



HAL
open science

A Simplified and Controllable Model of Mode Coupling for Addressing Nonlinear Phenomena in Sound Synthesis Processes.

Samuel Poirot, Stefan Bilbao, Richard Kronland-Martinet

► **To cite this version:**

Samuel Poirot, Stefan Bilbao, Richard Kronland-Martinet. A Simplified and Controllable Model of Mode Coupling for Addressing Nonlinear Phenomena in Sound Synthesis Processes.. 2023. hal-04355480

HAL Id: hal-04355480

<https://hal.science/hal-04355480>

Preprint submitted on 20 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Simplified and Controllable Model of Mode Coupling for Addressing Nonlinear Phenomena in Sound Synthesis Processes.

Samuel Poirot^{1*}, Stefan Bilbao² and Richard Kronland-Martinet¹

^{1*}Aix-Marseille Univ, PRISM, CNRS, 31 chem. Joseph Aiguier,
Marseille, 13009, France.

²Acoustic and Audio Group, University of Edinburgh, 12 Nicolson
Square, Edinburgh, EH8 9DF, United Kingdom.

*Corresponding author(s). E-mail(s): poirot@prism.cnrs.fr;
Contributing authors: s.bilbao@ed.ac.uk; kronland@prism.cnrs.fr;

Abstract

This paper introduces a simplified and controllable model for mode coupling in the context of modal synthesis. The model employs efficient coupled filters for sound synthesis purposes, intended to emulate the generation of sounds radiated by sources under strongly nonlinear conditions. Such filters generate tonal components in an interdependent way, and are intended to emulate realistic perceptually salient effects in musical instruments in an efficient manner. The control of energy transfer between the filters is realized through a coupling matrix. The generation of prototypical sounds corresponding to nonlinear sources with the filter bank is presented. In particular, examples are proposed to generate sounds corresponding to impacts on thin structures and to the perturbation of the vibration of objects when it collides with another object. The sound examples presented in the paper and available for listening on the accompanying site illustrate that a simple control of the input parameters allows the generation of sounds whose evocation is coherent, and that the addition of random processes yields a significant improvement to the realism of the generated sounds.

Keywords: modal synthesis, nonlinear, procedural audio

1 Introduction

Modal synthesis operates according to the decomposition of the complex dynamic behavior of a vibrating object into contributions from modes, each oscillating independently at a single frequency. This approach, applicable to linear and time-invariant systems, is widely used and forms the basis for various physical modelling synthesis software packages [1] [2] and is closely related to sound synthesis methodologies employing filter banks [3] [4] [5].

For vibrating objects incorporating nonlinear effects, the modal interpretation must be generalized to include energy transfer between different modes and other effects such as, e.g., frequency shifting of modes over time. It may cause the delayed and sustained appearance of tonal components that cannot be generated by a linear model. This complex phenomenon, widely studied for the typical case of thin plates and shells [6] [7], can be modelled and solved under certain conditions. The numerical solution of the Föppl-von Kármán system [8, 9] that governs the underlying dynamics of nonlinear thin plates at moderate vibration amplitudes yields realistic and convincing sound synthesis [10], but at heavy computational cost. Ducceschi and Touzé [11] propose the modal resolution of the system with the offline calculation of coupling coefficients. They manage under certain approximations to significantly reduce the computation time without being able to achieve real-time sound synthesis (about 8 times real-time on a CPU) [12]. As of 2023, real-time performance is available for limited plate sizes [13]. Another typical case of coupling between modes induced by nonlinear phenomena results from collisions in musical instruments [14] and has been the subject of various studies, including on modal interactions [15]. Computational cost for synthesis can also be heavy in such cases.

For synthesis purposes, and particularly if real-time performance is the ultimate aim, it can be useful to depart from strict physical models, and examine modal interactions from a perceptual point of view—closer in spirit to so-called “procedural audio” approaches [16]. Skare and Abel [17] perform real-time modal synthesis of crash cymbals with a GPU-accelerated modal filterbank. Their method consists in identifying the modal parameters (including a rough approximation of the couplings) on recorded sounds, although the energy transfer mechanism is unspecified.

In this paper, we propose a simple model for energy transfers between modes. Then we design coupled filters based on the design proposed by Mathews and Smith [18] and adapted by Skare and Abel [17] to incorporate energy transfer. This study is a direct extension of previously conducted work [19], incorporating additional effort to ensure that the design of coupled filters is more coherent with the underlying physical system. In particular, we propose an equivalence between the power of the signal of the filters and the energy of a vibration mode from an equivalent physical system to ensure energy conservation during transfers. Inter-modal energy transfer is encoded in a matrix containing all the coupling coefficients. The aim of this paper is not to propose a synthesis model performing an accurate simulation of a physical system. Instead, we seek to develop a framework allowing direct modelling of sounds targeted to the way they are perceived. This results in an efficient way to generate sounds evoking nonlinear sources and can yield real-time event-driven synthesis of sounds in virtual or augmented reality environments, a particularly active field of research [20] [21].

Some background on modal synthesis is given in Section 2 and the energy transfer model is presented in Sec.3. Then, the design of the coupled filters is detailed in Sec.4, the definition of the matrix containing the coupling terms is proposed in Sec.5, and methods to enhance the efficiency and randomize the process are presented in Sec.6. Various example systems used to generate prototypical sounds are presented in Sec.7. Sound examples are available online [22].

2 Modal synthesis for the linear case

The modal resolution of a linear partial differential equation (PDE) system describing the vibrations of a resonant object is well-described in various texts [23]. Solutions are of the following form for the displacement w depending on a spatial coordinate \mathbf{r} and time t :

$$w(\mathbf{r}, t) = \underbrace{w_h(\mathbf{r}, t)}_{\text{homogeneous solution}} + \underbrace{w_p(\mathbf{r}, t)}_{\text{particular solution}}, \quad (1)$$

where

$$w_h(\mathbf{r}, t) = \sum_{i=1}^{\infty} e^{-\alpha_i t} [A_i \cos(\omega_{d_i} t + \varphi_i)] \phi_i(\mathbf{r}) \quad (2a)$$

$$w_p(\mathbf{r}, t) = \sum_{i=1}^{\infty} (g_i(t) * h_i(t)) \phi_i(\mathbf{r}), \quad (2b)$$

Here, $*$ represents a convolution operation, and the impulse response $h_i(t)$ of the following form:

$$h_i(t) = \frac{1}{\omega_{d_i}} e^{-\alpha_i t} \sin(\omega_{d_i} t) \quad (3)$$

the function $\phi_i(\mathbf{r})$ is the i^{th} mode's shape or basis function, ω_{d_i} and α_i are the angular frequency and the damping coefficient of the i^{th} mode, respectively. One can note that the angular frequency differ from the angular natural frequency ω_i :

$$\omega_{d_i}^2 = \omega_i^2 - \alpha_i^2. \quad (4)$$

The constants A_i and φ_i derive from the initial conditions and $g_i(t)$ is the modal excitation (formally derived from a PDE system by the projection of an excitation source term $g(\mathbf{r}, t)$ onto the modal basis functions $\phi_i(\mathbf{r})$).

3 Inter-modal energy transfer

3.1 Definitions and approximations

For a linear system, the mechanical energy of the i^{th} mode $E_m^i(t)$ can be calculated by adding its potential energy $E_p^i(t)$, its kinetic energy $E_k^i(t)$, and the accumulated

energy $E_s^i(t)$ supplied by the source up to time t :

$$E_p^i(t) = \frac{1}{2} K_i q_i(t)^2 \quad (5a)$$

$$E_k^i(t) = \frac{1}{2} M_i \left(\frac{dq_i(t)}{dt} \right)^2 \quad (5b)$$

$$E_s^i(t) = \int_0^t g_i(\Xi) \frac{dq_i(\Xi)}{d\Xi} dt \quad (5c)$$

$$E_m^i(t) = E_p^i(t) + E_k^i(t) + E_s^i(t), \quad (5d)$$

Here, K_i is the modal stiffness, M_i is the modal mass and $q_i(t)$ is the modal displacement, defined as follows:

$$M_i = \int_{\mathbf{r} \in \Omega} m(\mathbf{r}) \phi_i^2(\mathbf{r}) d\mathbf{r} \quad (6a)$$

$$K_i = \omega_i^2 M_i \quad (6b)$$

$$q_i(t) = e^{-\alpha_i t} [A_i \cos(\omega_{d_i} t + \varphi_i)] + g_i(t) * h_i(t). \quad (6c)$$

The function $m(\mathbf{r})$ is the density, and Ω is the closed space containing the vibrating object.

A simple approximation to the mechanical energy follows from the assumption that it is proportional to the square of the modal displacement amplitude (see Fig.1). Thus, we can approximate the mechanical energy of the i th mode by computing the power of the modal displacement signal denoted as $P_i(t)$ (where the power of a sinusoidal signal is equal to its squared amplitude divided by 2). It is important to note that the power referred to here is not mechanical power expressed in Watts but rather signal power (this will be useful for the design of the coupled filters).

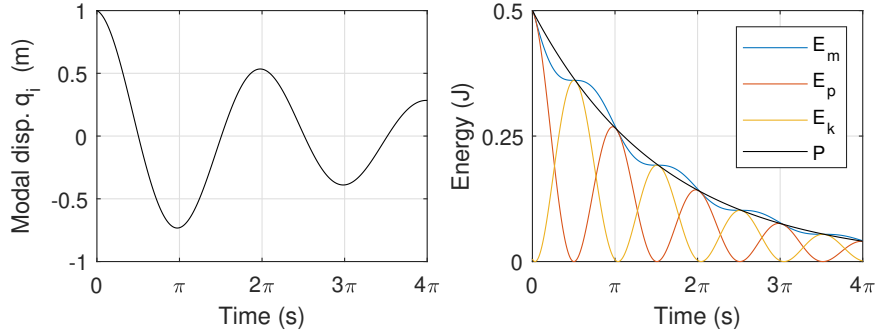


Fig. 1 Displacement and energies of a mode with modal mass $M_i = 1$, modal stiffness $K_i = 1$, for a damping coefficient $\alpha_i = 0.1$, an initial amplitude $A_i = 1$, and an initial phase $\varphi_i = 0$. This example is without the source ($g_i(t) = 0$). We can see that the power of the signal, proportional to the square of the modal displacement amplitude $P(t) = (e^{-\alpha_i t} A_i)^2 / 2$ is a rough approximation of the mechanical energie $E_m(t)$.

Additionally, assuming identical modal masses for all modes (this follows from a uniform density and modal orthogonality) allows us to establish that the mechanical energy of a given mode is also proportional to the square of the mode's angular frequency:

$$E_p^i(t) = \frac{1}{2} K_i q_i^2(t) \propto \omega_i^2 M_i q_i^2(t) \quad \Rightarrow \quad E_m^i(t) \propto \omega_i^2 P_i(t) \quad (7)$$

To establish a simple and controllable model, we neglect the influence of the phase on energy transfers. We introduce a term $\Pi_T^i(t)$ to induce transfers of energy between distinct modes. In the absence of an external source, the energy of a mode is expressed as its initial value, with a decrease over time due to the cumulative losses and modified by the transferred energy with other modes:

$$E_m^i(t) = \underbrace{E_0^i}_{\text{initial energy}} - \underbrace{\int_0^t 2\alpha_i E_m^i(\Xi) d\Xi}_{\text{losses}} + \underbrace{\int_0^t \Pi_T^i(\Xi) d\Xi}_{\text{energy transfers}}. \quad (8)$$

One can note that E_m^i appears in the loss term because we assumed an exponential decay for the modes, as it commonly done in modal models.

3.2 Energy transfer model

The challenge is to arrive at a model simple enough to be controllable (i.e., to be able to predict the sound outcome of a manipulation of the parameters) and complete enough to allow the matching of modal trajectories to a range of nonlinear phenomena. We define the transfer term as following:

$$\Pi_T^i(t) = \underbrace{-\lambda [E_m^i(t) - \tau_i]_+}_{\text{energy transferred to other modes}} + \underbrace{\sum_j \lambda c_{ij} [E_m^j(t) - \tau_j]_+}_{\text{energy transferred to mode i from other modes}}, \quad (9)$$

Here, $[\cdot]_+$ indicates the “positive part of”, i.e., $[\zeta]_+ = \frac{1}{2}(\zeta + |\zeta|)$, τ_i is the threshold beyond which the energy of mode i is transferred to other modes, λ is the proportion of energy above the threshold that is redistributed and c_{ij} is a coupling coefficient ($\sum_i c_{ij} \leq 1$).

Thus, the transfer terms are proportional to the excess energy above a threshold and the terms c_{ij} define the proportions distributed and received by each other component. Note that this relation is not an immediate consequence of a physical model but is a heuristic means of capturing salient phenomena in a physical system. Our focus is on the design of a synthesis process with a predictable sound outcome rather than on the simulation of a physical system. Nevertheless, our model remains physically informed and consistent with the conservation of energy in the associated mechanical system.

We can define the following differential equation that governs the energy variations of the modes, excluding the effect of the source:

$$\frac{dE_m^i(t)}{dt} = -2\alpha_i E_m^i(t) - \lambda \underbrace{[E_m^i(t) - \tau_i]_+ + \sum_j \lambda c_{ij} [E_m^j(t) - \tau_j]_+}_{\Pi_T^i(t)}. \quad (10)$$

It is possible to analytically solve this equation by with initial conditions $E_m^i(0)$ and distinguishing cases where the energy of each mode is above or below the corresponding threshold τ_i . For example, if $E_m^i(0) > \tau_i$ and $E_m^j(t) < \tau_j \forall j \neq i$ the solution takes the following form (see Fig.2):

$$E_m^i(t) = \begin{cases} \left(E_m^i(0) - \frac{\lambda \tau_i}{\lambda + 2\alpha_i} \right) e^{-(\lambda + 2\alpha_i)t} + \frac{\lambda \tau_i}{\lambda + 2\alpha_i} & \text{if } t < t_0 \text{ (i.e. } E_m^i(t) > \tau_i) \\ \frac{\tau_i}{e^{-2\alpha_i t_0}} e^{-2\alpha_i t} & \text{else} \end{cases} \quad (11)$$

$$\text{with } t_0 = -\frac{1}{\lambda + 2\alpha_i} \left[\ln \left(\tau_i \left(1 - \frac{\lambda}{\lambda + 2\alpha_i} \right) \right) - \ln \left(E_m^i(0) - \frac{\lambda \tau_i}{\lambda + 2\alpha_i} \right) \right].$$

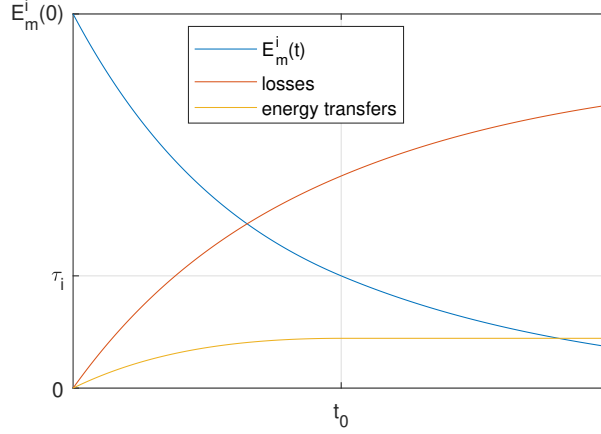


Fig. 2 Evolution of mechanical energy (in blue) for an initial value $E_m^i(0)$ exceeding the transfer threshold $\tau_i = 0.3E_m^i(0)$ and in the absence of source or transfer contributions from other modes for $\lambda = 1, \alpha_i = 1$ (see Eq.(11)). The energy lost due to dissipation is shown in red, and the energy to be transferred to other modes is presented in yellow. One can note that the energy transferred remains the same for $t > t_0$ ($t > t_0 \Rightarrow E_m^i < \tau_i$)

We can rewrite Eq.(10) in terms of signal power (see Eq.(7)), which will be useful for the implementation of the coupled filters (Sec.4):

$$\frac{dP_i(t)}{dt} = -2\alpha_i P_i(t) - \lambda [P_i(t) - \tau_i]_+ + \underbrace{\sum_j \lambda c_{ij} \frac{\omega_j^2}{\omega_i^2} [P_j(t) - \tau_j]_+}_{T_i(t)}. \quad (12)$$

3.3 Energy and stability

A sufficient condition for the stability of the system which is consistent with the physics (energy conservation) is to impose a non-positivity constraint for the sum of energy transfers:

$$\sum_i \Pi_T^i(t) \leq 0 \quad (13)$$

This condition impedes the creation of energy during transfer between modes. It is ensured by the following condition on the coupling coefficients:

$$\sum_i c_{ij} \leq 1 \quad \forall j \quad (14)$$

4 Design of the coupled filters

4.1 Linear filtering for modal synthesis

In the linear case, a straightforward approach to numerical solution at a sample rate f_s in Hz is to use recursive filters with an exponentially-damped sinusoidal impulse response. The filter proposed by Mathews and Smith [18] has this property. The implementation of this filter consists in calculating, for each time step n , the imaginary part of a complex number $z(n)$ whose rotation speed in the complex plane is constant and corresponds to the angular frequency ω of the exponentially damped sinusoid:

$$y(n) = \text{Im}(z(n)) \quad \text{where} \quad z(n+1) = Zz(n) + u(n) \quad (15)$$

with $u(n)$ the source of the filter, and Z the constant modification of the phase and modulus per time step:

$$Z = e^{-\alpha/f_s} e^{j\omega/f_s} = X + jY \quad (16)$$

with $X = e^{-\alpha/f_s} \cos(\omega/f_s)$ and $Y = e^{-\alpha/f_s} \sin(\omega/f_s)$.

The recurrence equation on the complex sequence $z(n)$ is computed by the following system including a recurrence equation for the real part $x(n) = \text{Re}(z(n))$ and a recurrence equation for the imaginary part $y(n) = \text{Im}(z(n))$, which is the output of the filter:

$$\begin{aligned} x(n+1) &= \text{Re}(z(n+1)) = Xx(n) - Yy(n) + u(n) \\ y(n+1) &= \text{Im}(z(n+1)) = Yx(n) + Xy(n) \end{aligned} \quad (17)$$

for a real source $u(n) \in \mathbb{R}$.

4.2 Principle and implementation of the coupling

Consider N filters defined as in Sec.4.1 in parallel to be coupled through the methodology outlined in Sec.3.2. We note $z_i(n)$ the complex sequence corresponding to the i^{th} filter, with $x_i(n)$ its real part and $y_i(n)$ its imaginary part (corresponding to the output signal of the filter). The source for the i^{th} filter, corresponding to the projection of the source of the filter bank $u(n)$ onto the i^{th} modal basis function, is noted $u_i(n)$.

The continuous equation (12) can be solved in discrete time using the following recurrence relation:

$$P_i(n+1) = P_i(n) \underbrace{e^{-2\alpha_i/f_s}}_{\text{losses}} + \underbrace{T_i(n)}_{\text{transfer}} \quad (18)$$

with $P_i(n)$ the power of the tonal component defined as following:

$$P_i(n) = \frac{|z_i(n)|^2}{2} = \frac{1}{2}(x_i(n)^2 + y_i(n)^2) \quad (19)$$

$x_i(n), y_i(n) \in \mathbb{R}$.

Thus, we can express the variation of the modulus of $z_i(n)$ due to energy transfer between two time steps:

$$|z_i(n+1)| = \sqrt{|z_i(n)|^2 e^{-2\alpha_i/f_e} + 2T_i(n)} \quad (20)$$

We can define an amplitude ratio between the modulus for two consecutive time steps if $|z_i(n)| \neq 0$:

$$\frac{|z_i(n+1)|}{|z_i(n)|} = \sqrt{e^{-2\alpha_i/f_e} + \frac{2T_i(n)}{|z_i(n)|^2}} \quad (21)$$

Thus, we can modify the recurrence equation defined in the previous section (see Eq.(15)) by incorporating the modulus variations due to energy transfers. It gives the following recurrence relation for z_i , including the source and phase variations:

$$z_i(n+1) = \begin{cases} \sqrt{2T_i(n)} + u_i(n) & \text{if } z_i(n) = 0 \\ \sqrt{e^{-2\alpha_i/f_e} + \frac{2T_i(n)}{|z_i(n)|^2}} e^{j\omega_i/f_s} z_i(n) + u_i(n) & \text{else} \end{cases} \quad (22)$$

Finally, we can write the system of equations for the implementation of the filters:

$$\begin{aligned} x_i(n+1) &= \begin{cases} \sqrt{2T_i(n)} + u_i(n) & \text{if } z_i(n) = 0 \\ X_i x_i(n) - Y_i y_i(n) + u_i(n) & \text{else} \end{cases} \\ y_i(n+1) &= \begin{cases} 0 & \text{if } z_i(n) = 0 \\ Y_i x_i(n) + X_i y_i(n) & \text{else} \end{cases} \end{aligned} \quad (23)$$

with $X_i = \sqrt{e^{-2\alpha_i/f_e} + \frac{2T_i(n)}{|z_i(n)|^2}} \cos(\omega_i/f_s)$ and $Y_i = \sqrt{e^{-2\alpha_i/f_e} + \frac{2T_i(n)}{|z_i(n)|^2}} \sin(\omega_i/f_s)$

In this way, power can be transferred among the different filters without affecting the phases. The coupling intervenes in the calculation of the transfer terms $T_i(n)$ which ultimately involve the other filters.

5 Distribution matrix

This section presents a formalism for the calculation and control of the coupling between filters.

Now define the column vectors $\mathbf{p}(n) = [P_1(n), \dots, P_N(n)]^T$ and $\mathbf{t}(n) = [T_1(n), \dots, T_N(n)]^T$. The power transfers between the tonal components $\mathbf{t}(n)$ at a given time step n are defined as:

$$\mathbf{t}(n) = \mathbf{M}[\mathbf{p}(n) - \boldsymbol{\tau}]_+ . \quad (24)$$

An $N \times 1$ column vector $\boldsymbol{\tau}$ containing the thresholds τ_i , $i = 1, \dots, N$ at which transfers are activated for each tonal component has also been introduced here.

Thus, the calculation of the transfer terms is performed by the matrix product of an $N \times N$ redistribution matrix \mathbf{M} with the vector resulting from the positive part of the difference between the power of each frequency component $\mathbf{p}(n)$ and the associated threshold $\boldsymbol{\tau}$. In other words, the transfer terms $T_i(n)$ are proportional to the excess power above the corresponding threshold and the terms of the matrix \mathbf{M} define the proportions distributed to and received by each component.

The diagonal entries M_{jj} of the matrix \mathbf{M} define the proportion of power of the j th mode that will be redistributed to other modes and the other terms of the column M_{ij} define the quantity that the i th mode will receive from this redistribution.

An efficient way to define the coefficients of the matrix is to use the following expression:

$$M_{ij} = \eta \frac{\lambda}{f_s} \frac{a_{ij}}{\underbrace{\sum_{i=1}^N a_{ij}}_{c_{ij}}} \frac{\omega_j^2}{\omega_i^2} - \frac{\lambda}{f_s} \delta_{ij} \quad (25)$$

Where a_{ij} is a coefficient weighting the redistribution from the j th mode to the i th mode. In this formulation, the stability of the filter bank is ensured for arbitrary a_{ij} , provided that at least one value per column is non-zero, that $\lambda \leq f_s$ and that $0 \leq \eta \leq 1$. η corresponds to the efficiency of the transfers ($\eta = 1 \Rightarrow \sum_i E_T^i(n) = 0$). λ is the proportion of power above threshold transferred to other modes at each time step ($0 \leq \lambda \leq f_s$). The values of the off-diagonal elements M_{ij} of the matrix \mathbf{M} are the proportion of energy transferred by the mode j that will be received by the mode i .

The i^{th} transfer term $T_i(n)$ can be expressed as follows:

$$T_i(n) = \underbrace{\eta \frac{\lambda}{f_s} \sum_{j=1}^N \left[\frac{a_{ij}}{\sum_{i=1}^N a_{ij}} \frac{\omega_j^2}{\omega_i^2} [P_j(n) - \tau_j]_+ \right]}_{\text{positive contribution } T_{i+}(n)} - \underbrace{\frac{\lambda}{f_s} [P_i(n) - \tau_i]_+}_{\text{negative contribution } T_{i-}(n)} \quad (26)$$

6 Efficiency and randomization of the process

It is possible to perform the energy transfer process at a lower rate than the sample rate without degrading to audio quality. If we perform the transfers every N_0 samples, the expression of the transfers (26) become:

$$T_i(n) = \begin{cases} \frac{\lambda N_0}{f_s} \left(\eta \sum_{j=1}^N \left[\frac{a_{ij}}{\sum_{i=1}^N a_{ij}} \frac{\omega_j^2}{\omega_i^2} [P_j(n) - \tau_j]_+ \right] - [P_i(n) - \tau_i]_+ \right) & \text{if } n \bmod N_0 = 0 \\ 0 & \text{else} \end{cases} \quad (27)$$

As this step constitutes the heaviest computation in the algorithm as a whole, avoiding the need to calculate it at each time step can significantly enhance the efficiency of the code.

On the other hand, since energy transfers are highly regular and predictable (and precisely why this type of model was chosen), one might wish to introduce randomness into the transfer processes to induce more chaotic variations. A simple and effective method involve introducing a random variable $\chi(n)$ following a uniform distribution between 0 and 2π . This randomness can be incorporated to vary the phase of the positive contribution of transfers in the implementation of the filters (see Eq.(23)):

$$\begin{aligned} x_i(n+1) &= \begin{cases} \operatorname{Re}(e^{j\chi(n)})\sqrt{2T_{i+}(n)} + u_i(n) & \text{if } z_i(n) = 0 \\ \tilde{X}_i x_i(n) - \tilde{Y}_i y_i(n) + \operatorname{Re}(e^{j\chi(n)})\sqrt{2T_{i+}(n)} + u_i(n) & \text{else} \end{cases} \\ y_i(n+1) &= \begin{cases} \operatorname{Im}(e^{j\chi(n)})\sqrt{2T_{i+}(n)} & \text{if } z_i(n) = 0 \\ \tilde{Y}_i x_i(n) + \tilde{X}_i y_i(n) + \operatorname{Im}(e^{j\chi(n)})\sqrt{2T_{i+}(n)} & \text{else} \end{cases} \end{aligned} \quad (28)$$

with $\tilde{X}_i = \sqrt{e^{-2\alpha_i/f_e} + \frac{2T_{i-}(n)}{|z_i(n)|^2}} \cos(\omega_i/f_s)$ and $\tilde{Y}_i = \sqrt{e^{-2\alpha_i/f_e} + \frac{2T_{i-}(n)}{|z_i(n)|^2}} \sin(\omega_i/f_s)$

One can note that this method tends to dissipate more energy due to the introduction of a component with a randomized phase, which has the potential to diminish the amplitude of a filter if their phase are opposed. Also, the modulations created with this approach are affected by N_0 .

7 Examples

Nonlinear vibration leads to complex phenomena that can produce subtle and chaotic variations in radiated sound. We can reduce the complexity of the model and propose a heuristic that attempts to maintain the essential perceptual attributes of an object vibrating under nonlinear conditions. The resulting synthetic sound is nevertheless less realistic and versatile than sounds generated by the direct resolution of physical models (such as, e.g., the Föppl-von Kármán system) although the synthesis quality can be improved by using random processes in the implementation of the algorithms.

The coupled filter bank proposed here is dependent on various parameters: the number of filters N , the oscillation frequencies ω_i and damping α_i for each filter, the coefficients a_{ij} and the parameters λ and η for the definition of the redistribution matrix \mathbf{M} , and the thresholds τ_i . Strategies for setting these parameters are presented

in two cases of musical interest. In the case of nonlinear plate vibration, energy is transferred to filters of nearby frequency in order to generate a gradual cascade of energy towards the high-frequency range. In the case of a string colliding with a rigid object, in contrast, there is simultaneous transfer of energy to many frequency components.

7.1 Energy cascade in thin plates

Consider a thin rectangular plate (according to the Kirchhoff model [24]), with mass density ρ $\text{kg}\cdot\text{m}^{-3}$, thickness H m, and flexural rigidity D in $\text{kg}\cdot\text{m}^2\cdot\text{s}^{-2}$, and side lengths L_x and L_y in m. If the plate is simply supported on all its edges, the modal frequencies ω_{lm} and modal shapes $\phi_{lm}(x, y)$ can be expressed as follows [25]:

$$\omega_{lm} = \frac{\pi^2}{L_x^2} \sqrt{\frac{D}{\rho H}} (l^2 + \nu^2 m^2) \quad \phi_{lm}(x, y) = \sin(l\pi x) \sin(m\pi y) \quad (29)$$

Here, $\nu = L_x/L_y$ is the plate aspect ratio, or the ratio between the length and width of the plate. The integer indices $l, m \geq 1$ correspond to the number of vibration extrema ($l - 1, m - 1$ correspond to the number of vibration nodes) in the main directions of the rectangular plate (Cartesian coordinates (x, y)) with x and y being normalized by the length of the plate in the corresponding direction (so that $0 \leq x, y \leq 1$).

For a point excitation force located at (x_e, y_e) , we can compute the modal forces using the mode shapes evaluated at the excitation point as $\phi_{lm}(x_e, y_e)$. We define the source of the l, m th filter as follows:

$$u_{lm}(n) = \sin(l\pi x_e) \sin(m\pi y_e) u(n) \quad (30)$$

where $u(n)$ is the global excitation function.

We use a raised sinusoid for the excitation force (as proposed in [4] and [26]) to simulate an impact:

$$u(n) = \begin{cases} A \sin^2(\pi n/N_{ex}) & \text{if } n \leq N_{ex} \\ 0 & \text{else} \end{cases} \quad (31)$$

For typical plate strikes, the strike duration N_{ex}/f_s in seconds is on the order of 1-4 ms.

The damping coefficients are chosen according to an exponential law, as proposed by Aramaki et al.[27], with parameters that are set to evoke a metallic object:

$$\alpha_{lm} = e^{(\alpha_G + \omega_{lm} \alpha_R)} \quad (32)$$

with $\alpha_R = 4 \times 10^{-5}$ and $\alpha_G = 0.33220$. This set of parameters permits direct modal synthesis for linear plate vibration. To each pair of indices (l, m) we associate an index i (chosen in terms of increasing modal frequency) corresponding to the filter number used to generate the corresponding tonal component.

In order to produce the cascade of energy towards higher frequencies, we carry out transfers between filters whose frequencies are close. Indeed, the energy supplied by

the impact is localised at low frequencies and transfers directed towards the neighbouring modes allow the progressive appearance of higher frequency components. The weighting coefficients a_{ij} can be set as follows (see Figure 3):

$$a_{ij} = \left[1 - \frac{|f_j - f_i|}{\Delta f} \right]_+ \quad (33)$$

with $f_i = \frac{\omega_i}{2\pi}$

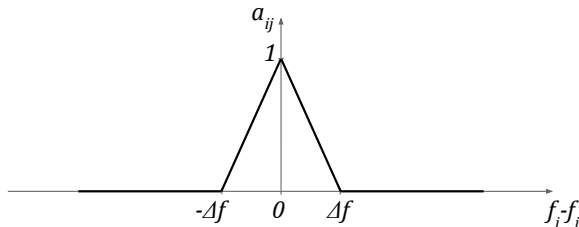


Fig. 3 Value of the coefficient a_{ij} as a function of the frequency difference between filter i and j .

We set $\eta = 1$ (ensuring conservation of energy during the redistribution). The cascade can be mainly controlled by λ , or by the definition of thresholds τ_i (see Figures 4 and 5).

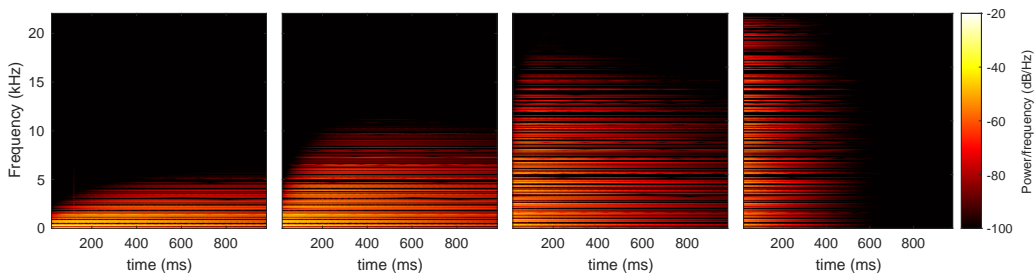


Fig. 4 Spectrograms of output for filters whose frequency corresponds to the modal frequency of a thin plate for different values of λ ($\tau_i = 0$). From left to right: $\lambda = 0.001$, $\lambda = 0.01$, $\lambda = 0.1$, $\lambda = 1$. We can observe that the energy cascade spreads faster and higher in frequency with the increase of λ . Transfers are performed at each time step.

In the case of wave turbulence in plates [28], couplings between modes can lead to rapid variations in amplitude and frequency leading to a chaotic regime. In the chaotic regime, the resulting signal is noisy, and difficult to reproduce by a set of tonal components. One way to reproduce this phenomenon with the coupled filters presented in this paper is to randomize the phases of the positive contributions of the transfer term in the source, as proposed in Sec.6. In this way, the tonal components are subject to rapid random amplitude modulations that can evoke the chaotic phenomenon occurring during wave turbulence in the plates (see Figure 6).

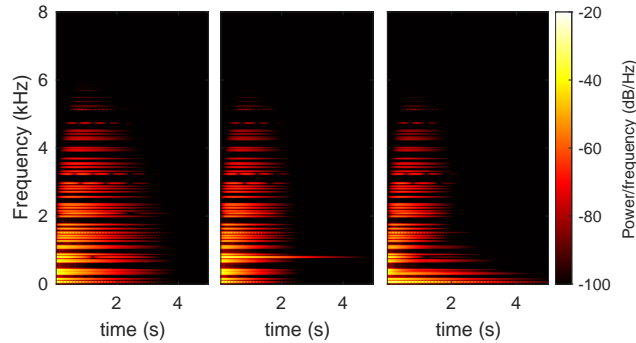


Fig. 5 Spectrograms of output for filters whose frequency corresponds to the modal frequency of a thin plate for different thresholds τ_i . Left: $\tau_i = 0$; middle: $\tau_i = 0$ except for $i = 10$ where $\tau_{10} = 1$; right: τ_i is half the excitation amplitude. All tonal components decay simultaneously when the thresholds are zero (left). A component emerges and decays more slowly when its threshold is non-zero (middle). When all thresholds are different from zero, we observe a usual exponential decay after the delayed appearance of the high frequency component (right).

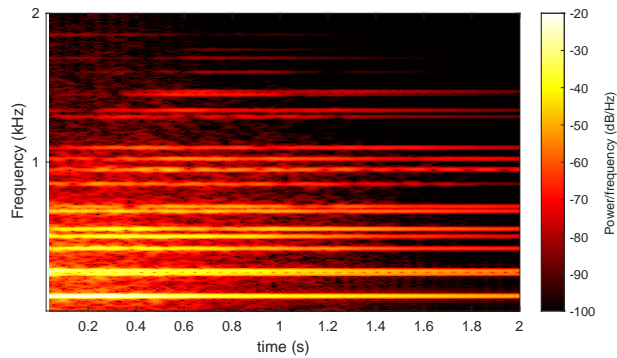


Fig. 6 Spectrogram of output for filters whose frequency corresponds to the modal frequency of a thin plate. The random modulation of the redistributions induces rapid variations in the amplitude of the tonal components which generate noise and beating in the signal. Transfers are performed very $N_0 = 4$ time step.

7.2 Collisions in sound production

The perturbation of the vibrations of an object when colliding with an obstacle can lead to different types of sound events. In the typical case of a guitar, the player can choke the string, mute it, play a natural harmonic. The string can also interact with the soundboard (slap, string buzz).

The model of a vibrating string with simply supported boundary conditions gives the following modal frequencies and shapes:

$$\omega_i = i\omega_1 \quad \phi_i(x) = \sin(i\pi x) \quad (34)$$

where here, the spatial coordinate x is normalized by the length of the string ($0 \leq x \leq 1$). For a point excitation force located at x_e , the source of the i^{th} filter can be defined as:

$$u_i(n) = \sin(i\pi x_e)u(n) \quad (35)$$

We use the same excitation force and damping model as previously (see Eqs.(31) and (32)).

The evocation of an obstacle disturbing the vibrations of the string requires the definition of thresholds that correspond to the location of the obstacle. We propose thresholds corresponding to the maximum amplitude of modal displacements without colliding with a virtual obstacle positioned at x_c, y_c , where y_c is the vertical displacement of the obstacle relative to the string:

$$\tau_i = \frac{1}{2} \left(\frac{y_c}{\sin(i\pi x_c)} \right)^2 \quad (36)$$

We define a redistribution matrix with all columns being identical in order to cause a simultaneous redistribution to a set of tonal components. The coefficients a_{ij} are defined as follows:

$$a_{ij} = |\sin(i\pi x_c)|\xi_i(f_i) \quad (37)$$

with $\xi_i(f_i)$ a parameter depending on the frequency allowing weighting of the redistribution according to the filter frequency. We define $\xi_i(f_i)$ as the Fourier transform of the raised cosine, an approximation of the force profile caused by a collision (as defined for the source, Eq.(31)):

$$\xi_i = \text{sinc}(f_i\gamma) + \frac{1}{2} (\text{sinc}(f_i\gamma - 1) + \text{sinc}(f_i\gamma + 1)) \quad (38)$$

with f_i the frequency of the i^{th} filter and γ a parameter corresponding to the duration of the raised cosine. This results in a cutoff frequency beyond which there is no more

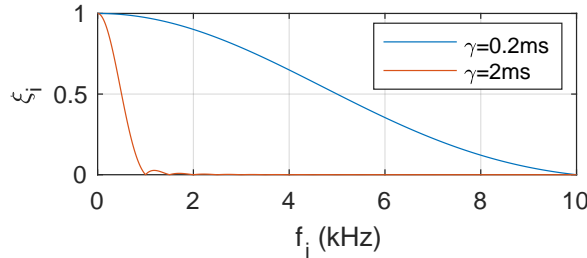


Fig. 7 Value of ξ_i as a function of the frequency of filter i .

transfer (see Figure 7). Various examples of sound outputs for different configurations are presented—see Figures 8 and 9.

Collisions in musical instruments may be the source of more subtle phenomena than a simultaneous appearance of various frequency components. The cases of

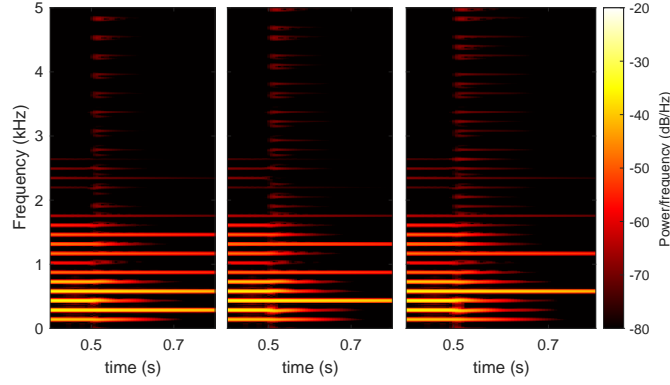


Fig. 8 Spectrograms of output for filters whose frequency are harmonic for different values of x_c ($y_c = 0$, $\gamma = 2 \times 10^{-4} s$, $\lambda = 0.25$, $\nu = 0.5$). Transfers are performed every 294 samples for times greater than 500ms, which corresponds to a collision every 6.67ms (150Hz). From left to right: $x_c = 1/2$, $x_c = 1/3$, $x_c = L/4$. We can observe that the transfer does not affect even harmonics (resp. multiples of 3 and 4) for $x_c = 1/2$ (resp. $x_c = 1/3$ and $x_c = L/4$), which allows the reproduction of a natural harmonic played on a guitar.

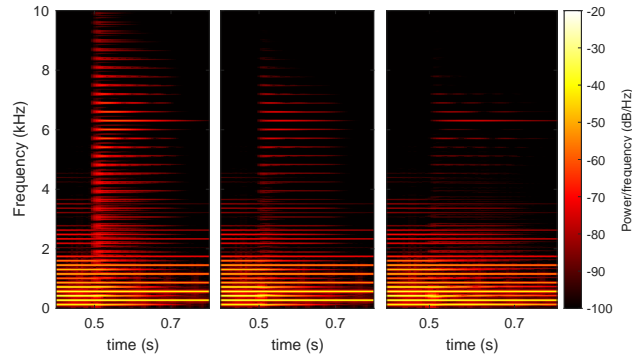


Fig. 9 Spectrograms of output for filters whose frequency are harmonic for different values of ν and γ ($x_c = 1/2$, $y_c = 0$). From left to right: ($\nu = 0.5$, $\gamma = 2 \times 10^{-4} s$), ($\nu = 0.15$, $\gamma = 2 \times 10^{-4} s$), ($\nu = 0.5$, $\gamma = 2 \times 10^{-3} s$). Transfers are performed every 294 samples for times greater than 500ms, which corresponds to a collision every 6.67ms (150Hz). There is a lower increase in the high-frequency components and a faster dissipation of all the tonal components involved in the redistribution as the efficiency decreases. As γ increases, there is also less energy distributed to the high-frequency components, but this energy is not dissipated and remains in the low-frequency components.

string buzz and tanpura can be approached by introducing random processes into the redistribution, as has been done for chaotic phenomena in plates (see Figure 10).

It is possible to apply the same principle for the generation of sounds corresponding to collisions with 2D objects. For example, we can generate muted plate sounds (see Figure 11).

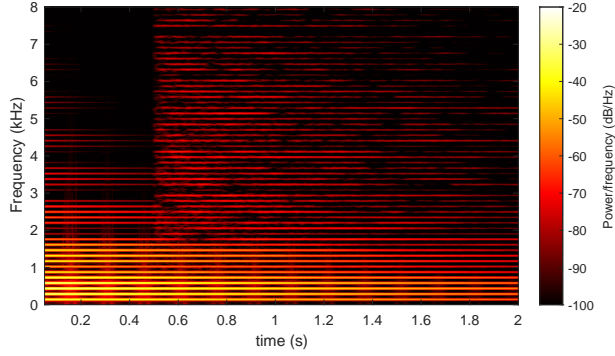


Fig. 10 Spectrograms of output for filters whose frequency are harmonic with the introduction of random processes during the redistribution ($\lambda = 0.001$, $\nu = 0.9$, $x_c = 0.38$, $y_c = 0.001$, $\gamma = 2 \times 10^{-4} s$). Transfers are performed every 294 samples for times greater than 500ms, which corresponds to a collision every 6.67ms (150Hz).

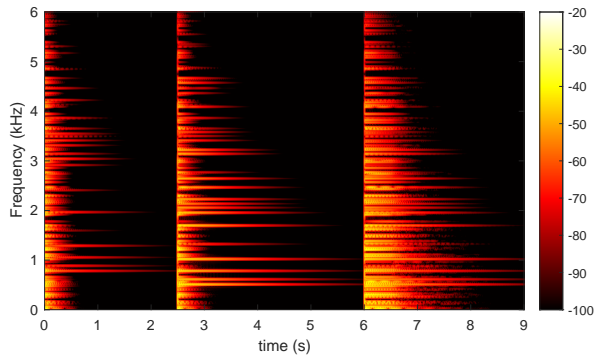


Fig. 11 Spectrograms of output for filters whose frequency corresponds to the modal frequency of a thin plate. here $\nu = 0$ and we observe the quick dissipation of certain tonal components for three distinct impacts, which creates a sensation of choking.

8 Conclusions

In this paper, we have presented a model for mode coupling and the design of coupled resonant filters geared towards the emulation of mode coupling effects in nonlinear vibrating structures. This filter bank allows efficient and real-time sound synthesis even for a large number of filters. The coupling, performed without modifying the phase, introduces predictable and controllable effects on the output signal. The terms controlling the coupling between the different filters are grouped in a matrix whose definition is the main challenge. The setting of the parameters of the sound synthesis process is presented through various examples corresponding to sources whose behavior is nonlinear. A simple setting allows the generation of typical sounds, though sometimes with an unnatural character—the introduction of random processes in the energy redistribution aids a great deal with both realism and plausibility.

Further work will be concerned with determining which sound morphologies are important from a perceptual point of view for the recognition of sound events [29] corresponding to nonlinear phenomena in order to reproduce them with this coupled filter bank. This could lead to the development of environmental sound synthesizers and virtual musical instruments (e.g. tanpura, cymbal ...), or to non-linear audio effects (such as the nonlinear reverberation of a snare drum due to the wires held under tension against the lower drumskin). The filter bank presented in this paper can also be used as an abstract sound generation tool. In this context, the challenge would be to design intuitive control for use in a musical or sound design context.

Declarations

- Availability of data and materials
No associated data
- Competing interests
The authors declare that they have no competing interests
- Funding
Not applicable
- Authors' contributions
SP: main contribution to the conception of the work and drafted the paper,
SB: design of the work and substantial revision,
RKM: Project initiator, design and work revision.

References

- [1] Adrien, J.-M.: The missing link: Modal synthesis. In: Representations of Musical Signals, pp. 269–298 (1991)
- [2] Morrison, J.D., Adrien, J.-M.: Mosaic: A framework for modal synthesis. *Computer Music Journal* **17**(1), 45–56 (1993)
- [3] Rocchesso, D.: The ball within the box: a sound-processing metaphor. *Computer Music Journal* **19**(4), 47–57 (1995)
- [4] Van Den Doel, K., Kry, P.G., Pai, D.K.: Foleyautomatic: physically-based sound effects for interactive simulation and animation. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, pp. 537–544 (2001)
- [5] Conan, S., Thoret, E., Aramaki, M., Derrien, O., Gondre, C., Ystad, S., Kronland-Martinet, R.: An intuitive synthesizer of continuous-interaction sounds: Rubbing, scratching, and rolling. *Computer Music Journal* **38**(4), 24–37 (2014)
- [6] Legge, K., Fletcher, N.H.: Nonlinearity, chaos, and the sound of shallow gongs. *The Journal of the Acoustical Society of America* **86**(6), 2439–2443 (1989)

- [7] Chaigne, A., Touzé, C., Thomas, O.: Nonlinear vibrations and chaos in gongs and cymbals. *Acoustical science and technology* **26**(5), 403–409 (2005)
- [8] Föppl, A.: *Vorlesungen über Technische Mechanik*. Druck und Verlag von B.G. Teubner, Leipzig (1907)
- [9] Kármán, T.: Festigkeitsprobleme im maschinenbau. *Encyklopädie der Mathematischen Wissenschaften* **4**(4), 311–385 (1910)
- [10] Bilbao, S.: A family of conservative finite difference schemes for the dynamical von karman plate equations. *Num. Meth. PDE* **24**(1), 193–216 (2008)
- [11] Ducceschi, M., Touzé, C.: Modal approach for nonlinear vibrations of damped impacted plates: Application to sound synthesis of gongs and cymbals. *Journal of Sound and Vibration* **344**, 313–331 (2015)
- [12] Ducceschi, M., Touzé, C.: Simulations of nonlinear plate dynamics: an accurate and efficient modal algorithm. In: *18th International Conference on Digital Audio Effects (DAFx-15)* (2015)
- [13] Bilbao, S., Wang, Z., Webb, C., Ducceschi, M.: Real-time gong synthesis. In: *Proceedings of the 26th International Conference on Digital Audio Effects, Copenhagen, Denmark* (2023)
- [14] Bilbao, S., Torin, A., Chatziioannou, V.: Numerical modeling of collisions in musical instruments. *Acta Acustica united with Acustica* **101**(1), 155–173 (2015)
- [15] Issanchou, C., Bilbao, S., Le Carrou, J.-L., Touzé, C., Doaré, O.: A modal-based approach to the nonlinear vibration of strings against a unilateral obstacle: Simulations and experiments in the pointwise case. *Journal of Sound and Vibration* **393**, 229–251 (2017)
- [16] Farnell, A.: *Designing Sound*. MIT Press, Cambridge Massachusetts, USA (2010)
- [17] Skare, T., Abel, J.: Real-time modal synthesis of crash cymbals with nonlinear approximations, using a gpu. In: *Proc. 22nd Int. Conf. Dig. Audio Effects* (2019)
- [18] Mathews, M., Smith, J.O.: Methods for synthesizing very high q parametrically well behaved two pole filters. In: *Proceedings of the Stockholm Musical Acoustics Conference (SMAC 2003)(Stockholm)*, Royal Swedish Academy of Music (August 2003) (2003)
- [19] Poirot, S., Kronland-Martinet, R., Bilbao, S.: A coupled resonant filter bank for the sound synthesis of nonlinear sources. In: *26th International Conference on Digital Audio Effects (DAFx23)* (2023)
- [20] Pruvost, L., Scherrer, B., Aramaki, M., Ystad, S., Kronland-Martinet, R.: Perception-based interactive sound synthesis of morphing solids’ interactions. In:

SIGGRAPH Asia 2015 Technical Briefs, pp. 1–4 (2015)

- [21] Gan, C., Schwartz, J., Alter, S., Mrowca, D., Schrimpf, M., Traer, J., De Freitas, J., Kubiłius, J., Bhandwaldar, A., Haber, N., et al.: Threedworld: A platform for interactive multi-modal physical simulation. arXiv preprint arXiv:2007.04954 (2020)
- [22] Poirot, S.: Sound examples. Available at <https://www.prism.cnrs.fr/publications-media/EURASIPPoirot/> (accessed December 17, 2023)
- [23] Fu, Z.-F., He, J.: Modal Analysis. Elsevier, ??? (2001)
- [24] Graff, K.F.: Wave Motion in Elastic Solids. Dover Publications, New York, USA (1991)
- [25] Fletcher, N.H., Rossing, T.D.: The Physics of Musical Instruments. Springer, ??? (2012)
- [26] Bilbao, S.: Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics. John Wiley & Sons, ??? (2009)
- [27] Aramaki, M., Besson, M., Kronland-Martinet, R., Ystad, S.: Controlling the perceived material in an impact sound synthesizer. *IEEE Transactions on Audio, Speech, and Language Processing* **19**(2), 301–314 (2010)
- [28] Ducceschi, M., Touz , C., Thomas, O., Bilbao, S.: Dynamics of the wave turbulence spectrum in vibrating plates: A numerical investigation using a conservative finite difference scheme. *Physica D* **280**, 73–85 (2014)
- [29] Kronland-Martinet, R., Ystad, S., Aramaki, M.: High-level control of sound synthesis for sonification processes. *AI & society* **27**(2), 245–255 (2012)