Detection of missing modal frequencies

Christophe Stoelinga and Robert Lutfi

University of Wisconsin-Madison, Department of Communicative Disorders, 1975 Willow drive, Madison, WI 53706, USA
stoelinga@wisc.edu

The amount of information about a source that can be recovered from sound is naturally limited by the ear's ability to resolve individual modal frequencies unique to that source. To measure these limits, listeners were asked to detect, in a standard two-interval, forced-choice task with feedback, which of two sounds on each trial contained a missing partial. The frequencies of the partials corresponded to the ideal modes of a simply-supported, rectangular plate. Plate surface areas and height-wide ratios were chosen to produce the same bandwidths (125-1125, 250-2250 or 500-4500 Hz) for different numbers of partials (11, 16 or 24). Overall level of the sounds was roved to discourage detection based on simple level differences. Detection of five highly-practiced listeners was largely independent of the frequency of the lowest partial, being best for partials 1-3, 1-4 or 1-6 for 11, 16 or 24 partials in total, respectively. Mutual masking among the higher-number partials is given as the likely cause. The results are discussed in terms of their implications for the identification of rudimentary source attributes from sound. [Work supported by NIDCD grant 5 R01 DC006875.

1 Introduction

Listeners have been shown to be capable of distinguishing the size and shape of a plate from the sound it produces upon impact [4][5]. However, the acoustical properties of the plate sound listeners use as basis to make such judgements are unclear. The sound emitted by a struck plate is effectively a sum of decaying partials, corresponding to the resonating modes of the plate. If this sound cannot be descriminated from the same sound with one partial missing, then that partial seems unlikely to convey unique information about the plate geometry to the listener. In the current study we test listener sensitivity to such missing partials, and thereby identify which partials are likely to convey unique information regarding the attributes of the plate. This in turn will allow us to evaluate what acoustic information associated with these partials might be available to the listener to judge plate attributes.

One obvious factor expected to influence detection of a missing partial is mutual masking among partials; however, the results from several studies suggest that this may not be the only factor. Moore and Ohgushi [6] studied whether listeners can "hear out" individual partials of a tone complex. To equate for mutual masking, they used tone complexes that consist of partials spaced evenly on an ERB scale. In their study, the probe and the tone complex were presented in sequence. The probe was close in frequency to one of the partials in the complex, but was mistuned slightly downward on half the trials (at random) and mistuned slightly upward on the other half. The task of the subject was to indicate whether the probe was higher or lower in frequency than the nearest partial in the complex. They showed that even for tone complexes that consist of partials spaced evenly on an ERB scale, discremination performance was better for the lower frequency partials. In a review of the literature, Plomp [7] suggests that the even harmonics are easier to hear out than the odd harmonics of a harmonic tone complex. These observations are inconsistent with the idea that components are resolved unless they are masked.

The previously mentioned studies focussed on the ability of listeners to "hear out" certain partials in a complex. This is somewhat different from our study where we test if listeners can detect whether a given partial is missing. Our method leaves open the possibility that detection could be mediated by a timbral difference between two sounds. In this respect, a study by Ellermeier [3] uses a methodology that is more like ours. The stimuli used were multitone complexes consisting of 3, 5, 7, 11, or 15 sinusoidal components equally spaced on log frequency, all occupying the same range between 200 and 3200 Hz. Thresholds were measured for an increase or decrease of the amplitude of the 800 Hz central component. For the closest spacing of components, the two frequencies nearest to the 800-Hz central component fell well outside the critical band. Overall level was roved to prevent detection based on simple level difference. The resulting threshold ranged from a difference of about 3 to 6 dB, depending on the sign of the change and the number of components. There was a slight trend for thresholds to improve as the frequency ratio between the adjacent components was increased. More important, however, detection of increment was generally better thatn detection of decrement; and observation that cannot be explained by mutual masking alone.

2 Methods

2.1 Stimuli

The frequencies of the partials corresponded to the ideal modes of a simply-supported, rectangular thin plate, as given by Rossing and Fletcher [8],

$$f_{mn} = 0.453 c_L h \left[\left(\frac{m+1}{L_x} \right)^2 + \left(\frac{n+1}{L_y} \right)^2 \right], \quad (1)$$

where m and n are non-negative integers denoting the number of nodal lines in the vertical, x, and horizontal, y, direction of the plate, c_L is the longitudinal wave velocity, h is the plate thickness and L_x and L_y are the lenght and width of the plate. Plate surface areas and height-wide ratios were chosen to produce the same bandwidths (125-1125, 250-2250 or 500-4500 Hz) for different numbers of partials (11, 16 or 24). Overall level of the sounds was roved to discourage detection based on simple level differences.

The sounds, $s(t)$ were made to decay in amplitude over time by multiplying the sinusoids of the partials with a frequency-dependent exponential:

$$s(t) = \Sigma_{n=1}^{N} \sin(2\pi f_n t) \mathrm{e}^{-1/\tau(f_n)t},$$

with f_n the modal frequency and $\tau(f_n)$ the corresponding decay for that mode. This corresponding decay is calculated so that $f_n \times \tau(f_n) = 200$. This envelope was chosen rather than the more conventional square envelope because it bears more resemblance to naturally resonating objects. The sounds were one second in duration, and they were gated on and off with a 4 ms raised cosine ramp. The sample rate was 48 kHz.

2.2 Procedure

The sounds were presented to the listeners via headphones (Beyerdynamic DT 990), that were connected via a headphone amplifier (Rolls RA62) to the soundcard of the PC (midiman Delta-1010). The listeners were seated in a double-walled, sound-attenuation chamber. Listeners were asked to detect, in a cued two-interval, forced-choice task with feedback, which of two sounds on each trial contained a missing partial. The actual partials that was missing was selected at random on each trial within a block of trials. Note that in real terms, one or more partials may not be not excited if the plate is hit at a point where the particular mode or modi have a nodal line. This way we have created two valid plate sounds, that is, they could both be generated by the same, real, plate.

3 Results

Figure 1 shows the listeners sensitivity to missing one of the partials of a tone complex. The data from our experiment is shown alongside the results from the simulations. Various pannels show the resulst for the three different numbers of partials and the three different $F0$'s. From this figure we can see that the detection of missing partials for the five highly-practiced listeners was largely independent of the frequency of the lowest partial, being best for partials 1-3, 1-4 or 1-6 for 11, 16 or 24 partials in total, respectively. The effect of changing $F0$ is relatively minor. To predict the effect of mutual masking of partials, we undertook a simulation using the AIM model of [1]. In the AIM model, an auditory filterbank is used, followed by a bank of haircell simulators is used to abttain the Neural Activity Pattern (NAP) that the sound would produce at the level of the auditory nerve or cochlear nucleus. The difference in NAP was taken for the sound with and without a missing partial. This difference was then converted to a value of d' according to the excitation pattern model of Buus and Florentine [2]. The results of the simulations using the AIM model predict reasonably well the number of the lowest components that are best detected when missing. Therefore, mutual masking among the higher-number partials seems the likely cause for the reduced sensitivity to their absence.

3.1 Discussion

The results of the present experiment indicate that partials above model number 5-6 are rarely detected when missing and so are unlikely to have a significant impact on listener judgment regarding physical attributes of plates. To evaluate the acoustic information available in the lower-frequency partials we will refer to Figure 2.

We first explain the origin of the different lines and reasons why they curve in different ways. Analyzing

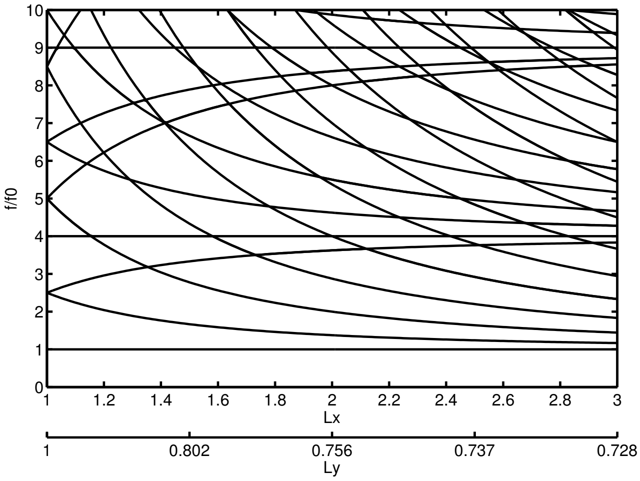

Figure 2: Modal frequencies as a function of plate dimensions. Due to changing geometries of a simply supported plate the frequencies of the modes varies. On the ordinate we find the normalized modal frequencies, on the abscissa the height-wide ratio of the plate. On the left is a square plate, that becomes narrower and longer when reading left to right. The geometry of the plate changes going from left to right on the horizontal axis. Lx and Ly denote the length and width of the plate, thus, the left, the plate is rectangular plate wheras it is long and narrow on the right. By counting the number of modes between 1 and 9 times F0, we can see that the square plate on the left has a less dense spectrum than the long and narrow one on the right.

the lowest harmonics higher than $F0$, we see the modal frequencies $f_{m,n=2,1}$ and $f_{m,n=1,2}$ move in opposite directions. Note that the values for (m, n) cannot be read from the figure but need to be calculated from Equation 1. The first harmonic becomes lower when the plate becomes narrower. The other, $f_{m,n=1,2}$, rises and is equal to $f_{m,n=3,1}$ for $L_x/L_y = 1.35/0.83$. Therefore, $f_{m,n=1,2}$ is the second harmonic for $1/1 < L_x/L_y < 1.35/0.83$ but it is the third harmonic for $1.35/0.83 < L_x/L_y < 3.5/0.77$. Two modes that have the same oscillation frequency are said to be degenerate. Degenerate modes shift a little in frequency, but this effect is not included in our figure. The three horizontal lines represent modes $m, n = 1, 1$ (which is $F0$) , $m, n = 2, 2$ and $m, n = 3, 3$.

As can be seen from Equation 1, the geometry also has an influence on the fundamental frequency, $F0$. The fundamental frequency is one of the most basic potential cues for the plate size. Bigger plates have a lower F0. The fundamental frequency ($F0$) can be controlled via the term in front of the equation, $c_L h$ in a simple linear way. We chose, therefore, to normalize the plot by changing L_x and L_y so that the fundamental frequency is constant. This leaves us in a better position to study the effects of the geometry on the other, higher, modes. As a result, however, both L_x and L_y vary along the abscissa. There is no geometrical constant, such as the plate surface, along the abscissa.

As was mentioned, the first harmonic (F1) monotonically decreases when the plate becomes longer and

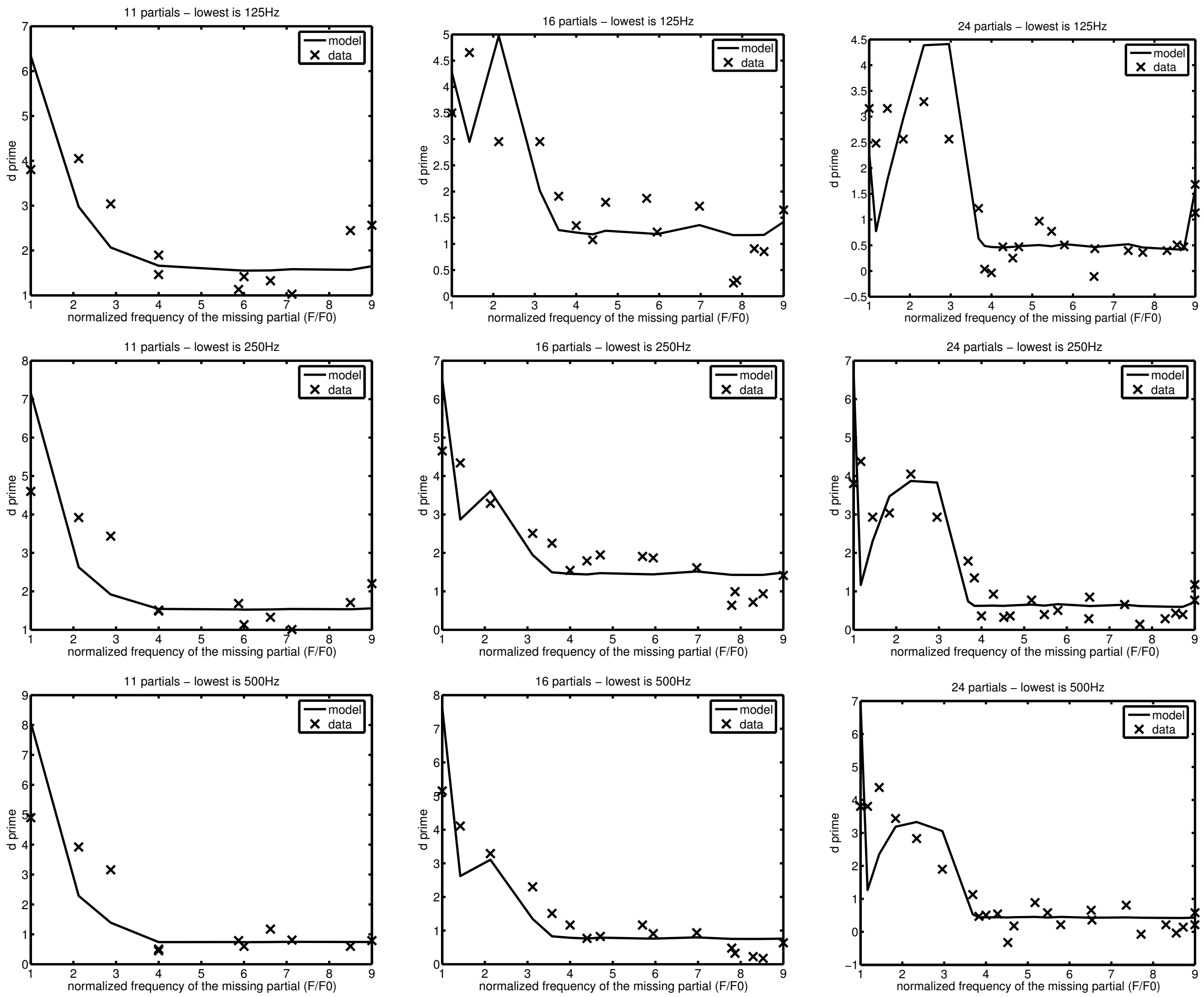

Figure 1: The nine panels show the sensitivity to missing partials of a plate sound, for three different fundamental frequencies and three different number of partials. On the abscissa is the normalized frequency of the missing partial, and the ordinate shows d', a measure of sensitivity. The experimental results are indicated with crosses, the lines result from simulation.

narrower. The frequency of F1 can, therefore, serve as a cue for the plate geometry. Although listeners are typically very poor in absolute frequency determination, they are much better in judging frequency ratios [10]. For the rest of the modes, such simple relations do not exist. For the third harmonic $m, n = 1, 2$ for more square like plates while for more rectangular plates with $Lx/Ly > 1.35/0.83$, the mode with $m, n = 3, 1$ is the third harmonic. This results in the third harmonic first increasing and then decreasing in frequency when the plate dimension goes from square to long and narrow.

Our results indicate that the higher harmonics are not heard out individually very well. They could, however, still provide information to the listener. In earlier work we showed that listeners are capable of estimating the spectral density of a sound, that is, they could distinguish the number of components in a given bandwidth [9]. For spectral-density discrimination, we derived a weber fraction of 0.3. We can estimate the minimal difference of the plate height wide ratio that results in an audible difference in density of the sound. Considering the band from $F0$ to $9F0$, we see that there are 10 partials for a plate with $L_x/L_y = 1.05/0.95$ a plate would have at least 13 partials which is the case for plates with $L_x/L_y > 1.45/0.8$

Our results show that the listeners can reliably detect a missing first harmonic. We argued that this first harmonic provides the listener with information about the plate geometry. Viemeister and Fantini [10] tested two subjects for their threshold in frequency ratio discrimination. The subjects had to judge the larger interval out of two intervals, both having a different average frequency. They tested frequencies of 400 to 600 Hz. The reference frequency ratio r was 1.25. Two subjects had thresholds $\delta r/r$ 1.65 and 0.83 %. A plate with $F1/F0 = 1.25$ has $L_x/L_y = 2.45/0.74$ and if we add 2 percent, we find a discriminable plate geometry of $L_x/L_y = 2.34/0.74$. Therefore, this potential cue has the potential of resulting in very precise geometry judgements.

While the present results suggest that such cues are audible to listeners, future work needs to determine whether this information is indeed used by listeners in the judgment of plate attributes.

References

[1] Stefan Bleeck, Tim Ives, and Roy D. Patterson, *Aim-mat: The auditory image model in matlab*, Acta Acustica united with Acustica **90** (2004), 781–787.

[2] Søren Buus and Mary Florentine, *Psychometric functions for level discrimination*, J. Acoust. Soc. Am. **90** (1991), no. 3, 1371–1380.

[3] Wolfgang Ellermeier, *Detectability of increments and decrements in spectral profiles*, J. Acoust. Soc. Am. **99** (1996), no. 5, 3119–3125.

[4] Andrew J. Kunkler-Peck and M.T. Turvey, *Hearing shape*, Journal of Experimental Psychology: Human Perception and Performance **26** (2000), no. 1, 279–294.

[5] Stephen Lakatos, Stephen McAdams, and René Caussé, *The representation of auditory source characteristics: Simple geometric form*, Perception & Psychophysics **59** (1997), no. 8, 1180–1190.

[6] Brian C.J. Moore and Kengo Ohgushi, *Audibility of partials in inharmonic complex tones*, J. Acoust. Soc. Am. **93** (1993), no. 1, 452–461.

[7] R. Plomp, *The ear as a frequency analyzer*, J. Acoust. Soc. Am. **36** (1964), no. 9, 1628–1635.

[8] Thomas D. Rossing and Neville H. Fletcher, *Principles of vibration and sound*, Springer, 1995.

[9] Christophe N.J. Stoelinga and Robert A. Lutfi, *Spectral density discrimination*, ARO Midwinter Meeting, 2008.

[10] Neal F. Viemeister and Deborah A. Fantini, *Auditory processing of complex sounds*, ch. 5: Discrimination of Frequency Ratios, pp. 47–56, Lawrence Erlbaum Associates, 1987.