COMPLEX NONLINEARITIES FOR AUDIO SIGNAL PROCESSING

Jatin Chowdhury

Center for Computer Research in Music and Acoustics
Stanford University
Palo Alto, CA
jatin@ccrma.stanford.edu

ABSTRACT

We present an ongoing study of new and interesting nonlinear structures for audio signal processing, intended to be used for audio effects and synthesis. We give a brief discussion of each structure, and present a series of open-source audio plugins that implement the structures.

1. INTRODUCTION

In digital audio signal processing it is common to find audio effects that use nonlinear elements to add harmonic content to the signal being processed, or to achieve a "distortion" type of effect. Typically, this is done either as part of an analog model, or using a static memoryless nonlinear element.

The goal of this research project is to develop structures for nonlinear audio signal processing that go beyond the traditionally used simple nonlinearities. While the structures developed here may be used for analog modelling and may be inspired by analog effects, they do not come about from direct physical modelling of an analog system, nor do they require knowledge of analog systems such as circuits to be understood and implemented.

1.1. Simple Nonlinearities

We refer to the desired nonlinear structures as "Complex Nonlinearities", as such we should take a moment to define what constitutes a "simple" nonlinearity, particularly since these will make up the building blocks of the complex nonlinearities that follow.

1.1.1. Saturators

The most commonly used nonlinearity in audio signal processing is the saturating nonlinearity, where the input "clips" to a constant value as the input gain increases. This class of nonlinearity includes functions such as the hard clipper, cubic soft clipper, and $tanh$ nonlinearities [1], which are described by the following functions:

$$f_{\text{hard clip}}(x) = \begin{cases} -1 & x < -1 \\ x & -1 \leq x \leq 1 \\ 1 & x > 1 \end{cases} \tag{1}$$

$$f_{\text{soft clip}}(x) = \begin{cases} -\frac{2}{3} & x < -1 \\ x - \frac{x^3}{3} & -1 \leq x \leq 1 \\ \frac{2}{3} & x > 1 \end{cases} \tag{2}$$

$$f_{\text{tanh}}(x) = \tanh(x) \tag{3}$$

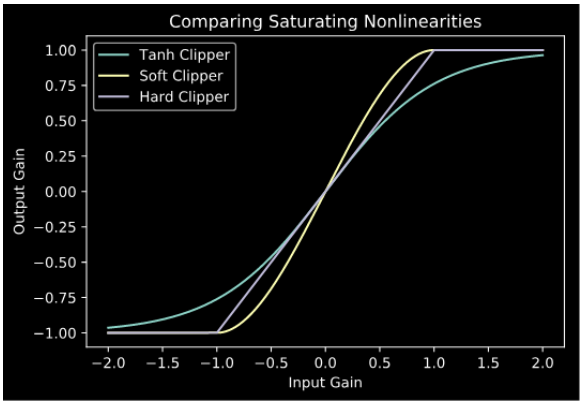

Figure 1: *Saturating Nonlinearities*

1.1.2. Rectifiers

Sometimes for audio effects such as compressors and limiters, it is useful to have a rectified signal (i.e. a signal that only contains non-negative values). The two most simple rectifying nonlinearities are the full-wave rectifier and the half-wave rectifier:

$$f_{\text{FWR}}(x) = |x| \tag{4}$$

$$f_{\text{HWR}}(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases} \tag{5}$$

The above rectifier equations have a downside in that they do not have continuous derivatives. As a potential alternative, we present another half-wave rectifier equation loosely modelled from a Shockley diode rectifier [2]:

$$f_{\text{Diode}}(x) = \beta \left(e^{\alpha x} - 1\right) \tag{6}$$

where α and β are tunable parameters.

1.1.3. Dropout

Another nonlinearity used in audio effects is the "dropout" nonlinearity, notably used for modelling underbiased analog tape recording [3]. A dropout nonlinearity is characterized by the fact that small input values are snapped to zero. Here we present a simple cubic dropout function:

$$f_{\text{Dropout}}(x) = \begin{cases} x + B - \left(\frac{B}{a}\right)^3 & x < -B \\ \left(\frac{x}{a}\right)^3 & -B \leq x \leq B \\ x - B + \left(\frac{B}{a}\right)^3 & x > B \end{cases} \tag{7}$$

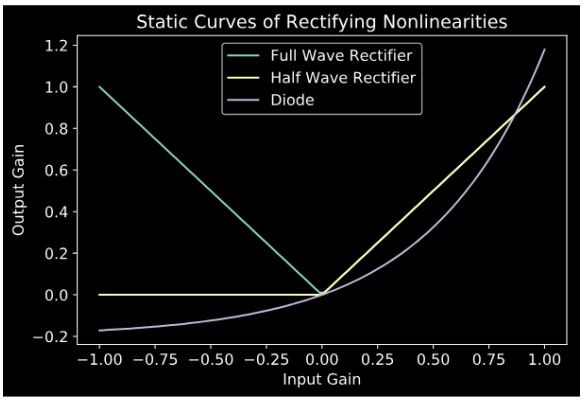

Figure 2: *Rectifying Nonlinearities. For the diode nonlinearity, we use* $\beta = 0.2$, $\alpha = 1.93$.

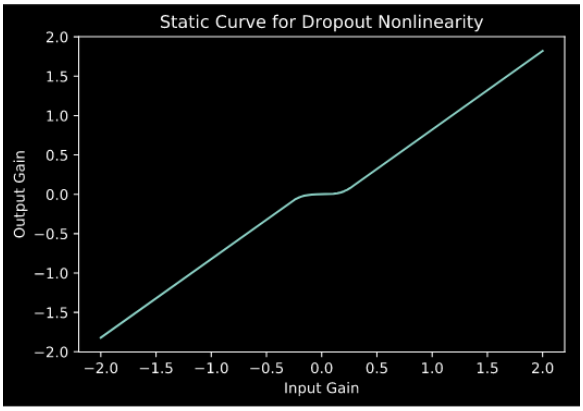

Figure 3: *Dropout nonlinearity with* $a = 0.6$.

where a is a tunable parameter, and $B = \sqrt{\frac{a^3}{3}}$.

2. DOUBLE SOFT CLIPPER

The Double Soft Clipper (see fig. 4) is a sort of combination between a saturating nonlinearity and dropout nonlinearity. The nonlinear function is given as:

$$f_{\text{DSC}}(x) = \begin{cases} 1 & u \geq 1 \\ (3/4) * (u - u^3/3) + 0.5 & 0 < u < 1 \\ (3/4) * (u - u^3/3) - 0.5 & -1 < u < 0 \\ -1 & u \leq -1 \end{cases} \quad (8)$$

where

$$u(x) = \begin{cases} x - 0.5 & x > 0 \\ x + 0.5 & x < 0 \end{cases} \quad (9)$$

The resulting static curve is essentially two cubic soft clipping functions stacked on top of each other (see fig. 5). One advantage of this nonlinear function is that it is highly parameterizable. Possible parameters include upper/lower clipping limit, linear slope, upper/lower skew, and dropout width (see fig. 6).

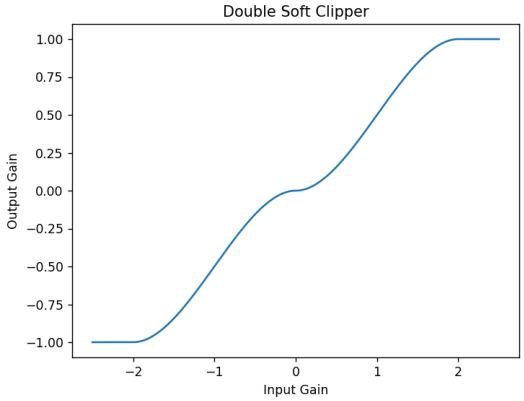

Figure 4: *Double soft clipper.*

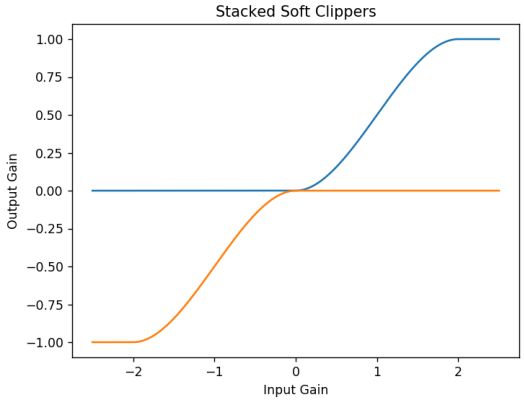

Figure 5: *Stacked soft clippers.*

3. HARMONIC EXCITER

A harmonic exciter is an audio effect often used by mixing engineers to add "brightness" to a track or a mix, for example the Aphex Aural Exciter [4]. Typically, digital exciter effects are implemented as circuit models of an original analog effect (e.g. [5]). The goal of this work is to create a generalized exciter model that can be implemented without knowledge of circuit theory, much in the way that [6] describes a generalized model of a dynamic range compressor.

Based loosely on the Aphex Aural Exciter, we propose the exciter architecture shown in fig. 7. For the generator component, we propose the processing architecture shown in fig. 8.

For the level detector component, we propose using a rectifying nonlinearity followed by a lowpass filter with a very low cutoff frequency (e.g., $f_c > 30$ Hz). Figure 9 shows the output of a level detector with the rectifying nonlinearities described in §1.1.2 and a first-order lowpass filter with a $f_c = 10$ Hz. For the nonlinearity component, any saturating nonlinearity of the type described in §1.1.1 will suffice. The output of the generator is very harmonically rich, with both even and odd harmonics, as seen in fig. 10.

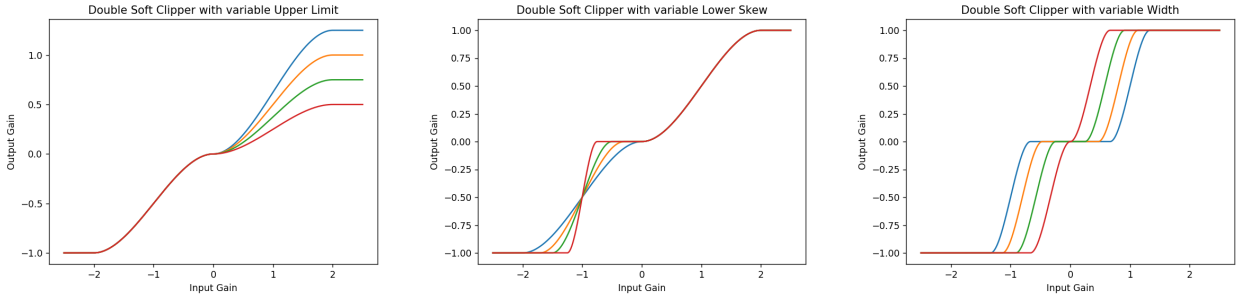

Figure 6: *Double soft clippers with variable upper clipping limit (left), variable lower skew (middle), and variable dropout width (right).*

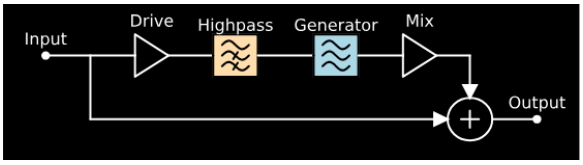

Figure 7: *Exciter architecture.*

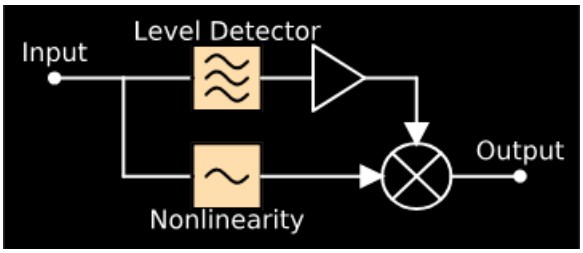

Figure 8: *Exciter generator architecture.*

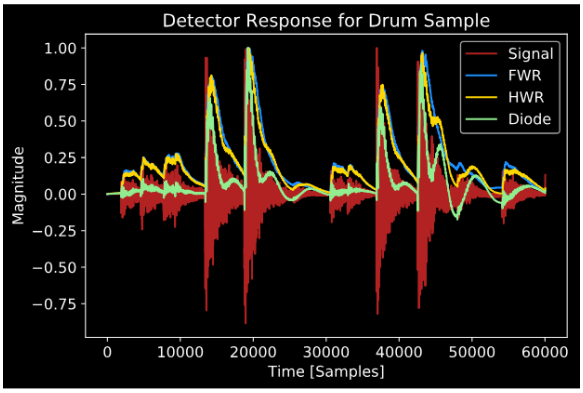

Figure 9: *Exciter level detector.*

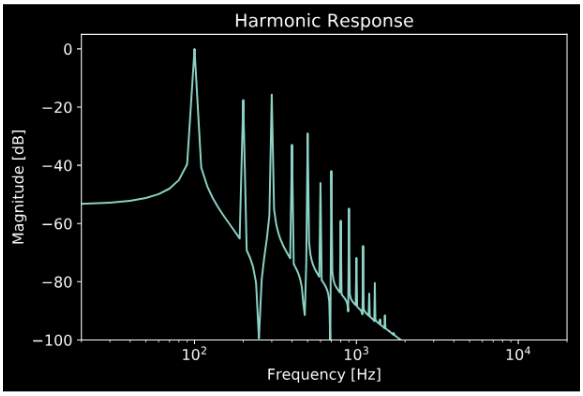

Figure 10: *Exciter generator harmonic response.*

4. HYSTERESIS

Hysteresis is an interesting complex nonlinear behavior that describes the magnetising of magnetic materials, as well as concepts in other fields including chemistry, structural engineering, even economics. [3] uses an adaptation of the Jiles-Atherton magnetic hysteresis model to recreate the sound of an analog tape machine. In this study, we attempt to generalize this hysteresis model to be used and implemented without understanding of electromagnetic physics, as well as adding useful parameters to the hysteresis nonlinearity.

The hysteresis model for input x is defined by the differential equation:

$$\dot{y} = \frac{\frac{(1-c)\delta_y(SL(Q)-y)}{(1-c)\delta_x k-\alpha(SL(Q)-y)}\dot{x} + c\frac{S}{a}\dot{x}L'(Q)}{1 - c\alpha\frac{S}{a}L'(Q)} \tag{10}$$

where $\dot{y}$ denotes the time derivative of the output, and k and a are constant values. δ_x and δ_y are defined as

$$\delta_x = \begin{cases} 1 & \text{if } x \text{ is increasing} \\ -1 & \text{if } x \text{ is decreasing} \end{cases} \tag{11}$$

$$\delta_y = \begin{cases} 1 & \text{if } \delta_x \text{ and } L(Q) - y \text{ have the same sign} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

$L(x)$ denotes the Langevin function:

$$L(x) = \coth(x) - \frac{1}{x} \tag{13}$$

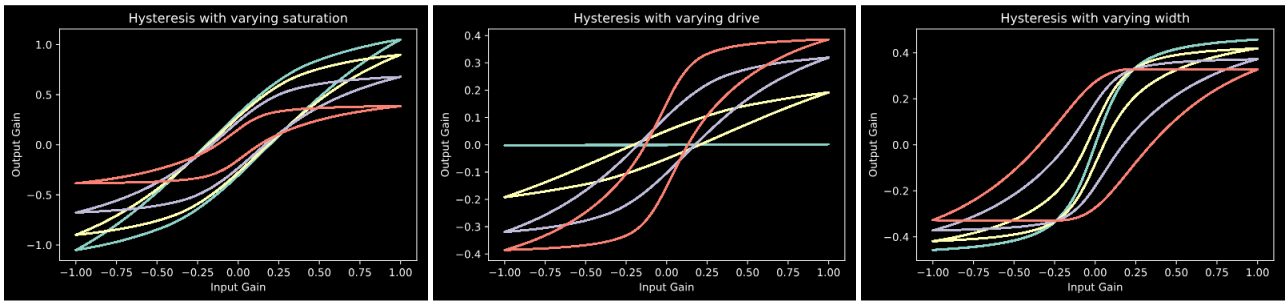

Figure 11: *Hysteresis curves with variable saturation (left), drive (middle), and width (right).*

and Q is defined as

$$Q(x, y) = \frac{x + \alpha y}{a} \tag{14}$$

where α is a constant value, and x and y are the system input and output respectively.

The variables S, a, and c from eq. (10) can be used as parameters to control the hysteresis function saturation, drive, and width respectively. Figure 11 shows the effects of modulating these parameters.

5. NONLINEAR BIQUADS

Transposed Direct Form II (TDF-II) (see fig. 12) is a standard method for implementing biquad filters in signal processing. We present two methods for adding nonlinear elements to the TDF-II structure that can create interesting sonic effects without affecting the filter's stability.

For the first method we propose adding nonlinear elements before the delay elements in the filter structure, as shown in fig. 13. When saturating nonlinearities are used for these elements, the filter exhibits "nonlinear resonance" similar to many analog filters. The filter frequency response at various operating points is shown in fig. 15.

For the second method we propose adding nonlinear elements to the feedback paths of the filter, as shown in fig. 14. Again, with saturating nonliearities, this structure causes the filter poles to "sweep" to different frequencies as the input level of the filter varies. The filter frequency response at various operating points is shown in fig. 16.

In order to maintain stability, we propose that the nonlinear elements must satisfy the constraint that $|f'_{NL}(x)| \leq 1$ (note that this implies that the derivative of the nonlinear function must exist everywhere in the range of x). A proof of this constraint using Lyapunov stability, as well as a more in-depth analysis of these nonlinear filter structures, including the potential use of these structures for analog modelling, is given in [7].

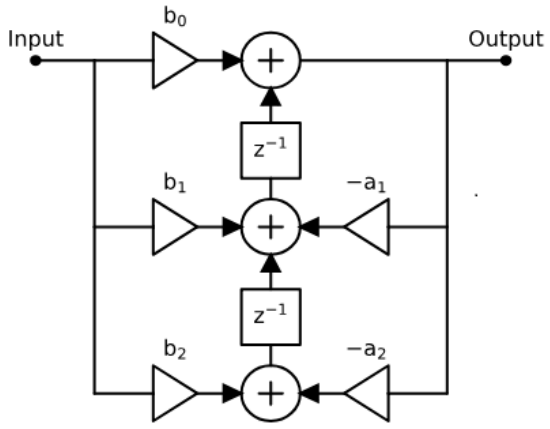

Figure 12: *Transposed direct form II filter structure.*

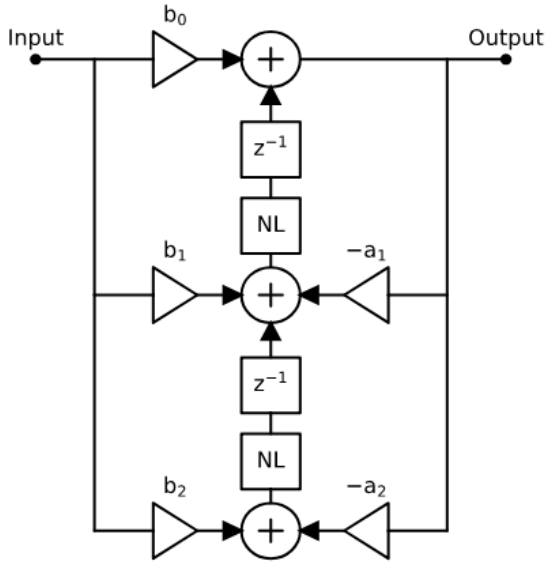

Figure 13: *Transposed direct form II with nonlinear resonance.*

6. WAVEFOLDING

Wavefolding is a sonically interesting nonlinear effect, where as the input level of the waveform increases, the wave seems to "fold

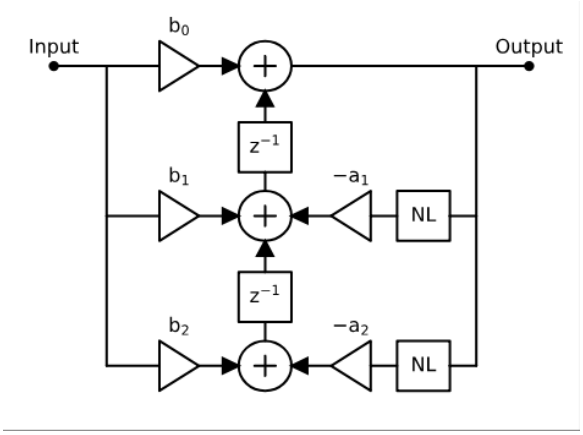

Figure 14: *Transposed direct form II with nonlinear feedback.*

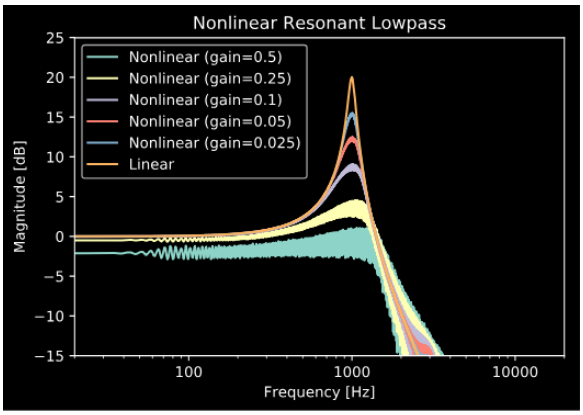

Figure 15: *Frequency response of nonlinear resonance filter at various operating points.*

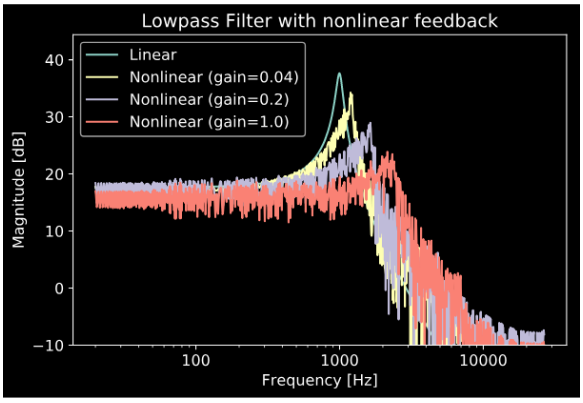

Figure 16: *Frequency response of nonlinear feedback filter at various operating points.*

over" on itself. This effect is especially popular in certain analog synthesizers. Circuit models of wavefolding effects have previously been developed in [8, 9]. In this work we attempt to create a purely digital wavefolder, with some intruiging extensions on the traditional wavefolding effect.

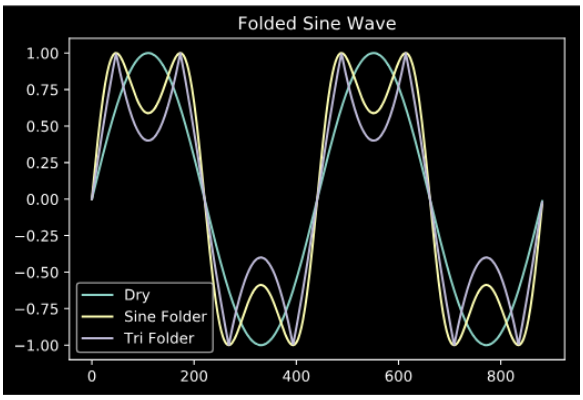

Figure 17: *The output of simple sine and triangle wavefolders.*

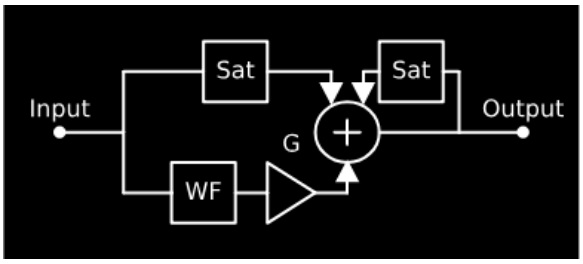

Figure 18: *Architecture for a saturating feedback wavefolder.*

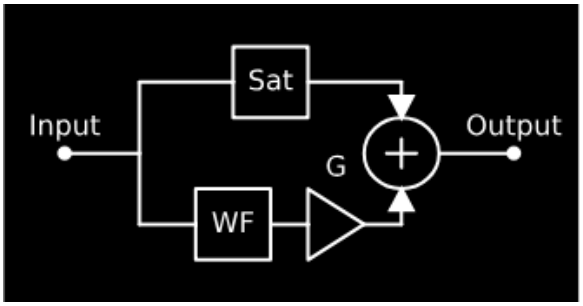

Figure 19: *Architecture for a saturating wavefolder.*

We begin with a well-known oversimplified wavefolding nonlinearity: $f_{NL}(x) = \sin(x)$ (other waveforms could replace the sin wave used in this formula, e.g. a triangle wave). The output of a sine and triangle wavefolder are shown in fig. 17.

6.0.1. Saturating Wavefolder

We can improve upon the simple wavefolder by summing its output with the output of a saturating nonlinearity, with the gain of the wavefolder output smaller than the gain of the saturator output. This architecture (including a variable gain G after the wavefolder) is shown in fig. 19; the static curve of the saturating wavefolder nonlinearity with $G = -0.2$ is shown in fig. 20. Finally, in fig. 21 we show the response of the simple sine wavefolder and the saturating wavefolder for a very large input signal. Note that while the simple wavefolder allows the signal to flip all the way from

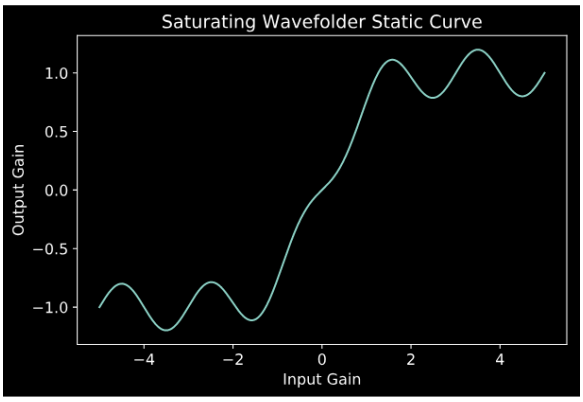

Figure 20: *Static curve for a saturating wavefolder.*

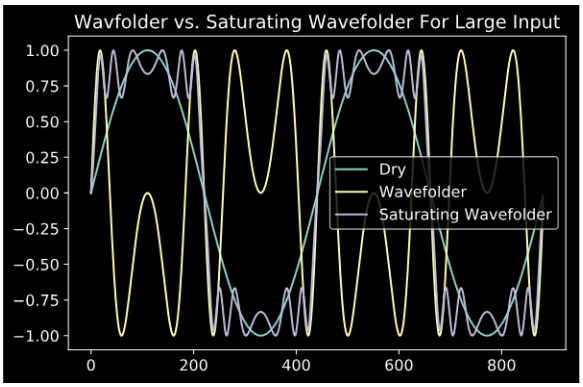

Figure 21: *Comparison of wavefolder outputs for large input.*

positive to negative within a half cycle of the dry signal, the saturating wavefolder does not. In this way, the saturating wavefolder exhibits behavior more similar to traditional analog wavefolders, and is sonically less "harsh".

6.0.2. Saturating Feedback Wavefolder

We can modify our wavefolder architecture further, by adding a feedback path with another saturating nonlinearity, as shown in fig. 18. The resulting nonlinear function has an interesting textural characteristic, almost a sonic "squishiness" caused by the feeding back of the already folded wave. This behavior is hinted at by the dynamic curve of the saturating feedback wavefolder (see fig. 22), but for a better understanding, we recommend listening to the audio examples presented in §9.

7. SUBHARMONICS GENERATOR

From a frequency-domain perspective, most of the nonlinear processors discussed thus far serve to create higher frequency content than that of the input signal, often a useful aspect of many distortion and exciter effects. In this section, we look at developing a nonlinear system for generating lower frequency content, or subharmonics. Most existing methods for generating subharmonics often rely on spectral domain processing (e.g. [10]), which can be difficult to use as a real-time effect, or rely on isolation of the fundamental frequency for which to generate the subharmonic as in

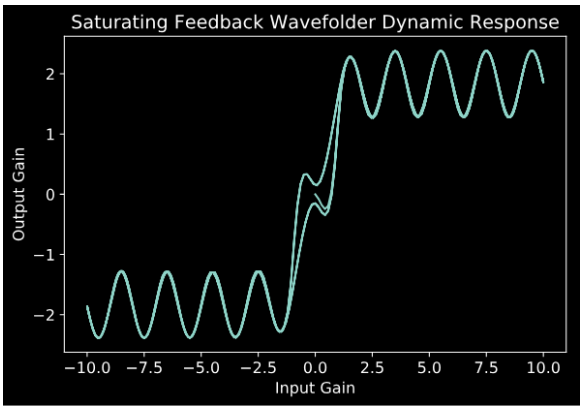

Figure 22: *Dynamic curve response of saturating feedback wavefolder.*

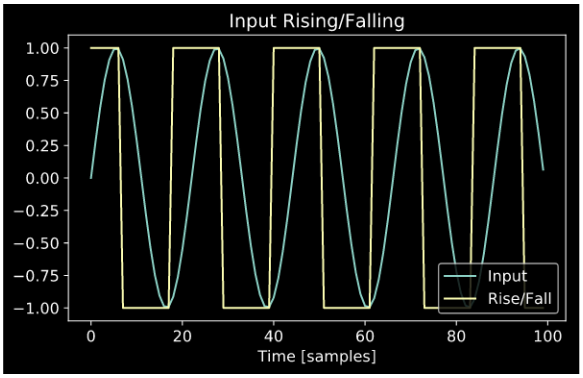

Figure 23: *Output of switching detector for an input sine wave.*

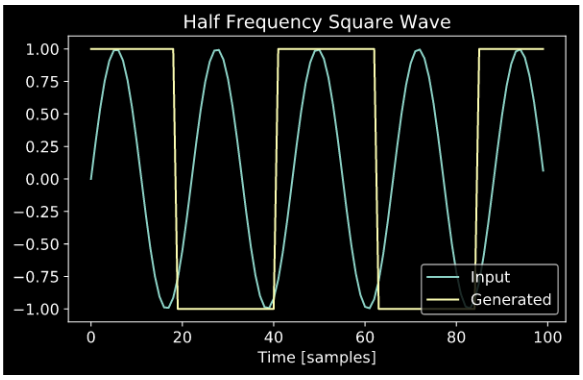

Figure 24: *Output of half frequency square wave generator for an input sine wave.*

[11]. Here we strive to develop a simpler, yet more general technique that can run in real-time, and be useful on variety of input signals (i.e. even when we are not able to isolate the fundamental).

First, we develop a simple switching detector, that can detect when the input signal switches directions. If we have our switch detector output a 1 when the signal is increasing, and a -1 when the signal is decreasing, we can generate a square wave of the same

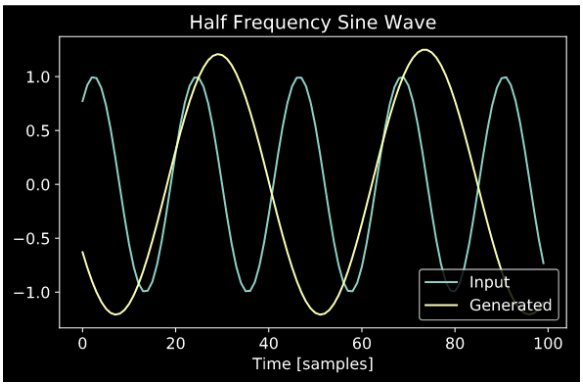

Figure 25: *Output of half frequency sine wave generator for an input sine wave.*

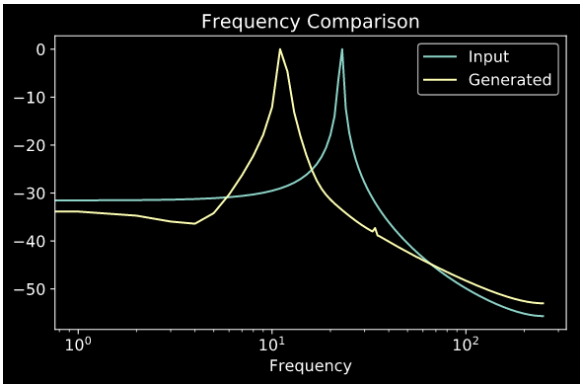

Figure 26: *Comparison of input and output sine waves.*

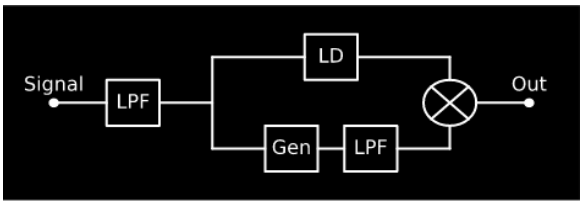

Figure 27: *Full architecture of a subharmonic generator effect.*

frequency as the input, as shown in fig. 23. Next we can adjust our generator to only flip the signal every *other* time a switch is detected in the incoming signal, thereby generating a square wave with half the frequency of the original signal (see fig. 24). Finally, we can use a lowpass filter to smooth our output signal into something more like a sine wave (see fig. 25). To verify our subharmonic generator we show the frequency spectra of the input and output signals in fig. 26.

To make our subharmonic generator more usable as an audio effect, we can attach a level detector and an input lowpass filter as shown in fig. 27. The level detector ensures that the generator output follows a similar amplitude envelope to the input signal, although in practice adjusting the attack and release of the level detector (similar to a dynamic range compressor) can create an interesting effect where the subharmonic outlasts the original signal, adding a "boomy" quality to the overall sound. The input lowpass filter is meant to limit any high frequency noise that might cause the switching detector to generate signal above the desired subharmonic frequencies. In practice, this effect can be used to add low frequency content to musical signals including drum sets, bass guitars, keyboards/synthesizers, and more.

8. GATED RECURRENT DISTORTION

The field of machine learning, and particularly deep learning, has benefitted greatly from the clever use of nonlinear functions including the hyperbolic tangent, sigmoid function, rectified linear unit (ReLU), and more. In recent years, a unit known as the Gated Recurrent Unit (GRU) has become popular for use in recurrent neural networks [12]. In this section we examine the use of a simplified form of the GRU, known as a "minimal gated unit", as a parameterizable distortion effect.

The operation of the minimal gated unit, as introduced in [13], can be described by the following equations:

$$\begin{aligned} \Gamma_f &= \sigma(W_f x[n] + U_f y[n-1] + b_f) \\ y[n] &= \Gamma_f y[n-1] \\ &\quad + (1 - \Gamma_f) \tanh(W_h x[n] + U_h \Gamma_f y[n-1] + b_h) \end{aligned} \tag{15}$$

where x is the input, y is the output, and $W_h, W_f, U_h, U_h, b_h, b_f$ are all free parameters. $\sigma(x)$ denotes the sigmoid function:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{16}$$

Typically, when used in a recurrent neural network, the elements of the above equations are vectors and matrices, and several of the multiplications are instead Hadamard products, however given the use of the minimal gated unit presented here, we can assume all elements to be scalars, and all products to be scalar products. We will also choose $b_h = 0$ to avoid adding unwanted DC noise to our signal. The final architecture of the minimal gated unit presented here is shown in fig. 28.

Adjusting the parameters of the gated recurrent unit can lead to a versatile set of dynamic nonlinear behaviors. The dynamic response of the effect when varying each parameter individually is shown in fig. 30. Note the great variety of distortion curves that are possible using a gated recurrent architecture, from simple saturation to hysteresis-like behavior. The harmonic response of the gated recurrent distortion can be equally versatile, with an adjustable mix of odd and even harmonics. One example harmonic response is shown in fig. 29.

9. IMPLEMENTATION AND PRESENTATION

The work done in this study was designed to be informative and inspiring both to recording engineers and musicians, as well as programmers and aspiring DSP engineers looking to design audio effects. With this audience in mind, it was determined that the results of this work should be written and presented for a relatively non-technical reader, and published in a location where these readers would most easily find it. To that end, the results of this work have been published as a series of articles on the popular blog website Medium[1]. According to statistics on the site, the articles have

[1] `https://medium.com/@jatinchowdhury18`

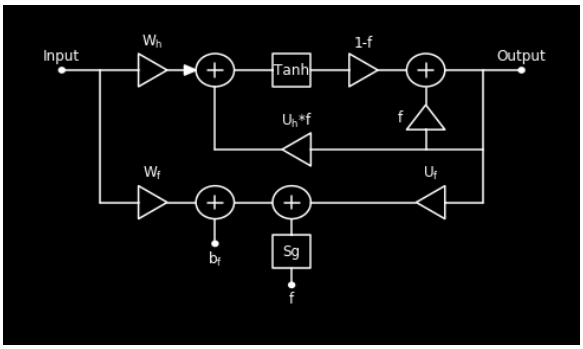

Figure 28: *Signal processing architecture of a minimal gated unit, modified for use as an audio distortion effect.*

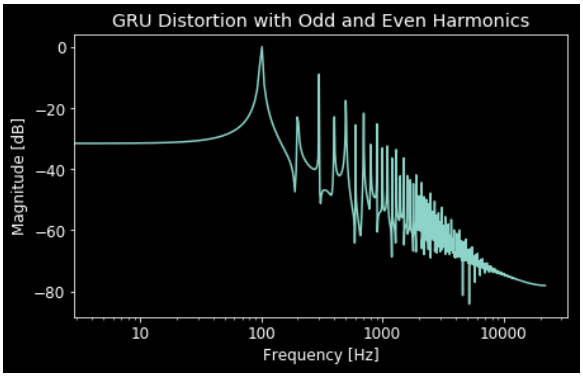

Figure 29: *Harmonic response of a gated recurrent distortion effect that generated both odd and even harmonics.*

been read approximately 400 times as of this writing. Additionally, the structures developed in this study have been implemented as a series of audio plugins using JUCE/C++. The plugins and their source code are freely available on GitHub[2]. Finally, video and audio demos of the plugins are published on YouTube, to demonstrate the sonic characteristics and practical usefulness of the effects created in this study[3].

10. CONCLUSION

In this paper, we have presented a series of complex nonlinear signal processing structures, with the intention of being used for audio effects and synthesis software. While this work has developed a number of useful and sonically rich structures, there is great potential for many more developements in this area of audio signal processing. We look forward to seeing more audio effects being built using these methods, as well as to more nonlinear DSP methods being developed in the vein of this work.

[2] https://github.com/jatinchowdhury18/ComplexNonlinearities

[3] https://www.youtube.com/playlist?list=PLrcXtWXbPsj1SS-3zfcaiRpuETqMYy6Hw

11. ACKNOWLEDGMENTS

The author would like to thank Julius Smith, Kurt Werner, Viraga Perera, Dave Berners, and the GASP working group for providing inspiring discussions and indispensible feedback.

12. REFERENCES

[1] D.T. Yeh, *Digital Implementation of Musical Distortion Circuits by Analysis and Simulation*, Ph.D. thesis, Stanford Univeristy, 06 2009.

[2] W. Shockley, "The theory of p-n junctions in semiconductors and p-n junction transistors," *The Bell System Technical Journal*, vol. 28, no. 3, pp. 435–489, July 1949.

[3] Jatin Chowdhury, "Real-time physical modelling for analog tape machines," in *22nd International Conference on Digital Audio Effects*, Birmingham, UK, 2019, p. 3.

[4] Aphex Systems Ltd., *Aphex Aural Exciter Type B: Operating Guide*.

[5] Priyanka Shekar and Julius Smith, "Modeling the harmonic exciter," 10 2013.

[6] Dimitrios Giannoulis, Michael Massberg, and Joshua D. Reiss, "Digital dynamic range compressor design - a tutorial and analysis," *J. Audio Eng. Soc*, vol. 60, no. 6, pp. 399–408, 2012.

[7] Jatin Chowdhury, "Stable structures for nonlinear biquad filters," dec 2019.

[8] Fabian Esqueda, Henri Pontynen, Julian D. Parker, and Vesa Valimaki, "Virtual anaglog bucha 259 wavefolder," in *20th International Conference on Digital Audio Effects*, Edinburgh, UK, 2017, p. 82.

[9] Fabian Esquda, Henri Pontynen, Julian D. Parker, and Stefan Bilbao, "Virtual analog models of the lockhart and serge wavefolders," *Applied Sciences*, vol. 7, no. 12, 2017.

[10] Alex Loscos and Jordi Bonada, "Emulating rough and growl voice in spectral domain," in *7th International Conference on Digital Audio Effects*, Naples, Italy, 2004, p. 49.

[11] Dieter Leckschat and Christian Epe, "An approach to the generation of subharmonic frequencies in audio applications," in *Audio Engineering Society Convention 136*, Apr 2014.

[12] Kyunghyun Cho, Bart van Merrienboer, Çaglar Gülçehre, Fethi Bougares, Holger Schwenk, and Yoshua Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *CoRR*, vol. abs/1406.1078, 2014.

[13] Guo-Bing Zhou, Jianxin Wu, Chen-Lin Zhang, and Zhi-Hua Zhou, "Minimal gated unit for recurrent neural networks," *CoRR*, vol. abs/1603.09420, 2016.

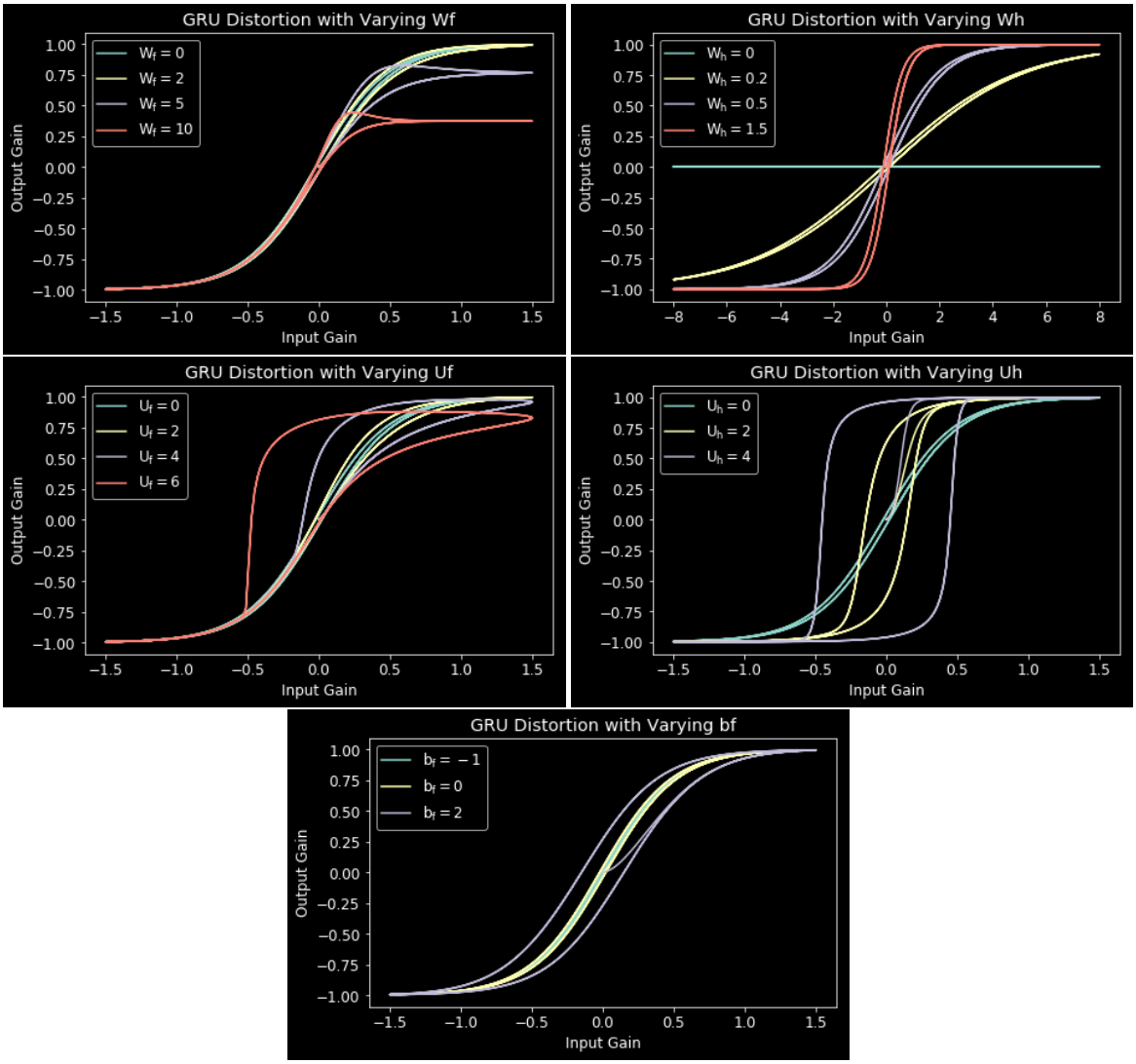

Figure 30: *Dynamic response of a gated recurrent unit distortion effect, varying each parameter individually.*