Perceptibility of Inharmonicity in the Acoustic Guitar

Hanna Järveläinen, Matti Karjalainen
Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland

Summary
Audibility of inharmonicity in acoustic guitar tones was studied through formal listening experiments separately for steel- and nylon-stringed guitars. Test tones were resynthesized from real recordings through high-resolution parametric modeling of acoustic guitar sounds, which allows accurate control of the inharmonicity of the partials. Inharmonicity was detected fairly easily by the subjects for the lowest strings especially for the steel-stringed guitar. For the nylon-stringed guitar, inharmonicity was also detected as long as the plucking transient was left out of the sound, but detection was much harder for the full sound with transient. Detection thresholds were obtained for two pitches at the low range of the nylon-stringed guitar. Mean thresholds were close to, though above, typical amounts of inharmonicity in the guitar. Implications to digital sound synthesis are discussed.
PACS no. 43.75.Gh, 43.66.Lj

1. Introduction

Inharmonicity of a nearly harmonic sound may be perceived in different ways, depending on the type and the amount of inharmonicity and the type of sound. A mistuned partial in the higher harmonics of an otherwise harmonic complex is detected by means of beats, but if the mistuned component is one of the lower harmonics, it is rather heard standing out from the complex [1].

In addition to basic research in music perception, inharmonicity in string instruments is an important question to sound synthesis. It is known that in the low range of the piano, inharmonicity is definitely perceived and should be synthesized. In the guitar, however, the importance of inharmonicity to the quality of the sound has been unknown. According to current knowledge, small amounts of inharmonicity are perceived based on a timbral effect, not on segregating partials or a pitch change. Whether this effect is present in the acoustic guitar, is the main question in the current study.

In real strings, stiffness contributes to the restoring force along with tension, which results in slightly inharmonic partials. It should be remembered that friction also makes a contribution to the total force, although the effect is small compared to that of stiffness in typical guitar strings [2]. Thus stiffness is considered here the main source of inharmonicity. The partial frequencies f_n are given by [3]:

$$f_n = n f_0 \sqrt{1 + Bn^2}, \qquad (1)$$
$$B = \pi^3 Q d^4 / 64 l^2 T, \qquad (2)$$

where n is partial number and B is inharmonicity coefficient, composed of Young's modulus Q, diameter d, length l and tension T of the string. It is seen from (1) that higher partials are stretched relatively more than lower ones. Thus even a small B value can produce a prominent amount of mistuning in the higher partials of a low-pitched tone. Since equation (2) includes string length and tension, B factor formulated this way increases with the fret being played (see Figure 1).

Inharmonicity coefficients were analyzed in this study from recorded guitar tones for the free strings and the first 17 frets. The inharmonicity coefficients increased with fundamental frequency and fret number. In an earlier study [4], perception thresholds of inharmonicity were investigated by listening tests for generic synthetic string instrument -like tones as a function of fundamental frequency. This threshold is presented in Figure 1 together with the measured inharmonicity coefficients for real guitar tones. In all tones inharmonicity clearly exceeds the perception threshold. The coefficients varied for string one between 2.9 ... 4.6 times B at the threshold, for string two 9.7 ... 17.1, for string 3 34.5 ... 58.3, for string 4 7.8 ... 15.4, for string 5 13.1 ... 22.1, and for string 6 34.4 ... 77.6 times B at the respective perception threshold. Thus inharmonicity should be most clearly audible for strings 3 and 6 and probably also for other strings. While the above mentioned study was made using generic string sounds with identical decay times for each harmonic, the present study carefully focuses on the detailed sounds of a high-quality classical acoustic guitar. Therefore, it is expected that the previous results may not fully predict the current case. The aim of this study is to find characteristic values for the perceivability of inharmonicity in typical nylon and metal string guitars. In specific cases the results vary according to details of strings used, body construction, and plucking style.

Received 1 February 2006, revised 16 June 2006,
accepted 23 June 2006.

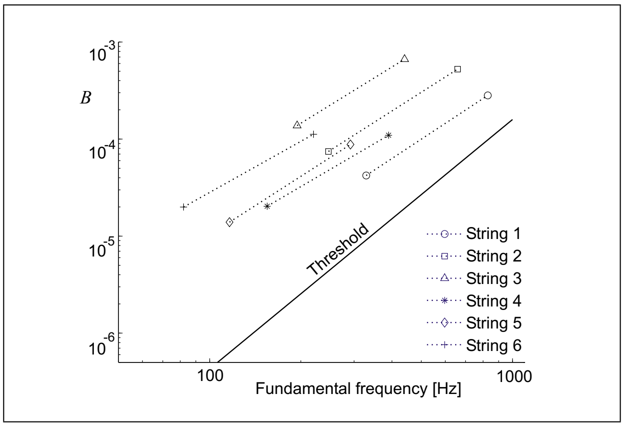

Figure 1. Perception thresholds from [3] (solid line) versus inharmonicity coefficients B for frets 0-17 of each string as a function of fundamental frequency.

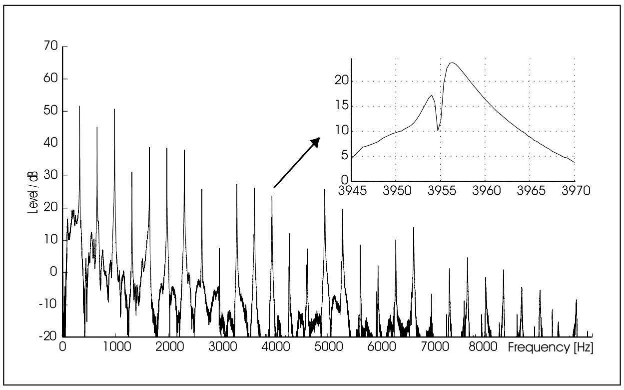

Figure 2. Spectrum of plucked string sound of a classical acoustic guitar, open string E_4. Spectrum zoomed to a single partial is shown in the subplot.

2. Manipulation of harmonicity

In order to modify the inharmonicity of a guitar sound the frequencies of the partials have to be shifted without changing any other perceptual features of the signal. This is a demanding task and should be based on careful parametric modeling of recorded sounds. One method that could be used is sinusoidal plus transient modeling [5]. Another choice is applying frequency warping techniques to the guitar pluck sound [6]. A third possibility is to use parametric modeling, such as all-pole or pole-zero modeling. Based on practical experiments, we achieved best results by using high-resolution pole-zero plus residual modeling, probably because it has a closer correspondence to the physical basis of the instrument sounds than the other two methods. Each partial is modeled as a set of modes that can be scaled in frequency for resynthesis, and the original residual of the sound attack is added back to the modal components in order to yield a new version of the sound with modified partial frequencies.

2.1. High-Resolution Mode Analysis

If the pluck response of the guitar were an AR (autoregressive or all-pole) process, it could be analyzed by straightforward AR modeling techniques, such as the linear prediction (LP) [7], resulting in an all-pole filter. This does not work in practice, however, because pole positions are very close to the unit circle and make the analysis numerically too critical. Also the allocation of the number of poles for each partial cannot be easily controlled. It is important to notice that two polarizations of string vibration with slightly different frequencies make each partial to be a sum of two decaying sinusoids, exhibiting beating or two-stage decay of the envelope [3] or two separate spectral peaks as illustrated in the subplot of Figure 2. There can also be additional phantom partials as described in [8]. Thus at least two pole pairs are needed to represent one partial.

Since the string's response to excitation is not a minimum-phase signal, the AR modeling approach is even in theory a wrong tool. ARMA (autoregressive moving average) models (also called pole-zero modeling) are better capable to match with such responses, but due to iterative solutions, model orders higher than 20-100 may not converge to a stable and useful result.

2.2. FZ-ARMA analysis

In order to avoid the problems in resolution and computational precision discussed above we have developed a subband technique called FZ-ARMA (frequency-zooming ARMA) analysis [9]. Instead of a single model, global over the entire frequency range, the signal is modeled in subbands, i.e., by zooming to a small enough band at a time, thus allowing a filter order low enough and individually selectable to each subband. This helps in resolving modal resonances that are very sharp or close to each other in frequency. The FZ-ARMA analysis consists of the following steps:

(i) Select a frequency range of interest, for example a few Hz wide frequency region around the spectral peak of a partial.

(ii) Modulate the target signal (shift in frequency by multiplying with a complex exponential) to place the center of the frequency band, defined in (i), at the origin of the frequency axis by mapping

$$h_{\mathrm{m}}(n) = \mathrm{e}^{-\mathrm{j}\Omega_{\mathrm{m}} n} h(n), \qquad (3)$$

where $h(n)$ is the original sampled signal, $h_{\mathrm{m}}(n)$ the down-modulated one, n is the sample index, and Ω_{m} the (normalized) modulation frequency. This rotates the poles i of transfer function by $\Omega_{i,\mathrm{rot}} = \Omega_i - \Omega_{\mathrm{m}}$.

(iii) Apply lowpass filtering to the complex-valued modulated signal in order to attenuate its spectral content outside the zoomed band of interest.

(iv) Decimate (down-sample) the lowpass filtered signal according to its new bandwidth. This zooms system poles z_i by

$$z_{i,\mathrm{zoom}} = z_i^{K_{\mathrm{zoom}}} = |z_i|^{K_{\mathrm{zoom}}} \mathrm{e}^{\mathrm{j}(\Omega_i - \Omega_{\mathrm{m}}) K_{\mathrm{zoom}}}, \qquad (4)$$

where K_{zoom} is the zooming factor, and $z_{i,\mathrm{zoom}}$ are the mapped poles in the zoomed frequency domain.

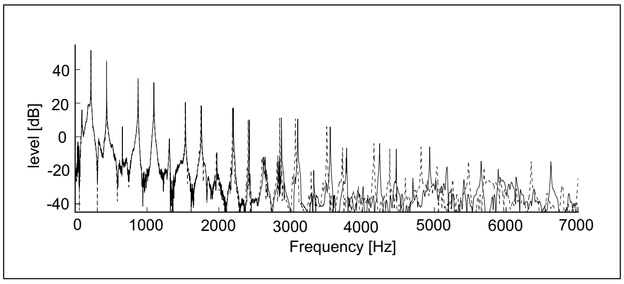

Figure 3. Example of an inharmonic plucked string spectrum (solid line) and a harmonic resynthesized version of it (dashed line).

(v) Estimate an ARMA (pole-zero) model for the obtained decimated signal in the zoomed frequency domain. For this we have applied the iterative Steiglitz-McBride algorithm (function `stmcb.m` in Matlab).

(vi) Map the obtained poles back to the original frequency domain by operations inverse to the above-presented ones. Zeros cannot be utilized as easily, thus we don't use them in this study for the final modeling. There may also be poles that correspond to the truncated frequency band edges, thus needing to be excluded. Therefore only relevant poles are directly useful parameters.

When applying pole-zero modeling, the selection of the number of poles has to be made appropriately according to the characteristics of the problem. The number of poles should correspond to the order of the resonator to be modeled. Only the poles with positive imaginary component are used in FZ-ARMA analysis, representing the complex-conjugate pole pairs needed for real-valued filter model. As mentioned above, a partial ('harmonic') of string vibration is composed of two polarizations, thus the partial may exhibit more than one peak in the frequency domain and beating or two-stage decay in the time-domain envelope.

A proper number of zeros in FZ-ARMA modeling is also needed to make it able to fit the phases of the decaying sinusoids as well as modeling of the initial transient. Often this number is not very critical, and it can be somewhat higher than the number of poles.

The zooming factor K_{zoom} can be selected so that the analysis bandwidth contains most of the energy of the resonances to be modeled, keeping the order (number of poles) manageable.

2.3. Modification of inharmonicity

Based on the FZ-ARMA analysis we can now modify the frequencies of the partials in the following way.

(i) Analyze each partial of the given recording by FZ-ARMA analysis in the zoomed frequency domain. Tune the fundamental frequency and the inharmonicity factor so that the main spectral peak of the partial remains in the middle of the zoomed band.

(ii) Map the partial back to the original frequency domain but by moving the modal frequency to a new desired frequency. Repeat this for each partial and collect the resynthesized partial components.

(iii) Resynthesize the partials also at the original frequencies by the method in (i)-(ii) and subtract this from the original signal to get a residual signal containing the pluck and the body response. Clip the most prominent part of the attack as the residual for resynthesis.

(iv) Resynthesize the new modified version from the frequency-shifted partials and the residual for a new inharmonicity (or a fully harmonic version).

Figure 3 depicts an example of plucked string spectra for the original inharmonic recorded sound and the resynthesized harmonic version of it.

3. Listening tests

Perception of inharmonicity was studied in formal listening tests separately for the nylon- and steel-stringed instruments. The first experiment explored the general perceptibility of inharmonicity in acoustic guitar tones versus similar but harmonic tones. In an additional experiment for the nylon-stringed guitar, detection thresholds of inharmonicity were obtained for two pitches.

The test subjects had previous experience from psychoacoustic experiments, and none of them reported a hearing defect. All played some string instrument. For training, the subjects were presented all the test material twice before experiment 1. The tests were performed in a silent room using headphones.

Test stimuli were synthesized using FZ-ARMA models of a high-quality classical guitar with nylon strings and a steel-stringed guitar. Original recordings of the guitar tones were made in an anechoic chamber using a AKG C 480 B microphone and a PC laptop with Digigram VX Pocket soundcard. The microphone was placed at one meter from the guitar in front of the tone hole. The nylon-stringed guitar had a set of D'Addario Pro-Arte Light Tension strings (diameters .0275, .0317, .0397, .028, .033, .042 inches) that were excited by the player's finger. The steel-stringed guitar had Galli Lucky Star strings (.011, .015, .024, .032, .042, .052), and it was excited by a thick (1.3 mm) plectrum. The plucking point was at the middle of the tone hole, approximately 12 cm from the bridge.

Inharmonicity coefficients B were varied such that the resulting inharmonicity either matched to, or was below or above, the original one. A harmonic reference tone was generated for each inharmonic stimulus with $B = 0$.

3.1. Nylon-stringed guitar

3.1.1. General perceptibility test

Eleven subjects participated the first experiment, including four professionally trained guitarists and author HJ. The ability to detect the amount of inharmonicity typical of the acoustic guitar was explored at six pitches: frets 1 and 12 of strings 1, 3, and 6. The inharmonicities of the test tones were similar to the original recorded tones (see Table I). According to initial listening, detection was extremely hard if the residual (the pluck and body response) was included. The plucking transient was therefore left out in the first experiment.

Table I. Nylon-stringed test tones and their inharmonicity coefficients.

String, fret	f_0 [Hz]	B
String 1, fret 1	349	0.000047
String 1, fret 12	659	0.00017
String 3, fret 1	207	0.00015
String 3, fret 12	392	0.00052
String 6, fret 1	87	0.0000205
String 6, fret 12	165	0.00007

The test followed the 2AFC (two-alternative forced choice) paradigm. In each trial, subjects were presented two pairs of tones, one consisting of two harmonic tones and the other one of the harmonic tone and the respective inharmonic tone. The task was to identify, which pair contained the inharmonic tone. The inharmonic tone was randomly placed in either sound pair. Two pairs were used instead of just two sounds in the 2AFC to prevent mix-ups in identifying inharmonic and harmonic tones. For each pitch, the trial was repeated 16 times. The percentage of correct responses was recorded from the last 12 trials.

The results are presented in Figure 4. Detection is clearly above chance level (50% correct) for strings 3 and 6: the means range between 82% correct for the 1st fret of string 6 and 94% correct for the 12th fret of string 3. For string 1, the results are generally under 75% correct, which suggests that the subjects have been mainly guessing.

The experiment was re-run for the 12th fret of strings 3 and 6, for which inharmonicity was detected most clearly, this time including the residuals for maximal correspondence to the original ones. The performance dropped from nearly 100% correct to nearly chance level (see Figure 5). Thus it seems that inharmonicity is mainly perceived during the very beginning of the tone. An explanation may be that in string instrument tones the higher partials, which are shifted most from their harmonic places, decay faster than the lower ones, thus making the initial transient mask them easily.

3.1.2. Perception threshold for individual tones

Since detection of inharmonicity remained uncertain even at the lowest pitch range, detection thresholds were studied for the 12th frets of strings 3 and 6. The 2AFC paradigm was now used with the adaptive 1 up, 2 down staircase procedure, giving the 70.7% correct point on the psychometric function [10]. After two correct responses for the inharmonicity factor, inharmonicity was reduced in the next trial, while after only one correct or an incorrect response, it was increased and the algorithm was reversed. The procedure stopped after 12 reversals, and the threshold was computed from the mean of the last eight reversals. B was varied in linear steps of about 1/7 of the original measured B. Seven subjects participated in the threshold experiment.

Figure 6 presents the results in relation to the B coefficients measured for all frets of strings 3 and 6, and the thresholds obtained in [4], as a function of fundamental frequency. While the current thresholds are far above those obtained previously, they are relatively close to the measured inharmonicity coefficients. The mean threshold for fret 12 of string 3 was $B = 0.00097$ and for fret 12 of string 6 $B = 0.00015$. These are 1.87 and 1.65 times the respec-

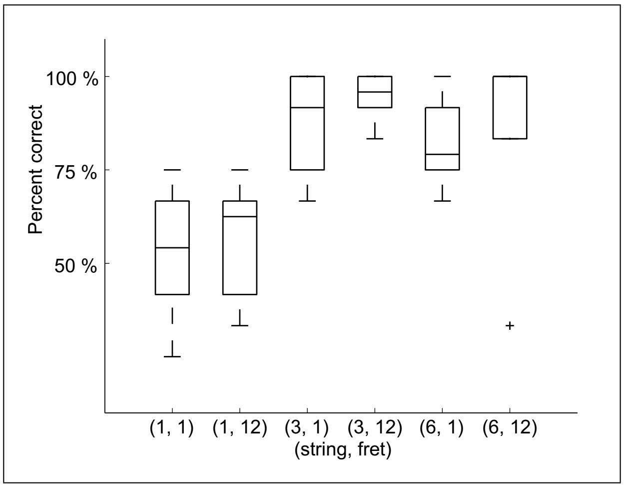

Figure 4. Boxplot of the results of the general inharmonicity perception test for the nylon-stringed guitar.

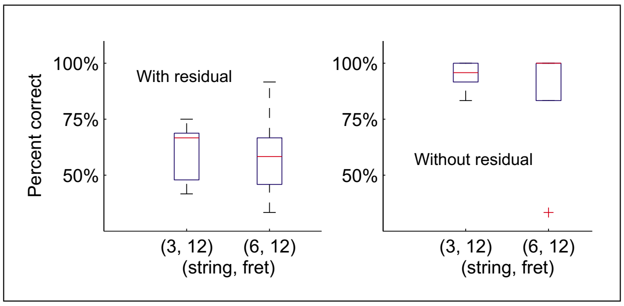

Figure 5. Results of the 12th fret of strings 3 and 6 with and without the residual for the nylon-stringed guitar (left and right panel, respectively).

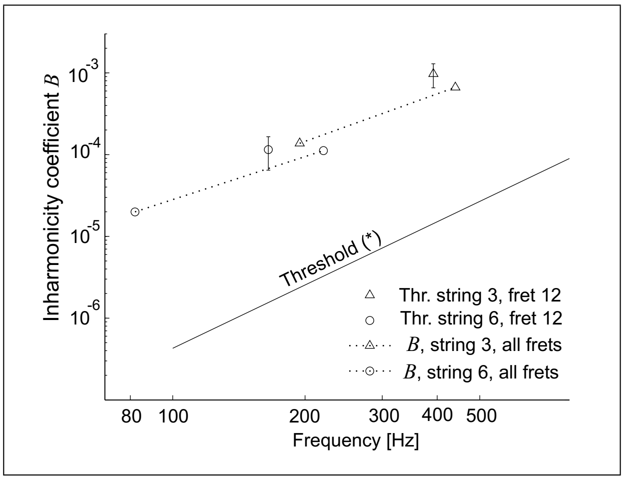

Figure 6. Currently obtained thresholds for the nylon-stringed guitar (with error bars of one standard deviation up and down), respective B values for guitar strings, and the threshold measured in [3] (*).

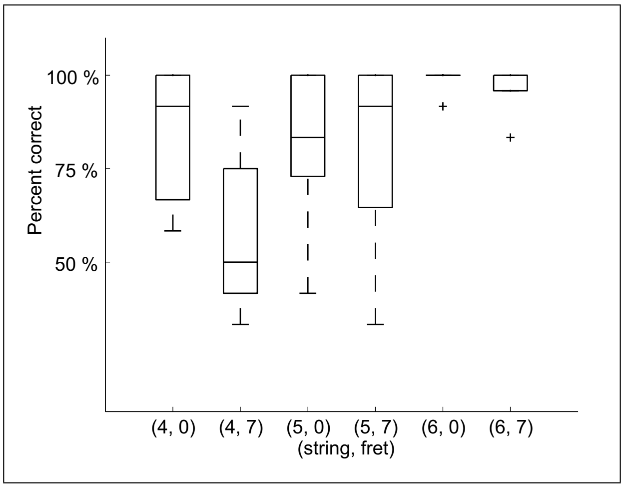

Figure 7. Boxplot of the results of the general inharmonicity perception test for the steel-stringed guitar.

Table II. Steel-stringed test tones and their inharmonicity coefficients.

String, fret	f_0 [Hz]	*B*
String 4, fret 0	146	0.000050
String 4, fret 7	219	0.000105
String 5, fret 0	110	0.000066
String 5, fret 7	164	0.00015
String 6, fret 0	82	0.000108
String 6, fret 7	123	0.000242

tive *B* coefficients measured from guitar tones. However, since two of the subjects performed clearly better than the rest five, it might be useful to re-run the experiment with expert listeners to see if their thresholds are significantly lower.

3.2. Steel-stringed guitar

3.2.1. General perceptibility test

Since inharmonicity seemed to have little importance to the quality of the higher strings in the classical guitar, the lowest three strings were investigated for the steel-stringed guitar. Nine subjects participated in the experiment. The chosen pitches were the open string and fret position 7 of strings 4, 5, and 6. Their fundamental frequencies and *B* coefficients are given in Table II. The initial transient was present in the tones.

The test was run as in the nylon-stringed condition, but the number of training trials at the beginning of each pitch condition was increased from four to six. Thus for each pitch, there were 18 trials, and the percentage of correct responses was recorded from the last 12 ones, as before.

The results are presented in Figure 7. The median results are above 75% correct for all pitch conditions except fret 7 of string 4. However, for strings 4 and 5 individual results range sometimes under 75%, which suggests that detection is uncertain for other strings than the lowest one. For both pitches of string 6, the median percent correct is 100. The difference in detection for the lowest string between the nylon-stringed and steel-stringed guitars is apparently due to the higher *B* coefficient in the steel-stringed guitar. For other strings the difference was smaller.

4. Summary and discussion

The perceptibility of inharmonicity in the acoustic guitar was studied in two formal listening experiments by using synthetic guitar tones whose inharmonicity could be accurately controlled. The first experiment studied general perceptibility of inharmonicity at six pitches along the pitch range of the nylon- and steel-stringed guitars. For the nylon-stringed guitar, inharmonicity was best detected at strings 3 and 6. However, perception seemed to be strongest at the very beginning of the sound and was interfered by the presence of the plucking transient. This is understandable because the most inharmonic higher partials decay fast in string instrument tones and become easily masked by the plucking transient. For the steel-stringed guitar however, detection in the lowest string seemed to be easy and was not impaired by the plucking transient. The difference in detection for the lowest string in favor of the steel-stringed guitar may be due to the higher *B* coefficient. For other strings the difference was smaller.

In the second experiment, detection thresholds were obtained for inharmonicity for the 12th fret of strings 3 and 6 of the nylon-stringed guitar. The mean thresholds were close to, although above, typical amounts of inharmonicity in the nylon-stringed guitar. However, the threshold is clearly below the *B* coefficient that was measured for string 6 of the steel-stringed guitar. In the first experiment, inharmonicity was also clearly detected in the lowest string of the steel-stringed guitar.

Furthermore, the currently obtained thresholds were significantly higher than the ones obtained in the previous study for generic, string instrument-like sounds. As the test tones in the previous study decayed at the same rate for all partials, the partials of the current tones decayed at increasing rate at higher frequencies, making the effect of inharmonicity fade out quickly.

Two of the subjects detected much lower inharmonicities than the others. This suggests that some expert listeners might detect even a weak timbral effect while normal listeners would not notice any difference. It is therefore recommended that in accurate synthesis of guitar tones, inharmonicity should be included for the lowest 4 strings. This is especially important for the steel-stringed guitar. For the nylon-stringed guitar or for less critical applications of sound synthesis, the effect of inharmonicity could be ignored. The current result is useful in sound synthesis, for instance in design of dispersion filters for physical modeling [11].

Acknowledgement

The work was supported by the Academy of Finland, project numbers 105651 and 106953. The authors are thankful to Mr. Jukka Rauhala for analysis tools and to Dr. Mikael Laurson, Mr. Jussi Pekonen, and Mr. Henri Penttinen for recordings of guitar sounds.

References

[1] B. C. J. Moore, R. W. Peters, B. R. Glasberg: Thresholds for the detection of inharmonicity in complex tones. J. Acoust. Soc. Am. **77** (1985) 1861–1867.

[2] H. Fletcher: Normal vibration frequencies of a stiff piano string. J. Acoust. Soc. Am. **36** (1964) 203–209.

[3] H. Fletcher, E. D. Blackham, R. Stratton: Quality of piano tones. J. Acoust. Soc. Am. **34** (1962) 749–761.

[4] H. Järveläinen, V. Välimäki, M. Karjalainen: Audibility of the timbral effects of inharmonicity in stringed instrument tones. Acoustics Research Letters Online **2** (2001) 79–84. http://ojps.aip.org/arlo/.

[5] X. Serra: Spectral modeling synthesis: Past and present. http://www.iua.upf.edu/~xserra/presentacions/spectral-modeling-synthesis-past-and-present.pdf.

[6] A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, J. Huopaniemi: Frequency-warped signal processing for audio applications. J. Audio Eng. Soc. **48** (2000) 1011–1031.

[7] J. D. Markel, A. H. Gray Jr.: Linear prediction of speech signals. Springer, Berlin, 1976.

[8] J. Woodhouse: Plucked guitar transients: Comparison of measurements and synthesis. Acta Acustica united with Acustica **90** (2004) 945–965.

[9] M. Karjalainen, P. Esquef, P. Antsalo, A. Mäkivirta, V. Välimäki: Frequency-zooming ARMA modeling of resonant and reverberant systems. J. Audio Eng. Soc. **50** (2002) 1012–1029.

[10] H. Levitt: Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. **49** (1971) 467–477.

[11] J. Rauhala, V. Välimäki: Tunable dispersion filter design for piano synthesis. IEEE Signal Processing Letters **13** (2006) 253–256.