Psychoacoustics of Soft Asymmetric Clipping: Validation through Time and Frequency

Mario Alberto Vallejo Reyes
School of Physics, Engineering and Technology, University of York
York, United Kingdom
ckb535@york.ac.uk

[1] ***Abstract— This study aimed to determine whether time and frequency analyses align with existing listening test results on different types of distortions, with a particular focus on soft clipping, which is often perceived as a clean form of distortion. By integrating theories on various distortion types and using existing listening test results as a basis for subjective perception, MATLAB was employed to synthesize and analyse the distortions. Waveform and spectrogram analyses confirmed that soft clipping maintains a significant resemblance to the original signal, supporting its perception as a clean distortion.***

Keywords—Time/Frequency Analysis, Soft Clipping, Sound Design, Audio Distortion, MATLAB, Psychoacoustics.

I. INTRODUCTION

This research explored the perception of soft clipping as a 'clean' form of distortion compared to other types of audio distortion. In this context, 'clean' refers to distortion that listeners perceive as being closest to an original, completely undistorted audio signal. The study began with a review of relevant literature, focusing on different types of audio distortions and listening tests conducted in this field. MATLAB was then used to synthesize examples of various distortions. Finally, time and frequency domain analyses of these examples were performed to demonstrate how soft clipping maintains its 'cleanliness,' even when applied at high intensities.

II. LITERATURE REVIEW

A. *Some Types of Distortion*

1) Hard Clipping: As discussed by Reiss and McPherson (2014), hard clipping occurs when an audio signal exceeds the maximum processing capacity of digital or analog systems, causing portions of the signal that go beyond these limits to be effectively 'clipped,' as shown in Equation 1.

$$f(x) = \begin{cases} -1 & \text{if } Gx \leq -1 \\ Gx & \text{if } -1 < Gx < 1 \\ 1 & \text{if } Gx \geq 1 \end{cases} \tag{1}$$

where *x* denotes the input signal and *G* represents the applied gain. The function returns -1 when $Gx\leq-$, and 1 when $Gx\geq1$. For values of *Gx* between -1 and 1, the output is *Gx*, indicating that no distortion occurs.

This phenomenon causes a sudden and pronounced transition in the waveform of the signal, from unclipped to clipped parts, as shown in Figure 1. This abrupt change contributes to a harsher and more piercing sound character, which is often perceived as unpleasant or rough.

2) Soft Clipping: As explained by Creasey (2016), unlike hard clipping, which introduces a sudden discontinuity in the signal once a certain threshold is exceeded, soft clipping gradually transitions the signal as it approaches and surpasses the threshold, resulting in less abrupt distortion. Equation 2 characterizes a mathematical model for soft clipping:

$$y = \frac{\tanh(cx)}{\tanh(c)} \quad \text{where} \quad -1 \leq x \leq 1 \quad \text{and} \quad c > 0 \tag{2}$$

Mario Alberto Vallejo Reyes works at the Music Research Center at the University of York in York, United Kingdom (email: ckb535@york.ac.uk).

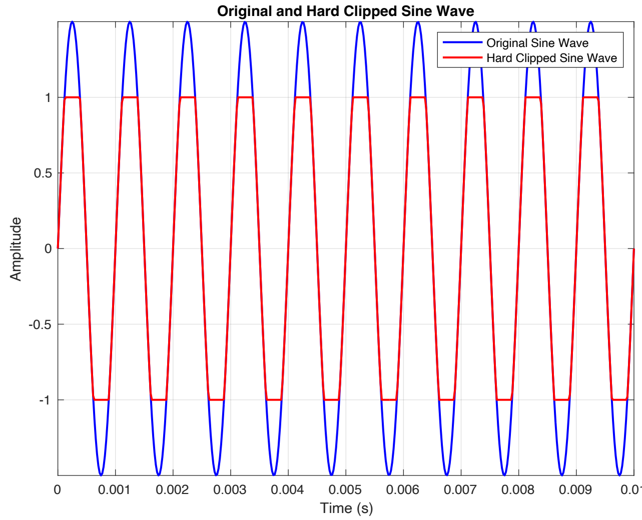

Fig. 1. Visual representation of hard clipping.

where y represents the output signal after soft clipping has been applied to the input signal x. The hyperbolic tangent function, tanh, is used to smoothly limit the amplitude of the output signal. For values of x within the range of $-1 \leq x \leq 1$, the function $\tanh(cx)$ produces an output y that closely follows the input. However, as x moves outside this range, the hyperbolic tangent function ensures that y approaches 1 for $x>1$ and approaches -1 for $x<-1$, thus avoiding the abrupt cutoff characteristic of hard clipping.

The constant c in $\tanh(cx)$ controls how quickly the output signal saturates. A higher value of c results in a more abrupt transition from the unclipped to the clipped signal, which can sound more aggressive, resembling hard clipping. Conversely, a lower value of c produces a smoother transition, preserving more of the signal's dynamic range before saturation. Finally, dividing by $\tanh(c)$ effectively scales the output. This ensures a more controlled and predictable response to varying values of x and c.

In some contexts, depending on the input gain and the specific shape of the transfer function used, this gradual quality of soft clipping is desirable because it allows for more detailed control over the distortion effect, as illustrated in Figure 2.

3) Center Clipping: Center clipping, as described by Giannakopoulos and Pikrakis (2014), is a signal processing technique where an audio signal is modified based on a predefined threshold. This technique involves analyzing and retaining only those samples whose absolute value meets or exceeds the threshold. Samples that fall below this threshold are set to zero. The application of center clipping is mathematically represented in Equation 3:

$$x_c(n) = \begin{cases} x(n) - T_h, & \text{if } |x(n)| \geq T_h, \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

where $x_c(n)$ represents the resulting signal after the application of center clipping at the n-th sample. The function $x(n)$ denotes the value of the original signal at sample n, and T_h is the threshold that determines the level at which clipping is activated. If the absolute value of $x(n)$, indicated by $|x(n)|$, is less than the threshold T_h, then the output $x_c(n)$ is set to zero. Conversely, if the magnitude of the input signal meets or exceeds T_h, the signal passes through unaltered, as shown in Figure 3.

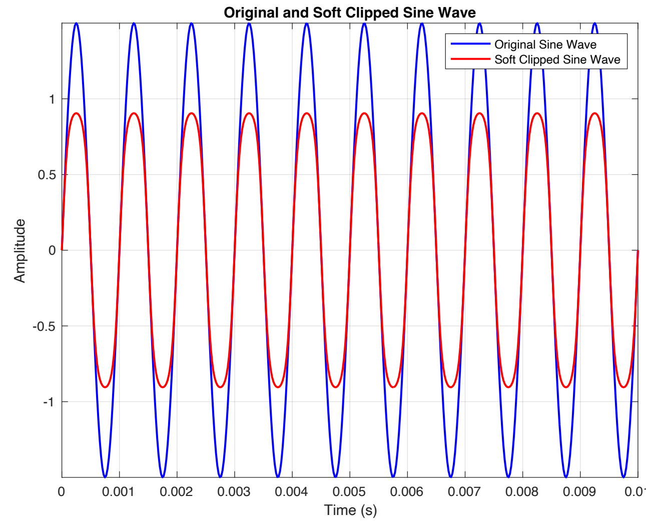

Fig. 2. Visual representation of soft clipping.

4) *Wave Shaping and Full-Range Distortion:* Wave shapers, as explored by Izhaki (2023), are fundamental audio processing tools that apply a transfer curve to a signal according to a specific predefined function, offering an immediate transformation of amplitude. One application of wave shapers is the creation of full-range distortion. Equation 4 represents a common form of this transfer function:

$$y(t) = \alpha x(t)^{\beta} \tag{4}$$

where $x(t)$ is the input signal, $y(t)$ is the output signal, α is a scaling factor, and β is the exponent that determines the degree of non-linearity. The above equation shows how input amplitudes are non-linearly mapped to output amplitudes, as illustrated in Figure 4.

B. Symmetrical vs Asymmetrical Clipping

The concepts of symmetrical and asymmetrical clipping, key forms of signal modification in audio distortion, are studied in Kevin Robinson's analysis (2020). Symmetrical clipping uniformly reduces the peaks of an audio signal. On the other hand, asymmetrical clipping applies uneven clipping to the signal waveform, clipping one side more heavily than the other. This results in a mixture of odd and even harmonics, giving the sound a distinct character. These techniques open various experimental possibilities for sound manipulation, each with unique sonic qualities.

C. Perception of Different Types of Distortions

In their study, Tan, Moore, and Zacharov (2003) explored the impact of different types of nonlinear distortion on how listeners perceive the quality of speech and music signals. The researchers utilized various types of distortions, including hard and soft symmetrical and asymmetrical clipping, center clipping, and full-range waveform distortion. This full-range distortion involved altering the waveform by raising its instantaneous absolute value to a power while preserving the waveform's sign.

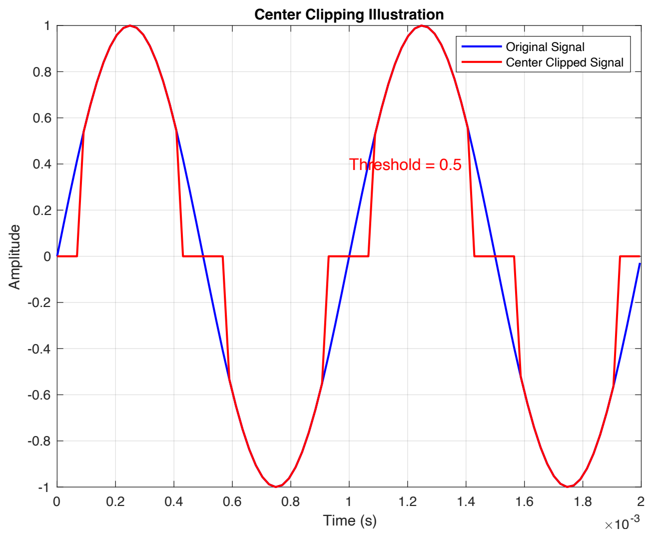

Fig. 3. Signal before and after Center Clipping with a threshold of 0.5

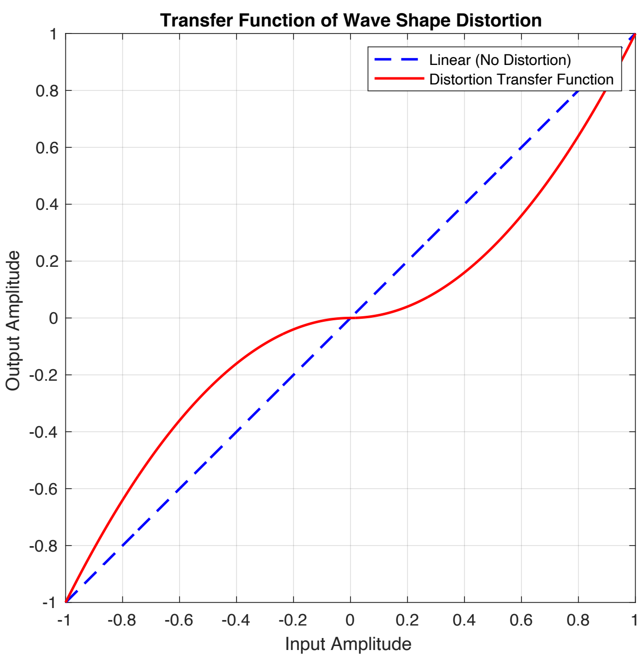

Fig. 4. Transfer function of a wave shaper.

Participants in the study were asked to rate the level of perceived distortion on a scale from 1 to 10, with 1 indicating the most distortion and 10 the least. The research was conducted in two phases: In the first, distortions were applied to broadband signals, while in the second, they were applied to signal sub bands.

The study involved applying various distortions over the audio signals at different intensities. For this work on the psychoacoustics of soft clipping, the focus was on the condition where distortions were applied at their maximum intensity for this study. Under these conditions, as seen in Figure 5, the findings revealed that among the different types of distortion, soft clipping was perceived as the cleanest, causing only minimal alterations in the ratings. This contrasted with full-range distortion, which was perceived as the most severe.

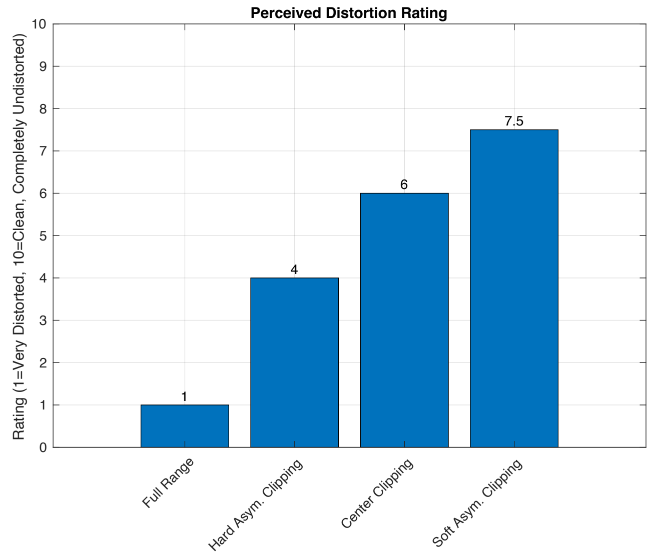

Fig. 5 Perceived distortion ratings.

The subjective assessments made by the listeners aligned well with objective distortion measurements. These objective measures, denoted as DS, were based on the output spectrum of each nonlinear system in response to a multitone signal. The study found a high negative correlation between these objective measures and the subjective ratings, indicating that larger values of DS (denoting more distortion) corresponded to lower subjective ratings (indicating a perception of more distortion).

An additional experiment further confirmed the relationship between subjective perceptions and objective measurements. In this phase, the stimuli with nonlinear distortion were produced by recording the outputs of real transducers, which were then digitally filtered to minimize amplitude-frequency response irregularities. The results showed a moderately strong negative correlation between subjective ratings and the objective DS measure.

II. Design and Implementation

A. *Synthesizing Distortions in MATLAB*

Once the theory on different types of distortion and previous listening tests were analysed, MATLAB was used to synthesize various types of distortions applied to the musical piece 'Sakura' (2023) by artist Ruddi Nizz. This section details the technical operation of the MATLAB functions coded to synthesize the different types of distortions.

Each function begins by reading and normalizing the audio file to standardize amplitudes. After applying the specific distortion, the signal is normalized again to keep amplitudes within appropriate limits, and a new file is saved. Distortion parameters, such as thresholds and impact percentages, were established to replicate those used by Tan, Moore, and Zacharov.

`applySoftAsymmetricalClipping` was the function developed to apply soft distortion to the positive amplitudes of an audio signal that exceed a calculated threshold. This threshold is determined based on the percentage of samples intended to exceed it, identified by `exceedPercentage`, using quantile calculation on the positive samples of the normalized signal.

Once the threshold is established, the function modifies the samples that exceed it. Soft clipping is performed by adding to the threshold the result of the arctangent function applied to the difference between the sample and the threshold. This operation adjusts only those samples above the threshold, smoothing the peaks by transforming the excess amplitude non-linearly.

The `applyHardAsymmetricalClipping` function applies hard clipping only to the highest parts of the audio signal that exceed a determined threshold. This threshold is defined so that only a specific percentage of the highest positive samples is affected by clipping.

To calculate the threshold, the `exceedPercentage` parameter is used, representing the percentage of samples desired to exceed the threshold. To identify this threshold, the corresponding quantile of the set of normalized positive samples is calculated. Once the threshold is established, the next step is to modify the samples that exceed it. These samples are adjusted directly to the threshold, thus effecting hard clipping.

The `applyFullRangeDistortion` function was designed to modify the amplitude of audio signals by applying distortion based on an alpha parameter. This parameter defines how the amplitude of each sample is altered: the amplitude is adjusted by raising each sample to the power of alpha, directly affecting the signal's dynamics.

If alpha is greater than 1, the signal peaks become more pronounced, intensifying the differences in the signal's dynamics. Conversely, an alpha less than 1 reduces the variability of amplitudes, compressing the signal. This distortion method doesn't select specific parts of the signal for modification but uniformly affects all samples, applying the same degree of transformation throughout the entire signal.

The `applyCenterClipping` function executes the process of center clipping distortion on audio signals. This process begins by calculating the RMS value of the signal, which is used to establish the threshold. The threshold is defined using the `clippingPercentage` parameter, which determines the proportion of the RMS value that will delineate the range of amplitudes to be clipped.

Center clipping is applied directly to the signal: all samples with absolute amplitudes less than this calculated threshold are set to zero, while samples with amplitudes above this threshold remain unchanged.

B. Analysis in the time and frequency domains.

Two functions were coded in MATLAB to analyse the impact of different distortion techniques on audio signals. The first function facilitates spectral analysis, while the second allows for a visual comparison of waveforms.

The spectral analysis function evaluates the frequency alterations in the audio signal induced by the distortions. This function normalized the signals to the RMS value, inverted the phase of the distorted signals, and summed them with the original signal to highlight the induced differences. Spectrograms were generated by adjusting parameters such as window size and overlap to obtain a detailed representation of the frequency changes.

Simultaneously, the waveform comparison function was implemented to directly visualize the modifications in amplitude and the temporal structure of the signals. This function extracted and visually compared specific segments of the original signal and the distorted signals in subplots, displaying the variations introduced by each type of distortion.

III. Results

Regarding the spectral analysis, Figure 6 presents four subplots. Each subplot shows the resulting spectral differences obtained by adding the original audio mix with its distorted counterpart (with inverted phase).

The subplot comparing the original mix with the soft clipping version (top left) reveals minimal spectral differences, particularly in the transients. The psychoacoustic effect of soft clipping being perceived as relatively clean is further reinforced compared to the other subplots. These additional graphs demonstrate the differences between the original mix and the versions subjected to hard clipping, full-range distortion, and centre clipping. In these cases, the spectral deviations from the original mix are more pronounced, highlighting how these distortions are more readily perceived as distorted.

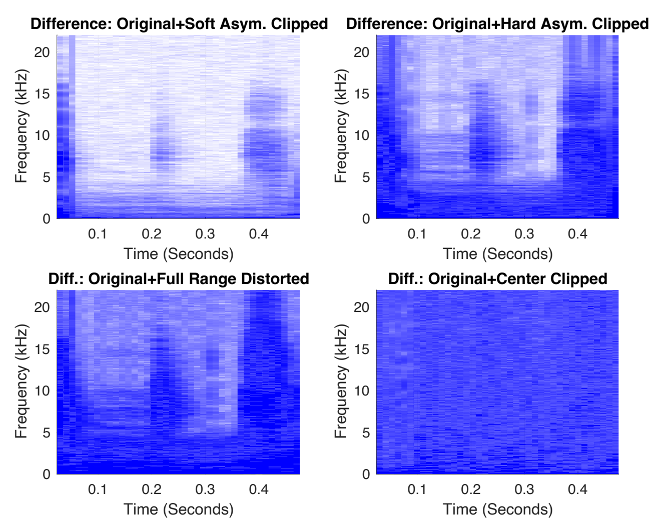

Fig 6. Spectrograms showing the difference between original and distorted audio

Another approach to demonstrating the relatively non-intrusive nature of soft asymmetric clipping involves analysing waveform alterations through various distortion processes. Figure 7 shows four graphs, each representing a segment of the original mix's waveform (blue line) juxtaposed with its processed counterpart (red line) using different distortion techniques.

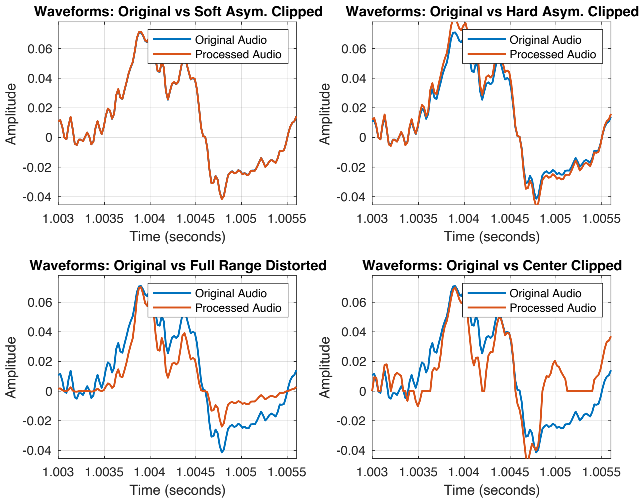

Fig 7. Waveform comparisons between original and distorted audio

The comparison between the original mix and the soft clipping mix is illustrated in the top left graph. Here, the waveforms are almost indistinguishable at first glance, with only slight discrepancies appearing upon close inspection, particularly at the peaks where they do not perfectly overlap. The top right graph contrasts the original mix with the hard clipping version. Although the waveforms appear generally similar, the hard-clipped waveform exhibits notable variations, especially at points far from the zero-amplitude line.

The bottom left graph shows the original mix alongside the full-range distorted version. This comparison reveals a marked contrast between the two waveforms, with many segments compressed towards the zero line, resulting in a significant loss of dynamic range. Finally, the bottom right graph compares the original mix with the signal subjected to center clipping. This graph demonstrates the most significant divergence between the original and distorted signals. Large portions of the waveform are altered, with segments consistently flattened to a zero amplitude and the emergence of new shapes, significantly altering the characteristics of the signal.

IV. CONCLUSIONS

This study examined the correlation between time and frequency analyses with the perception of different forms of distortion in audio signals, with a particular focus on soft clipping, which is often appreciated as a clean form of distortion. Using MATLAB to synthesize and analyse distortions, and drawing on results from previous listening tests, it was confirmed that soft clipping retains a notable resemblance to the original signal, supporting its perception as a less invasive form of distortion.

Time and frequency analyses showed that soft clipping primarily affects transient peaks without significantly altering the fundamental characteristics of the sound, supporting its use in applications where preserving the quality of the original sound is desired. In contrast, other forms of distortion, such as hard clipping and full-range distortion, introduce more pronounced changes that alter the perceived sound quality.

These results verify previous hypotheses about soft clipping and provide a foundation for future research in psychoacoustics and sound design. This work emphasizes the importance of integrating objective measurements with subjective evaluations to gain a comprehensive understanding of the effects of distortion and other phenomena that occur in audio.

V. REFERENCES

Tan, C.-T., Moore, B., & Zacharov, N. (2003). *The effect of nonlinear distortion on perceived quality of music and speech signals*. Journal of the Audio Engineering Society, 51, 1012-1031.

Giannakopoulos, T., & Pikrakis, A. (2014). *Introduction to audio analysis: A MATLAB approach*. Academic Press.

Reiss, J. D., & McPherson, A. (2014). *Audio effects: Theory, implementation and application*. CRC Press.

Creasey, D. (2016). Audio processes: Musical analysis, modification, synthesis, and control. London, England: Routledge.

Robinson, K. (2020). *Practical audio electronics*. CRC Press, London, England.

Izhaki, R. (2023). *Mixing audio: Concepts, practices, and tools* (4ta ed.). Oxford, England: Focal Press.

Ruddi Nizz (2023). Sakura [Song]. YouTube Music. https://music.youtube.com/watch?v=RWkBvb5KTcc&si=eiVhQlHbEa3Qf19T