitched sounds. These ratios are defined by musical systems. One such system, twelve-tone equal temperament (12-TET) [11], is used as the most common basis for interval measurement in studies regarding musical instruments, and is often considered the one most widespread in the Western music. However, apart from sound synthesizers, it may not be the case, and acoustic instruments often employ more or less different systems in order to obtain perceptually better results, and take into account physical properties of particular instrument. Such systems can "sacrifice" tuning of certain less commonly used structures in order to obtain better tuning of more important ones. With this regard 12-TET is classified as one of "compromise" systems [11]. The other group of systems, exploiting "natural" intervals, i.e. with frequency ratios expressed by natural numbers, cannot be utilised directly in standard instruments due to several reasons. They produce a number of variants of each pitch, differing by certain "comma", and require a selection to obtain only 12 pitches within an octave. In turn, it makes only some of intervals natural, with the remaining ones noticeably worse due to frequency ratios far exceeding small natural numbers.

When a musical system is imposed onto an acoustic instrument with detuned-from-natural ratios between spectral components to produce a multi-pitch chord, it results in a complex spectral structure, with components in various, difficult to predict ratios. These ratios are the basis for the sensation of consonance or dissonance.

There are various interpretations for the meaning of consonance. In music however, it is commonly agreed that a consonant structure is stable and does not introduce sensation of tension. Thus it can be used as finishing structure. A dissonant structure, on the contrary, is unstable, introduces a sensation of tension and needs a "resolution", i.e. it tends to progress into another structure, a consonant one. The simplest structures characterised as consonant or dissonant are intervals – pairs of pitched sounds. In larger structures intervals between all pitches have to be consonant in order for the chord to be consonant. Even one dissonant interval makes a whole chord dissonant [10]. However, this categorisation is often too simplified, and due to above mentioned frequency ratio complications, may require a more detailed approach.

It has been proposed to introduce a gradation of consonance or dissonance in musical structures on the basis of sensory dissonance evoked by beating and roughness – both related to amplitude modulation and resulting from proximity of certain spectral components [12, 13]. Clearly, for two pitches in a distance of natural interval, and with component spectral frequencies of each pitch forming perfect harmonic series, there is no beating and roughness, because all frequencies are either equal, or sufficiently distant. But in most of real situations various amount of sensory dissonance is present, and this amount is a source of consonance gradation.

3. Concept and assumptions of the effect

The concept is based on results of a previous semi-formal, unpublished study, where five variants of C-major chord were prepared and presented for comparison in pairs. One variant consisted of pitches in 12-TET system reproduced using violin samples. Four remaining variants were based on the same recording, but with certain spectral components filtered out. Components to be removed were selected in such a manner, as to reduce amplitude modulation effects. The principle was to locate the strongest local components, and remove all weaker components in their proximity that would cause fast beating or roughness, i.e. weaker components were left intact if they were closer than a given distance expressed in hertz (beating), or farther than a given distance expressed in percent of equivalent rectangular bandwidth (ERB). Stop-bands were applied both above, and below strong components. Particular values are presented in Table 1.

Table 1. Bands of removed spectral components in proximity of a strong component.

Variant	Close limit [Hz]	Far limit [% of ERB]
1	11	25
2	32	25
3	54	25
4	54	40

The group of five listeners, all students of the Academy of Music in Krakow, used closed headphones in a listening room. They were presented with 700 pairs consisting of various (repeating) combinations of variants, and were asked to point a preferred variant of a chord in each pair (see Fig. 1). The results had no statistical significance, but variants 1 and 4 were chosen above others. Later, the same examples were presented multiple times, under various conditions (headphones, or loudspeakers) to different groups, in an informal manner, and with a different task – to point the more consonant variant. In these cases almost all listeners pointed to variants with "roughness" bands removed when compared to non-modified chord.

These informal results pointed to a conclusion that even though it may be difficult to "prefer" more consonant or more dissonant chord, because of unclear base for preference, the difference in consonance gradation even within the same chord, caused only by removal of roughness bands, was relatively easy to perceive. The concept then was to exploit this phenomenon, and the most straightforward way would be to implement it in a sound synthesizer, because unlike in acoustic instruments, there are sound synthesis methods that allow to precisely adjust individual components of generated signal. Once implemented, it would serve as a controllable, timbre-related effect, operating, firstly, by finding spectral components of a chord within the range of beating or roughness, and secondly, by modifying selected components to strengthen or weaken both phenomena.

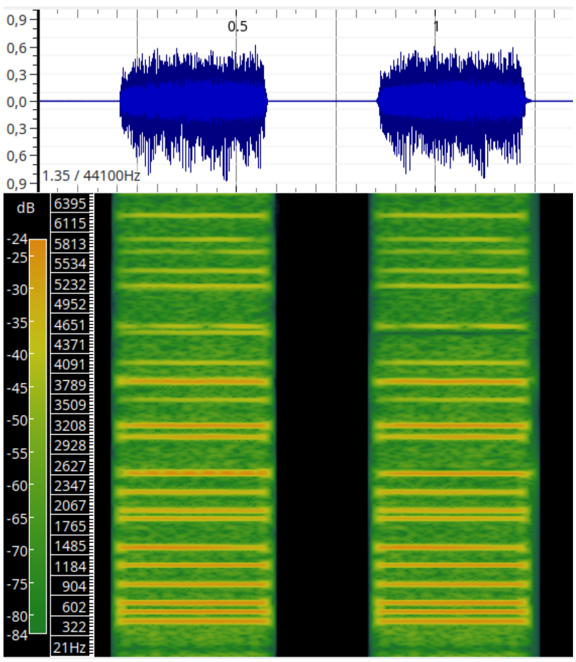

Figure 1. An example pair of chords used in informal tests (top – waveform with time [s] on horizontal and amplitude [-] on vertical axis, bottom – spectrogram with time [s] on horizontal and frequency [Hz] on vertical axis); the first in pair is unmodified, the second one has weaker components removed within frequency band between 11 Hz and 0.25 ERB around stronger components.

A method particularly well-suited for this kind of operation is the additive synthesis [14, 15], where a signal is produced by mixing sine components. Each component can have its frequency and amplitude controlled individually, using a set of fixed values or envelopes (see Fig. 2). Alternatively, in simpler designs, common envelopes can be used for groups of components, or even for the whole sound.

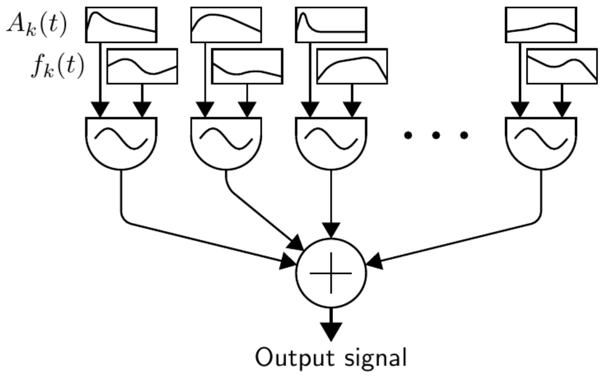

Figure 2. A general diagram of additive synthesis method with individual amplitude and frequency envelopes $A_k(t)$, $f_k(t)$ for each sine component, after [14].

Additive synthesis is one of several methods allowing a straightforward resynthesis, i.e. synthesis using parameters obtained from analysis of existing signals, such as recordings of musical instruments. Moreover, the method does not assume any particular frequency ratios between components, therefore it can reproduce sound of acoustic instruments where ratios are close to, but not precisely natural.

A generalised concept for the algorithm implementing the effect would be outlined as follows:

- track parameters (A, f – amplitude, frequency) of all sine components of all active notes,
- sort all spectral components of all simultaneously sounding notes in order of f,
- in sorted list search for pairs of components falling within a given frequency band $\Delta f(f)$,
- indicate a weaker (with regards to A) component in each pair, completely or partially masked by the stronger one, and add it to a list,
- allow a user to simultaneously attenuate all components from the list by a controllable factor.

Alternatively, instead of attenuating components from the list, one could increase amplitude contrast in found pairs of components in order to maintain total signal level, or control a frequency instead of amplitude not to weaken, but to avoid beating and roughness by moving components outside of proximity region.

4. Implementation

The effect has been implemented as a hybrid solution. The additive synthesizer has been programmed in PureData visual audio programming language [16, 17], due to its good compatibility with various personal computer operating systems, and MIDI controllers, as well as good interoperability with other languages. The effect algorithm has been programmed in C language, as a standalone program, due to its efficiency, and again – ability to be compiled and run on any personal computer operating system. Both parts communicate bi-directionally using UDP internet protocol [18] in a client-server model.

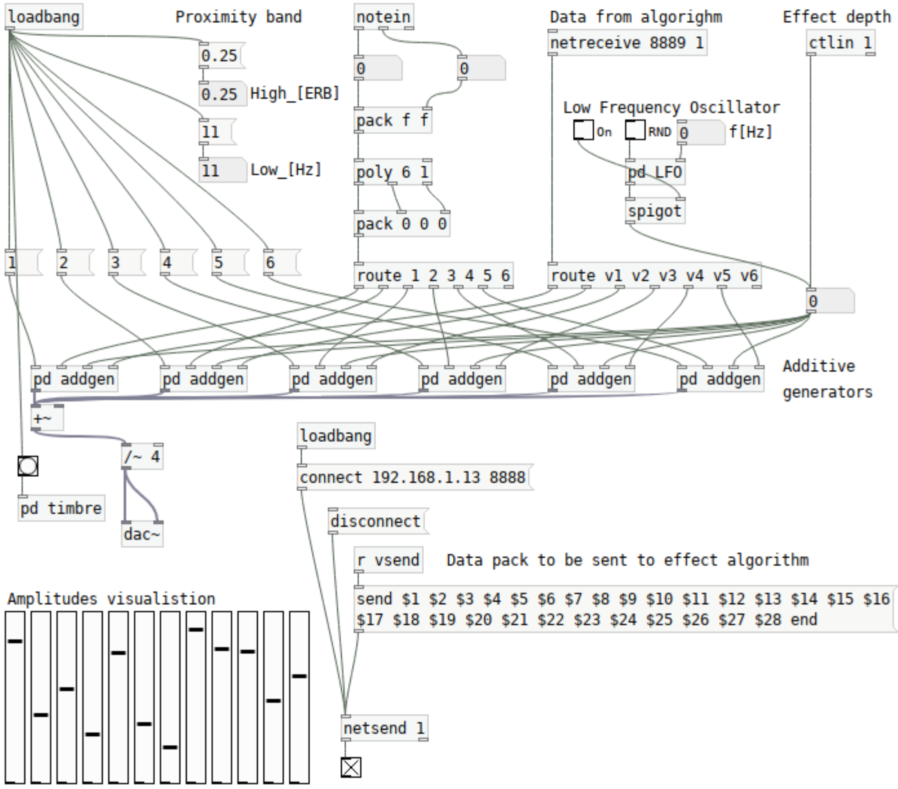

Figure 3. Synthesizer program implemented in PureData – MIDI input, network data transfer routines, and polyphony.

Current implementation assumes that UDP messages are sent on demand. They are triggered by events that change state of the synthesizer with regards to generated signal, i.e., when a voice is starting or ending a note (MIDI messages *note-on* and *note-off*), which adds or removes spectral components to the signal. Message is also sent in case the depth of the effect has changed. All these situations require the algorithm to update information regarding components that can be modified, and send it back to the synthesizer.

Figures 3 and 4 present main parts of the synthesizer. Due to clarity of PureData language and its graphical structure mirroring the actual synthesis algorithm, the synthesizer is presented directly, as implemented, without reverting to abstract diagrams. There is a six-voice polyphony (see Fig. 3), i.e. possibility to independently control six separate pitches, required for playing chords, and observing the effect of changing consonance.

The synthesizer applies a simple variant of additive method (see Fig. 4), with no individual envelopes for sine components. There is only a simple fade-in/fade-out amplitude envelope applied to the total signal, after components mixing. A single pitch representing a note consists of twelve sine components with fixed values for amplitudes and frequencies. Frequencies are initially assumed to be in perfectly natural ratios, although the synthesizer is designed to handle offsets to these values.

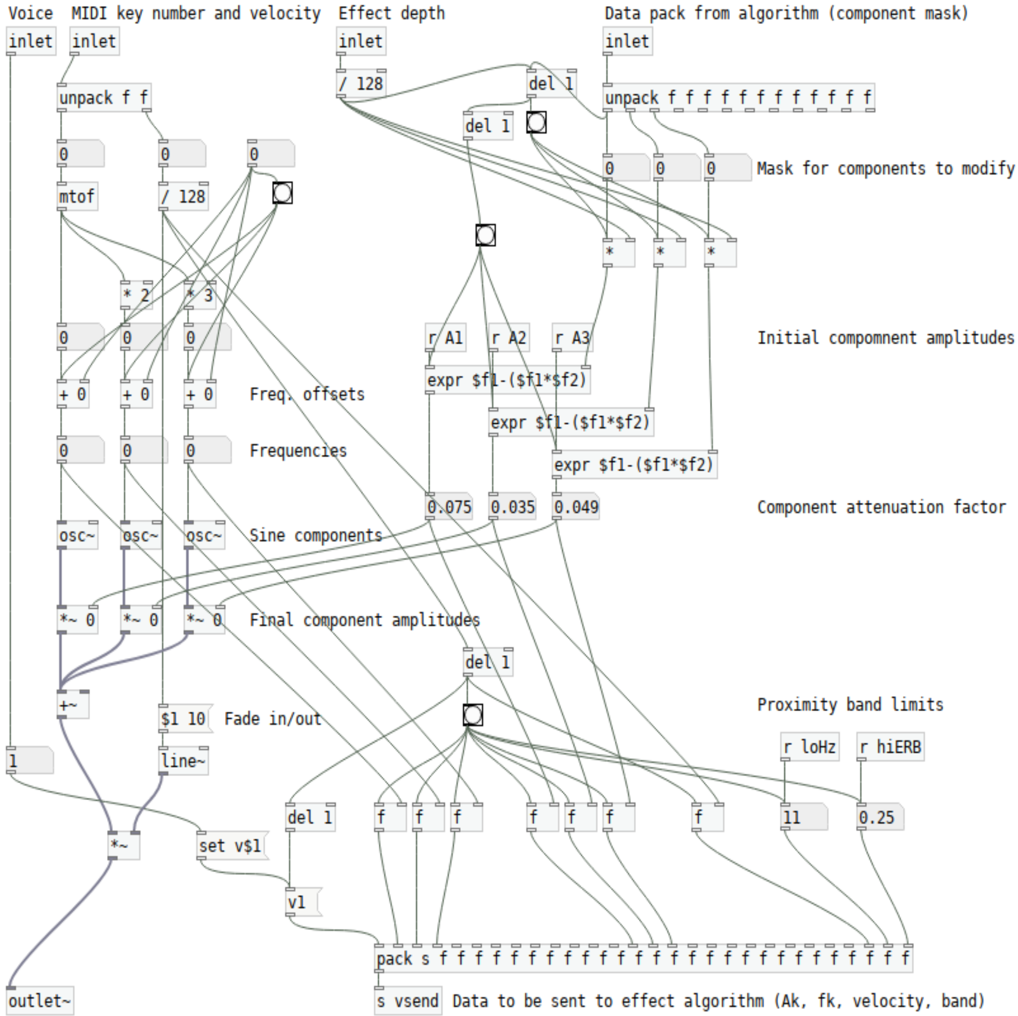

Figure 4. Synthesizer program implemented in PureData – an example of additive signal generator, with higher 9 out of 12 components removed for clarity of presentation.

Amplitudes have initial values set according to timbre setting, but they are further modified using component mask received from the effect algorithm. The mask is a vector of binary values, with value "1" representing a component that will be modified. The actual amplitudes are calculated using initial values and attenuation factor dependent on effect depth and mask. After calculation, these values are sent to the effect algorithm.

Effect depth is controlled either directly by a user through MIDI modulator wheel, or automatically with LFO (Low Frequency Oscillator), with depth modulation frequency defined by a user (see Fig. 3). There is also a random setting that periodically changes depth modulation frequency.

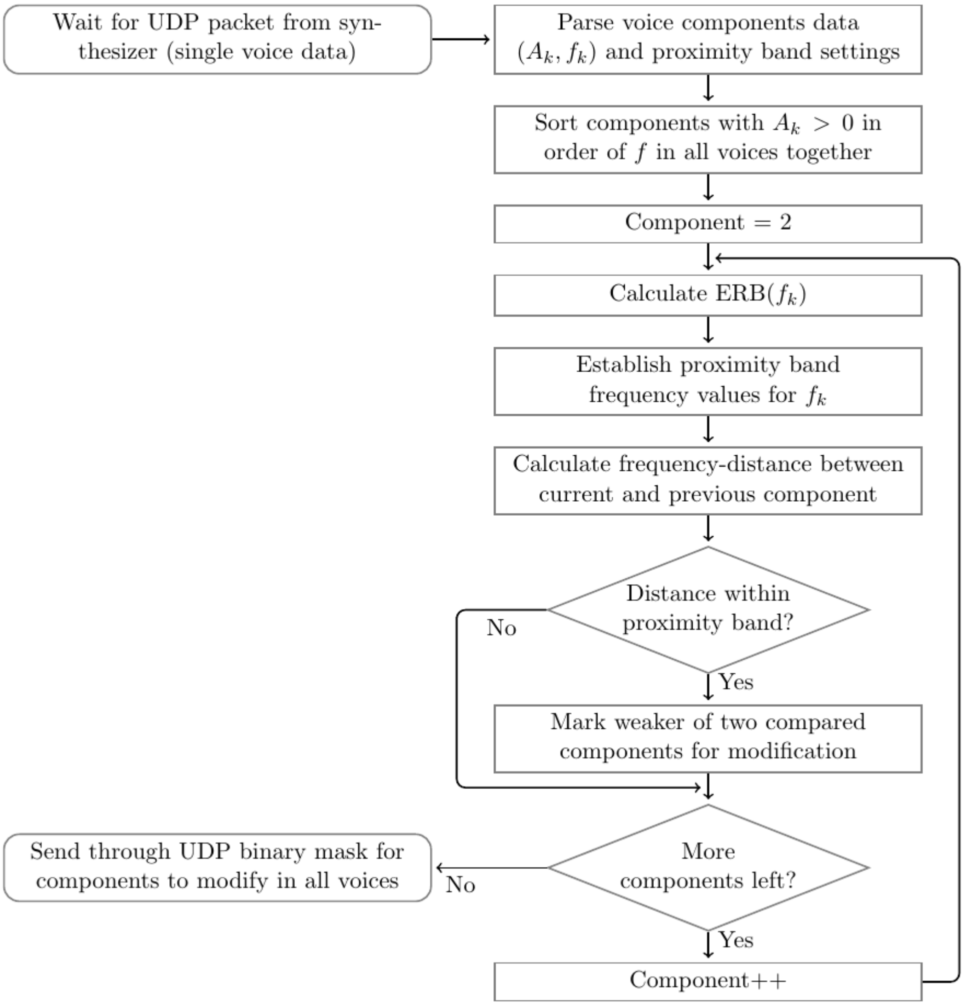

Figure 5. The effect algorithm.

Figure 5 presents the algorithm implemented in C language that supplements the synthesizer and calculates a binary mask of components to be modified. It receives a single voice data from the synthesizer as a list of pairs (A_k, f_k), as well as proximity band settings. The program keeps the data regarding all currently used voices with their component amplitudes and frequencies. On every change a new voice data is added to the pool, or data of a voice that had been disabled is removed. For new arrangement of active components proximity is recalculated to prepare a new mask.

5. Results

The audibility of the effect, i.e. changing amount of dissonance, depends on the played chord. It is much stronger in case of dissonant structures, but is still noticeable in consonant ones if 12-TET tuning system is applied, in which case intervals between pitches deviate slightly from exact natural ratios, and beating occurs.

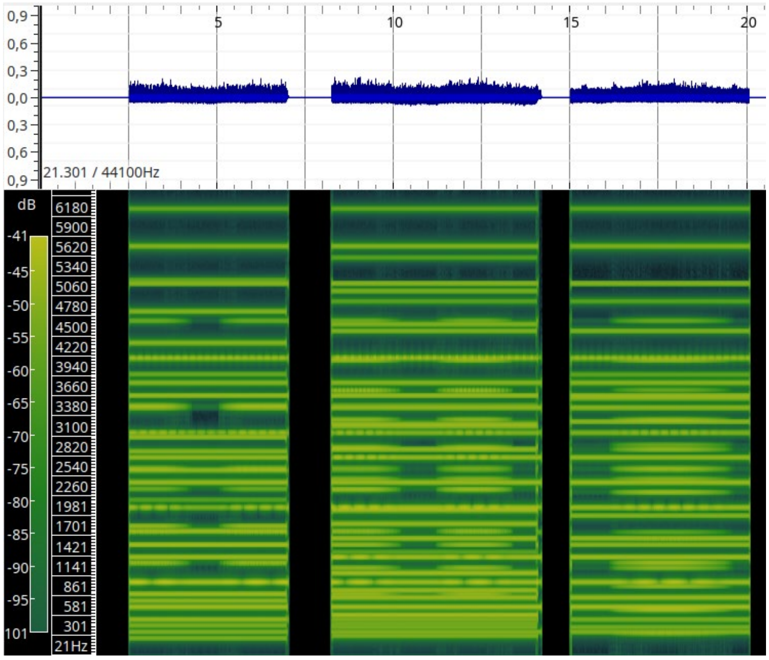

Figure 6. Waveform (X-axis – time [s], Y-axis – amplitude [-]) and spectrogram (X-axis – time [s], Y-axis – frequency [Hz]) of the effect controlled manually with modulator wheel on three subsequent chords; on each chord the wheel has been moved between extreme positions; attenuation of components in dense parts of spectrum is visible.

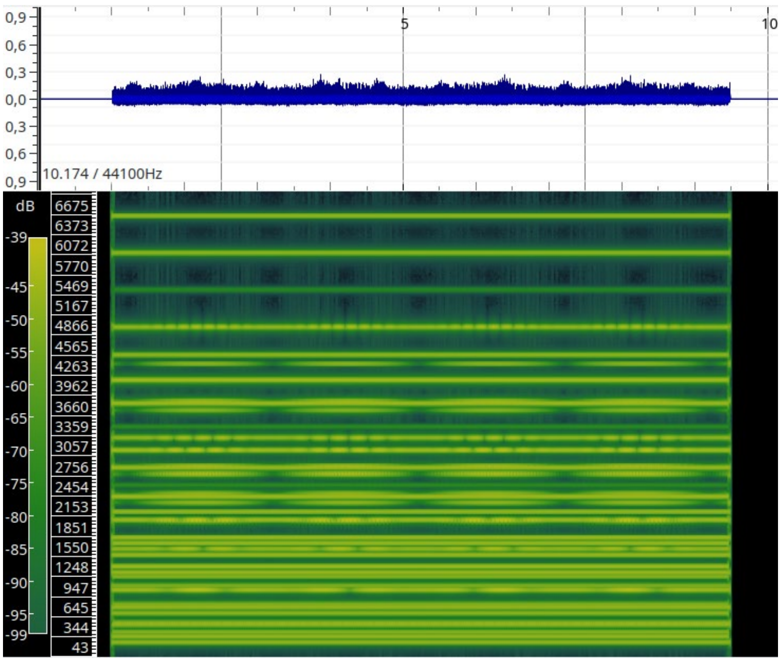

Figure 7. Waveform (X-axis – time [s], Y-axis – amplitude [-]) and spectrogram (X-axis – time [s], Y-axis – frequency [Hz]) of the effect controlled with LFO; visible beating on some components due to close proximity of component pairs is periodically and gradually removed.

Figures 6 and 7 present waveforms and spectrograms of two cases: one in which the effect is controlled manually, through modulator wheel operated by a musician, and the other with effect depth controlled by LFO. In the manual variant the wheel has been turned between extreme positions several times, resulting in areas with some components in dense parts of spectrum completely removed, along with beating of others. The variant with LFO allows to more closely observe attenuation of beating that occurs gradually and periodically on close component pairs.

6. Conclusions

The study presented the concept, assumptions, and implementation of the effect of control over consonance sensation in sound synthesizer based on the additive principle. The effect is accomplished by controlling spectral components of simultaneous pitches consisting a chord, and attenuating by an adjustable factor amplitudes of weaker components in component pairs close enough to evoke sensory dissonance.

In presented implementation it is possible to control both: depth of attenuation of close components, as well as limits of frequency band considered such as to evoke sensory dissonance. Changes of the depth of the effect can be adjusted manually, or through LFO. Each of the effect parameters can be included in synthesizer modulation matrix.

The effect can be classified as one affecting sound timbre. It is distinct enough to be noticeable not only in dissonant chords, but in consonant ones as well, provided a compromise musical tuning system, such as 12-TET, had been used to establish interval frequency ratios, and some beating is present. It is particularly well-suited for application in a pad-class instruments, to introduce new kind of timbre variability on long chords. It is important to note that the algorithm is universal and does not assume any particular tuning system, or any particular frequency ratios between components.

In current form the effect algorithm is executed on demand, whenever there is a change in chord structure, such as addition or removal of pitch. This kind of approach is sufficient if the synthesizer produces constant components, and changes happen only due to user actions. However, if the sound changes over time due to, e.g., application of variant of additive synthesis with separate envelopes of individual components, the algorithm shall be executed periodically, with in-between attenuation values interpolated, so that it follows changes in component structure.

Finally, there are more possibilities regarding action performed to remove sensory dissonance. Instead of attenuating components within a distance of sensory dissonance, one could increase amplitude contrast in pairs of close components in order to maintain total signal level. It is also possible to control frequency of components instead of their amplitude, and therefore not weaken, but avoid beating and roughness by moving components outside of proximity region.

Acknowledgments

The article was published as part of the research subsidy 16.16.130.942 of the Department of Mechanics and Vibroacoustics AGH-UST Krakow.

Additional information

The author declare: no competing financial interests and that all material taken from other sources (including their own published works) is clearly cited and that appropriate permits are obtained.

References

1. P. N. Vassilakis; Perceptual and Physical Properties of Amplitude Fluctuation and their Musical Significance; Doctoral Dissertation, Los Angeles: University of California, Los Angeles, 2001
2. P. N. Vassilakis; Auditory roughness as a means of musical expression; Selected Reports in Ethnomusicology, 2005, 12, 119-144
3. E. Terhardt; On the Perception of Periodic Sound Fluctuations (Roughness); Acustica, 1974, 30 (4), 201-13
4. W. A. Sethares; Tuning, Timbre, Spectrum, Scale (2nd ed.); Springer, London, 2005
5. N. H. Fletcher, T. D. Rossing; The Physics of Musical Instruments, second edition; Springer-Verlag GmbH, 1998
6. B. Moore; An Introduction to the Psychology of Hearing; Brill Academic Pub, 2013
7. E. Ozimek; Dźwięk i jego percepcja. Aspekty fizyczne i psychoakustyczne; Wydawnictwo Naukowe PWN, 2002
8. A. H. Benade; Fundamentals of Musical Acoustics; Dover Publications, 1990
9. G. Wieliczko; Badanie harmoniczności składowych dźwięku wytwarzanego przez strunę gitary elektrycznej; Engineering thesis, AGH University of Science and Technology, Krakow, 2018
10. M. Pluta; Zasady muzyki i notacja muzyczna; Wydawnictwa AGH, Krakow, Poland, 2012
11. M. Toporowski, M. Pilch; Dawne temperacje. Podstawy akustyczne i praktyczne wykorzystanie; Wydawnictwo Akademii Muzycznej im. Karola Szymanowskiego w Katowicach, 2014

12. N. Di Stefano, P. Vuust, E. Brattico; Consonance and dissonance perception. A critical review of the historical sources, multidisciplinary findings, and main hypotheses; Physics of Life Reviews, 2022, 43, 273-304; DOI: 10.1016/j.plrev.2022.10.004
13. T. Pankovski, E. Pankovska; Emergence of the consonance pattern within synaptic weights of a neural network featuring Hebbian neuroplasticity; Biologically Inspired Cognitive Architectures, 2017, 22, 82-94; DOI: 10.1016/j.bica.2017.09.001
14. M. Pluta; Sound synthesis for music reproduction and performance; Wydawnictwa AGH, Krakow, Poland, 2019
15. C. Roads; The Computer Music Tutorial; The MIT Press, 1996
16. M. Puckette; The Theory and Technique of Electronic Music; World Scientific, Singapore, 2007
17. J. Kreidler; Loadbang: Programming Electronic Music in Pure Data; Wolke Verlag, Hofheim, 2009
18. STD 6, RFC 768, IETF; User Datagram Protocol, 1980; DOI: 10.17487/RFC0768