

# AUDIBILITY OF INHARMONICITY IN STRING INSTRUMENT SOUNDS, AND IMPLICATIONS TO DIGITAL SOUND SYNTHESIS

*Hanna Järveläinen, Vesa Välimäki, and Matti Karjalainen*

Helsinki University of Technology  
Laboratory of Acoustics and Audio Signal Processing  
P.O.Box 3000, FIN-02015 HUT, Finland  
hanna.jarvelainen@hut.fi, vesa.valimaki@hut.fi  
matti.karjalainen@hut.fi  
<http://www.acoustics.hut.fi>

## ABSTRACT

Listening tests were conducted in order to find out the audibility of inharmonicity in musical sounds produced by string instruments, such as the piano or the guitar. The threshold of the audibility of inharmonicity was measured as a function of the inharmonicity coefficient  $B$  at five fundamental frequencies. It was found that the detection of inharmonicity is strongly dependent on the fundamental frequency. A simple model is presented for estimating the threshold as a function of the fundamental frequency. The need to implement inharmonicity in digital sound synthesis is discussed.

## 1. INTRODUCTION

The frequencies of the partials of string instrument sounds are not exactly harmonic. This is due to stiffness of real strings, which contributes to the restoring force of string displacement together with string tension. The stretching of the partials can be calculated in the following way [1]:

$$f_n = n f_0 (1 + B n^2)^{0.5} \quad (1)$$

$$B = \frac{\pi^3 Q d^4}{64 l^2 T} \quad (2)$$

where  $n$  is the partial number,  $Q$  is Young's modulus,  $d$  is the diameter,  $l$  is the length and  $T$  is the tension of the string, and  $f_0$  would be the fundamental frequency of the string completely without stiffness.

Inharmonicity is not necessarily unpleasant. Fletcher, Blackham, and Stratton [1] pointed out that a slightly inharmonic spectrum added certain "warmth" into the sound. They found that synthesized piano tones sounded more natural when the partials below middle C were inharmonic.

The effect of mistuning one spectral component in an otherwise harmonic complex is well known. Moore and his colleagues [2] reported that the thresholds for

detecting mistuning decreased progressively with increasing harmonic number and increasing fundamental frequency. In their experiment, mistuning was expressed as percentage of the harmonic frequency, and the test tones were complex tones with 12 harmonics at equal levels.

Moore's group also showed that mistuning is heard in different ways depending on the harmonic number [2]. Shortening the stimulus duration produced a large impairment in performance for the higher harmonics, while it had only little effect on the performance for the lower harmonics. It was reasoned that particularly for long durations beats provide an effective cue, but for short durations many cycles of beats cannot be heard. For the lower harmonics, beats were generally inaudible, and the detection of mistuning appeared to be based on hearing the mistuned component stand out from the complex. The thresholds varied only weakly with duration.

Scalcon et al. [3] studied the bandwidth of correct positioning of the partials of synthesized piano tones. They found cutoff frequencies above which it was unnecessary to imitate inharmonicity. For low tones the relevant bandwidth was smaller than for higher tones, but on the other hand, many more partials were included in the frequency range at low tones. They also stated that the effect of inharmonicity was unimportant on the highest part of the keyboard.

We studied the audibility of inharmonicity as a function of the inharmonicity coefficient  $B$  with fundamental frequency and sound duration as parameters. The aim is to find general rules for the need to implement inharmonicity in digital sound synthesis. If inharmonicity were ignored, computational savings could be achieved, since for instance in digital waveguide modeling, an additional allpass filter is needed to implement inharmonicity [4], [5], [6].

## 2. LISTENING TESTS

Subjects were required to distinguish between a complex tone whose partials were exactly harmonic and an otherwise identical complex tone whose partials were mistuned. The threshold of audible inharmonicity was measured as a function of  $B$  at fundamental frequencies 55 Hz ( $A_1$ ), 82.4 Hz ( $E_2$ ), 220 Hz ( $A_3$ ), 392 Hz ( $G_4$ ), and 1108.7 Hz ( $C\#_6$ ).

### 2.1. Test sounds

The test signals were generated using additive synthesis to enable accurate control of the frequency and amplitude of each partial. The sampling rate was 22.05 kHz. The spectrum of the tones had a lowpass character of the form  $1/\text{frequency}$ , which is similar to spectra of many string instruments. The decay of all partials was exponential with a time constant  $\tau = 0.5$  seconds. The initial phase of each partial was chosen randomly.

The tones contained all partials of the fundamental frequency  $f_0$  up to 10 kHz. A constant cutoff frequency was chosen because it was impossible to use the same number of partials for every fundamental frequency. Up to 50 partials would be important to the perception of inharmonicity at low bass tones [3]. However, less than ten partials could be generated for the highest tone before meeting the Nyquist limit. Realizing that the variable number of partials might affect the audibility of inharmonicity, we reasoned that a constant cutoff would still be the most practical solution.

A constant upper limit was also needed to keep the spectral width of all test sounds equal. Galembo and Cuddy showed [7] that without control of the spectral width, inharmonicity changes the balance between high and low frequencies and creates an impression of sharpening.

Based on initial listening, we found that the threshold of audibility varies over frequency. Thus, the range of the inharmonicity factor  $B$  must also vary as a function of frequency due to the chosen test method. A range of  $B$  was found for each fundamental frequency to cover the probable threshold of audibility. The values of  $B$  were uniformly spaced within the range.

A pitch increase due to inharmonicity was heard at the highest tone. The subjects might listen to pitch differences instead of timbre and ignore the effect of inharmonicity unless there was an audible change of pitch. To prevent this, the harmonic references of  $C\#_6$  were slightly tuned up to match the pitch of each inharmonic test sound. The pitches were adjusted until no annoying difference was heard.

The effect of duration of the tone on the audibility of inharmonicity was also studied. Two tone lengths were tested, 1.5 s and 300 ms. The short samples were generated by cutting off most of the decay phase of the longer sounds. The time constant remained the same for both durations.

### 2.2. Subjects and test method

Four subjects participated the final listening test. The listeners were 20-30 years old, and all of them had musical training in some string instrument, either the piano or the guitar. None of them reported any hearing defects. One of the listeners was the author HJ. The sound samples were played through headphones from a Silicon Graphics workstation using the GuineaPig software [8]. Before the test the subjects were allowed to practise until they made firm judgments.

The method of constant stimuli [9] was used. The subjects heard pairwise a perfectly harmonic reference sound and a possibly inharmonic sound, and the task was to decide whether they sounded equal or different. Eight values of  $B$  (including  $B = 0$ ) were used for each fundamental frequency, and each sound pair was judged four times. The playback order of the sound pairs was randomized, and the harmonic reference was the first sound within a pair twice and the second one twice.

## 3. TEST RESULTS

The psychometric function was approximated by counting how many times each subject had regarded a sound pair as “different” [9]. If there were zero “different” judgments out of the four trials, the subject had noticed the possible inharmonicity 0% of the time, and by making four “different” judgments 100% of the time. The 50% threshold was chosen for the threshold of audibility. If the threshold was not directly seen in the data, it had to be interpolated between higher and lower percentages. If it was spread over two values of  $B$ , the mean was calculated between them.

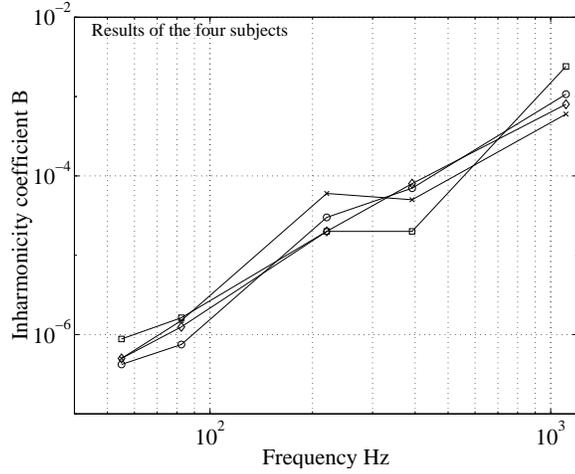
The audibility thresholds are shown as a function of fundamental frequency for each subject in fig. 1. The sample duration was 1.5 seconds. The answers of the four subjects were roughly normally distributed. The mean thresholds over the listeners at different fundamental frequencies are shown in table 1.

**Table 1:** Inharmonicity coefficient  $B$  at mean thresholds averaged over the subjects for the different fundamental frequencies.

$f_0$	$B$ at mean threshold	Standard deviation $\sigma$
$A_1$	0.0000058	0.00000021
$E_2$	0.0000013	0.00000038
$A_3$	0.000033	0.000019
$G_4$	0.000055	0.000027
$C\#_6$	0.0012	0.00081

The thresholds show a strong linear trend when a logarithmic scale is used both for the frequency and the  $B$ . The highest note  $C\#_6$  was judged with least accuracy, and the effect of inharmonicity was also small-

est. At the mean threshold the value of  $B$  was more than 2000 times higher than that for the lowest tone.



**Figure 1:** The individual thresholds for the four listeners at  $A_1$ ,  $E_2$ ,  $A_3$ ,  $G_4$ , and  $C\#_6$ . Sample duration was 1.5 seconds.

A straight line was fitted to the mean threshold values in the least-squares sense. This way a simple formula was derived that could be used to model the audibility threshold as a function of fundamental frequency:

$$\log B = 2.54 \log f_0 - 24.6 \quad (3)$$

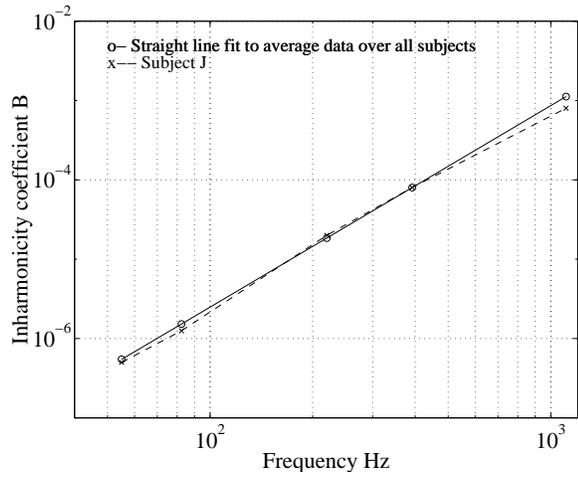
The natural logarithm was used. The fitted line is illustrated in fig. 2. The thresholds of one subject were especially near the estimated line (subject J, see figure 2). The thresholds of two other subjects differ from the line at  $G_4$ , where they were able to detect smaller inharmonicity than at  $A_3$ . The nonmonotonic behavior is not surprising, since the subjects reported that they used several different cues to detect inharmonicity. The performance can depend on the existence of certain cues at different fundamental frequencies and the subject's sensitivity to the particular cue such as beating or roughness.

### 3.1. Effect of duration

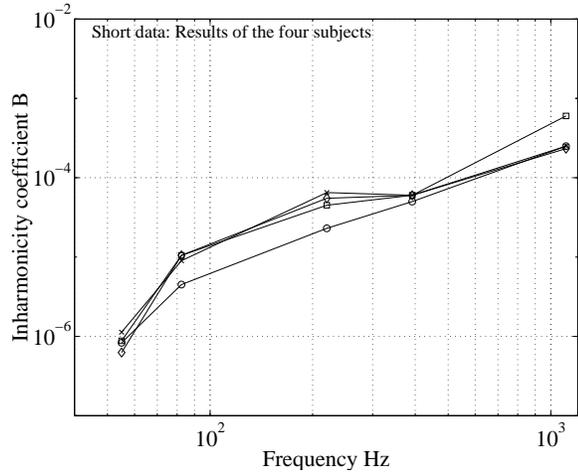
The test was repeated using shorter sound samples (300 ms). The thresholds of the four subjects (see fig. 3) showed again a linear trend, and a straight line was fitted to the results as before. The following model was derived:

$$\log B = 1.78 \log f_0 - 20.2 \quad (4)$$

Both models as well as the average thresholds over all subjects in both cases are shown in fig. 4. The slopes of the models suggest that at low fundamental frequencies the detection of inharmonicity becomes harder and at high fundamental frequencies somewhat



**Figure 2:** The straight line fitted to average data over all subjects (solid line), and the thresholds of subject J (dashed line).

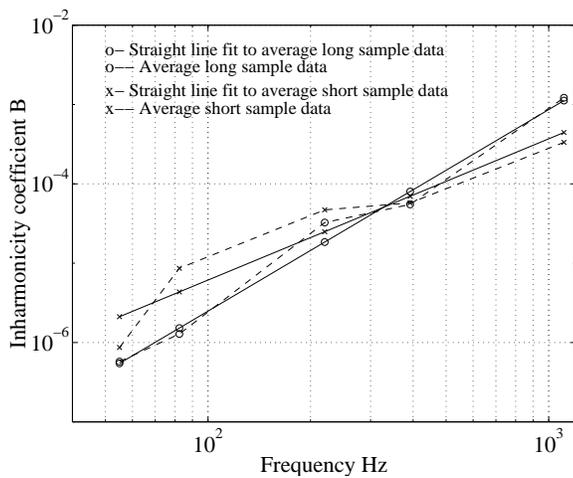


**Figure 3:** The individual thresholds for the four listeners at  $A_1$ ,  $E_2$ ,  $A_3$ ,  $G_4$ , and  $C\#_6$ . Sample duration was 300 ms.

easier when the samples become shorter. To test the significance of the linear models, t-tests were performed on the slopes [10]. The steeper slope of the long duration model was tested against the slope of the short duration model, and the slope of the short duration model was tested against zero slope. Both results were highly significant, suggesting that the two slopes are greater than zero ( $P(\text{slope}=0) = 0.0028$ ) and that the slope of the long duration model is steeper than that of the short duration model ( $P(\text{slopes are equal}) = 0.0008$ ).

## 4. DISCUSSION

Typical values of  $B$  for piano strings lie roughly between 0.00005 for low bass tones and 0.015 for the high treble tones [11]. Though the sounds are less in-



**Figure 4:** The linear models of audibility of inharmonicity as a function of fundamental frequency for long and short durations (solid lines), and the corresponding mean thresholds (dashed lines).

harmonic in the bass range, inharmonicity is detected more easily than at the highest tones. Compared to the audibility thresholds, inharmonicity would be clearly audible at low fundamental frequencies. In the treble range, audibility is questionable. At  $C\#_6$ , the threshold of audibility is closer to the possible values of  $B$  in real instruments. Furthermore, the highest tone was judged less accurately than the others and even non-monotonously, i.e., the performance weakened locally when inharmonicity was increased.

Thus it would be necessary to implement the effect of inharmonicity in digital sound synthesis systems at least in the bass range. In the treble range computational savings could be achieved by omitting the all-pass filter responsible for the effect of inharmonicity.

There can be several causes for the better performance at lower frequencies. The subjects told that they were using beats as a cue. Beats were mostly audible at low fundamental frequencies. When most of the decay phase was cut off, the performance weakened in the bass range but improved in the treble range, where another cue was possibly used.

The test results are consistent with the general assumption that the effect of inharmonicity is greater when more partials are present [7]. This could also be a cause for the differences in performance. However, the number of partials in our hearing range decreases with increasing fundamental frequency also in sounds of real musical instruments, so it would make no sense to isolate these factors from each other. Inharmonicity can be influenced also by many other features [7]. To study all these in a formal listening experiment would be too laborious.

Our future objective is to find more general rules for the need to implement the effect of inharmonic-

ity in digital sound synthesis systems. More accurate models are needed for the effect of duration, number of partials and spectral width, relative level of the partials, and different decay rates between them.

## 5. ACKNOWLEDGMENTS

This work was supported by the Academy of Finland and the Pythagoras Graduate School. The authors express sincere thanks to the test subjects and also to Alexander Galembo for kindly providing his recent publications.

## 6. REFERENCES

- [1] H. Fletcher, E. D. Blackham, and R. Stratton, "Quality of piano tones," *J. Acoust. Soc. Am.*, vol. 34, no. 6, pp. 749–761, 1962.
- [2] B. C. J. Moore, R. W. Peters, and B. C. Glasberg, "Thresholds for the detection of inharmonicity in complex tones," *J. Acoust. Soc. Am.*, vol. 77, no. 5, pp. 1861–1867, 1985.
- [3] F. Scalcon, D. Rocchesso, and G. Borin, "Subjective evaluation of the inharmonicity of synthetic piano tones," in *Proc. Int. Comp. Music Conf. ICMC'98*, pp. 53–56, 1998.
- [4] D. A. Jaffe and J. O. Smith, "Extensions of the Karplus-Strong plucked-string algorithm," *Computer Music Journal*, vol. 7, no. 2, pp. 56–69, 1983.
- [5] A. Paladin and D. Rocchesso, "A dispersive resonator in real time on MARS workstation," in *Proc. Int. Comp. Music Conf. (ICMC'92)*, (San Jose, CA), pp. 146–149, Oct. 1992.
- [6] S. Van Duyne and J. O. Smith, "A simplified approach to modeling dispersion caused by stiffness in strings and plates," in *Proc. Int. Comp. Music Conf. ICMC'94*, pp. 407–410, 1994.
- [7] A. Galembo and L. Cuddy, "String inharmonicity and the timbral quality of piano bass tones: Fletcher, Blackham, and Stratton (1962) revisited." Report to the 3rd US Conference on Music Perception and Cognition, MIT, Cambridge, MA, July - August 1997.
- [8] J. Hynninen and N. Zacharov, "GuineaPig – a generic subjective test system for multichannel audio." Presented at the 106th Convention of the Audio Engineering Society, May 8-11 1999, Munich, Germany, preprint no. 4871, 1999.
- [9] J. P. Guilford, *Psychometric Methods*. McGraw-Hill, 1956.
- [10] J. Milton and J. C. Arnold, *Introduction to probability and statistics*. McGraw-Hill, 1990.
- [11] H. A. Conklin, "Generation of partials due to nonlinear mixing in stringed instruments," *J. Acoust. Soc. Am.*, vol. 105, no. 1, pp. 536–545, 1999.