Physics- and Psychoacoustics-Based Analysis of Viscoelastically Damped Membranes

Cristhiam Fidel Martínez Orellanos[1], Rolf Bader[2]
[1, 2]*Institute for Systematic Musicology, University of Hamburg, 20354 Hamburg, Germany*
E-Mail: crismartinez@t-online.de

Abstract

Viscoelastic Damping (VD), a type of internal damping characterized by its memory effect, i.e. the ability to store and reintroduce damped energy back into the system, plays an important role in materials like leather and polymers, often used in musical membranes, therefore influencing their sound. However, the relation between the parameters of VD and the perceptual characteristics of the resulting sound remains unclear. Unraveling this relation is, nonetheless, paramount: it could enable us to use VD as a sound-shaping element, fostering the development of new materials, interfaces and tools for musical expression. To this end, this study introduces a physics- and psychoacoustics-based framework for analyzing percussive sound transients, including visualization methods for key psychoacoustic features. The framework is then applied to a physics-informed dataset. We find, spectral dissimilarities are more pronounced in viscoelastically damped frequencies near the fundamental frequency. Measurements also indicate that dissimilarities caused by viscoelastically damping frequencies in the high-frequency region are also perceivable.

1 Introduction

VD causes beating, amplitude and frequency modulation, as well as a nonlinear amplitude response of neighboring frequencies concerning modifications in the parameters that govern its behavior [3, 6]. Despite this complex dynamic, its use in musical situations to shape the sound of a drum [1] or to alter the harmonic relations among a set of drums [2] points towards a clearly perceivable influence of this material property upon the sound. Furthermore, it also demonstrates that the auditory system is able to extract patterns from among such complexities that, put together, are perceived as timbre. However, this ability has been investigated mostly in relation to instrument identification [9] and mostly in relation to harmonic instruments, and not regarding systematic modifications of the material properties of the instrument. To bridge this gap in the literature, this study introduces an analysis framework based on physical characteristics of VD in musical membranes gained from previous research, as well as on temporal and spectral cues relevant for the perception of musical sounds found in the literature. This framework is explained in the following.

1.1 Physics- and Psychoacoustics-Informed Analysis Framework (PPF)

In previous research [3, 6], we have identified important physical characteristics of VD in musical membranes, forming the basis of the analytical approach in this study. Specifically, (a) the amplitude decay of Viscoelastically Damped Frequencies (VD_{freq}), exhibits a non-exponential behavior that begins around 40–50 ms, (b) the amplitude increase of neighboring frequencies around the VD_{freq} due to the memory effect is not linear w.r.t. to changes in the parameter space of VD, (c) the effect of VD has been found to decrease after 120 ms—though this threshold might be frequency-dependent—, meaning that VD influences the sound mostly during its transient, and (d) when modeling VD, there is a minimum number of periods required in the simulation for it to completely unfold.

On the other hand, the perception of a sound relies on the integration of information delivered by the basilar membrane to the brain within particular time intervals, wherein specific acoustic characteristics are extracted and interpreted. The duration of these intervals is determined by the mechanics of the basilar membrane and the physical properties of both the sound and its source. These properties are perceived mainly during the transient of the sound, which in percussion instruments is regarded as highly complex [7]. Since VD acts during the transient, the perceptual cues happening in this time window are relevant for the present analysis. Although psychocoustic analyses of the transient of percussive isntruments are scarce, psychoacoustically relevant time intervals for non-percussive instruments are found in the literature. These temporal cues serve as guidelines for the present study and are explained in the following. The first interval relates to the state of inertia of the ear. Research shows that it takes about 10 ms for the auditory system to define the width of the frequency filters along the basilar membrane. Beyond this threshold changes in the envelope can be perceived, while under it amplitude changes are perceived as changes in timbre [8]. The second interval oscillates between 10-50 ms and is related to the perception of pitch [8]. Also, once the frequency filters in the basilar membrane have been set, their sensibility plays an important role in the perception of timbre. A threshold of 500 Hz has been established [10] below which frequency variations as small as 1.8 Hz are perceivable, while above, the variation limit increases depending on frequency. A third time interval oscillates between 100-150 ms. Above this limit, the sensibility to fluctuations in the frequency components of the sound does not increase significantly [8]. However, if the appearance of shorter fluctuations extends to about 250 ms, the sensation of pitch is not clear and the sound is perceived as noisy. Accordingly, to investigate the influence of VD on perceptual characteristics of the sound of musical membranes, the present work focuses on the transient of the sound (up to 120 ms), with special at-

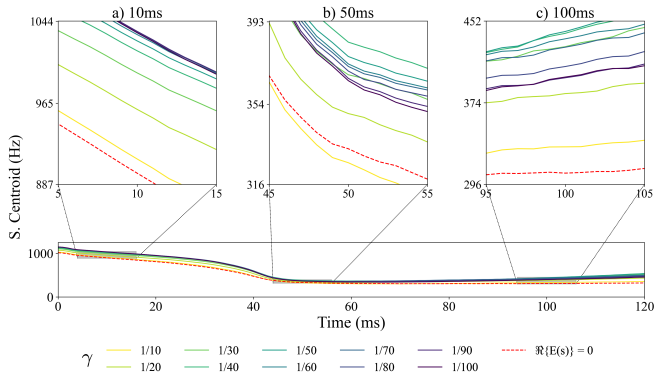

Figure 1: Spectral Centroid of VD_{freq} = 700 Hz - $\Re\{E(s)\} = 0.003$. The color gradient represents the decay rate (γ), with darker colors indicating slower decay. The insets (top row) are placed at t = 10 (a), 50 (b) and 100 ms (c). Although relationships between centroid behavior over time and changes in the rate of decay are visible, this approach does not focus on the time points established in the PPF, and only presents one damping coefficient at a time.

tention to three time points t = 10, 50 and 100 ms.

2 Methods

2.1 Physics-Informed Data Set

A Finite-Differences Time Domain (FDTD) model for a viscoelastically damped membrane was introduced in [3], with four parameters controlling VD behavior (Table 1).

Table 1: Parameters of the FDTD.

Parameter	Description
ω_k	Viscoelastically Damped Frequency (VD_{freq})
$\Re\{E(s)\}$	Damping coefficient of the material
γ	Rate of decay of the memory effect
T	Integration time. Length of the memory effect

A physics-informed data set produced by modifying this parameter space was introduced in [6], simulating a circular membrane with a tension of 2284 Pa, density of 300 kg/m^3, thickness of 3 mm, radius of 10 cm, and fixed boundary conditions ($u_{rim} = 0$). For this study, a subset of 220 time series was selected (see Table 2), each lasting 500 ms with a sampling rate of 96 kHz. The $\Re\{E(s)\} = 0$ is equivalent to the no damping case. A Short Time Fourier Transform (STFT) was applied with window size = 85.3 ms, hop length = 1 ms and a Bartlett window function. Since the highest simulated frequencies in the FDTD are about 15 kHz, for the calculation of the psychoacoustic features for each VD_{freq} (see Section 2.2) only frequency bins up to 15 kHz were considered.

Table 2: Parameter space of the data set.

VDF(Hz)	T(ms)		$Re\{E(s)\}$	γ
700	10	Min.	0	1/100
6400		Max.	0.003	1/10
		Step	0.0003	1/10
		N. of steps	11	10

2.2 Psychoacoustics Features

Spectral Centroid

It is defined as the center of gravity of a frequency spectrum and correlates with the sensation of brightness of a sound. For a discrete signal obtained through a STFT, it is calculated as

$$C(n) = \frac{\sum_{i=1}^{N} X_i k_i}{\sum_{i=1}^{N} X_i}, \tag{1}$$

where n is a discrete time point, and X_i is the magnitude of the frequency bin k_i. Still, this work focuses on the changes caused by modifications in the parameter space at the time points defined in the PPF (see Sect. 1.1). Therefore, an analysis of the entire transient (see Fig. 1) is inefficient because (a) information not related to the PPF must be visually discarded, and (b) it only allows to observe one damping value at a time, losing sight of the possible relations between the entire parameter space. To improve this, a different way of visualizing the spectral centroid is proposed by mapping the centroid values at the required time points to a heatmap as presented in Fig. 2. The blue rectangles found in the top row of Fig. 2 correspond to the inset plots of Fig. 1, demonstrating the way in which information is extracted and compressed for better visualization, allowing a broader comparison across the entire parameter space of each VD_{freq}. Once this method is applied, it becomes easier to visually extract and compare trends within the data. The next step is to understand how VD influences the spectral centroid by analyzing the differences between the frequency spectrum of each sample at each time point from the PPF and the no damping case, for which the spectral flux is used.

Spectral Flux per Mel Band

Also known as spectral dissimilarity, it evaluates differences along the frequency spectrum between frames [7] and is computed as

$$SF(n) = \sum_{m=1}^{m=M} H\left(X_{log,filt}(n,m) - X^{\max}_{log,filt}(n-\mu,m)\right), \tag{2}$$

where $H(x) = \frac{x+|x|}{2}$ is the half-wave rectifier function, $X_{log,filt}$ the logarithmically filtered Mel spectrogram, $X^{\max}_{log,filt}$ the max filtered logarithmic spectrogram, and μ the offset between compared frames [4]. This approach provides a broad overview of spectral dissimilarities during the transient, facilitating the detection of sudden spectral changes. Still, a normalized dissimilarity index like the one produced by Eq. 2 does not offer exact information as to where in the frequency spectrum is VD affecting the magnitude in such a way that it explains the changes in the centroid found in the heatmaps, nor if the changes in amplitude within a particular frequency range are perceivable. To extract this information, Eq. 2 is altered such that the dissimilarity value calculated for each Mel band is kept, both positive and negative dissimilarities are preserved, and the logarithmic compression is changed to the decibel scale to analyze changes in terms of perceivable loudness, thus resulting in

$$SF(n,m) = X_{dB,filt}(n,m) - X_{dB,filt}(n-\mu,m), \tag{3}$$

where $\mu = 1$, which implies a calculation of dissimilarity with respect to the previous time frame $n-1$. For the

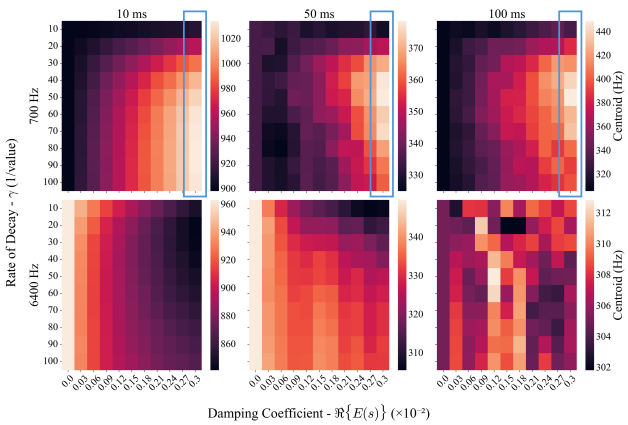

Figure 2: Centroid analysis for VDF$_{freq}$ = 700 Hz and 6400 Hz. Columns correspond to t = 10, 50 and 100 ms (left to right). Heatmaps display damping coefficients increasing left to right and memory effect decay rates decreasing top to bottom. Blue rectangles correspond to the insets from Figure 1, for comparison regarding visual compression of relevant information. At t = 10 ms, damped frequencies exhibit inverse behavior: at 700 Hz, the centroid increases from low damping/fast decay to high damping/slow decay, while at 6400 Hz, it decreases from low damping/slow decay to high damping/fast decay. At t = 50 ms and 100 ms, 700 Hz maps show a stable centroid increase towards high damping/mid decay rates, whereas 6400 Hz maps continue to show a centroid decrease, with smaller reductions for $\gamma \geq 1/60$, though damping coefficient also plays a key role, and a highly nonlinear relationship between the parameter space and the centroid emerges at t = 100 ms for this VD$_{freq}$.

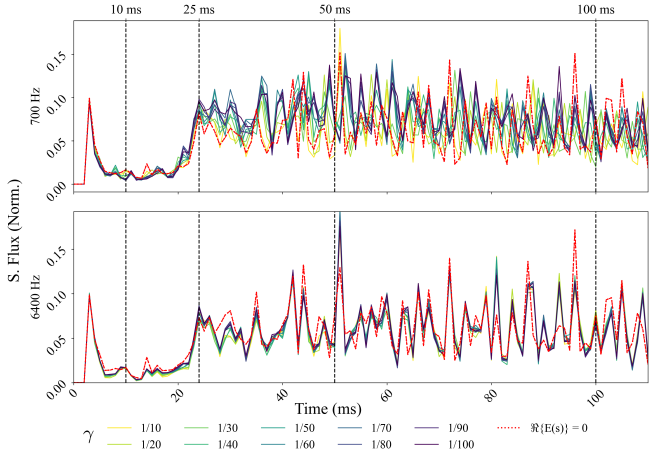

Figure 3: Spectral flux during the transient for VD$_{freq}$ = 700Hz and 6400Hz - $\Re\{E(s)\} = 0.003$. Changes in the rate of decay w.r.t. to the no damping case (red dashed line) produce a stronger flux in the 700Hz case, as represented by the color gradient (the darker the color the slower the rate of decay). At t = 10 and 50 ms (the first and third vertical dashed lines from left to right), both cases exhibit similar flux values. However, the highest increase in flux is driven by differing rates of decay in each case. At t = 100 ms (right most dashed line) the 700 Hz case presents a lot more variability. A significant increase is observed at t = 25 ms (second dashed line from left to right) for both cases.

present study, 1/4 tone Mel bands were applied. These changes result in a flux calculation per Mel band measured in decibels.

In summary, the spectral centroid, overall spectral flux, and spectral flux per mel band are analyzed across the entire data set, specifically at the time intervals defined in the PPF given their psychoacoustic relevance. The results of each analysis are presented below.

3 Results and discussion

3.1 Spectral Centroid

Results are shown in Fig. 2. At t = 10 ms, both VD$_{freq}$ show an inverse behavior; at 700 Hz, there is an increase in centroid from low damping coefficient with a fast rate of decay (top left) towards high damping coefficient with a slow rate of decay (bottom right), while at 6400 Hz there is a decrease in centroid from low damping/fast decay to strong damping/slow decay. Furthermore, at t = 50 and 100 ms, a stable pattern of increased centroid towards the high damping/mid decay space is observed in the 700 Hz maps, while at 6400 Hz there is a decrease in the centroid values again from low damping/fast decay to high damping/slow decay, where samples with $\gamma \geq 1/60$ present mostly a smaller decrease, specially around mid damping values. Lastly, a highly nonlinear relation of the parameter space w.r.t. the centroid is observed at t = 100 ms for the 6400 Hz case.

Overall, these results suggest that the distance between the VD$_{freq}$ and the fundamental frequency plays an important role on how the centroid is influenced. This could be explained by the fact that there is more energy present in the lower modes, leading to more energy stored by the memory effect and consequently more energy injected back into the system. Furthermore, the trend observed in the 700 Hz of increasing centroid towards the middle right zone of the heatmap (high $\Re\{E(s)\}$ and middle γ) could also be explained by the memory effect: a higher damping coefficient, which generates more energy to be stored, combined with a middle rate of decay, which manages the injection of energy in such a way that it stays longer within the system, seems to heavily influence the spectral content of the sound. The opposite trend observed in the 6400 Hz case suggests that VD leaks the energy from the high frequency range towards a lower part of the spectrum in a different dynamic, namely by first generating as much energy as possible (high $\Re\{E(s)\}$) and then injecting it in a very fast way (low γ), thus lowering the centroid.

3.2 Spectral Flux

Fig. 3 shows the spectral dissimilarity for VD$_{freq}$ = 700 Hz and 6400 Hz during the first 120 ms of the sound. Differences in color gradient between both cases reveal that changes in the rate of decay lead to stronger spectral dissimilarities if the VD$_{freq}$ is closer to the fundamental frequency, supporting the results obtained in the centroid analysis. Regarding the time intervals established int the PPF, data suggests that as VD develops both cases exhibit increasing spectral dissimilarity up until t = 50 ms approx., maintaining a high dissimilarity value until t = 100 ms, although at VD$_{freq}$ = 6400 Hz the rate of decay causes less variability w.r.t. to the no damping case. This increase in dissimilarity over time may be explained by the growth in damped energy caused by a strong damping ($\Re\{E(s)\} = 0.003$) and the interdependence between strain and stress, characteristic of viscoelastic materials [5]. Moreover, a significant rise is observed at t = 25 ms in both cases (Fig. 3, second dashed line), resulting in the inclusion of this time point in the PPF. Results of the spectral flux per Mel band method proposed (see Sec. 2.2) are shown in Fig. 4. The highest dissimilarity value corresponds to the 700 Hz case at t = 100 ms (+13 dB

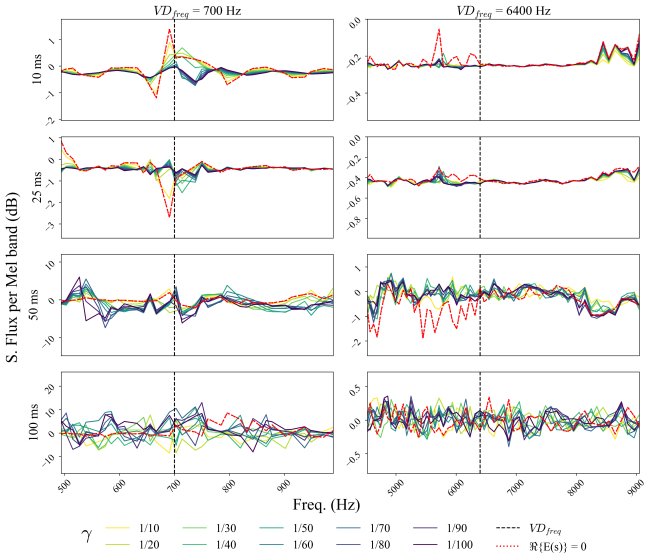

Figure 4: Spectral Flux per Mel band for VD_{freq} 700Hz and 6400Hz - $\Re\{E(s)\} = 0.003$. Time increases from top to bottom. In both cases, the frequency range of the analysis is one octave around the VD_{freq} (black dashed line). In accordance with results shown in Fig. 3, dissimilarity increases along time, showing a progressively nonlinear behavior in both cases. Dissimilarity values in the 700 Hz case reach +13 dB at $t = 100$ ms, and +0.5 dB in the 6400 Hz case at $t = 50$ ms. In the latter, the total dissimilarity value when compared to the no damping case (red line) is +2 dB approx. For $VD_{freq} = 700$ Hz at $t = 50$ ms, while the centroid analysis from Fig. 2 establishes the longest centroid distance between the no damping case and the $\gamma = 1/50$ / $\Re\{E(s)\} = 0.003$ sample, the dissimilarity analysis within the analyzed frequency range presents the highest dissimilarities at $\gamma = 1/60$ and $1/70$.

approx.), while for 6400 Hz, the highest value is +0.5 dB at $t = 50$ ms. However, in the latter case, when compared to the dissimilarity value of the no damping case at the same time point, the total dissimilarity is approximately +2 dB. Furthermore, for $VD_{freq} = 700$ Hz at $t = 50$ ms, while the centroid analysis from Fig. 2 indicates the greatest centroid distance relative to the no damping case at $\gamma = 1/50$, the spectral flux per Mel band analysis shows the highest dissimilarities at $\gamma = 1/60$ and $\gamma = 1/70$ within the analyzed frequency range (one octave around the VD_{freq}), which suggests that the width of the frequency range influenced by viscoelastic damping could be larger than previously found [6]. Additionally, the highly nonlinear dissimilarities observed at $t = 100$ ms for $VD_{freq} = 6400$ Hz match the changes in spectral centroid shown in Fig. 2 for this time point.

In general, the increase in spectral flux per Mel band with increasing time displayed in both cases matches the increment in spectral flux found in Fig. 3. Furthermore, measuring the dissimilarities in decibels across the frequency spectrum reveals that, while damped frequencies that are closer to the fundamental have a stronger impact on the sound, applying viscoelastic damping to higher frequencies can produce a more subtle yet still perceivable alteration in the frequency regions below the damped frequencies.

4 Conclussions

A framework for analyzing the influence of viscoelastic damping in musical membranes has been presented, integrating its acoustic behavior, relevant time intervals for the perception of musical instrument sounds, and tailored analysis strategies for psychoacoustic assessment. Patterns linking the memory effect's energy input to spectral centroid variations were identified. Furthermore, spectral flux increase over time may explain spectral centroid behavior, particularly in high frequencies. Additionally, results suggest a relationship between the distance of the damped frequency from the fundamental frequency and the intensity of spectral dissimilarity, which could be due to how the damping coefficient and the rate of decay store and manage the damped energy. Measurements also indicate that dissimilarities caused by damped frequencies in the high-frequency region are also perceivable. However, in future work, further analysis of other parts of the frequency spectrum as well as other psychoacoustic features and comparisons against listening tests are needed to assess the validity of these hypotheses.

References

[1] Alexandraki, C., Starakis, M., Zervas, P., & Bader, R.: Inferring drumhead damping and tuning from sound using finite difference time domain (FDTD) models. *Acoustics* 5 (3) (2023), 798–816. https://doi.org/10.3390/acoustics5030047

[2] Bader, R.: Finite-Difference model of mode shape changes of the Myanmar pat wain drum circle using tuning paste. *Proc. Mtgs. Acoust.* 29 (1) (2016). https://doi.org/10.1121/2.0000450

[3] Bader, R.: Spectrally-Shaping Viscoelastic Finite-Difference Time Domain (FDTD) model of a membrane. International Journal of Informatics Society 12 (2) (2020), 81–93.

[4] Böck, S., & Widmer, G.: Maximum Filter Vibrato Suppression for Onset Detection. Proceedings of the 16th International Conference on Digital Audio Effects (DAFx-13), 55–61 (2013).

[5] Christensen, R. M.: Theory of viscoelasticity. An introduction. 2nd edition. Academic Press. New York, 1982.

[6] Martínez Orellanos, C. F., & Bader, R.: Analysis of nonlinear behavior of viscoelastic damping in musical membranes using physics-informed self-organizing maps. *APL Machine Learning* 3 (1) (2025). https://doi.org/10.1063/5.0242985

[7] Masri, P.: Computer Modeling of Sound for Transformation and Synthesis of Musical Signals. Ph.D. thesis, University of Bristol, 1996.

[8] Reuter, C.: Der Einschwingvorgang nichtperkussiver Musikinstrumente. Vol. 148. Peter Lang, 1995.

[9] Siedenburg, K.: Specifying the perceptual relevance of onset transients for musical instrument identification. *The Journal of the Acoustical Society of America* 145 (2) (2019), 1078–1087. https://doi.org/10.1121/1.5091778

[10] Zwicker, E., & Feldtkeller, R.: Das Ohr als Nachrichtenempfänger. 2nd edition. S. Hirzel, 1967.