BNAM 2018
Baltic-Nordic Acoustics Meeting
15-18 April 2018
Harpa, Reykjavík, Iceland

Octave stretching phenomenon with complex tones

Jukka Pätynen
Aalto University, Department of Computer Science, FI-00076 AALTO, Finland, jukka.patynen@aalto.fi

Jussi Jaatinen
University of Helsinki, Finland, jussi.o.jaatinen@helsinki.fi

The octave interval is one of the central concepts in musical acoustics. In a physical sense, it indicates a 2:1 relation between two tones, and it is common for most tuning systems. In contrast to the physically based frequency ratio, the frequency ratio that is subjectively perceived as octave has been found to differ from the physical octave. This phenomenon has originally been identified before the 20th century, and more studies have on this topic has been carried out particularly in the 1950s and 1970s. However, most of the research has been conducted from a psychoacoustical perspective with pure tones, whereas complex tones would represent natural instrument sounds more accurately. This paper presents results from a listening experiment where the subjects adjusted pairs of complex tones to match their perception on the subjective octave over a wide range of frequencies. The analysis compares the current outcome to the well-known octave stretching effect found in the conventional tuning of pianos.

1 Introduction

The octave (frequency ratio of 2:1 or 1200 cents) is a special interval in both musical and scientific sense. In music, it divides the musical scale of tones to the pitch chroma, which is the framework for Western tonal music system [1]. As a scientific tool, it has particular benefits: the octave is a perfect consonant interval and represents identical frequency ratio independent of different tuning systems, such as equal temperament, Pythagorean, or Just Intonation. The evaluation of the octave interval is comparably easy for human listener due to chromatic pitch class equivalence. Several psychoacoustic studies [2-13] have reported relatively small, yet significant deviation from the mathematically correct (1200 cents) physical octave (PO). As the general trend, this stretched subjective octave (SO) is slightly larger than the physical octave. Earlier works on this topic have applied either sinusoidal or synthetic complex tones as stimuli. This paper presents initial observations from controlled experiments that explore the octave stretching with sampled orchestra instrument stimuli.

A practical consequence of the SO is that the whole musical scale stretches to accommodate the enlargement from PO. That is, the size of semitones should gradually increase as one progresses farther away from the reference tone. Thus, the bigger an interval between two tones, the bigger is the deviation of the interval from the mathematical tone scale (or a reference scale used in traditional tuning machines). This also means that the stretched musical scale is not perfectly cyclic like the mathematical scale. The most common example of stretched musical tone scale is present in tuning of pianos [14] due to the inharmonicity of the overtones [15]. The magnitude of this phenomenon is illustrated in Fig. 1. However, inharmonicity explains only stretch in low register in a grand piano. High register strings of grand pianos are claimed not to have significant inharmonicity, but the scale stretches in a manner similar to the bass register [16]. Sundberg argues that scale stretching and inharmonicity in pianos is unwittingly reminiscent [17].

Stretched musical scale is a crucial part of music perception and musical experience. However, its existence is mostly obscure due to lack of knowledge. For a musically aware listener able to distinguish pitch with average precision, music played with unstretched (i.e. mathematically tuned) scale may sound out of tune. An experienced musician may use stretched tuning unconsciously [18-20]. Besides soloists, this effect applies naturally for ensembles as well. The second author (JJ) has an extensive career of over 25 years of professional oboe playing in Helsinki Philharmonic Orchestra. According to long-term observations, in a symphony orchestra it is mandatory to play high register notes higher than in mathematical tone scale. Same phenomenon can be seen in low register with downward tuning adjustment in order to avoid sounding out-of-tune. The higher a note, the more stretching upwards is needed. A tuning reference, A4 or 440Hz (442Hz in our study) divides the scale to upper and lower half. In practice, all notes below A4 should be lower than in mathematical scale and, vice versa, all notes above A4 should be higher. This principle has also been reported by Terhardt [21]. The stretched scale represents the lowest layer in intonation system. Above that concept is harmonic and melodic intonation layer (similar to temperaments in piano). The stretched scale layer is assumed mainly constant, whereas the intonation layer is highly musical context dependent.

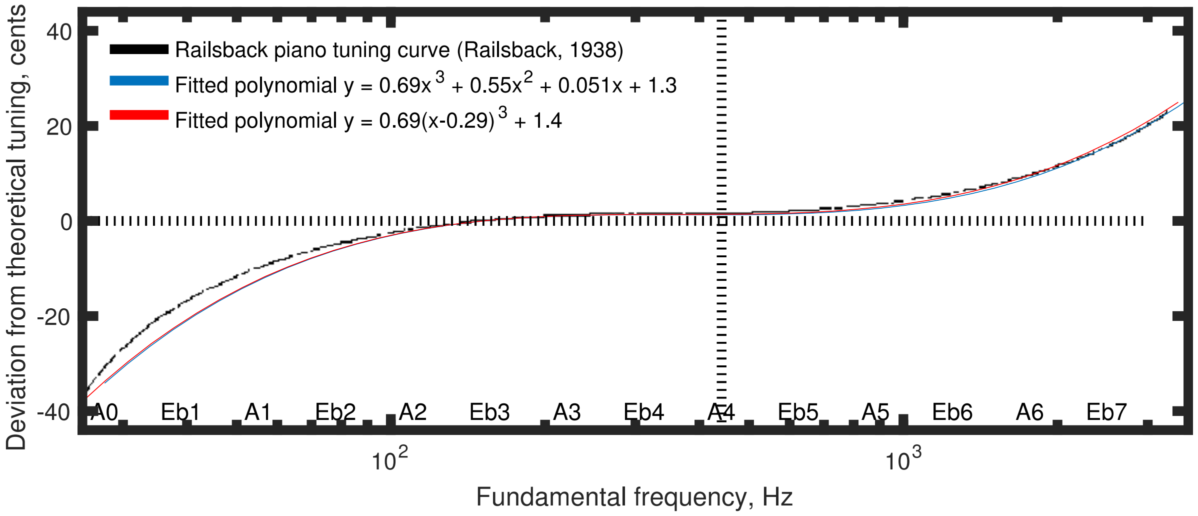

Figure 1: Railsback piano tuning curve and two cubic polynomial approximations. Railsback data has been traced with image processing from published facsimile. Polynomial variable *y* denotes the deviation from theoretical tuning in cents, whereas *x* is the pitch distance from A4 in octaves.

2 Methods

The subjective octave enlargement was explored with a listening test. The subjects listened to pairs of complex tones alternating every second. The task for the subjects was to adjust the higher tone according to their perceived SO. The tone pairs ranged all possible octave multiples from single octaves up to six-octave interval as enabled by different orchestra instrument types, while the analysis focuses on the single octave intervals.

2.1 Listening experiment

The listening test was conducted with dichotic listening in sound-isolated booths. The tuning adjustment was implemented as keyboard presses which switched between 41 pre-calculated increments over the range of [-72 –+72] cents with respect to the mathematically perfect octaves. In order to provide a wider adjustment range without increasing the number of tunings, the increments were slightly larger outside the range of [-36 +36] cents, where the resolution was 3 cents. Between 36 and 52 cents the increment was 4 cents, and 5 cents beyond that. This non-linear adjustment resolution was not disclosed to the subjects. The test procedure including the user interface and audio playback was programmed in Max environment (Cycling'74, CA, USA). The test routine prompted for a short break to be taken every 60 trials. A designated duration for the entire experiment, including the audiometry and breaks, was approximately 3 hours.

2.2 Stimuli

The stimuli represented five groups of typical orchestra instruments: flutes, single reed woodwinds (i.e. clarinets), double reed woodwinds (i.e. oboe and bassoons), brass, and strings. The presented tones spanned the typical compass of each instrument group, including the low- and high-register versions of different instruments. The spectral content of the instrument sound varies with playing dynamics. This effect was taken into consideration by including three nominal dynamic levels (*pp*, *mf*, and *ff*) from each instrument.

The stimuli for the listening experiments were built upon a steady-state wavetable synthesis. The signals were mostly gathered from a database compiled for professional music production (Vienna Symphonic Library GmbH, Austria). The database contains sampled voices from common orchestra instruments. The samples are recorded professionally in an acoustically controlled studio (manufacturer reports a 0.8 s reverberation time) over a wide range of pitches. Additional samples for the experiment were captured by the authors with professional musicians. The generation of the signals consisted of the following procedure: First, the waveform of a particular tone was upsampled to the sample rate of 384 kHz, after which a single wave period was isolated with Wavelab software (Steinberg GmbH, Germany). Second, the waveforms were imported to Matlab environment for accurate regulation of the pitch. The fundamental frequency was adjusted to the nominal equal tempered pitch *re.* A4=442 Hz by replicating the single wave period to 2-second length and iteratively resampling the waveform. The entire process resulted in an accuracy of <0.003 cents. Similar approach was used to create the detuned versions for the range of adjustment. Signal amplitudes were equalized with C-weighting to the same value as the note A4 with flute in the respective playing dynamics. The prepared signals were resampled to the presentation sample rate of 48 kHz, truncated to a length of 1 s, and onset and offset transients were tapered with 35 ms fade-in and fade-out.

The instrument tones included in the final experiment configuration were A0-A6 and Eb1-Eb6 (brass), A1-A6 and Eb1-Eb7 (clarinet), A0-A6 and Eb1-Eb6 (double reeds), A3-A7 and Eb4-Eb8 (flutes), and A0-A7 and Eb1-Eb8 (strings). All octave combinations within particular instruments, A and Eb tones and three dynamic levels yielded a total number of 552 trials. The average equivalent A-weighted sound pressure level over all presented stimuli was 72 dB.

2.3 Subjects

Due to the inherent difficulty of distinguishing smallest deviations in the sizes of octaves, some musical background is necessary. The experiments employed a total of N=36 both professional as well as amateur musicians (7 females, 29 males, ages 20-64, mean 45.6 years, std 12.3 years). Half of the participants were professional and other half amateur musicians. Both authors were included in the amateur and professional groups, respectively. Professional musicians were from Finnish symphony orchestras of the highest level, or work as teachers or professors in Sibelius Academy. All of the hobbyist musicians had long instrument background and good knowledge in music theory and terminology. Most of them play in amateur symphony orchestras and few of them have been graduated from the second grade music schools. The standard audiometry did reveal some issues which, however, can be regarded typical for the participants' occupation, instrument, or individual background.

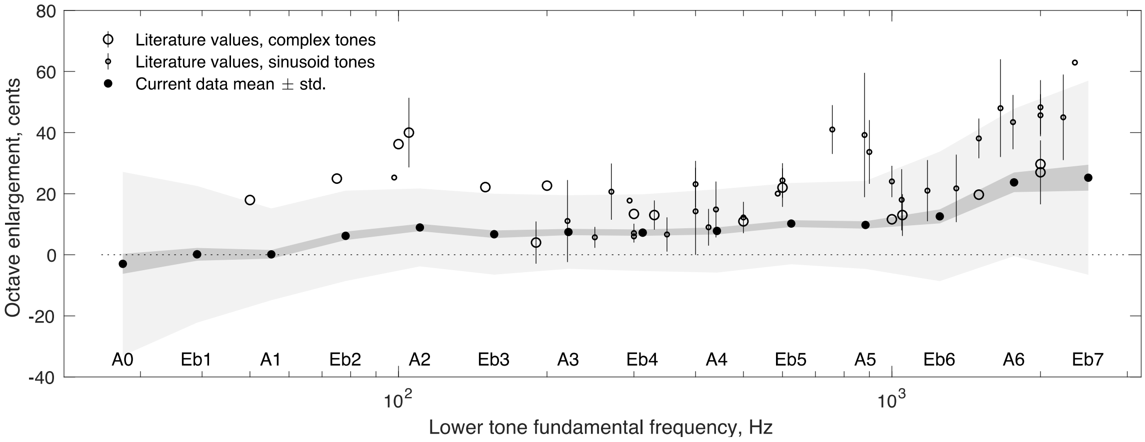

Figure 2: Overall results of the listening experiment accompanied with data extracted from a literature survey [2-8, 11,12, 22]. Dots indicate the mean value in cents for single octaves above each tone marked on the horizontal axis. The light shaded area denotes one standard deviation above and below the mean values, and darker area shows the 95% confidence intervals. Whiskers for the literature values display the reported standard deviation if reported for respective experiments.

3 Results

The experiment produced a total of 19872 data points which includes all single and multi-octave intervals. However, in the scope of the current paper we focus on the single-octave intervals. Octaves with the lower tone exceeding the pitch of Eb6 (fundamental frequency of 2500 Hz) appeared particularly difficult to judge and adjust. For this reason, the inspection was restricted to the octave intervals with the lower tone between A0 – Eb6. This portion of the data contains 5940 tone pairs. The main results of the SO enlargement are shown in Fig. 2. First, we observe that the mean SO matches the PO only at the lowest octaves. Beginning from Eb1 the mean tuning value exceeds the PO by approximately 7 cents. Above A4 the enlargement of SO follows a steadily increasing trend towards the high register. The standard deviation over all independent variables is fairly large. The highest standard deviation occurs at the pitch extrema, as noted above. However, a considerably high number of responses yield narrow confidence intervals for the population mean estimate.

Figure 1 also includes data reported in comparable literature [2-8, 11,12, 22]. Due to absence of actual number values, some data has been estimated from published figures, or converted from various frequency ratio units to cents. A tentative comparison suggests that the current results provide a substantially more stable pattern for SO across frequencies. In addition, current results on the SO are relatively conservative in comparison to data shown in earlier publications. The octave enlargement obtained with complex tones appear to correspond to current data better than experiments conducted with sinusoid stimuli, especially in the middle and high registers. It is worth noting that the current experiment contains prominently more subjects and repetitions for each tone than the preceding publications. The highest number of subjects in studies applying complex tones has been reported by Terhardt (N=6) [5]. In contrast, the data points in Fig. 2 comprise a minimum of 216 values (2 instruments x 3 dynamics x N=36). Some of the cited studies have similar standard deviation over the results as the current data. Here, it is possible that the relatively high variability of subjects can inflate the standard deviation around the mean.

3.1 Tuning curve for orchestra instrument spectra

The conducted experiment yields a series of adjusted octave tunings using the lower tone as the reference. This raw data provides information on stretching of specific octave intervals. On the other hand, the aim is also to estimate a tuning curve based on the tuning values obtained with various orchestra instruments. Estimating the tuning curves over the tested range of tones is based on the assumed additivity property of octave tuning as formulated by Ward [3]. This means that the tuning values for adjacent SO can be combined for multi-octave intervals, i.e. $SO_{A4\text{-}A5}+SO_{A5\text{-}A6}=SO_{A4\text{-}A6}$. Hence, the stretching of single octaves becomes accumulated for the respective multi-octave interval.

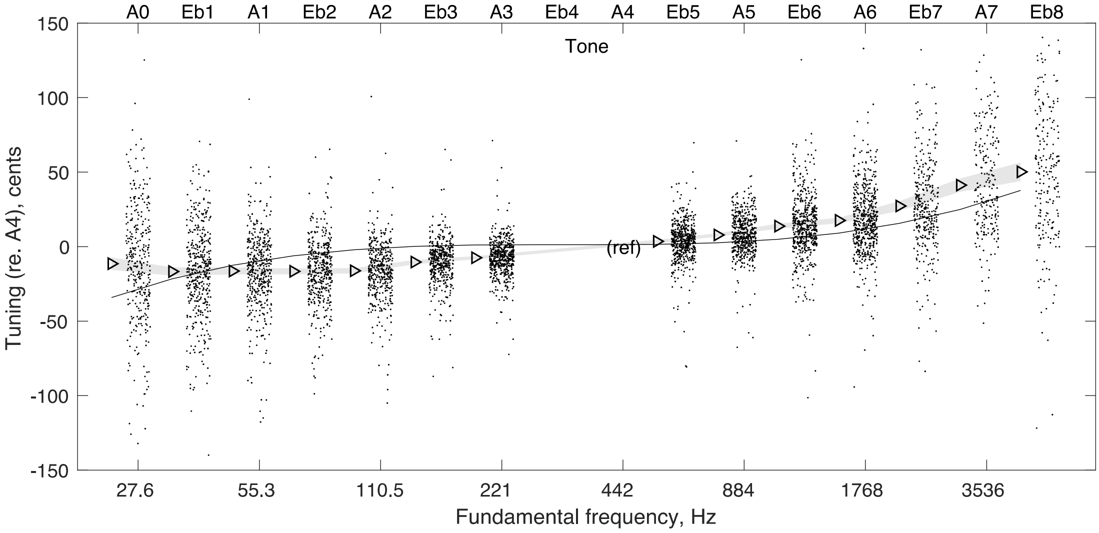

Figure 3: Detailed scatter plot of the results converted to an average tuning curve according to the octave additivity. Individual data points have been treated with horizontal and vertical jitter within the frequency and tuning adjustment resolution for readability. Triangles show the grand mean for each respective fundamental frequency. Shaded area expresses the 95% confidence interval for the sample means. The solid line shows the Railsback piano tuning curve [14] as the fitted cubic polynomial (see Fig. 1) for comparison to the experiment data.

The stretching curve is calculated upon a reference point, i.e. the tuning center. Since A4 is the standard pitch reference, it is a natural choice as this point. Single-octave interval tuning values above and below the center are stacked according to the additivity assumption. As all trials in the experiment were conducted by adjusting the higher tone, the tuning curve values for pitches below the reference tone are inverted to always accumulate the octave stretching away from the tuning center. The algorithm starts at the tuning center and progresses to the low and high octave extrema according to the assumed additivity of tuning values. This procedure is repeated separately for data on each combination of subject, instrument, and dynamics. That is, each condition of independent variables has its own tuning curve.

Since octaves of one tone only give a relatively sparse sampling of the overall compass of pitches, the above procedure is repeated equally for both A and Eb –based octaves, where the tuning center for the latter is Eb4. In order to incorporate the Eb-octaves to the stretching curves of A-octaves, the tuning values obtained for all Eb notes are adjusted by a constant, which is obtained as the interpolation between the tuning values between the adjacent values for A-notes, that is, A3 and A4 (tuning center). This approximation assumes that the octave stretching occurs linearly within the region of interpolation, which appears reasonably valid based on the gathered data (see Fig. 2).

Figure 3 illustrates the tuning curve profile across tested frequencies. In essence, the mean values indicate how much the notes played by a musician should deviate from the exact mathematical tuning in order to agree with the subjectively perceived octaves. It should be noted again that the single points shown in the diagram represent different combinations of instrument timbres, dynamic levels, as well as individual subjects. In the initial survey of gathered data it was seen that different subjects yielded highly varying trends in their tuning curves. This effect largely explains the wider spread of values visible in the low and high register.

Compared to the Railsback piano tuning curve, the mean values suggest a different behaviour in the low register. This observation is expected, as the lowest piano strings exhibit substantial inharmonicity. However, the method used in producing the stimuli in the current experiment ensures a perfect harmonic spectrum. In contrast, middle and high registers of the current results follow the general profile of the Railsback curve. Assuming that the polynomial model for Railsback data, compared in Fig. 1, is fairly accurate approximation of the true Railsback curve, we can observe that the conventional piano tuning does not entirely fit within the confidence interval for stretched tuning now seen suitable for orchestra instruments.

4 Conclusions

This paper presented the overall outcome of a listening experiment studying the octave stretching phenomenon with complex tones that correspond to spectral envelopes of different orchestra instrument groups. In comparison, preceding studies have applied either sinusoidal or synthetic complex tones. The listening test produced an unparalleled amount of data on subjective perception of octave tuning. The data was compared to numerous literature sources reporting comparable experiments in the past. Currently presented experiments yielded a more conservative effect for the subjective octave enlargement than literature sources. Conversely, the results were more consistent over an extended range of frequencies.

The experiment data were converted into an average tuning curve for orchestra instruments in a manner of traditional piano tuning curve. Tentative comparison suggested that the Railsback curve is not accurately applicable for orchestra instruments. In the bass register the stimuli did not have an inharmonic overtone series like found in the lowest piano strings, which offers a natural explanation for the difference between the tuning curves. However, above the middle register the increasing stretching trend was nearly comparable to the Railsback curve. Earlier studies [23, 24] suggest that the piano tones above C4 would have increasing degree of inharmonic behavior, which would then explain the tuning curvature. However, a similar trend was observed here with harmonic stimuli. This discrepancy provides an avenue for investigating the neurological phenomena related to pitch perception [25]. Unfortunately, the original Railsback data appears to be lost, based on the present authors' investigation.

The presented results are encouraging regarding further exploring the octave enlargement effect. The collected data offers a fruitful ground for future research, where the effects of multiple variables could be modeled for better understanding of the subjective octave and the preferred tuning curve for the orchestra instruments.

References

[1] A. Bachem, "*Tone height and tone chroma as two different pitch qualities*," Acta Psychologica 7, 80–88 (1950).

[2] C. Stumpf and M. Meyer, "*Massbestimmungen über die Reinheit consonanter Intervalle.*," Beiträge zur Akustik und Musikwissenschaft 2, 84–167 (1898).

[3] W. Ward, "*Subjective musical pitch*," J. Acoust. Soc. Am. 26(3), 369–380 (1954).

[4] K. Walliser, "*Über die Spreizung von empfundenen Intervallen gegüber matematisch harmonischen Intervallen bei Sinustönen*," Frequenz 23 (1969).

[5] E. Terhardt, "*Die Tonhöhe Harmonischer Klange und das Oktavintervall*," Acustica 24, 126–136 (1971).

[6] J. E. Sundberg and J. Lindqvist, "*Musical octaves and pitch*," J. Acoust. Soc. Am. 54(0001-4966), 922– 929 (1973).

[7] W. Hartmann, "*On the origin of the enlarged melodic octave*," J. Acoust. Soc. Am. 93(6), 3400–3409 (1993).

[8] P. A. Dobbins and L. L. Cuddy, "*Octave discrimination : An experimental confirmation of the "stretched" subjective octave*," J. Acoust. Soc. Am. 72(2), 411–415 (1982).

[9] E. Terhardt and M. Zick, "*Evaluation of the tempered tone scale in normal stretched and contracted intonation*," Acustica 32, 268–274 (1975).

[10] G. Van Den Brink, "*Octave and fifth settings for pure tones and residue sounds*," in Psychophysics and Physiology of Hearing, edited by E. Evans and J. Wilson (Academic press, London, 1977), pp. 373–379.

[11] L. Demany and C. Semal, "*Harmonic and melodic octave templates*," J. Acoust. Soc. Am. 88(5), 2126–2135 (1990).

[12] B. S. Rosner, "*Stretching and compression in the perception of musical intervals*," Music Perception 17(1), 101–114 (1999).

[13] C. von Maltzew, "*Das Erkennen sukzessiv gegebener musikalischer Intervalle in den äusseren Tonregionen*," Zeitschrift für Psychologie 64, 161–257 (1913).

[14] O. Railsback, "*Scale temperament as applied to piano tuning*," J. Acoust. Soc. Am. 9(3), 274–274 (1938).

[15] O. Schuck and R. Young, "*Observations on the Vibrations of Piano Strings*," J. Acoust. Soc. Am. 15(1) (1943).

[16] O. Jorgensen, *Tuning: Containing the Perfection of Eighteenth-Century Temperament, the Lost Art of Nineteenth-Century Temperament and the Science of Equal Temperament* (Michigan state university press, East Lansing, 1991), pp. 739–745.

[17] J. Sundberg, *The science of musical sounds (Cognition and perception)* (Academic Press, San Diego, 1991), p. 274.

[18] J. F. Corso, "*Scale Position and Performed Melodic Octaves*," Journal of Psychology: Interdisciplinary and Applied 37(2), 297– 305 (1954).

[19] F. Fransson, J. Sundberg, and P. Tjernlund, "*Statistical computer measurements of the tone-scale in played music*," STL- QPSR 2-3, 41–45 (1970).

[20] A. Rakowski, "*Intonation Variants of Musical Intervals in Isolation and in Musical Contexts*," Psychology of Music 18, 60–72 (1990).

[21] E. Terhardt, "*Stretch of the musical tone scale*," (2000), http://www.mmk.ei.tum.de/fileadmin/w00bqn/www/Personen/Terhardt/ter/top/scalestretch.html.

[22] K. Ohgushi, "*The origin of tonality and a possible explanation of the octave enlargement phenomenon*," J. Acoust. Soc. Am. 73(5), 1694–1700 (1983).

[23] N. Giordano, "*Explaining the Railsback stretch in terms of the inharmonicity of piano tones and sensory dissonance*," J. Acoust. Soc. Am. 138(4), 2369-2366 (2015).

[24] R. R. Rasch and V. Heetvelt, "*String Inharmonicity and Piano Tuning*," Music Perception: An Interdisciplinary Journal 3(2), 171-189 (1985).

[25] A. J. Oxenham, "*Pitch Perception*," The Journal of Neuroscience 32(39), 13335–13338 (2012).