Bibliographical reference

Auditory Evaluation of Sounds Radiated from a Vibrating Plate Inside a Damped Cavity: Adjustment of the Frequency Resolution of Vibro-Acoustical Computing, A. Trollé, C. Marquis-Favre and N. Hamzaoui, Acta Acustica **vol. 98** (Number 3), 2012, pp. 441-450

DOI
https://doi.org/10.3813/AAA.918528

Acta Acustica united with Acustica
Vol. 98 (2012) 441 – 450

DOI 10.3813/AAA.918528

Auditory Evaluation of Sounds Radiated from a Vibrating Plate Inside a Damped Cavity: Adjustment of the Frequency Resolution of Vibro-Acoustical Computing

A. Trollé[1)], C. Marquis-Favre[1)], N. Hamzaoui[2)]
[1)] Université de Lyon, Lyon, F-69003, France, École Nationale des Travaux Publics de l'État (ENTPE), Départment Génie Civil et Bâtiment (DGCB) FRE 3237, 3 rue Maurice Audin, 69120 Vaulx-en-Velin, France. Catherine.MARQUIS@entpe.fr
[2)] Université de Lyon, Lyon, F-69003, France, Institut National des Sciences Appliquées de Lyon, Laboratoire Vibrations Acoustique, Bâtiment St. Exupéry, 25 bis avenue Jean Cappelle, 69621 Villeurbanne Cedex, France

Summary
The issue of adjusting vibro-acoustical computational cost is especially important when enhancing the sound quality of a mechanical structure using simulation. Simulation parameters must be specified so as to: i) reduce computational cost, and ii) satisfy perceptual constraints. For a simple plate-cavity system, the adjustment of computational cost is investigated in the framework of evaluating the influence of variable structural parameters (cavity absorption, thickness and tightening conditions of the plate) on the auditory perception of sounds radiated inside the cavity. By concentrating on such a simple system, we eventually aim at achieving some recommendations adaptable to different vibrating structures of our environment – in the areas of transport and building – made up of plane elements coupled with enclosures. Considering stationary sounds, this study deals with the adjustment of a major frequency computation parameter, i.e. frequency resolution. The specification of the frequency resolution for vibro-acoustical simulations has a great influence on the spectral envelope of the sounds synthesized afterwards. Here, we modify the spectral envelope of real sounds in order to reproduce the envelope of spectra of synthesized sounds that would be computed in a simulation using a given frequency resolution. The adjustment process pursues the following objective: determining a frequency resolution for which certain perceptual outcomes relative to modified real sounds remain similar to those concerning the original real sounds. These perceptual constraints are pertinent regarding the issue of adjusting vibro-acoustical computational cost. These constraints notably concern the merits of the stimuli and are translated into two adjustment criteria. The results of a first experiment, an auditory evaluation of partial corpora of modified stimuli, highlight that, for the plate-cavity system under consideration, 4 Hz may be an adjusted frequency resolution. The consequences of using this frequency resolution are then further assessed through a second experiment, an auditory evaluation of the complete corpus of modified stimuli. A wider range of perceptual results (perceptual spaces, preference spaces and merits) obtained for the modified sounds and the original sounds are compared. It is shown that both adjustment criteria are fulfilled for the adjusted frequency resolution equal to 4 Hz.

PACS no. 43.20.Rz, 43.20.Rj, 43.66.Lj, 43.40.At

1. Introduction

Investigations based on psychophysical methods are often pursued in different industrial sectors (e.g., automotive industry, electrical goods industry, etc.), as early as the design stage to improve the perceived quality of the sounds generated by the products (e.g., [1, 2, 3]).

In line with this approach, certain academic works have focused on assessing the influence of variable physical parameters of different simple vibro-acoustical systems on the auditory perception of the sounds radiated from them. The systems studied until now include the bar [4], a baffled single plate [5, 6, 7, 8, 9, 10], and a plate coupled with a parallelepipedal air cavity [11, 12]. The results of these works have turned out to be useful for drawing up recommendations capable of improving the sound quality of such vibro-acoustical systems from the design stage, as they form the basic elements of structures found in the fields of transport and building.

At the design stage, priority is often given to vibro-acoustical simulations in order to carry out a parametric study of the design of the system concerned. The issue

Received 22 November 2010,
accepted 29 March 2012.

of adjusting vibro-acoustical computational cost is becoming more complex whenever it is necessary to perform an additional sound quality study using synthesized sounds. Indeed, simulation parameters have to be defined to fulfil two somewhat conflicting goals simultaneously: i) reduce computational cost as much as possible, and ii) satisfy perceptual constraints. The latter can vary according to the approach to sound quality adopted by the researcher [13]. For example, perceptual constraints could consist in synthesizing artificial sounds as close as possible to real sounds when listening. This would require generating a refined time and frequency structure for the artificial sounds, which would then require allocating costly computational resources to sound synthesis. Several academic works have focused on these adjustment issues [6, 12, 14], by applying different strategies. In every case, the specification of simulation parameters must satisfy minimum perceptual constraints: formulating recommendations on the structural design of the system from the auditory evaluation of synthesized sounds that are faithful to those that would be formulated from the auditory evaluation of real sounds. This original strategy is applied in the present study for a simple vibro-acoustical system as a first step.

Our research aims to adjust vibro-acoustical computational cost in the framework of performing a sound quality evaluation for a plate-cavity system. By focusing on such a simple system, we eventually aim at achieving some recommendations adaptable to different vibrating structures of our environment – in the areas of transport and building – made up of plane elements coupled with enclosures.

We consider a stationary mechanical source with a pink-noise spectrum. In the case of stationary excitations, the vibro-acoustical response of a system is often computed in the frequency domain. Recourse to numerical finite element (FEM) and boundary element (BEM) methods is common; thus vibro-acoustical computational cost is intimately linked to two important computation parameters: maximum frequency of computation and frequency resolution. Specifying the maximum frequency determines the extent of the frequency content of synthesized sounds; specifying the frequency resolution influences their spectral envelope, in particular with a smoothing effect on spectral components for increasingly large values of the frequency resolution. Consequently we aim to determine appropriate values of both parameters for which the perceptual outcomes resulting from the auditory evaluation of synthesized sounds and real sounds remain analogous.

The first stage consisted in formulating perceptual outcomes from the auditory evaluation of real sounds [15]. To obtain a corpus of real sounds, an experiment on a plate-cavity system was carried out. The sound radiated from the vibrating plate inside the cavity was recorded for different structural configurations of the experimental system. These configurations differed in the cavity sound absorption properties, plate thickness and tightening conditions. The auditory evaluation, based on the method of paired comparisons, consisted in asking listeners to utter dissimilarity and preference judgments on pairs of stimuli. It was then possible to extract perceptual outcomes, i.e. information on the way structural variations influence auditory perception and on the structural configurations preferred by the listeners. Concerning the latter point, a preference ranking was built up from the merits of all the stimuli (merits are values on a linear scale of preference). It could be seen that the cavity sound absorption properties predominantly influenced the preference ranking: all the stimuli referring to a structural configuration with strong cavity sound absorption properties occupied the first places in the ranking, forming a leading group. It could also be shown that the plate tightening conditions moderately influenced the preference ranking: for weaker cavity sound absorption properties, all the stimuli referring to a structu ral configuration with weak tightening conditions occupied places in the middle of the ranking, forming an intermediate group, and all the stimuli referring to a structural configuration with stronger tightening conditions occupied the last places in the ranking, forming a queue group. Moreover, a simple linear regression model based on Zwicker's loudness was found to be sufficient to accurately predict the merits.

In the second stage, the maximum frequency was adjusted [15]. To do this, we used real sounds rather than synthesized sounds, in order to be free from other approximations due to modeling the system. The adjustment was carried out with the final aim of preserving the perceptual outcomes relative to the merits. This perceptual constraint was translated into two criteria: i) the leading group in the preference ranking had to remain unchanged, and ii) the metrics on which the preference model was based had to remain the same without significant deterioration of the goodness-of-fit of this model. The first criterion is motivated by the fact that it is important e.g. for industrialists to obtain the same best design solutions whatever the approximations in the modeling. The second criterion is justified as follows: in sound quality studies, one usually aims at finding out a good predictor of preference and, for the sake of consistency, the predictions based on this predictor should remain the same, whatever the approximations in the modeling. It was found that, for the plate-cavity system under consideration, a maximum frequency equal to 2500 Hz could fulfil these two criteria.

In order to further reduce vibro-acoustical computational cost, the third stage consisted in adjusting the frequency resolution Δf, by considering the conclusions of the previous adjustment of the maximum frequency f_{max} [15]. That is to say we worked using the real sounds lowpass-filtered at 2500 Hz, which corresponds to the adjusted maximum frequency f^*_{max} [15]. Thus all the stimuli presented during the listening tests carried out in the present study were real sounds lowpass-filtered at 2500 Hz. Afterwards, the spectral envelope of these sounds was modified in order to simulate that of synthesized sounds that would be computed with a given frequency resolution Δf. These modified stimuli are henceforth denoted by the expression *modified filtered stimuli*. The non modified and non lowpass-filtered stimuli are

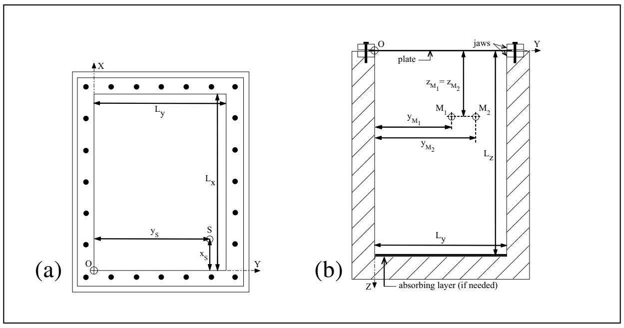


Figure 1. The experimental device. (a) Top view, (b) Cross-section view.

henceforth referred to as the *original stimuli*. Then, the adjustment process aimed to determine a frequency resolution value for which the perceptual outcomes resulting from the auditory evaluation of the modified filtered stimuli remained similar to those resulting from the auditory evaluation of the original stimuli[1], previously highlighted in [15]. As with the adjustment of the maximum frequency f_{max} [15], adjusting the frequency resolution Δf was carried out with the final aim of preserving the perceptual outcomes relative to the merits. When comparing the newly formulated perceptual outcomes with those relative to the original stimuli, the two same criteria had to be fulfilled in order to ensure their adequacy: i) the leading group in the preference ranking had to remain unchanged, and ii) the metrics on which the preference model was based had to remain the same without significant deterioration of the goodness-of-fit of the model.

To facilitate understanding, we first recall the main features of the experiment carried out in order to elaborate a corpus of real sounds (see section 2). Then, we present two experiments that were carried out to achieve the adjustment of the frequency resolution Δf (see sections 3 and 4, respectively). Finally, conclusions on main results presented in this paper are given.

2. Sound corpus

2.1. Apparatus

The experimental device is illustrated in Figure 1. The dimensions L_x, L_y and L_z of the the parallelepipedal air-cavity were 0.6, 0.5 and 0.7 m, respectively. Five sides of the cavity enclosure were made of concrete 0.1 m thick. The upper side was composed of a steel plate with a vibrating surface equal to $L_x \times L_y$. The rim of the plate was held between two thick steel jaws screwed into the concrete enclosure.

The 24 screws used were evenly spaced all along the jaws and uniformly tightened with a torque wrench. The plate was mechanically excited with a shaker placed at point S of coordinates $(x_S, y_S, z_S) = (0.10, 0.425, 0)$ m. The point of application of the force was selected so that a compromise between a high radiated acoustic power and a high number of excited structural modes could be found. The stationary excitation was a pink noise.

2.2. Acoustical measurements

Three structural parameters were modified: cavity sound absorption properties, plate thickness and tightening conditions. Three levels were assigned to each structural parameter (see Table I).

Nine configurations matching different combinations of parameter levels were designed using a Taguchi's L_9 table [16] (labels "A" to "I" in Table II). Two additional configurations (labels "J" and "K" in Table II) were added to obtain a more substantial number of sound samples.

For each configuration (A to K), the sound radiated from the vibrating plate was simultaneously recorded at two inner receiver points M_1 and M_2, with the respective coordinates $(x_{M_1}, y_{M_1}, z_{M_1}) = (0.22, 0.26, 0.145)$ m and $(x_{M_2}, y_{M_2}, z_{M_2}) = (0.22, 0.38, 0.145)$ m (Cf. Figure 1b). This experimental setting aimed at better representing the differences in acoustic field that could be met at high frequencies for a plate radiating into a cavity. The sound was recorded using two B&K 1/2" omnidirectional microphones. A second recording was performed for configurations A and H (this second recording aimed at assessing the perceptual effects of variability in acoustic measurement with respect to those of the structural parameter variations; it was shown in [15] that the former were negligible in comparison with the latter). All the recordings were performed at a sampling frequency equal to 44.1 kHz. Stereophonic samples four seconds in length were extracted from the sound sequences recorded at the two receiver points. Afterwards, all the sound samples were corrected to compensate for the transfer function of the headphones used later for sound reproduction. The sound samples associated with the different structural configurations are shown in Table II.

3. Experiment 1

This section presents a first experiment aimed at assessing the perceptual outcomes relative to the merits resulting from the modeling approximations, by using different frequency resolution values: $\Delta f_1 = 2$ Hz, $\Delta f_2 = 4$ Hz, $\Delta f_3 = 5$ Hz and $\Delta f_4 = 10$ Hz. These values were selected on the basis of the results of preliminary discrimination tests which showed that a frequency resolution equal to 1 Hz could not be discriminated from the baseline frequency resolution (Δf_0 equal to 0.25 Hz).

For this first experiment, it was considered pertinent to carry out an auditory evaluation of a partial corpus of representative stimuli in order to roughly estimate, for each Δf_i value ($i = 1, \ldots, 4$), how the modeling approximations affected the perceptual outcomes relative to the merits. Using a partial corpus of representative stimuli made it possible to reduce the length of subjective testing. From these Δf_i values, a single frequency resolution value Δf_p, to be used in Experiment 2, could then be selected. This value Δf_p may correspond to the coarsest frequency resolution for which the perceptual outcomes concerning the

---

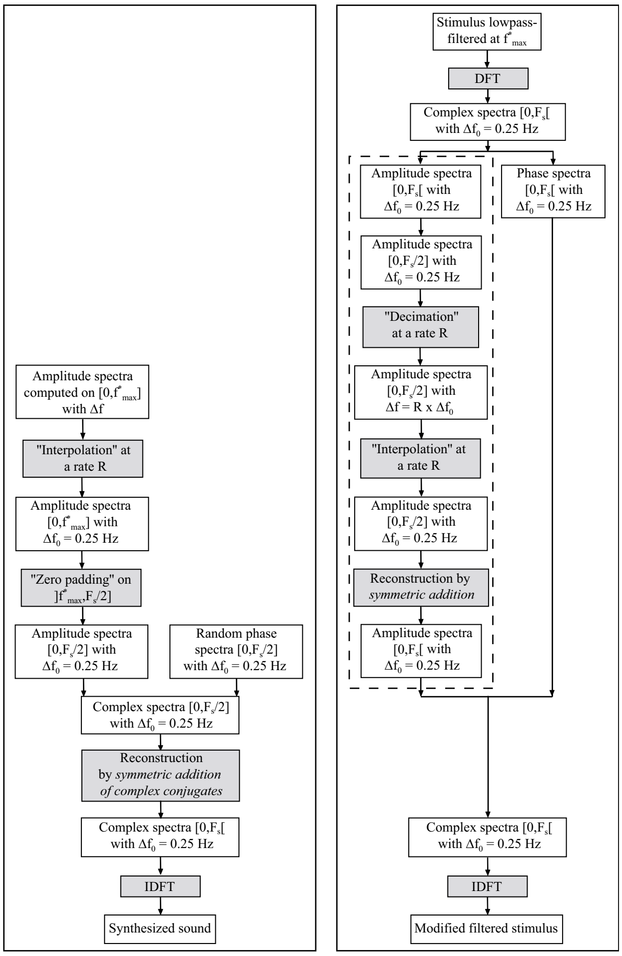
[1] NB: This approach does not involve modified filtered sounds perceived as identical to the original sounds.

Table I. Description of the levels of the structural parameters. α: plate thickness, β: cavity sound absorption properties, γ: plate tightening conditions, C: tightening torque.

Label	Level 1	Level 2	Level 3
α	$1.5{\cdot}10^{-3}$ m	$2{\cdot}10^{-3}$ m	$3{\cdot}10^{-3}$ m
β	weak (no absorption material)	medium (one carpet layer)	strong (one foam layer)
γ	weak ($C = 20$ N.m)	medium ($C = 50$ N.m)	strong ($C = 80$ N.m)

Table II. Plate-cavity configurations and associated sound samples.

Label	α	β	γ	Associated sound samples
A	1	1	1	A_1, A_2
B	1	2	2	B_1
C	1	3	3	C_1
D	2	3	1	D_1
E	2	1	2	E_1
F	2	2	3	F_1
G	3	2	1	G_1
H	3	3	2	H_1, H_2
I	3	1	3	I_1
J	1	1	3	J_1
K	3	3	3	K_1



(a) (b)

Figure 2. A parallel between the steps underlying (a) the synthesis of a stationary broadband noise from sound pressure amplitude spectra computed under simulation at the two receiver points up to maximum frequency f^*_{max} and with frequency resolution Δf, and (b) the generation of any modified filtered stimulus. Δf_0: frequency resolution corresponding to a stimulus length T equal to 4 s ($\Delta f_0 = 1/T$), R: "decimation"/"interpolation" rate ($R = \Delta f/\Delta f_0$), F_s: sampling frequency, $F_s/2$: Nyquist frequency, DFT: Discrete Fourier Transform, $IDFT$: Inverse Discrete Fourier Transform.

modified filtered stimuli remain a good approximation of those concerning the original stimuli.

3.1. Methods

3.1.1. Participants

Thirty-four trained listeners, 17 women and 17 men, aged from 20 to 55 years, participated in the listening test. Fifteen of the listeners who took part in Experiment 1 (around 44% of them) had evaluated the original stimuli [15]. All the listeners were students and members of the academic staff of ENTPE. They were remunerated for their participation.

3.1.2. Stimuli

The partial corpus consisted of 6 stimuli (against 13 stimuli for the complete corpus, see last column of Table II). The number of pairs was thus reduced to 15 (from 78 for the complete set of stimuli); the time necessary for the auditory evaluation of a partial corpus of stimuli was therefore quite short. In addition, in order to legitimately assess the adequacy of the perceptual outcomes relative to modified filtered stimuli with those relative to the original stimuli, the 6 stimuli selected within the complete corpus had to be sufficiently representative of the preference ranking of the 13 original stimuli, drawn up from preference judgments obtained through the method of paired comparisons [15]. In particular, the selection had to be performed so that any break-up of the leading group – due to substantial approximations – could be observed; the break-up of the leading group would mean that the structural configurations with strong cavity sound absorption properties would not all appear anymore at the first places of the preference ranking (see introduction). The following stimuli were selected: i) 4 stimuli in the leading group, i.e. D_1 (in first position), H_1, H_2 and K_1 (which have the last three positions), ii) 1 stimulus in the intermediate group, i.e. A_1 (in intermediate position), and iii) 1 stimulus in the queue group, i.e. I_1 (in last position).

For each Δf_i value, the 6 modified filtered stimuli were generated using a process developed by referring to the method [17, 18] used to synthesize a stationary broad-

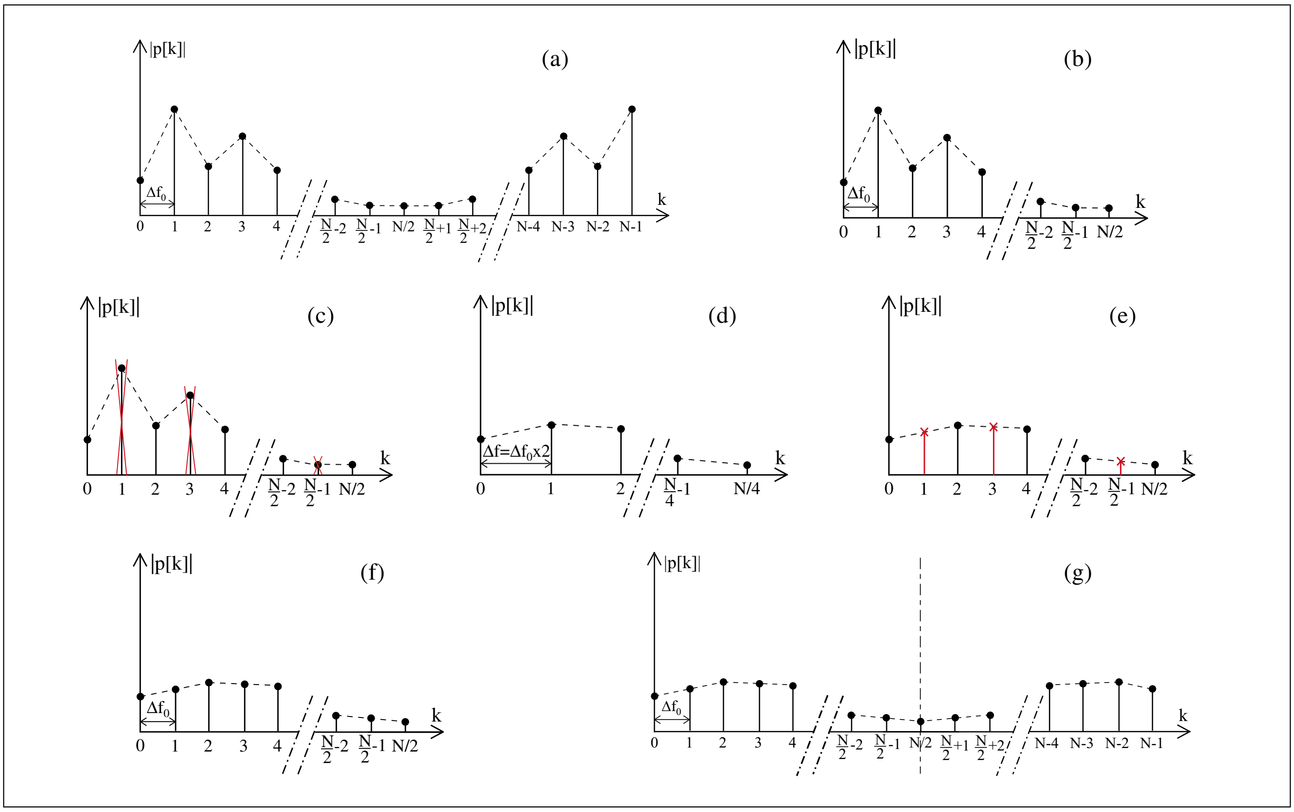


Figure 3. Illustration of several steps underlying the modification of the envelope of an amplitude spectrum of a stimulus lowpass-filtered at f^*_{max} (indicated by a dashed line in Figure 1b). N is the total number of samples ($N = F_s \times T$, T being the stimulus length in seconds); R, the "decimation"/"interpolation" rate, here is equal to 2. The envelope of an amplitude spectrum of a synthesized stimulus computed with a frequency resolution $\Delta f = 0.5$ Hz is thus simulated. (a) Amplitude spectrum $[0, F_s[$ with $\Delta f_0 = 0.25$ Hz, (b) Amplitude spectrum $[0, F_s/2]$ with $\Delta f_0 = 0.25$ Hz, (c) "Decimation" of the amplitude spectrum at a rate $R = 2$, (d) Amplitude spectrum $[0, F_s/2]$ with $\Delta f = \Delta f_0 \times 2 = 0.5$ Hz, (e) "Interpolation" of the amplitude spectrum at a rate $R = 2$, (f) Amplitude spectrum $[0, F_s/2]$ with $\Delta f_0 = 0.25$ Hz, (g) Reconstruction of the amplitude spectrum on $[0, F_s[$ with $\Delta f_0 = 0.25$ Hz by symmetric addition (i.e. a symmetry of the amplitude values in $]0, F_s/2[$ in relation to the vertical axis passing through abscissa $F_s/2$ is performed).

band noise from sound pressure amplitude spectra computed at the two receiver points inside the cavity up to maximum frequency f^*_{max} and with frequency resolution Δf. The different steps underlying this sound synthesis are summed up in Figure 2a. The different steps underlying the generation of any modified filtered stimulus – designed in relation to the steps of the sound synthesis – are presented in parallel in Figure 2b. In these processes, $\Delta f_0 = 0.25$ Hz is the baseline fine frequency resolution of the spectra of a stereophonic stimulus with a length $T = 4$ s (i.e. $\Delta f_0 = \frac{1}{T}$). Further details on the processing operations performed in each step can be found in [19]. Figure 3 illustrates some of the steps used to modify the envelope of an amplitude spectrum of a stimulus lowpass-filtered at f^*_{max} (indicated by a dashed line in Figure 2b), for a "decimation/interpolation" rate R equal to 2 (i.e. $\Delta f = 0.5$ Hz).

3.1.3. Apparatus

The listening test took place in a semi-anechoic room. Sound reproduction was performed using open Sennheiser HD600 headphones and a high quality Lynx ONE sound card. The mean reproduction levels for the four corpora of stimuli ($\Delta f_1 = 2$ Hz, $\Delta f_2 = 4$ Hz, $\Delta f_3 = 5$ Hz and $\Delta f_4 = 10$ Hz) were respectively equal to 71.8, 73.2, 72.5 and 72.1 dB(A) (respective ranges: 68.8 – 76.2 dB(A), 71.3 – 76 dB(A), 69.6 – 73.7 dB(A) and 69 – 73.3 dB(A)). The participants' answers were collected using a computer interface.

3.1.4. Design and procedure

For each corpus, the stimuli were pairwise presented to listeners (see [6] for details). The pairs were presented in a different random order for each listener. Within a pair, both stimuli were presented at random as "sound 1" and "sound 2". For each pair, listeners were asked to evaluate the dissimilarity between the stimuli on a seven-point numeric scale, ranging from "0" (labelled "very similar") to "6" (labelled "very different"). They were told to consider the scale as a metric one. For each pair, listeners were also asked to simultaneously indicate which stimulus they preferred, with a dichotomous choice. The listeners could listen to stimuli as often as they wanted. Preference judgments were linked in the process with dissimilarity judgments so that, if listeners had rated two stimuli "very similar", they did not need to choose which sound they preferred [6]. The 4 partial corpora of stimuli (i.e. one corpus per Δf_i value) were evaluated during a single listening test lasting less than 30 minutes.

3.1.5. Data analysis

For each Δf_i value, the data analysis aimed at drawing up a preference ranking by sorting the merits of the stimuli in ascending order.

The merits of the stimuli were computed by applying *Case V* of Thurstone's law of comparative judgment [20] to the preference data. 95% confidence intervals were constructed for the merits by using the bootstrap technique with 250 replications [6].

To check whether two stimuli within a particular preference ranking had significantly different merits (or not), we performed a statistical test using the bootstrap replications. For each bootstrap solution, we computed the difference between the merits of both stimuli. Afterwards, a 95% confidence interval for the difference between both merits was computed. If this confidence interval includes zero, one can state that the merits of both stimuli are not significantly different from each other, i.e. that both stimuli obtain ranks which are not significantly different from each other. Conversely, if this confidence interval does not include zero, one can state that the merits are significantly different from each other, i.e. that both stimuli obtain ranks which are significantly different from each other. Such statistical tests were carried out for examining the differences between merits within the preference ranking of the original stimuli and within a given preference ranking of modified filtered stimuli. Afterwards, a comparison of the stimulus ranks across the preference rankings of original stimuli and modified filtered stimuli was carried out. By using the preference ranking, it was possible to highlight the structural configuration(s) leading to an improvement of the quality of the sounds radiated inside the cavity. Finally, a preference model which allows the measured merits to be accurately predicted was built; to do this, a linear regression analysis was performed between the measured merits and the values of relevant metrics.

3.2. Results

For each Δf_i value, the merits measured for the 6 modified filtered stimuli, sorted in ascending order, were confronted with those previously measured for the original stimuli [15] (Cf. Figure 4).

For each Δf_i value, the ranks of the modified filtered stimuli are well correlated with those of the original stimuli ($0.89 <$ Spearman's $\rho < 0.94$, $p \leq 0.033$). Moreover, according to the results of the statistical tests performed (see section 3.1.5), it can be stated that the four preference rankings of modified filtered stimuli show no significant rank inversion (notably between stimulus K_1 and stimuli A_1 and I_1) that would be synonymous with a break-up of the leading group. Thus, the first criterion of adequacy (i.e. preservation of the leading group in the preference ranking) is fulfilled whatever the Δf_i value.

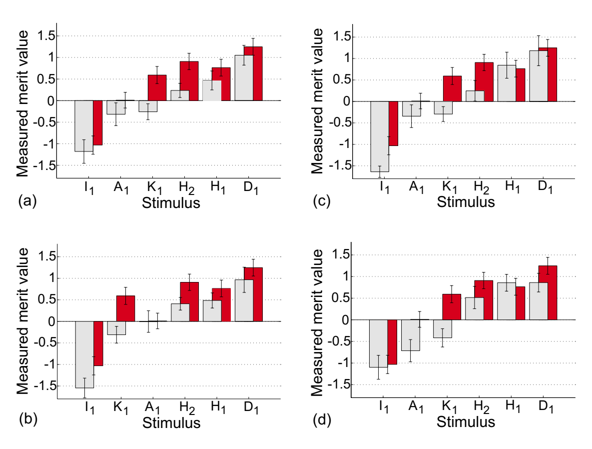


Figure 4. Comparison of the merits of the modified filtered stimuli with the merits of the original stimuli, for the different Δf_i values ($i = 1$ to 4). In light gray: merits measured for the modified filtered stimuli. In dark gray/red: merits measured for the original stimuli. I: 95% confidence intervals. (a) $\Delta f_1 = 2$ Hz, (d) $\Delta f_2 = 4$ Hz, (c) $\Delta f_3 = 5$ Hz, (d) $\Delta f_4 = 10$ Hz.

Table III. Different statistics for the preference models based on Zwicker's loudness built for the different Δf_i values. †reference case (original stimuli), $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Δf_i	r	r^2	$\hat{\sigma}_e^2$	$F_{4,11}$
0.25 Hz†	-0.96***	0.93	0.080	–
2 Hz	-0.95**	0.90	0.073	1.21
4 Hz	-0.96**	0.93	0.067	1.12
5 Hz	-0.83*	0.69	0.391	6.53*
10 Hz	-0.79*	0.63	0.337	5.63*

Table III presents different statistics for the preference models based on Zwicker's loudness built for the different Δf_i values: a measure of strength of association between Zwicker's loudness values and measured merits with the Pearson's linear correlation coefficient r, a measure of goodness-of-fit of the model with the proportion of explained variance r^2, and a measure of lack-of-fit with the residual variance $\hat{\sigma}_e^2$. As a reminder, these statistics are also shown for the preference model relative to the original stimuli.

Regarding the second criterion of adequacy (i.e. non-significant deterioration of the goodness-of-fit of the preference model based on Zwicker's loudness), a two-tailed F-test was performed to test for the equality between the residual variance for each preference model relative to modified filtered stimuli and the residual variance for the preference model relative to the original stimuli.

The results, i.e. the F-statistics (with 4 and 11 degrees of freedom[2]), are shown in the last column of Table III. For $\Delta f_1 = 2$ Hz and $\Delta f_2 = 4$ Hz, the result of the F-test is not significant, i.e. the goodness-of-fit of the preference model relative to the modified filtered stimuli is not significantly different from that of the preference model relative to the

---

[2] The number of degrees of freedom for each residual variance is equal to the number of stimuli to which the linear regression analysis applied minus two.

original stimuli. For $\Delta f_3 = 5\,\text{Hz}$ and $\Delta f_4 = 10\,\text{Hz}$, the result of the F-test is significant, i.e. the lack-of-fit (alternatively, goodness-of-fit) of the preference model relative to the modified filtered stimuli is significantly higher (alternatively, lower) than that of the preference model relative to the original stimuli. Thus, for the tested Δf values higher than 4 Hz, the second criterion is not fulfilled.

Thus, it appears that $\Delta f_2 = 4\,\text{Hz}$, i.e. the frequency resolution value for which both criteria were found to be fulfilled, may be the adjusted value Δf_p searched.

4. Experiment 2

This section presents a second experiment aimed at evaluating the impact of the use of the adjusted frequency resolution Δf_p =4 Hz on perceptual outcomes: perceptual spaces, preference spaces and merits. To this end, an auditory evaluation of the complete corpus of modified filtered stimuli was carried out in order to stringently assess how the modeling approximations due to the use of $\Delta f_p = 4\,\text{Hz}$ affected such perceptual outcomes.

4.1. Methods

4.1.1. Participants

Thirty-four trained listeners, 12 women and 22 men, aged from 20 to 65 years, participated in the listening test. Twenty-one of them (i.e. around 62%) had already participated in experiment 1. All the listeners were students and members of the academic staff of ENTPE. They were remunerated for their participation.

4.1.2. Stimuli

The 13 modified filtered stimuli (A_1 to K_1, A_2 and H_2) composing the complete corpus were considered and generated using the process described in section 3.1.2 with $\Delta f_p = 4\,\text{Hz}$.

4.1.3. Apparatus

The listening test took place in a semi-anechoic room. Sound reproduction was performed using Sennheiser HD600 open headphones and a high quality Lynx ONE sound card; the mean reproduction level was equal to 72 dB(A) (range = 64.7 – 76.2 dB(A)). The participants' answers were collected using a computer interface.

4.1.4. Design and procedure

The design and procedure for experiment 2 were identical to those described in section 3.1.4.

4.1.5. Data analysis

Dissimilarity and preference data were both analyzed.

Concerning the dissimilarity data, a preliminary check [15] that the listeners well considered the scale as a metric one was carried out using the method of successive intervals. The dissimilarity data were then processed with INDSCAL [21] – a metric MultiDimensional Scaling (MDS) model – in order to build a perceptual space of the stimuli, displaying the perceptual distances existing between them. 95% confidence ellipsoids were constructed for the stimulus points, using the bootstrap technique with 250 replications [22]. The dimensions of the INDSCAL space refer to salient auditory attributes whose variation probably led the listeners to differentiate between the stimuli. For each perceptual dimension, meticulous listening to the stimuli sorted according to their respective coordinates made it possible to identify the auditory attribute it referred to [23]. An analysis of the correlations between the values of acoustical and psychoacoustical metrics (computed by using dB-Sonic© software) and the coordinates of the stimuli along the perceptual dimensions was also performed to confirm the previous identification of auditory attributes.

Two complementary analyses of the preference data were planned. The dichotomous judgments were first processed with the MDS program MDPREF [24]. This program provides a preference space that displays stimulus points and unit subject vectors passing through the origin, located so that the projections of the points on the vectors are in maximum agreement with the initial judgments of the listeners. A vector endpoint represents the point of maximum preference of a subject. The direction in which a vector points indicates how a subject combines salient auditory attributes of the stimuli to utter her/his preference judgments [25]. In order to identify these attributes, acoustical and psychoacoustical metrics were fitted as unnormalized vectors into the preference space by a PREFMAP procedure [26].

Then, just like for Experiment 1, the preference data were processed using *Case V* of Thurstone's law of comparative judgment [20] in order to compute the merits of all the stimuli (see section 3.1.5). A preference model which allows the measured merits to be accurately predicted was built as well (see section 3.1.5).

4.2. Results

This section proposes a comparison between the perceptual results respectively obtained for the modified filtered and original stimuli. It starts with a comparison between perceptual spaces, continues with a comparison between preference spaces, and ends with a comparison between merits which are of great interest given the adjustment criteria chosen.

4.2.1. Perceptual spaces

The dissimilarity data were scaled in a two-dimensional perceptual space. This optimal dimensionality was selected according to the scree test method [27]. This method was applied to the plot of the cumulative proportion of explained variance against the number of dimensions. The two-dimensional model accounted for 49% of the variance in the individuals' scalar products matrices [21]. Figure 5 provides a comparison between the perceptual space of the modified filtered stimuli and that previously drawn up for the original stimuli [15].

The adequacy between both perceptual spaces turns out to be satisfactory ($r = 0.911$, $p < .001$ between their first dimensions; $r = 0.992$, $p < .001$ between their second dimensions). Moreover, it must be emphasized that there are only two stimuli, A_1 and G_1, for which the positions in both spaces are different from each other, in view of the non-overlapping of the 95% confidence ellipses.

As with the perceptual space of the original stimuli, the first dimension of the perceptual space of the modified filtered stimuli (accounting for 25% of the variance in the individuals' scalar products matrices) refers to loudness ($r = -0.828$, $p = .001$ between coordinates and values of Zwicker's loudness[3] [28]); the second dimension (accounting for 24% of the variance in the individuals' scalar products matrices) relates to the spectral balance ($r = 0.905$, $p < .001$ between coordinates and values of the spectral centroid[4] [29]).

4.2.2. Preference spaces

Figure 6 presents the factorial plan 1-2 of the three-dimensional preference space of the modified filtered stimuli, including stimulus points, unit subject vectors and metrics vectors. This factorial plan accounted for 91% of the total variance. The optimal dimensionality was selected according to the scree test method [27]. The test was applied to the plot of the eigenvalues against the number of dimensions. Several preference space data relative to the original stimuli are also displayed in Figure 6: the stimulus points, the endpoints of the extreme subject vectors, the endpoint of the average subject vector and the metrics vectors.

The preference space of the modified filtered stimuli was first made to fit that of the original stimuli by using Procrustes analysis [30] (see also [15] for further details about the practical analysis steps).

The adequacy between both configurations of stimulus points turns out to be satisfactory (after fit, $r = 0.956$, $p < .001$ between the coordinates along dimension 1 of the modified filtered stimuli and those of the original stimuli; $r = 0.827$, $p = .001$ between their respective coordinates along dimension 2). However, quite large disparities between the positions of the original and modified filtered stimuli can be observed (C_1, H_1, I_1 and J_1).

In addition, the beam of subject vectors relative to the modified filtered stimuli diverges slightly from that relative to the original stimuli, regarding the mean direction. In terms of dispersion, both beams remain rather similar; good consensus between the listeners' preferences still predominates.

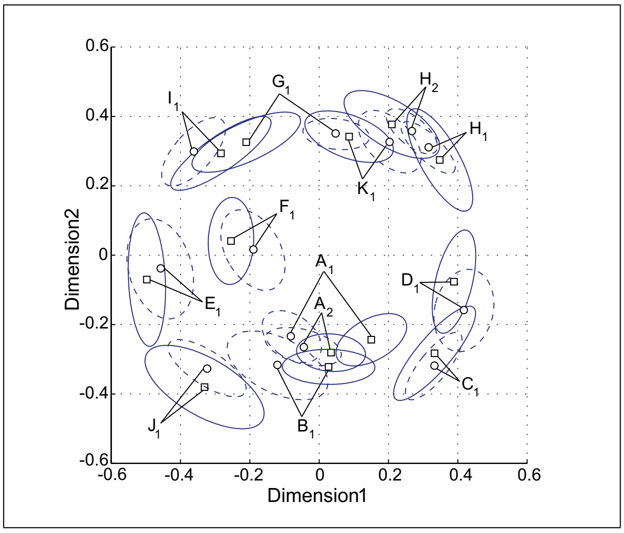


Figure 5. 2-D perceptual spaces of the stimuli, with their respective 95% confidence ellipses. □: modified filtered stimuli, ∘: original stimuli, in solid line: 95% confidence ellipses for the modified filtered stimuli, in dashed line: 95% confidence ellipses for the original stimuli.

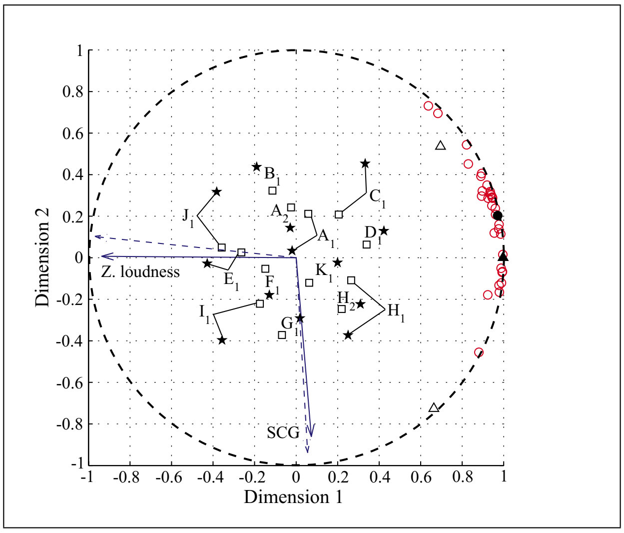


Figure 6. Factorial plan 1-2 of the preference space. □: stimulus points for the modified filtered stimuli, ∘: endpoints of the subject vectors relative to the modified filtered stimuli, •: endpoint of the average subject vector relative to the modified filtered stimuli, →: metrics vectors relative to the modified filtered stimuli, ⋆: stimulus points for the original stimuli, △: endpoints of the extreme subject vectors relative to the original stimuli, ▲: endpoint of the average subject vector relative to the original stimuli, - - →: metrics vectors relative to the original stimuli.

Furthermore, as with the original stimuli, loudness and spectral balance appear to be the salient auditory attributes whose variation probably guided the listeners in their preference choices: i) the goodness-of-fit of the metrics vectors relative to the modified filtered stimuli, i.e. Zwicker's loudness and spectral centroid, remains good (respectively $r = 0.941$, $p < .001$ and $r = 0.881$, $p < .001$), when compared with that of the metrics vectors relative to the original stimuli (respectively $r = 0.991$, $p < .001$ and $r = 0.943$, $p < .001$) and ii) each metrics vector relative to the modified filtered stimuli remains quite close to the respective one relative to the original stimuli.

4.2.3. Merits

Figure 7 shows the measured merits of the modified filtered stimuli (in gray), sorted in ascending order; the measured merits of the original stimuli are also reported (in dark gray/red). The ranks of the modified filtered stim-

---

[3] The values of Zwicker's loudness vary between 16.7 and 31.9 sones.

[4] The values of the spectral centroid vary between 3 and 8 Barks.

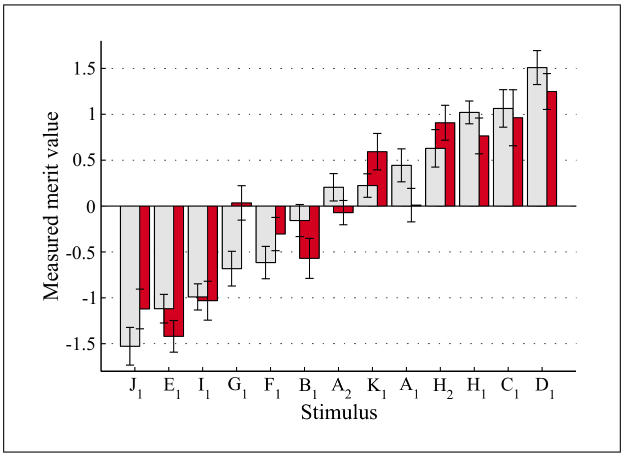


Figure 7. Comparison of the measured merits of the modified filtered stimuli with those of the original stimuli. In light gray: measured merits of the modified filtered stimuli. In dark gray/red: measured merits of the original stimuli. I: 95% confidence intervals.

uli are well correlated with those of the original stimuli (Spearman's $\rho = 0.92$, $p < 0.001$).

Within the preference ranking of the modified filtered stimuli, an inversion of rank between stimuli A_1 (intermediate group) and K_1 (leading group) can be seen. However, this inversion turns out to be non-significant. Indeed, a statistical test (see section 3.1.5) shows that the difference between their measured merits is not significant. Thus, the leading group including all the structural configurations with strong cavity sound absorption properties is kept.

The same statistics as those showed in Table III were computed for the preference model with $\Delta f_p = 4$ Hz: $r = -0.92$ ($p < 0.001$), $r^2 = 0.84$ and $\hat{\sigma}_e^2 = 0.150$. Regarding the second criterion of adequacy (i.e. non-significant deterioration of the goodness-of-fit of the preference model based on Zwicker's loudness), a two-tailed F-test was also performed to test for the equality between the residual variance for the preference model relative to the modified filtered stimuli and the residual variance for the preference model relative to the original stimuli. The result of the F-test is not significant ($F_{4,11} = 2.51$, $p > 0.05$), i.e. the goodness-of-fit of the preference model relative to the modified filtered stimuli is not significantly different from that of the preference model relative to the original stimuli.

4.2.4. Discussion

From the comparison between the perceptual and preference spaces, respectively relative to the original stimuli and to the modified filtered stimuli, it turned out that the modification of the spectral envelope of the real sounds related to the use of a frequency resolution equal to 4 Hz, in addition to their lowpass filtering at 2500 Hz, moderately affected the mechanisms of differentiation and of preference evaluation, slightly more than they were affected by lowpass filtering alone at 2500 Hz [15].

The comparison between the perceptual outcomes relative to the merits showed that both criteria of adequacy (conservation of the leading group in the preference ranking and non-significant deterioration of the goodness-of-fit of the preference model based on Zwicker's loudness) were fulfilled.

Afterwards, analysis of the signals, i.e. of the original stimuli (i.e. $\Delta f_0 = 0.25$ Hz) and of the modified filtered stimuli (i.e. $\Delta f_p = 4$ Hz), was carried out in order to try to explain these results. It notably showed that, concerning the loudness, passing from a frequency resolution equal to 0.25 Hz to one equal to 4 Hz led to a translation in the loudness values for all the sounds. This could explain why the first dimension of the preference space relative to the modified filtered stimuli remained well correlated with Zwicker's loudness. This could also explain why the preference ranking was not modified, and finally, why the goodness-of-fit of the preference model was not significantly deteriorated. Passing from a frequency resolution equal to 0.25 Hz to one equal to 4 Hz led to a modification of the spectral envelop (and thus the timbre) of the sounds, and thus led to a modification of the spectral centroid of the sounds. But no simple transformation of the values of this metrics was noticed.

5. Conclusions

This work dealt with the adjustment of vibro-acoustical computational cost in the framework of a sound quality evaluation of a simple plate-cavity system. Considering stationary sounds, and in line with the previous adjustment of the maximum frequency f_{max} of their spectra, the aim was to adjust the frequency resolution Δf in order to further reduce computational cost, while still satisfying certain perceptual constraints.

A process was proposed to modify the envelope of amplitude spectra of real sounds lowpass-filtered at the adjusted maximum frequency $f^*_{max} = 2500$ Hz so as to simulate the envelope of amplitude spectra of synthesized sounds that would be computed under simulation up to f^*_{max} and with frequency resolution Δf. The results of a first listening test led us to focus the adjustment process on a specific value of Δf: 4 Hz. An additional auditory evaluation was then carried out. This evaluation was identical in every respect to that carried out to establish the perceptual outcomes relative to the original stimuli [15]. Its goal was to investigate whether the highlighted frequency resolution could lead to fulfilling certain perceptual constraints. These constraints consisted in establishing perceptual outcomes resulting from the auditory evaluation of the modified filtered stimuli close to those resulting from the auditory evaluation of the original sounds (non lowpass-filtered and non modified). To assess their closeness, we wanted to fulfil both the criteria proposed previously in the framework of the adjustment of f_{max} [15]: i) conservation of the leading group in the preference ranking, and ii) non-significant deterioration of the goodness-of-fit of the preference model based on Zwicker's loudness. Both criteria were fulfilled. Thus, it can be claimed that an adjusted value has been obtained. A Δf value equal to 4 Hz provides a genuine advantage regarding computational cost; it allows reducing computation time by a factor of 8 in comparison to the Δf value equal to 0.25 Hz that would have been considered if this adjustment process had not been performed.

This adjusted frequency resolution value is valid for the plate-cavity system studied, a simple system used as a first study case in order to obtain trends and to understand them. Further steps have to be carried out by considering stationary sounds radiated from more complicated plate-cavity systems such as those encountered in the domain of transport, with various absorbing conditions, coupled plates, stiffeners, mounting conditions, etc. Different types of excitation spectra must also be considered to better assess the adjusted frequency computation parameter in link with the excitation content.

Acknowledgements

The authors are very grateful to the reviewers for their helpful comments on the earlier version of this paper.

References

[1] C. V. Beidl, W. Stücklschwaiger: Application of the AVL-annoyance index for engine noise quality development. Acta Acustica united with Acustica **83** (1997) 789–95.

[2] S. K. Lee, J. Gu, J. Kim, E. Y. Lee, J. Y. Yang: Development on the sound index for the evaluation of sound quality of a refrigerator using artificial neural network. Proceedings of the 33rd International Congress and Exposition on Noise Control Engineering, Prague, Czech Republic, 2004.

[3] U. Widmann: Three application examples for sound quality design using psychoacoustic tools. Acta Acustica united with Acustica **83** (1997) 819–26.

[4] S. McAdams, A. Chaigne, V. Roussarie: The psychomechanics of simulated sound sources: Material properties of impacted bars. Journal of the Acoustical Society of America **115** (2004) 1306–1320.

[5] G. Canévet, D. Habault, S. Meunier, F. Demirdjian: Auditory perception of sounds radiated by a fluid-loaded vibrating plate excited by a transient point force. Acta Acustica united with Acustica **90** (2004) 181–93.

[6] F. Faure, C. Marquis-Favre: Perceptual assessment of the influence of structural parameters for a radiating plate. Acta Acustica united with Acustica **91** (2005) 77–90.

[7] C. Marquis-Favre, J. Faure: Auditory evaluation of sounds radiated from a vibrating plate with various viscoelastic boundary conditions. Acta Acustica united with Acustica **94** (2008) 419–432.

[8] N. Hamzaoui, C. Sandier, E. Parizet, P. Wetta, C. Besseyrias: Subjective assessments of the acoustic radiation from steel structures: some effects of a few parametric variations. Proceedings of Forum Acusticum, Sevilla, Spain, 2002.

[9] S. McAdams, V. Roussarie, A. Chaigne, B. L. Giordano: The psychomechanics of simulated sound sources: Material properties of impacted thin plates. Journal of Acoustical Society of America **128** (2010) 1401–1413.

[10] S. Meunier, D. Habault, G. Canévet: Auditory evaluation of sound signals radiated by a vibrating surface. Journal of sound and Vibration **247** (2001) 897–915.

[11] C. Marquis-Favre, J. Faure, N. Hamzaoui: Auditory evaluation of sounds generated in a cavity covered with absorbing layers. Proceedings of the 12th International Congress on Sound and Vibration, Lisboa, Portugal, 2005.

[12] E. Guibert, D. Habault, F. Poisson, P. E. Gautier: Psychomécanique d'un système plaque/cavité: Application à une voiture de TGV. Actes du CFA '06, Tours, France, 2006.

[13] G. Lemaitre, P. Susini, S. Winsberg, S. McAdams: The sound quality of car horns: A psychoacoustical study of timbre. Acta Acustica united with Acustica **93** (2007) 457–68.

[14] F. Demirdjian, S. Meunier, D. Habault, G. Canévet: A comparative study of recorded and computed sounds radiated by vibrating plates. Proceedings of Forum Acusticum, Budapest, Hungary, 2005.

[15] A. Trollé, C. Marquis-Favre, N. Hamzaoui: Auditory evaluation of sounds radiated from a vibrating plate inside a damped cavity. Acta Acustica united with Acustica **95** (2009) 343–355.

[16] M. Pillet: Les plans d'expériences par la méthode Taguchi. Les Éditions d'Organisation, Paris, 1997.

[17] F. Faure: Influence des paramètres d'une plaque rayonnante sur la perception sonore. PhD Thesis, Institut National des Sciences Appliquées de Lyon, 2003.

[18] N. Hamzaoui, J. L. Guyader, H. Shiraiwa, A. Chaigne, N. Topalovic, B. Smith, S. McAdams: Prévision du bruit rayonné par les structures en vue d'une évaluation perceptive. Tech. Report, Ministère de l'aménagement du territoire et de l'environnement, Direction de l'évaluation environnementale et des études économiques, 1999.

[19] A. Trollé: Évaluation auditive de sons rayonnés par une plaque vibrante à l'intérieur d'une cavité amortie: ajustement des efforts de calcul vibro-acoustique. PhD Thesis, Institut National des Sciences Appliquées de Lyon. http://docinsa.insa-lyon.fr/these/pont.php?id=trolle, 2009.

[20] L. L. Thurstone: A law of comparative judgment. Psychological Review **34** (1927) 273–286.

[21] J. D. Carroll, J. J. Chang: Analysis of individual differences in multidimensional scaling via an n-way generalization of "eckart-young" decomposition. Psychometrika **35** (1970) 283–319.

[22] S. L. Weinberg, J. D. Carroll, H. S. Cohen: Confidence regions for INDSCAL using the Jacknife and Bootstrap techniques. Psychometrika **49** (1984) 475–491.

[23] P. Susini, S. McAdams, S. Winsberg: A multidimensional technique for sound quality assessment. Acta Acustica united with Acustica **85** (1999) 649–54.

[24] J. J. Chang, J. D. Carroll: How to use MDPREF: A computer program for multidimensional analysis of preference data. Tech. Report, Bell Telephone Laboratories, 1969.

[25] A. P. M. Coxon: The user's guide to multidimensional scaling. Heinemann Educational Books, London, 1982.

[26] J. J. Chang, J. D. Carroll: How to use PREFMAP and PREFMAP-2: Program which relate preference data to multidimensional scaling solutions. Tech. Report, Bell Telephone Laboratories, 1972.

[27] R. B. Cattell: The scree test for the number of factors. Multivariate Behavioral Research **1** (1966) 245–276.

[28] E. Zwicker, H. Fastl: Psychoacoustics, facts and models (second updated edition). Springer Verlag, Berlin, Heidelberg, 1999.

[29] D. J. Freed: Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events. Journal of Acoustical Society of America **87** (1990) 311–322.

[30] T. F. Cox, M. A. A. Cox: Multidimensional scaling. Chapman & Hall/CRC, Boca Raton, 2001.