ForumAcusticum 2005

Is inharmonicity perceivable in the acoustic guitar?

Matti Karjalainen and Hanna Järveläinen
Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, P.O.BOX 3000, FI-02015 HUT, Finland,
e-mail: **{matti.karjalainen, hanna.jarvelainen}@tkk.fi**,

Plucked and struck string instrument sounds are known to exhibit inharmonicity due to stiffness of the strings. Especially in the low piano tones this is very prominent so that the higher partials have frequencies higher than integer multiple of the fundamental frequency. Järveläinen et al (2001) studied the perceivability of inharmonicity for generic string instrument sounds where each harmonic decays with the same rate. While being useful to roughly estimate the threshold where inharmonicity is just noticeable, the importance of this phenomenon in the acoustic guitar has remained unknown. In the present study we have explored this question by listening tests and careful manipulation of the inharmonicity of authentic guitar sounds. A high-resolution parametric modeling technique called FZ-ARMA was applied to estimate the partials of recorded sounds, and these partials were scaled in frequency so that inharmonicity was freely controllable. The results of the listening tests show that the inharmonicity is close to the JND threshold for the four lowest strings while for the two highest strings it is not perceivable. This result was obtained for generally experienced listeners, but it seems that for well trained experts the inharmonicity of low notes is more easily perceived. This implies also that for highest quality sound synthesis of the guitar the inharmonicity must be taken into account, while for an average listener it is questionable if it makes any difference.

1 Introduction

Inharmonicity may be perceived in different ways, depending on its type. From the viewpoint of auditory organization, the ability to detect mistuning in one partial of an otherwise harmonic complex was subject to research in the 80's and 90's. It was found that detection of mistuning in the higher harmonics is primarily based on beats or roughness, whereas a mistuned component in the lower harmonics tends to stand out from the complex (Moore, Peters and Glasberg 1985). The roughness cue was found to depend strongly on stimulus duration: for short sounds the ability to detect mistuning by roughness drops dramatically for the higher harmonics (Moore et al. 1985), (Lee and Green 1994). Pitch matching experiments showed that also neural synchrony may be involved in detecting mistuning. In (Hartmann 1990), listeners lost their ability to hear out mistuned harmonics at frequencies higher than 2.2-3.5 kHz, the same frequency region where synchronous neural firing vanishes.

In string instruments, all partials are mistuned systematically due to string stiffness. The partial frequencies f_n are given by (Fletcher, Blackham and Stratton 1962):

$$f_n = n f_0 \sqrt{1 + Bn^2} \qquad (1)$$

$$B = \pi^3 Q d^4 / 64 l^2 T \qquad (2)$$

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.

The inharmonicity pattern given above affects both the timbre and the pitch of the complex. However, in string instruments the practical importance to the quality of sound is largely unknown. In the low register of the piano the effect is easily perceivable, since the lowest strings are quite massive and dispersive (Galembo and Cuddy 1997). Inharmonicity adds warmth to the sound (Fletcher et al. 1962), and accurate simulation of inharmonicity in frequency bands up to 1.5 - 4 kHz is considered important in synthesis of piano tones (Rocchesso and Scalcon 1999). The effect on pitch was recently studied in (Anderson and Strong 2005). In the piano bass range, the study demonstrates a much higher perceived pitch than would be expected based on frequencies of the lowest partials that are generally considered to dominate pitch perception. The pitch effect was studied in the whole piano range in (Järveläinen, Verma and Välimäki 2000), where a closer relation was found between pitch and the lowest harmonics.

The importance of inharmonicity in the acoustic guitar has been unknown, since the previous results cannot be generalized to other instruments. This study explores perception of inharmonicity in the guitar pitch range by two formal listening experiments. The first experiment measures the general audibility of inharmonicity in guitar tones, and the second experiment measures the detection thresholds for inharmonicity at two pitches.

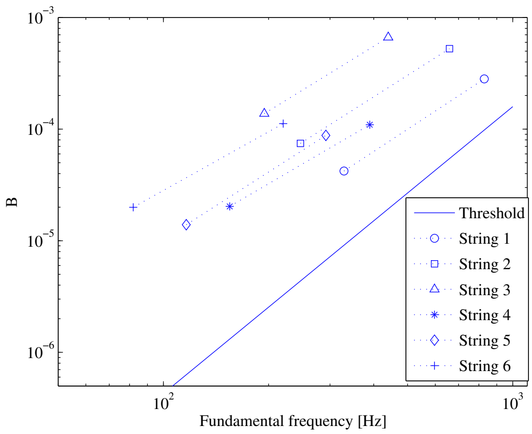

Figure 1: Audibility thresholds from (Järveläinen et al., 2001) (solid line) versus inharmonicity coefficients B for frets 0-17 of each string as a function of fundamental frequency.

2 Inharmonicity in guitar tones

Inharmonicity coefficients were measured from recorded classical guitar tones for the free strings and the first 17 frets. The inharmonicity coefficients increased with fundamental frequency and fret number. Audibility thresholds were measured in (Järveläinen, Välimäki and Karjalainen 2001) for synthetic string instrument -like tones as a function of fundamental frequency. This threshold is presented in Fig. 1 together with the measured inharmonicity coefficients for real guitar tones. In all tones inharmonicity clearly exceeds the audibility 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 audibility 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, the previous results may not fully predict the current case.

In general, slightly inharmonic tones differ from respective harmonic tones by a timbral effect. Increasing inharmonicity increases the pitch of the tone. In highly inharmonic tones some of the partials may segregate from the complex and be heard separately. Obviously, inharmonicity in string instruments causes only the timbral effect and possibly a very weak pitch change. Since the timbral effect is of interest here, the test tones were checked for pitch changes in preliminary listening but no perceivable effects were observed.

3 Manipulation of harmonicity

To modify the inharmonicity of a guitar sound the frequencies of the partials have to be modified without changing any other perceptual features of the sound. 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. We decided to use high-resolution pole-zero plus residual modeling, because it has a closer correspondence to the physical basis of the instrument sounds. Each partial is modeled as a set of modes that can be scaled in frequency for resynthesis, and the original residual is added to yield a new version of the sound with modified harmonic structure.

3.1 High-Resolution Mode Analysis

If the response of a string to excitation were an AR (autoregressive) process, it could be analyzed by straightforward AR modeling techniques, such as the linear prediction (LP) (Makhoul 1975), resulting in an all-pole filter. This is not manageable in practice because the pole positions close to the unit circle makes the analysis numerically too critical, and the allocation of the number of poles for each partial of the string cannot be easily controlled. It is important to notice that the two polarizations of string vibration with slighly different frequencies make each partial to be a sum of two decaying sinusoids, exhibiting beating or two-stage decay of the envelope (Fletcher et al. 1962) or two separate spectral peaks as illustrated in Fig. 2. This requires at least two pole pairs to represent one partial.

Since the string 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 are capable of matching such responses, but due to iterative solutions, models that are higher in order than 20–200 may not converge to a stable and useful result.

3.2 FZ-ARMA analysis

To avoid the problems in resolution and computational precision discussed above, we have developed a subband technique called FZ-ARMA (frequency-zooming ARMA) analysis (Karjalainen, Esquef, Antsalo and Välimäki 2002). Instead a single model, global over the entire frequency range, the signal is pole-zero 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 resonant modes that are very sharp or close to each other in frequency. The FZ-ARMA analysis consists of the following steps.

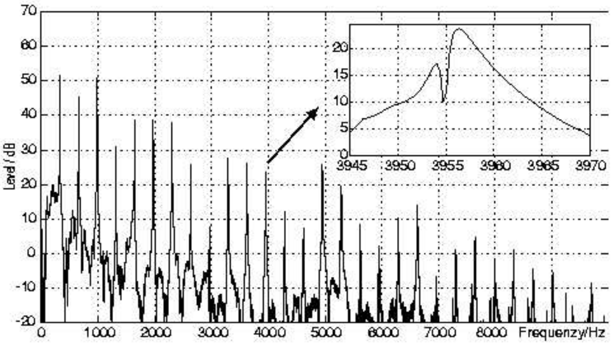

Figure 2: Spectrum of plucked string sound of a classical acoustic guitar, open string E_4. Spectrum zoomed to a single partial is shown in the subplot.

(i) Select a frequency range of interest, e.g. 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

$$h_m(n) = e^{-j\Omega_m n} h(n) \tag{3}$$

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

(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

$$z_{i,zoom} = z_i^{K_{zoom}} = |z_i|^{K_{zoom}} e^{j(\Omega_i - \Omega_m)K_{zoom}} \tag{4}$$

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

(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. For example 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. Figure 2 depicts the spectrum of a plucked guitar sound with a subplot zoomed into one partial having two spectral peaks.

A proper number of zeros in FZ-ARMA modeling is 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.

3.3 Modification of inharmonicity

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

(i) Analyze each partial of the given recording by the 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. Crop the most prominent part from the beginning as the residual for resynthesis.

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

4 Listening tests

Perception of inharmonicity was studied in two formal listening tests. The first experiment explored the general audibility of inharmonicity in acoustic guitar tones versus similar but harmonic tones. In the second experiment,

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

Table 1: Test tones and their inharmonicity coefficients in Experiment 1.

detection thresholds of inharmonicity were measured for two pitches based on the results of the first experiment.

Eleven subjects participated in both experiments, including author HJ. None of them reported a hearing defect, and all had previous experience from psychoacoustic experiments. All played some string instrument but only four were professionally trained guitarists. 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. Inharmonicity coefficients B were varied such that the resulting inharmonicity either matched to, or was below or above, the original. A harmonic reference tone was generated for each inharmonic stimulus with $B = 0$.

4.1 General audibility of inharmonicity in guitar tones

The ability to detect an 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 matched to the original recorded tones (see Table 1). 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.

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 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. For each pitch, the question was repeated 16 times. The percentage of correct responses was recorded from the last 12 trials.

The results are presented in Fig. 3. 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,

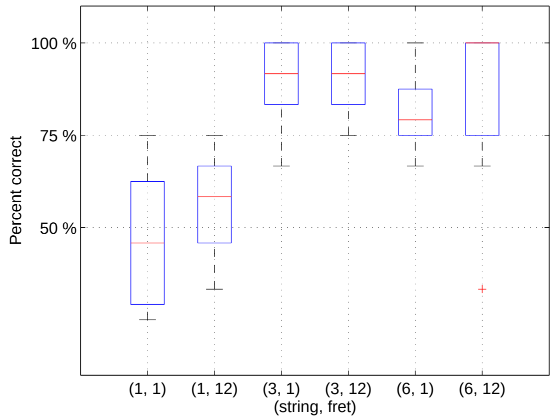

Figure 3: Boxplot of the results of experiment 1.

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. The performance dropped from nearly 100 % correct to nearly chance level (see Fig. 4). 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 causing the effect to be masked by the initial transient. A similar phenomenon was observed by (Moore et al. 1985), who found that shortening the stimulus duration produced an impairment in detection performance for higher harmonics.

4.2 Perception threshold for individual tones

Since detection of inharmonicity remained uncertain even at the lowest pitch range, detection thresholds were measured 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 (Levitt 1971). After two correct responses, 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 measurement 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.

Fig. 5 presents the results in relation to the B coefficients measured for all frets of strings 3 and 6, and the thresholds measured in (Järveläinen et al. 2001), as a function of fundamental frequency. While the current

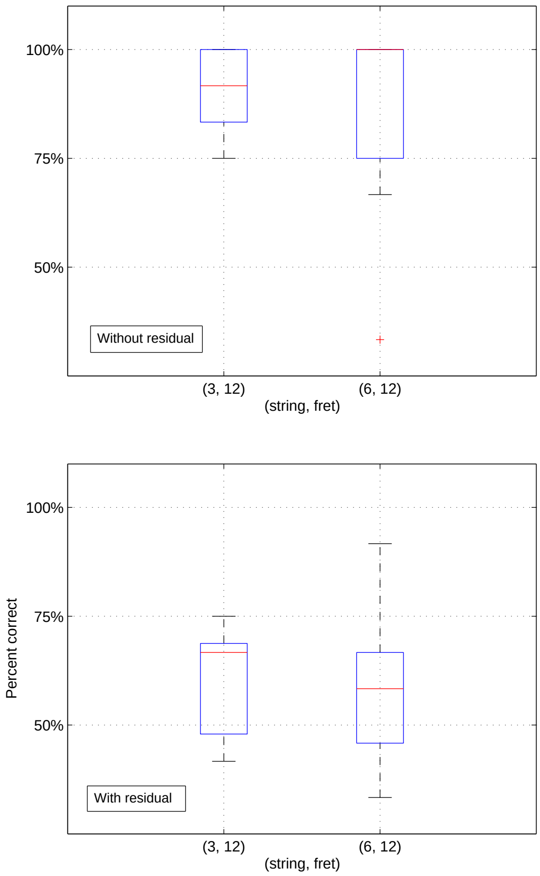

Figure 4: Results of the 12th fret of strings 3 and 6 without and with the initial transient (top and bottom panel, respectively).

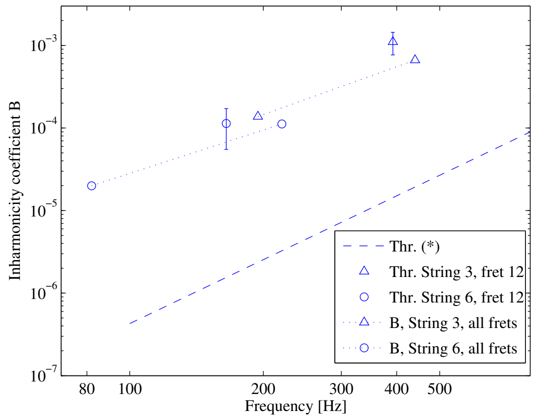

Figure 5: Currently measured thresholds (with error bars of one standard deviation up and down), respective B values for guitar strings, and the threshold measured in (Järveläinen et al., 2001) (*).

thresholds are far above those measured previously, they are relatively close to the realistic inharmonicity coefficients. The mean threshold for fret 12 of string 3 was $B = 0.0011$ and for fret 12 of string 6 $B = 0.00013$. These are 2.12 and 1.62 times the respective B coefficients measured from guitar tones. However, three of the subjects performed much better than the rest, and their thresholds were for at least one of the studied pitches below typical B values in real guitars.

In the adaptive procedure, inharmonicity was increased after each incorrect response. The drawback of the method was that after many incorrect responses the B value became so high that the stimuli started to sound unnatural. The difference was then easily perceivable but may have drawn the subjects' attention from the less perceivable effect (produced by lower B values) that was perceived by a few listeners and unnoticed by many.

5 Summary and conclusions

The effect of inharmonicity in the acoustic guitar was studied by two formal listening experiments by using synthetic guitar tones whose inharmonicity could be accurately controlled. The first experiment studied general audibility of inharmonicity at six pitches along the pitch range of the guitar from strings 1, 3, and 6. Inharmonicity was generally detected at strings 3 and 6 when the initial transient was not included in the synthesis. 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.

In the second experiment, detection thresholds were measured for inharmonicity for the 12th fret of strings 3 and 6. The mean thresholds were close to, although above, typical amounts of inharmonicity in the guitar. However, they were significantly higher than the previously measured thresholds for generic, string instrument-like sounds. As the previous test tones decayed at the same rate for all partials, the partials of the current tones decayed at increasing rate at higher frequencies, probably making the effect fade out quickly.

Two of the subjects detected much lower inharmonicities than the others. This suggests that expert listeners might detect even a weak timbral effect while normal listeners would not notice any difference. It is therefore recommended that for accurate synthesis of guitar tones, inharmonicity should be included for the lowest 4 strings. For less critical applications it could be well ignored.

The present study included only single notes of a string. It could be argued that in real playing with chords the beats between harmonics of two or more notes could make inharmonicity easier to perceive. In a single case of preliminary testing with a two-note interval we did not notice increased perceivability, but more work is needed to test it more generally. Another task of future research is to explore if it is possible to develop a computational model that can predict the perceivability of inharmonicity in more general cases.

6 Acknowledgements

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 Mr. Henri Penttinen for recordings of guitar sounds.

References

Anderson, B. E. and Strong, W. J.: 2005, The effect of inharmonic partials on pitch of piano tones, *J. Acoust. Soc. Am.* **117**(5), 3268–3272.

Fletcher, H., Blackham, E. D. and Stratton, R.: 1962, Quality of piano tones, *J. Acoust. Soc. Am.* **34**(6), 749–761.

Galembo, A. and Cuddy, L.: 1997, String inharmonicity and the timbral quality of piano bass tones: Fletcher, Blackham, and Stratton (1962) revisited, Report to the 3rd US Conf. Music Perception and Cognition.

Hartmann, W. M.: 1990, Hearing a mistuned harmonic in an otherwise periodic complex tone, *J. Acoust. Soc. Am.* **88**(4), 1712–1724.

Järveläinen, H., Välimäki, V. and Karjalainen, M.: 2001, Audibility of the timbral effects of inharmonicity in stringed instrument tones, *Acoustics Research Letters Online* **2**(3), 79–84. http://ojps.aip.org/arlo/.

Järveläinen, H., Verma, T. and Välimäki, V.: 2000, The effect of inharmonicity on pitch in string instrument sounds, *Proc. Int. Computer Music Conf.*, Berlin, Germany, pp. 237–240.

Karjalainen, M., Esquef, P., Antsalo, P. and Välimäki, V.: 2002, Frequency-zooming ARMA modeling of resonant and reverberant systems, *J. Audio Eng. Soc.* **50**(12), 1012–1029.

Lee, J. and Green, D. M.: 1994, Detection of a mistuned component in a harmonic complex, *J. Acoust. Soc. Am.* **96**(2), 716–725.

Levitt, H.: 1971, Transformed up-down methods in psychoacoustics, *J. Acoust. Soc. Am.* **49**(2), 467–477.

Makhoul, J.: 1975, Linear prediction: A tutorial review, *Proc. IEEE* **63**, 561–580.

Moore, B. C. J., Peters, R. W. and Glasberg, B. R.: 1985, Thresholds for the detection of inharmonicity in complex tones, *J. Acoust. Soc. Am.* **77**(5), 1861–1867.

Rocchesso, D. and Scalcon, F.: 1999, Bandwidth of perceived inharmonicity for musical modeling of dispersive strings, *IEEE Trans. Speech and Audio Proc.* **7**, 597–601.