Extending the Analysis and Synthesis Approach to Classes of Nonlinear Systems

Mike Collicutt

Master of Arts

Schulich School of Music

McGill University

Montreal, Quebec

February 2011

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Arts

ACKNOWLEDGEMENTS

I would first like to thank my parents Joyce and Gary for their constant support and encouragement for the duration of my Master's Degree. I would also like to thank my supervisor, Dr. Philippe Depalle for his great ideas, guidance, and patience; I could not have done it without you.

I must also acknowledge NSERC for their partial financial support.

My gratitude also goes out to all of my friends in Montreal and Hamilton for their kind words and support throughout my time at McGill. Lastly but certainly not least, I thank Katie for her love and friendship, without which my life would have been much more difficult.

ABSTRACT

This thesis investigates the properties of an analysis and re-synthesis method of a class of nonlinear systems, with an application to audio effects for guitar. The goal of this work is to develop a straightforward method to characterize certain types of nonlinear systems (the *analysis*), and subsequently use this characterization to create a generic structure for the model of the system. The model imitates the nonlinear system's behaviour such that the output of the model to a given input signal is the same as the output of the actual nonlinear system under study (the *synthesis*). A method for system identification of linear systems is first presented, and then the method is extended to analyze nonlinear systems as well. An in-depth presentation of how the method works is presented. The information extracted by the analysis is then used as parameters in a synthesis model to emulate a particular nonlinear system under study. The analysis/synthesis method is then tested on some simple memoryless nonlinear systems with simple inputs. Finally, three 'real-world' nonlinear systems are then used to validate the analysis/synthesis method developed in this work. The nonlinear systems are all distortion effects intended for electric guitar. Outputs of the model agreed well with the actual system output when the input was a simple sinusoid. The model's performance did however suffer when a wide-bandwidth musical signal was used as input. Outputs were lacking in higher harmonic content and overall gain. This is thought to be due to the limited bandwidth of the chirp used as well as a limitation on the number of harmonics that can be modeled.

ABRÉGÉ

Cette thèse porte sur l'étude des propriétés d'une technique d'analyse/synthèse appliquée à une classe particulière de systèmes non-linéaires, avec une application numérique aux effets audio pour guitare. L'objectif de ce travail est de développer une méthode simple de caractérisation de cette classe de systèmes (l'analyse), pour ensuite ré-injecter cette paramétrisation dans un modèle générique (la synthèse). Ce dernier va alors imiter le comportement du système étudié de sorte que leurs signaux de sortie soient identiques pour un signal d'entrée donné. Une méthode d'identification des systèmes linéaires est d'abord présentée, pour ensuite être étendue à l'analyse des systèmes non-linéaires. Le fonctionnement de la méthode est exposée en détail. Les informations extraites durant l'analyse sont injectées comme paramètres dans le modèle de synthèse pour simuler le système non-linéaire étudié. La méthode d'analyse/synthèse est ensuite appliquée à plusieurs modèles non-linéaires simples, sans mémoire, soumis à des signaux d'entrée élémentaires. Enfin, trois systèmes non-linéaires existants sont utilisés pour valider la méthode présentée. Ces dispositifs sont tous des effets de distorsion pour guitare électrique. Les sorties de chacun des systèmes et de leur modèle associé sont très similaires lorsqu'ils reçoivent une sinusoïde simple en entrée. Les performances du modèle sont cependant un peu moins bonnes dans le cas d'un signal musical à bande large. Le signal de sortie manque alors de contenu harmonique dans les hautes fréquences, et présente un gain global plus faible. Ce phénomène est sans doute lié

soit à la bande passante réduite du chirp utilisé, soit aux contraintes sur le nombre d'harmoniques présentes dans le modèle.

TABLE OF CONTENTS

LIST OF FIGURES

KEY TO ABBREVIATIONS

HIR: Harmonic Impulse Response

ID: Intermodulation Distortion

LTI: Linear Time-Invariant

LTV: Linear Time-Variant

MIR: Multiple Impulse Responses

MISO: Multiple Input Single Output

MLS: Maximum Length Sequence

NSUT: Nonlinear System Under Test

SF: Spectral Flatness

SNR: Signal-To-Noise Ratio

SUT: System Under Test

TDS: Time Delay Spectrometry

THD: Total Harmonic Distortion

Chapter 1
Introduction

The analysis and identification of audio systems is an important and common task in digital signal processing. Applications such as simulated room acoustics, enhanced digital audio effects, source/filter modeling, physical modeling of systems and instruments, improved loudspeakers, etc. can be understood by studying the temporal and spectral properties of these systems.

The study of the behavior of physical systems usually begins with a basic modeling process. A model is a way of predicting the behavior of a system through the use of functions and operators, and defining how the input and output of the system are related. The relation between the operators and signals can be illustrated by way of a block diagram [1]. As the need for greater accuracy in predicting or simulating a system's behaviour increases, the complexity of the model often increases as well. There are many approaches to selecting a model, and the choice of which model to use depends on a number of factors including desired accuracy, computational load or complexity, and personal preferences. General common approaches in signal processing for audio applications include physical modeling, statistical, and psychoacoustic models, among many others.

Once a model is chosen and implemented, the computation usually involves the calculation or estimation of model's parameters. This process is called *system identification*. Once a system has been characterized, the information can be used to learn more about the system under study, or try to simulate the system in hardware or software. Regardless of the type of model used, the characterization of a system, followed by a recreation of the system by some other means is an example of the *analysis/synthesis* approach. In general, *analysis* is defined as the procedure by which an intellectual or substantial whole is broken down into parts or components. *Synthesis* is defined as the opposite procedure: to combine separate elements or components in order to form a coherent whole. Analysis and synthesis, as scientific methods complement one another. Every synthesis is built upon the results of a preceding analysis, and every analysis requires a subsequent synthesis in order to verify and correct its results [2].

Farina [3][4] developed a simple novel method for characterizing a linear system which also shifted any nonlinear responses from the system earlier in time. While the initial application was for characterizing linear systems, Farina's method allows for the extraction of both linear and nonlinear responses of a system in a single measurement. These responses provide temporal and frequency domain information about how a nonlinear system processes an input signal.

The responses can then be used in a model to imitate the nonlinear system under study. The goal of the model is to imitate the nonlinear system's behaviour such that the output of the model to a given input signal is the same as the output of the actual nonlinear system under study.

This thesis describes an analysis and re-synthesis method of a class of nonlinear systems, with a particular focus on audio effects for guitar. The goal of this work is to develop a straightforward method to characterize certain types of nonlinear systems (the *analysis*), and subsequently use this characterization to create a generic software model of the system (the *synthesis*).

It is convenient at this point to first define a few more key terms that are often used in this dissertation.

(Nonlinear) system under test (SUT/NSUT) The SUT is the actual physical or virtual system is to be modeled.

Excitation An excitation signal $s(t)$ is a time varying signal $s(t)$ that is used for system identification. It is the test signal that is first sent to the SUT.

Response The response $r(t)$ of a system is the measured output of a system under test to an excitation signal. The analysis method used in this work (Chapter 2) uses both $s(t)$ and $r(t)$ to extract the important information about the SUT.

'Re-synthesized', or 'simulated' system The re-synthesis of a nonlinear system is the recreation or emulation of the behaviour of a 'real' nonlinear system through a model, either software or hardware based, such as the model summarized in this thesis. The re-synthesized version of the nonlinear system can be compared to the SUT by sending identical input signals to each system and comparing the output of each (See Figure 1–1).

Input signal The input is any signal $x(t)$ that is input to a nonlinear system that will produce an output y(t). The input can be directed into either the real physical SUT or the simulated system. An input is different from an excitation in that the excitation is a specific signal used for the analysis of a nonlinear system, and an input is simply a signal to be processed by the nonlinear system or its simulated counterpart.

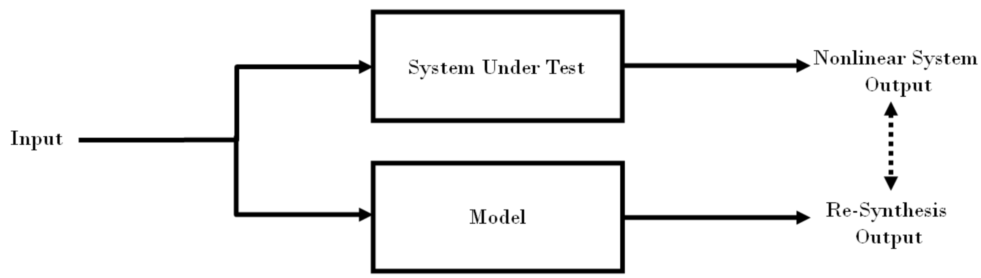


Figure 1–1: The re-synthesis model can be compared to the NSUT by sending identical signals to both and comparing the outputs.

In the context of audio, the analysis/synthesis approach aims to extract the important features of a system as parameters, then use these parameters as elements in a synthesis model. This model can be used to emulate the sound (i.e. output) of the real-world object or system in response to some input, or even to add flexibility to an audio system not possible in the 'real world'. The analysis/synthesis concept for systems whose input is known[1] is illustrated in Figure 1–2. Once the parameters are estimated and a synthesis model is in place, it can then be compared to the SUT by sending an identical signal to both systems and comparing the outputs in some way, as illustrated in Figure 1–1.

[1] The approach differs somewhat when the input signal to a system is unknown. In these cases, the analysis method is statistically-based and the input signal is often assumed to white Gaussian noise or some other signal derived from a random or pseudo-random process.

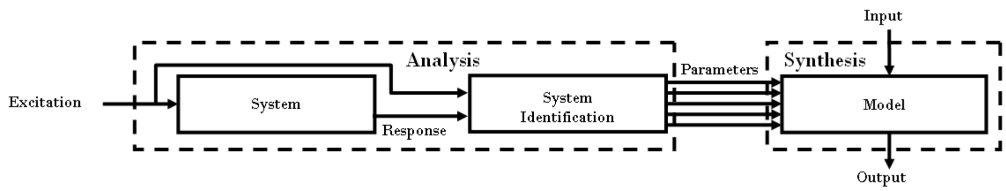


Figure 1–2: The analysis and synthesis framework.

1.1 Overview

This dissertation focuses on an analysis/synthesis approach to the modeling of certain nonlinear audio systems. The dissertation is organized into 4 chapters. Chapter 1 presents background information on the analysis of both linear and nonlinear systems, the differences between the two, and typical ways in which linear systems are measured in practice. Chapter 2 illustrates Farina's method [3] [4] for identifying linear and nonlinear systems and explains how the method works in detail. Chapter 3 describes a synthesis method for nonlinear systems called the MISO Polynomial Hammerstein Model, and how the information extracted from nonlinear convolution is used in this model. Finally, Chapter 4 summarizes the implementation of this analysis/synthesis system in MATLAB, and compares the results of real nonlinear systems to their re-synthesized versions using this method.

1.2 Background

Often analysis methods at some point make the assumption that the system under study is *linear*. Linear systems are relatively easy to understand mathematically and have some very useful properties. The defining property of linear systems is that they satisfy the principle of superposition. This states that for any given inputs to a linear system x_1 and x_2 with respective outputs y_1 and y_2, then any linear combination of the inputs $ax_1 + bx_2$ to the linear system will produce a proportionally scaled and summed output $ay_1 + by_2$. As a consequence, a signal consisting of many partials (as is the case with most sounds) passed through a linear system will retain all of these partials in the output signal, with only their relative magnitudes and phases changed. No new frequency content can be introduced into a signal by a linear system. This and the principle of superposition are in fact equivalent and merely express the same idea in the frequency and temporal domains, respectively.

A second property of a linear system is the time-dependence of its parameters. A linear system whose parameters do not change in time is called a Linear Time-Invariant (LTI) system, and a linear system that is time-dependent is called a Linear Time-Variant (LTV) system. The properties of LTI systems make their characterization straightforward. An LTI system is completely characterized by its *impulse response* (IR), denoted $h[n]$ (or $h(t)$ in continuous time). The IR is the time-domain output of a system to an impulse input signal. In discrete time, this input signal to the system is (ideally) a Kronecker function $\delta[n]$ which equals 1 at $n = 0$, and 0 for all other time samples n. Figure 1–3 illustrates an impulse in the digital domain, and the impulse response of a LTI.

Systems that do vary in time cannot be represented by a simple impulse response of the form $h_i(t)$. For LTV systems to be represented in terms of impulse responses, one must know the impulse response $h_v(\tau, t)$ at *each time instant* t to an impulse applied τ time before [6]. Thus in this notation, for an LTI system $h_v(\tau, t) = h_i(\tau)$.

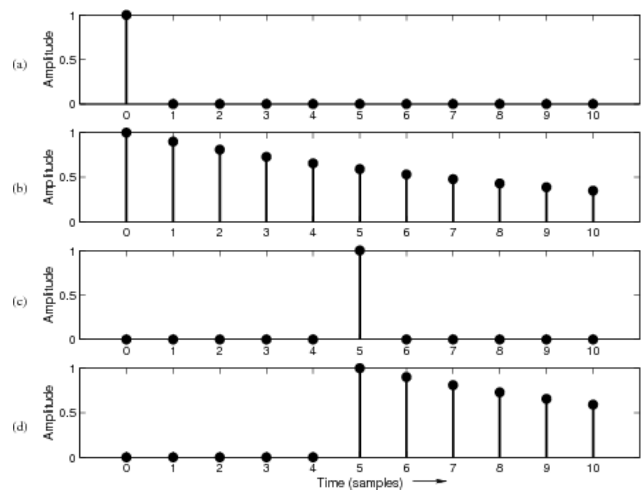


Figure 1–3: a. Unit impulse $\delta[n]$. b. Output impulse response $h[n] = 0.9^n$. c. Input delayed-impulse $\delta[n-5]$. d. Output delayed-impulse response $h[n-5]$. Taken from [5].

Another way of characterizing an LTI system is through its *frequency response*. The frequency response is defined as the spectrum of the output signal divided by the spectrum of the input signal. It specifies the attenuation of a signal as a function of the frequency of that signal. In addition, it also specifies the phase delay of a signal as a function of frequency. The frequency response and impulse response are related through the Fourier Transform. They are completely analogous and convey the same information in different domains. The Fourier Transform of the IR shown in Figure 1–3b is shown in Figure 1–4.

Convolution

Once the impulse response of the system is measured, the output of the system $y(t)$ to *any* input can be calculated using only the input $x(t)$ and the impulse response $h(t)$. The output is given by the *convolution* of the impulse response with the input signal as shown in Eq. (1.1). Because no new frequency

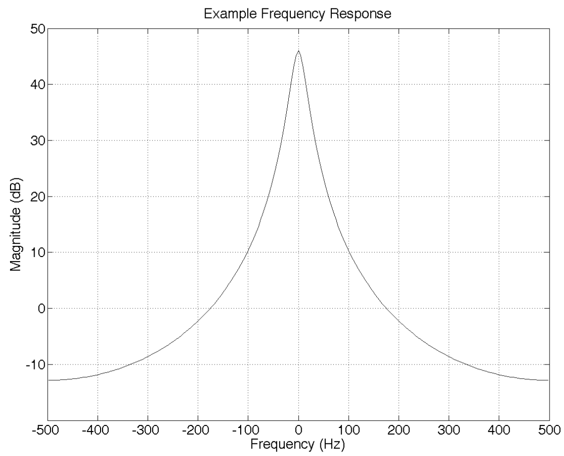


Figure 1–4: Frequency response of the filter shown in Figure 1–3b. with a sample rate of 1 kHz.

content is generated by the convolution process, convolution is a linear operation.

$$y\left(t\right) = \left(h * x\right)\left(t\right) = \int_{-\infty}^{\infty} x\left(\tau\right) h\left(t-\tau\right) d\tau \tag{1.1}$$

In order to give a more concrete description of the convolution in Eq. (1.1), the operation is illustrated graphically in Figure 1–5. First, $h(\tau)$ (in red) is reversed in time to give $h(-\tau)$, which flips the signal about $t = 0$. $h(-\tau)$ is then shifted left by a time t towards $-\infty$ to give $h(t-\tau)$. Since in practice the signals are of finite length the time-reversed signal $h(t-\tau)$ is shifted far enough left that it does not overlap with $x(\tau)$ (in blue). If $x(\tau)$ and $h(\tau)$ are causal, then merely time-reversing $h(\tau)$ is sufficient to eliminate any overlap between the two functions. Then as t is incrementally increased, $h(t-\tau)$ is shifted to the right, and the product of the two functions x and h is integrated (indicated by the yellow area) yielding the output y(t) (in black). This shifting of $h(t-\tau)$ and integrating is performed until there

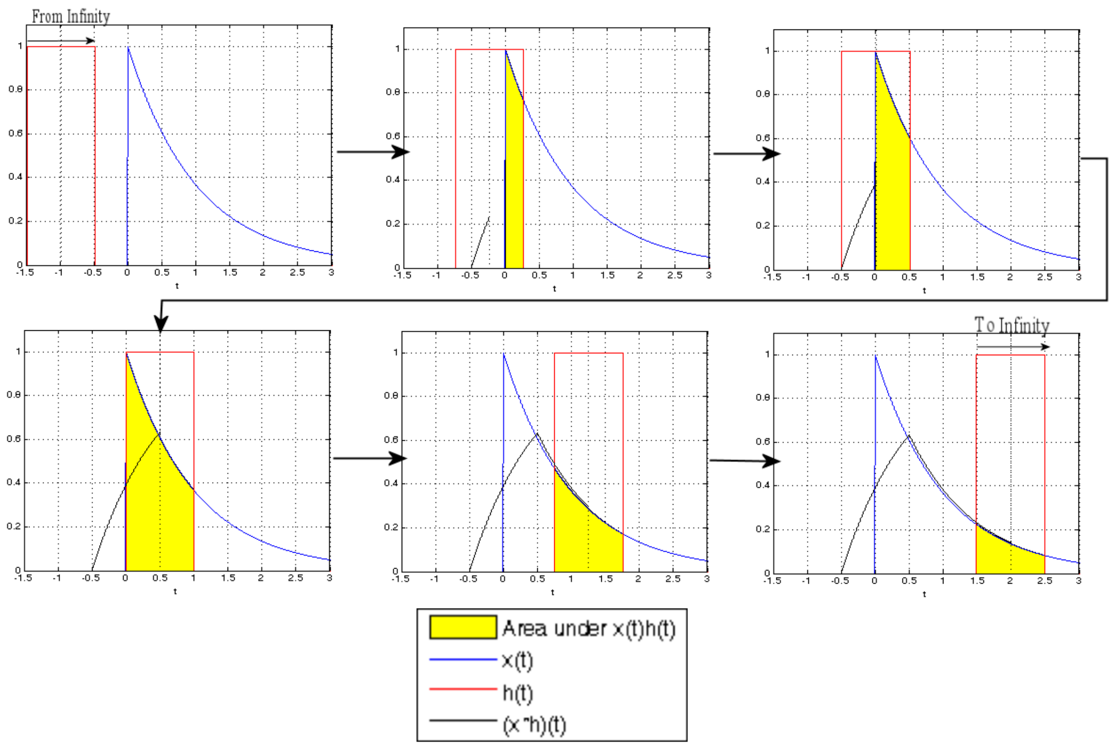


Figure 1–5: Sequenced plots of the convolution of a unit square function and a damped exponential function, for increasing values of the shift t. In each shift, the product of the two functions is integrated (yellow area), to give a value on the black curve at that shift value. Code from [7].

is no more overlap between the two functions. Thus the length of the convolution of signals x and h of lengths t_1 and t_2 respectively is $t_1 + t_2$. This graphical view of convolution is important for the explanation of Farina's method which will be discussed in Section 2.4.

Performing a convolution as in Eq. (1.1) can be quite computationally intensive for audio signals, due to the many multiply and add operations required for any signal of appreciable length. Fortunately, convolution is expressed much more simply in the Frequency Domain. The Convolution Theorem [8] states that given two time domain functions $x(t)$ and $h(t)$, and their respective spectra $X(f)$ and $H(f)$, denote $y(t) = (x * h)(t)$ and it's spectrum $Y(f)$. Then,

$$Y(f) = X(f)H(f) \tag{1.2}$$

Thus convolution of two sequences in the time domain is equivalent to the multiplication of their corresponding Fourier transforms. Because of the efficiency of the Fast-Fourier-Transform algorithm [9], it takes fewer operations to transform each of the signals to the frequency domain with the FFT, perform the complex-valued multiplication, and inverse FFT back to obtain the time-domain convolved result than to perform the convolution directly. A number of methods [8] [10] [11] [12] have been developed for performing 'fast convolution'. While they differ in the details, they all use the FFT to perform the convolution in the frequency domain.

Nonlinear Systems

Often real-world systems do not exactly obey superposition, and new frequency content can be introduced into the output that was not present in the input signal. These systems are said to be *nonlinear*.[2] Thus nonlinear systems cannot be simply characterized by an impulse response, since convolution is a linear operation.

[2] Linear Variant Systems can also introduce new frequency content in the output signal. However these systems are beyond the scope of this work. Attention will be restricted to time-invariant systems.

The generation of new frequencies in the output signal that were not present in the input signal is called *distortion*. Distortion can either be a desired or undesired consequence of a system. For example, the distortion of an electric guitar signal can produce very full and rich guitar sounds high in harmonic content. FM synthesis [13] and distortion by waveshaping [14] are other examples of nonlinear systems with musical applications, producing many frequencies from just two or more input frequencies. On the other hand, in the case of audio reproduction, distortion of the signal from poor quality speakers can give a harsh, unpleasant quality to the original sound.

Nonlinear systems will typically generate two types of distortion - harmonic distortion and intermodulation distortion. Harmonic distortion creates *harmonics* of the input signal in the output. A harmonic is a component of the signal that is an integer multiple of the fundamental frequency of the signal. Thus if the input signal has a fundamental frequency f_o, a harmonic distorting system will generate frequencies at $f_o, 2f_o, 3f_o$, etc. in varying strengths. The highest harmonic of significant power generated by a system is called the *order* of the nonlinearity, and so an n^{th} order system produces (up to) n harmonics of the original input.

Intermodulation distortion occurs when a signal containing two different frequencies f_1 and f_2 is introduced to a nonlinear system. The result is the generation of frequencies at sums and differences of the input frequencies. In general, components at frequencies $pf_1 + qf_2, p, q \in \mathbb{Z}$ will be produced, with typically decreasing strength as $p + q$ increases. Thus the most prominent frequencies will be at $f_1 + f_2, 2f_1 + f_2, ...$ and $f_1 - f_2, f_1 - 2f_2, ...$

Although many real systems are in fact nonlinear, when a system only introduces a relatively small amount of new frequency content into the output, it is said to be weakly nonlinear and a linear approximation can be sufficient. This simplification is for good reason: the fact that nonlinear systems can produce new frequencies makes them much more difficult to analyze than their linear counterparts. If however the distortions are too large to ignore, or they are of interest such as in the context of guitar distortion, the use of a linear model cannot produce accurate results.

On the other hand, nonlinear models can become computationally demanding very quickly even for models of low order. If a nonlinear model must be used, a balance must be found between its complexity and accuracy. The decision whether to use a linear or nonlinear model depends on the specific application and required level of accuracy of the model.

Harmonic and intermodulating distortion systems are the most common types of nonlinear system. Other types, such as chaotic, hysteretic, or quasi-stable/multi-stable exist [15] but are far less common and not considered in this dissertation.

1.3 Motivation

Having a unified representation of a class of nonlinear system is very valuable. As mentioned above, nonlinear systems are often approximated as linear due to the complexity of nonlinear models. However an accurate and relatively simple representation of these nonlinear systems could improve linear models without much increase in computational cost. This unified representation also allows for a way to compare nonlinear systems of this class using a single framework.

One application to this research is amplifier and effects box simulation. The sound of classic tube amplifiers and vintage effects boxes are highly desired by the guitar community [16] [17]. The development of digital simulations of classic audio components, often times involving nonlinear components, has been commercially popular over the last 10 years or so. Examples include the Guitar Rig product line [18] and Universal Audio suite of plugins [19]. Owning a large suite of such units or their software simulation counterparts can be very costly. On the other hand, having an analysis/synthesis system that can simulate many different nonlinear systems using the same framework, such as the saturating or clipping systems often present in guitar distortion circuits can be an attractive alternative to the high cost of owning these actual units. Furthermore, the ability to automatically characterize, modify, and recreate any new systems or effects allows for non-specialists to take a do-it-yourself approach to creating custom-made sounds.

1.4 IR and Distortion Measurement Techniques

As stated above, the IR of a system is the response of a system to an impulse. In practice, an ideal Dirac distribution is virtually impossible to create. An ideal impulse has a flat spectrum and provides enough energy to the system to give a reasonable signal-to-noise ratio. Sounds like electrical sparks, popping balloons, and pistol or cannon shots are often sufficiently loud sound sources, but lack the spectral flatness, repeatability and omni-directionality required for most applications [20]. On the other hand, when direct IR measurements were commonplace at least 30 years ago, the loudspeakers used could be sent an impulse from a digital source which was more spectrally flat, but lacked sufficient power to obtain a good signal-to noise ratio [21]. Alternatively, several indirect but more reliable methods have been developed to extract the impulse response of a linear system. These methods do not measure the impulse response of a room directly, but rather excite the system with a known signal and measure the response. The IR is then extracted through mathematical operations involving the input and response. The specific operations of course depend on the method chosen. Some of the most popular IR measurement methods are described briefly below. It will be shown in Chapter 2 that the sine sweep method described below can be applied to analyze both linear and nonlinear systems.

Maximum Length Sequences (MLS)

Schroeder [21] devised a method to measure the impulse response of linear systems using pseudo-random binary sequences as test signals to the system. An MLS is a periodic sequence of binary digits, usually -1 and +1. It can be generated by an n-stage shift register in a feedback loop, along with an 'exclusive-OR' operation. The most important property of an MLS is that the Fourier Transform is flat, like the Fourier Transform of an impulse. Furthermore, the circular autocorrelation function [8] of an MLS approaches a Kronecker as the length of the MLS increases. Thus to measure the impulse response of a system, the MLS is repeatedly sent through the system and its output is measured. When this is cross-correlated with the input MLS, a train of impulse responses results. A greater signal-to-noise ratio can be obtained by averaging these multiple impulse responses, and the measurement is very tolerant to background noise. Details and further explanations can be found in [21] and [22].

Time Delay Spectrometry (TDS)

TDS is a clever technique for measuring the frequency response of a linear system [23]. When performing acoustic measurements on a piece of equipment, the room that it is done in often becomes a part of the measurement as well. TDS allows for anechoic measurements of a system without the need of a specialized room. The two basic pieces of equipment in a TDS measurement are the audio generator and tunable receiver. The audio generator feeds the system to be tested with a sinusoidal chirp. The tunable receiver is a device that can be made to respond to signals only within some controlled bandwidth. When tuned manually, it is analogous to a radio tuner, except in the auditory frequency range. The generator and receiver can be coupled to one another, such that frequency that the generator creates is tracked and tunes the central frequency of the tunable receiver [24].

To measure a system, a microphone is connected to the tunable receiver and placed at a fixed, measured distance away from the system. The audio generator sends the test signal (a sinusoidal chirp) to the system under study, which creates pressure waves through the air. This is picked up by the microphone and the signal is sent to the receiver. As the frequency of the chirp increases, the receiver changes its central frequency along with the chirp, offset by an amount to account for the time-of-flight delay of the sound from the system to the microphone. The idea is that when the sound pressure wave directly from the system reaches the microphone, it will be tuned to accept that particular frequency, with almost all others blocked out. When the pressure waves from reflections off of the walls, etc. reach the microphone (taking a longer path length and a longer time), the frequency of these waves will already be outside of the bandwidth of the receiver and are suppressed. Thus it is possible to measure only the direct sound from a system, thanks to the difference in arrival time of the direct and reflected sound. By adjusting the offset, one can discard the direct sound and instead measure reflection/absorption characteristics of a particular spot in a room, or measure the acoustic properties of a test material [23] [24].

Sine Sweeps

Farina [3] popularized a method for measuring impulse responses using sinusoidal sweeps, or 'chirps' as the excitation signal to the system. MLS and TDS techniques are based on the assumptions that the system under study is linear and time-invariant. While both methods are sensitive, MLS in particular is very intolerant of any non-linear behaviour or time-variance of the system [3]. While the use of chirps in acoustical measurements was nothing new [20] [23], Farina's sine sweep technique for measuring systems became popular due to its simplicity. This technique came out of research attempting to overcome this limitation in MLS through the use of TDS measurements. A sinusoidal chirp (whose frequency increased linearly or exponentially) is input to the system and the output is recorded. When the output is convolved with a time-reversed version of the input signal, the resultant signal contains the linear impulse response, with any nonlinear artifacts shifted earlier in time. This method is comparatively simpler than MLS or TDS, and can also be used to investigate nonlinear systems. Thus [3] serves as a starting point for this research.

Basic Nonlinear System Measurements

One way of characterizing the degree to which a system is nonlinear is by measuring its *total harmonic distortion* (THD). The power THD of a system is the ratio of the sum of the power in all of the harmonics produced P_n to the power of the fundamental P_1 in a sinusoidal signal fed through the system.

$$THD = \frac{\sum_{n=2}^{\infty} P_n}{P_1} \tag{1.3}$$

Since individual harmonic amplitudes are measured for the THD, it is necessary to specify the frequency (or frequency range) of the test signal and its amplitude for a proper THD measurement. The THD of a system is generally specified with the input signal close to full-scale, although it can be specified at any level [25]. THD measurements are often given for loudspeakers, amplifiers and microphones.

As described on page 10, *intermodulation distortion* (ID) produces frequencies in the output that are the sum and difference multiples of frequencies in the input. Similar to THD, an intermodulation distortion measurement gives the power ratio of these various sum and difference frequencies to the power of the fundamental frequencies in the input signal.

A simple comparison derived from the THD measurement, called the spectral flatness (SF) can be made which is useful for comparing the outputs of a nonlinear system to its simulated model's. The spectral flatness plots the spectral amplitude ratio (or difference in dB) of the response of one system to another in response to a simple input signal. Mathematically the spectral flatness is given by

$$SF = 10\log\left(\frac{FFT(|\Theta|)}{FFT(|\hat{\Theta}|)}\right) \tag{1.4}$$

Where Θ and $\hat{\Theta}$ are the outputs of the two systems under comparison. Typically the input signal to both systems is a simple sinusoid or simple combination thereof in order to resolve all of the harmonics and/or intermodulation products produced.

The spectral flatness has the same frequency resolution as the FFT's under analysis. For simple signals, this is often less than helpful, because disagreement in unimportant areas of the spectrum (i..e areas where there is no frequency content) can overshadow those frequency bins which are the most important perceptually for recreation of the signal. Thus the spectral flatness is evaluated at the expected location of harmonics and intermodulation products to give a more local picture of the accuracy compared to a percentage THD or ID measurement.

Obviously, linear characterizations like the IR and nonlinear characterizations like the THD or intermodulation distortion are done separately. However Farina's [3] sine-sweep method allows characterizations of both linear and non-linear aspects of the system in a single measurement. Furthermore, the nonlinear system characterizations are far more informative than simple power ratios given by THD and ID measurements. This is described in detail in the following chapter.

Chapter 2
Nonlinear Convolution

2.1 Method

Farina [3][4] developed a simple novel method which he called *nonlinear convolution* for measuring both the linear and nonlinear responses of a system in a single measurement. This name is somewhat misleading as the actual convolution operation performed is the same linear operation as the one in Eq. (1.1). However, the signals used for the analysis of linear systems can also be used to extract information about nonlinear characteristics of the system. The general method will first be outlined qualitatively, followed by an explanation of how each stage works. At certain stages in the outline of the method, some details or explanations will be omitted until later in the chapter in order to provide a more general concept of the method first.

First, the nonlinear system is fed with an excitation signal $s(t)$. The excitation in this case is a chirp with an exponential frequency evolution and is defined as

$$s(t) = \cos\left[\frac{\omega_1 T}{\ln\left(\frac{\omega_2}{\omega_1}\right)}\left(e^{\frac{t}{T}\ln\left(\frac{\omega_2}{\omega_1}\right)} - 1\right)\right] \tag{2.1}$$

Where ω_1 and ω_2 are respectively the start and end angular frequencies of the chirp and T is the length of the chirp in seconds. To illustrate the concepts of nonlinear convolution, we will use a chirp with $\omega_1 = 1$ Hz, $\omega_2 = 4000$ Hz, and

$T = 10$ s. A spectrogram of this exponential chirp is given in Figure 2–1. An example nonlinear system is also chosen which generates 10 harmonics of the input signal, with the n^{th} harmonic having strength $1/n$. The response $r(t)$ of the nonlinear system contains the original signal plus harmonics of the input signal. The system's response to $s(t)$ is given in Figure 2–2. The linear response is the lowest curve on the plot, and the nonlinear responses are the higher curves. Notice that for an input signal with bandwidth $[\omega_1, \omega_2]$, an n^{th} order nonlinear system will have a bandwidth of $[\omega_1, n\omega_2]$. Thus one must be mindful of the bandwidth of the input signal to a nonlinear system in discrete time, as aliasing can easily occur from a high-order nonlinear system. After the response is measured, it

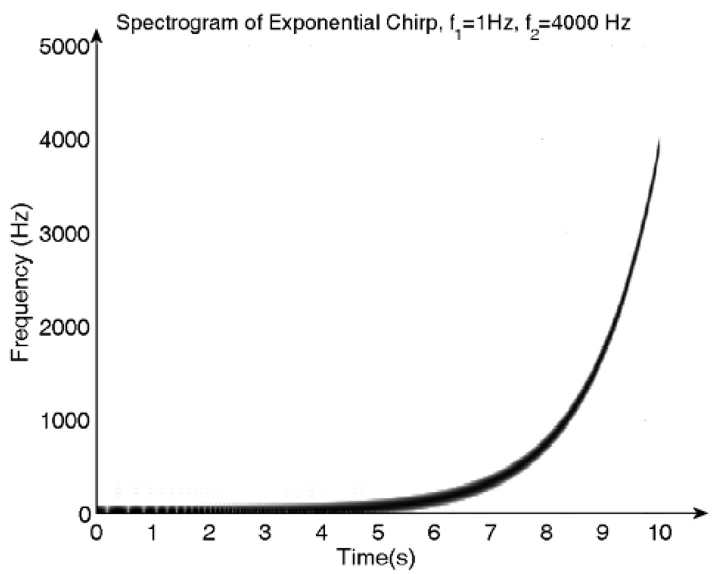


Figure 2–1: Spectrogram of an exponential chirp from $f_1 = 1$ Hz to $f_2 = 4000$ Hz.

is then convolved with the input's *inverse system*. The inverse system shall be defined as the signal $\tilde{s}(t)$ such that its convolution with $s(t)$ yields a scaled impulse

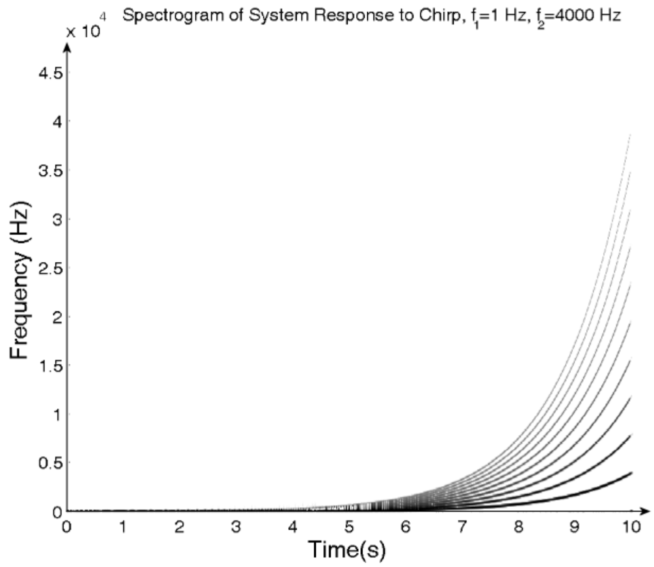


Figure 2–2: Response of a 10th order nonlinear system to a sinusoidal exponential chirp from $f_1 = 1$ Hz to $f_2 = 4000$ Hz.

shifted in time.[1] $\tilde{s}(t)$ is considered a system in this context because it is used in the same way as a filter's impulse response in convolution. Thus depending on the context, $\tilde{s}(t)$ will be called either a signal or a system (or in this case, a filter). For chirps, $\tilde{s}(t)$ is a time-reversed and shifted version of the input, possibly scaled by a temporal envelope depending on the type of chirp. The frequency evolution of $s(t)$ and $\tilde{s}(t)$ for the case of exponential chirps are given in Figure 2–3. The result of the convolution of the response $r(t)$ with $\tilde{s}(t)$ is a single time-domain sequence $h(t)$ containing multiple band-limited impulses, called Multiple

[1] Strictly speaking, the impulse occurs at $t = 0$, since the inverse system exists for $t < 0$ only when time-reversed. In a practical realization however, both signals start at $t = 0$ and thus the impulse is shifted in time.

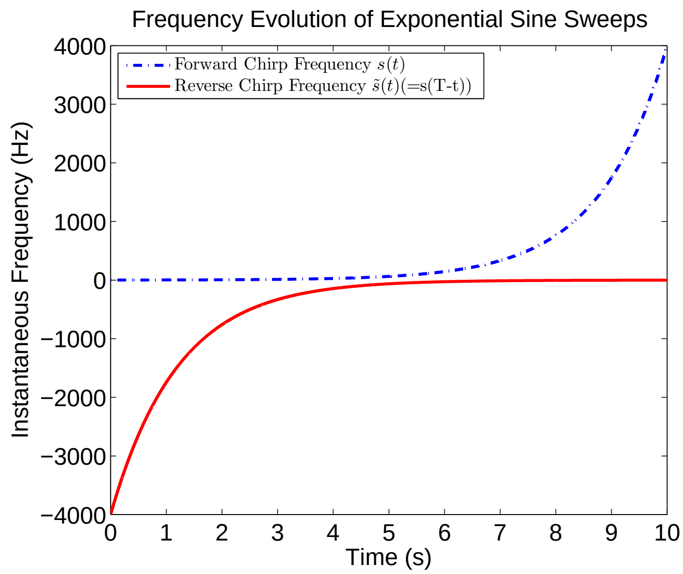


Figure 2–3: Frequency evolution of forward chirp $s(t)$ (blue) and inverse chirp $\tilde{s}(t)$ (red).

Impulse Responses (MIR), as shown in Figure 2–4, which are separated in time and occur at very precise delays after one another. The linear impulse response appears furthest to the right, and preceding it are (typically smaller) impulses. In fact, these impulses can be thought of as higher-order impulse responses. Higher order responses appear in the MIR with a smaller delay, and so the highest order response appears first in the MIR sequence, followed by progressively lower order responses until the linear response. Each Harmonic Impulse Response (HIR)[2] describes how one of the generated harmonics of the input are filtered, analogous

[2] MIR refers to the time-domain sequence output from the convolution containing several (or a single) band-limited impulse(s). HIR refers to an individual impulse in the MIR that is usually separated out by temporal windowing.

to how the impulse response describes how the input is filtered in a linear system. Thus nonlinear convolution characterizes the system by *multiple* impulse responses, one for each order of the nonlinearity. The impulses are easily separated into individual HIR, denoted here as $h_i(t)$ by temporal windowing[3] such that each impulse response occurs at the same time in its window. A flow diagram of the nonlinear convolution process is given in Figure 2–5.

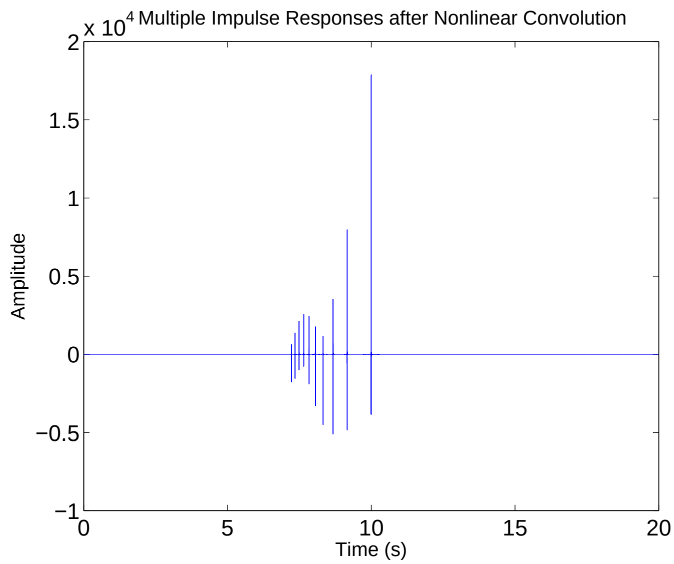


Figure 2–4: Multiple impulse responses after convolving the response in Figure 2–2 with the time-reversed input.

2.2 Why an exponential chirp?

Chirps for acoustical measurements were in use long before Farina's method. Heyser [23] used linear sweeps for performing Time Delay Spectrometry, which

[3] We consider here for the sake of simplicity that the impulse responses do not overlap in time.

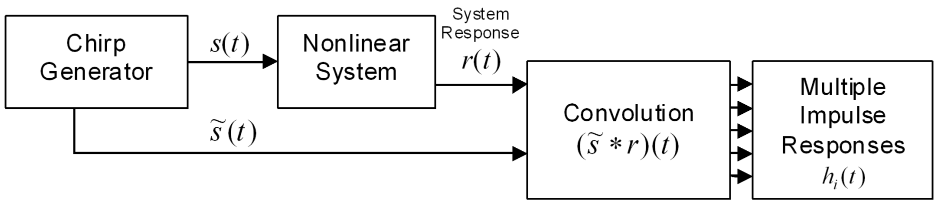


Figure 2–5: Flow diagram of nonlinear convolution.

was Griesinger's basis for his work with sweeps [20] in room measurement. He began to use sweeps with an exponential frequency evolution, and seemed to anticipate that they could be used to push out distortion when measuring impulse responses. For nonlinear convolution to produce multiple impulse responses, a chirp whose frequency varies exponentially must be used. However, it is not immediately obvious why this must be. This section aims to show what happens when different types of chirps are used as excitation signals for a nonlinear system, and the outputs produced when attempting to produce the MIR output by Farina's method. Some of the attractive features of exponential chirps will then be presented.

Linear Sweeps have a frequency evolution of the form:

$$\omega_{lin}(t) = \omega_1 + \frac{\omega_2 - \omega_1}{T} t \tag{2.2}$$

Where ω_1 and ω_2 are the start and end angular frequencies of the chirp respectively, and T is the chirp duration. Integrating this equation gives the phase $\phi_{lin}(t)$, and thus the equation for the linear chirp is:

$$s_{lin}(t) = \cos\left[\omega_1 t + \frac{\omega_2 - \omega_1}{2T} t^2 + \phi_o\right] \tag{2.3}$$

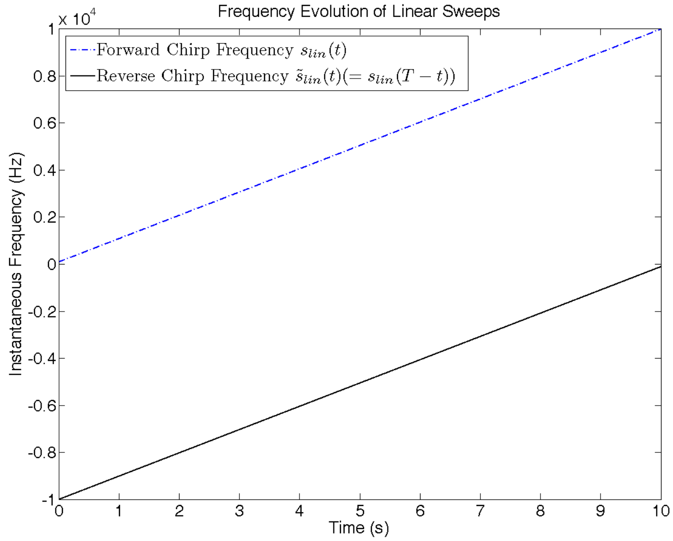


Figure 2–6: Frequency evolution of forward and reverse linear chirps over time.

Where ϕ_o may be tuned to specify the initial or final phase of the chirp. The inverse system $\tilde{s}_{lin}(t)$ is again the time-reversed version of s_{lin}. When $T - t$ is substituted for t in Eq. (2.3) or any chirp, the frequency evolves from $-\omega_2$ to $-\omega_1$. Figure 2–6 shows the frequency evolution of a linear chirp from 100 Hz to 10000 Hz.

When $s_{lin}(t)$ is fed through a nonlinear system and the response is convolved with $\tilde{s}_{lin}(t)$, we get what is supposed to be the MIR for the system. The nonlinear system generates 10 harmonics of the output, with the n^{th} harmonic having strength $1/n$. Thus we expect the MIR output to contain 10 band-limited impulses, one for each harmonic order. In this case the impulses should be band-limited to $[\omega_1, \omega_2]$ and have no reverberant character. The spectrogram is shown in Figure 2–7.

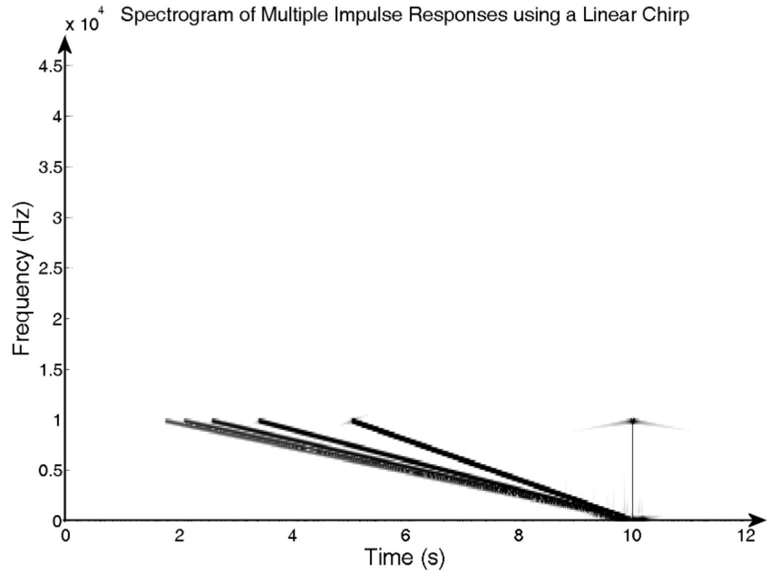


Figure 2–7: Spectrogram of deconvolved MIR when using linear chirps. The higher order IRs are smeared in time. The linear impulse response does appear as expected at 10 s (length of the chirp) with some artifacts surrounding it.

Here the higher order responses do not 'pack' at a precise time into impulses, but rather are smeared out over time before the linear impulse at 10 s. The linear impulse response is present however. Some artifacts also appear around the linear response.

Cosine sweeps have a frequency evolution of the form

$$\omega_{cos}(t) = \frac{\omega_1 + \omega_2}{2} + \left(\frac{\omega_1 - \omega_2}{2}\right) \cos \frac{\pi t}{T} \tag{2.4}$$

Where again the frequency evolves from ω_1 to ω_2 in a time T. Thus the equation for the chirp is

$$s_{cos}(t) = \cos \left[\frac{\omega_1 + \omega_2}{2} t + \left(\frac{\omega_1 - \omega_2}{2}\right) \frac{T}{\pi} \sin \frac{\pi t}{T} + \phi_o\right] \tag{2.5}$$

The frequency evolution of s_{cos} and the time reversed version $\tilde{s}_{cos}(t)$ are given in Figure 2–8. After sending $s_{cos}(t)$ through a nonlinear system, nonlinear convolution using cosine chirps yields the spectrogram given in Figure 2–9.

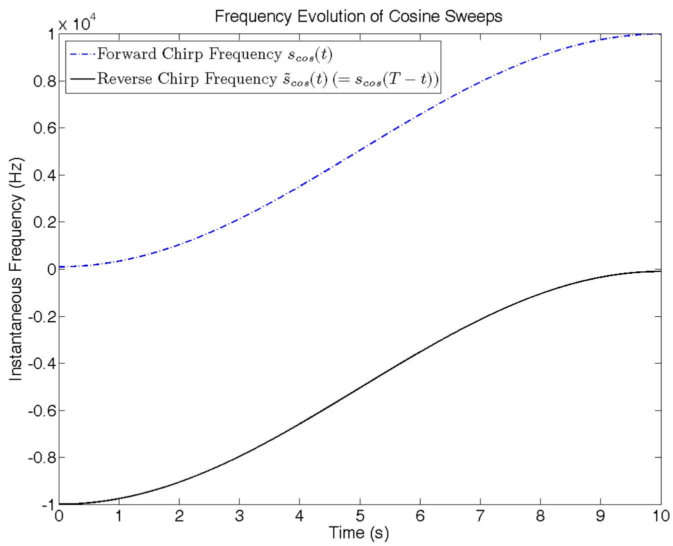


Figure 2–8: Frequency evolution of forward and reverse cosine chirps.

The situation is similar to that of the linear chirps. The higher order responses do not pack into impulses but rather are smeared in time before the linear impulse, and the linear impulse is present with some artifacts surrounding the impulse.

Exponential chirps have a different behaviour when used for nonlinear convolution however. The frequency evolution of a chirp is given by

$$\omega_{exp}(t) = \omega_1 e^{\frac{t}{T}\ln\left(\frac{\omega_2}{\omega_1}\right)} \tag{2.6}$$

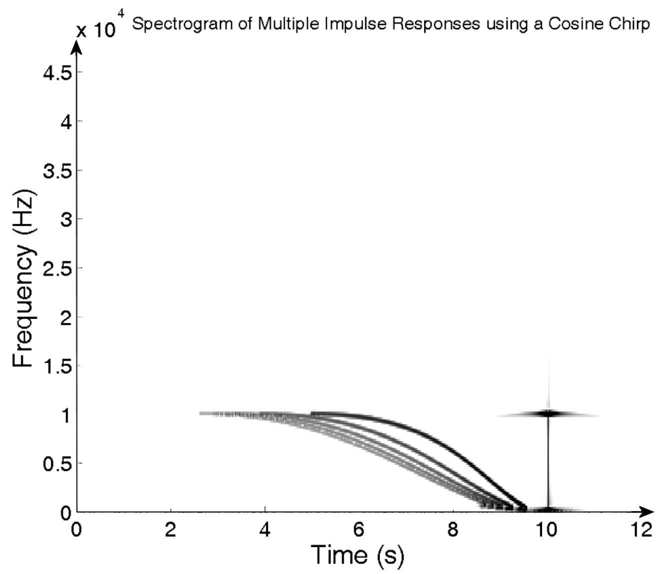


Figure 2–9: Spectrogram of deconvolved MIR when using cosine chirps. The higher order IRs are also smeared in time like the linear chirp case. Again, the linear impulse response does appear as expected at 10s (length of the chirp).

Eq. (2.1) gives the expression for the exponential chirp $s_{exp}(t)$. The frequency evolution of $s_{exp}(t)$ and $\tilde{s}_{exp}(t)$ are given in Figure 2–3. The spectrogram of the result is shown in Figure 2–10. Here, the higher order responses pack into band-limited impulses at precise anticipatory times before the linear response as expected. Some artifacts appear to be present, but in actuality they are of much lower energy than the impulses $\sim 5 \times 10^{-4}$.

There appears to be something special about the exponential chirp. While all the chirps produced an impulse at time T, only the exponential packed the higher order response of the nonlinear system into impulses.

Exponential sine sweeps spend an equal amount of time in each octave. For example, the chirp takes the same time to go from f_1 to $2f_1$ as it does to go from

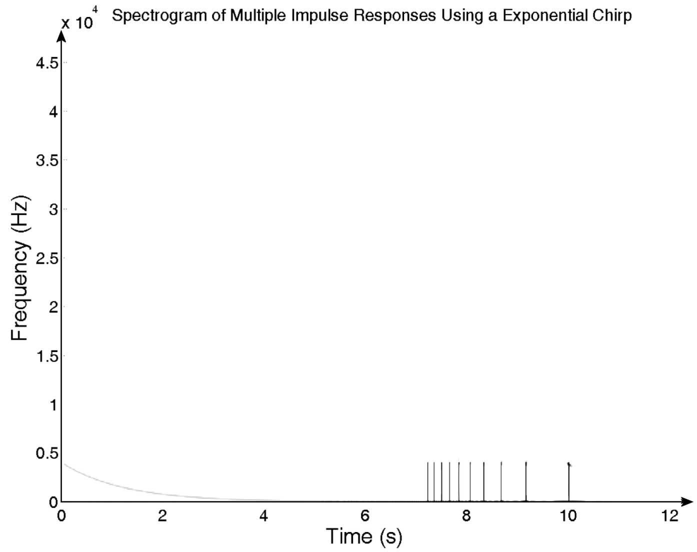


Figure 2–10: Spectrogram of deconvolved MIR when using exponential chirps. Each higher order response packs into an individual impulse at precise anticipatory times before the linear response.

$2f_1$ to $4f_1$. Thus the magnitude of the Fourier Transform of the chirp is not flat within the chirp boundaries and instead declines by 3 dB per octave. Every octave shares the same energy (in terms of the squares of the magnitudes of the frequency domain samples), but this energy spreads out over an increasing bandwidth. Therefore the energy at a particular frequency decreases as the frequency increases. The equal time per octave property of the exponential chirp is the key to why the higher responses pack into impulses. A thorough and more precise explanation will be provided in Section 2.4. Before this can be done, it is necessary to see how the convolution of the inverse filter with the system's response to a chirp produces an impulse response in a linear system.

2.3 Deconvolution

Section 2.2 showed that an exponential chirp could be used in the nonlinear convolution process to extract multiple impulse responses for a nonlinear system. While other types of chirp would not pack the higher order responses into impulses, they did however extract a *linear* impulse. The goal of this section is show how convolution of a time-reversed chirp with the system's response to the chirp produces an impulse response, and why the particular form of the chirp does not matter for extracting simply the linear response.

Since we are not exciting the system under study with an ideal impulse but rather a chirp, some work must be done to extract the impulse response $h(t)$ from the system response. The idea of *deconvolution* is to solve for a function h in the convolution equation, knowing the input $s(t)$ and the response $r(t)$:

$$(s * h)(t) = r(t) \tag{2.7}$$

In general, the problem of deconvolution is ill-posed, and can produce nonsensical results if the response is exactly zero for some output sample [26]. For the purposes of extracting one or more impulse responses, it suffices to find an inverse system $\tilde{s}(t)$ such that convolving it with $s(t)$ approximates a scaled, shifted impulse:

$$(s * \tilde{s})(t) \approx A\delta(t - t_o) \tag{2.8}$$

Where $A\delta(t - t_o)$ is the Dirac delta function delayed in time by an amount t_o and scaled by A. A derivation of the inverse system will now be presented, inspired by [1] with some details filled in. We will restrict the focus slightly such that the amplitude A of the Dirac delta function is $\delta(t)$.

For mathematical convenience, we shall use the analytic versions [8] $z_s(t)$ and $z_{\tilde{s}}(t)$ of $s(t)$ and $\tilde{s}(t)$ respectively. Thus

$$z_s(t) = a_s(t)e^{i\phi_s(t)} \tag{2.9}$$

$$z_{\tilde{s}}(t) = a_{\tilde{s}}(t)e^{i\phi_{\tilde{s}}(t)} \tag{2.10}$$

Both are Fourier Transformed to give $Z_s(f)$ and $Z_{\tilde{s}}(f)$. Therefore we seek $Z_{\tilde{s}}(f)$ such that, ideally

$$Z_s(f)Z_{\tilde{s}}(f) = 1 \tag{2.11}$$

Which is the frequency domain analog of Eq. (2.8).[4] If $Z_s(f)$ and $Z_{\tilde{s}}(f)$ are both expressed in the form

$$Z_s(f) = B_s(f)e^{j\Psi_s(f)} \tag{2.12}$$

then by Eq. (2.11),

$$B_s(f) = \frac{1}{B_{\tilde{s}}(f)} \tag{2.13}$$

$$\Psi_s(f) = -\Psi_{\tilde{s}}(f) \tag{2.14}$$

[4] In practice, we are actually seeking $Z_{\tilde{s}}(f)$ such that $Z_s(f)Z_{\tilde{s}}(f) = e^{-2\pi jft_o}$ to incorporate a possible delay t_o in the arrival of the impulse as per Eq. (2.8).

To find the amplitude $a_{\tilde{s}}(t)$ of $\tilde{s}(t)$, we first look at the expression for the Fourier Transform of $z_{\tilde{s}}(t)$:

$$Z_{\tilde{s}}(f) = \int_{-\infty}^{\infty} a_{\tilde{s}}(t) e^{-i[-\phi_{\tilde{s}}(t) + 2\pi f t]} dt \tag{2.15}$$

In the case of chirp signals, Eq.(2.15) can be viewed as an oscillatory integral. This integral can be approximated by the Method of Stationary Phase [27] [28] [29] [30]. A requirement of this method is that the phase $\Phi_{\tilde{s}}(t) = -\phi_{\tilde{s}}(t) + 2\pi f t$ is monotonically modulated, meaning that the phase is consistently changing in only one direction. Since we are dealing with chirps, this method is valid for approximating the integral in Eq. (2.15). We first assume that the variations in the phase $\Phi_{\tilde{s}}(t)$ are fast compared with the variations in the envelope $a_{\tilde{s}}(t)$. Rapid oscillations of the exponential mean that the integral is roughly zero in the areas where the oscillation is very fast. Thus the essential contribution to the integral is the region where $\Phi_{\tilde{s}}(t)$ varies slowly, i.e. a point of *stationary phase* t_{STAT} where[5]

$$\Phi'_{\tilde{s}}(t_{\text{STAT}}) = 0 \tag{2.16}$$

Because $a_{\tilde{s}}(t)$ is assumed to vary slowly compared to $\Phi_{\tilde{s}}(t)$, $a_{\tilde{s}}(t)$ is assumed to be constant around t_{STAT}. We then Taylor expand $\Phi_{\tilde{s}}(t)$ about t_{STAT} to 2^{nd} order:

$$\Phi_{\tilde{s}}(t) = \Phi_{\tilde{s}}(t_{\text{STAT}}) + \frac{1}{2}\Phi''_{\tilde{s}}(t_{\text{STAT}})(t - t_{\text{STAT}})^2 \tag{2.17}$$

[5] When the expression for $\Phi_{\tilde{s}}(t)$ is substituted in Eq. (2.16), it is clear that we are looking in the area around where $\phi_{\tilde{s}}(t)$ has almost linear phase, i.e. $\phi'_{\tilde{s}}(t_{\text{STAT}}) \approx 2\pi f$.

Plugging the above back into Eq.(2.15), we have

$$Z_{\tilde{s}}(f) = a_{\tilde{s}}(t_{\text{STAT}})e^{-i\Phi_{\tilde{s}}(t_{\text{STAT}})}\int_{-\infty}^{\infty} e^{-\frac{i}{2}\Phi''_{\tilde{s}}(t_{\text{STAT}})(t-t_{\text{STAT}})^2}dt \tag{2.18}$$

This has the form of a Gaussian integral over infinite bounds. The result is

$$Z_{\tilde{s}}(f) = B_{\tilde{s}}(f)e^{i\Psi_{\tilde{s}}(f)} \approx a_{\tilde{s}}(t_{\text{STAT}})\sqrt{\frac{2\pi}{\phi''_{\tilde{s}}(t_{\text{STAT}})}}e^{-i(-\phi_{\tilde{s}}(t_{\text{STAT}})+2\pi f t_{\text{STAT}})+i\frac{\pi}{4}} \tag{2.19}$$

Then we can define the instantaneous frequency and group delay as follows [31]:

$$f_{\tilde{s}}(t) = \frac{1}{2\pi}\frac{d\phi_{\tilde{s}}(t)}{dt} \tag{2.20}$$

$$t_{\tilde{s}}(f) = -\frac{1}{2\pi}\frac{d\Psi_{\tilde{s}}(f)}{df} \tag{2.21}$$

Evaluating Eq.(2.16) using Eq.(2.20) we can see that the instantaneous frequency at t_{STAT} is simply the frequency f as we expect. We can also use Eq.(2.21) to show that the group delay from Eq.(2.19) is simply t_{STAT}.

$$f_{\tilde{s}}(t_{\text{STAT}}) = f \tag{2.22}$$

$$t_{\tilde{s}}(f) = t_{\text{STAT}} \tag{2.23}$$

Thus Eq.'s (2.20) and (2.21) can be regarded as inverses of one another, and if one is known, then the other can be derived. Using the result of Eq.(2.14) in Eq.(2.21) we obtain a simple relation between the group delay of $s(t)$ and $\tilde{s}(t)$. This also allows us to calculate the phase relation between $\phi_s(t)$ and $\phi_{\tilde{s}}(t)$:

$$t_{\tilde{s}}(f) = -t_s(f) \tag{2.24}$$

$$\phi_{\tilde{s}}(t) = \phi_s(-t) \tag{2.25}$$

This tells us that the inverse filter $\tilde{s}(t)$ has a phase that is time-reversed from $s(t)$. The phase of the input signal $\phi_s(t)$ can be obtained from Eq.(2.1). To evaluate $B_{\tilde{s}}(f)$ in Eq.(2.19) we make the substitution $K \equiv \frac{T}{\ln \frac{\omega_2}{\omega_1}}$ and using Eq. (2.25), we have

$$\phi_s''(t_s) = \phi_{\tilde{s}}''(t_{\tilde{s}}) = \frac{2\pi}{K} f(-t_{\tilde{s}}) \tag{2.26}$$

Where $f(t) = f_1 e^{t/K}$. Now we substitute the above results into the magnitude portion of Eq. (2.19):

$$\begin{aligned} B_s(f) &= 1 \times \sqrt{\frac{2\pi}{\phi_{\tilde{s}}''(t_{\tilde{s}})}} \\ &= \sqrt{\frac{K}{f(-t_{\tilde{s}})}} \end{aligned} \tag{2.27}$$

$$\begin{aligned} B_{\tilde{s}}(f) &= a_{\tilde{s}}(t_{\tilde{s}}) \sqrt{\frac{2\pi}{\phi_{\tilde{s}}''(t_{\tilde{s}})}} \\ &= a_{\tilde{s}}(t_{\tilde{s}}) \sqrt{\frac{K}{f(-t_{\tilde{s}})}} \end{aligned} \tag{2.28}$$

Now because $B_s(f)$ and $B_{\tilde{s}}(f)$ are inverses of one another, we can equate Eq.'s (2.27) and (2.28) to solve for $a_{\tilde{s}}(t)$:

$$a_{\tilde{s}}(t) = \frac{f(-t)}{K} = \frac{f_1}{K} e^{-t/K} \tag{2.29}$$

Thus we see that the inverse filter has an exponential amplitude envelope. Substituting the results of Eq.'s (2.29) and (2.25) into the expression for $z_{\tilde{s}}(t)$:

$$z_{\tilde{s}}(t) = \frac{2\pi f_1}{K} e^{-t/K} e^{j\phi_s(-t)} \tag{2.30}$$

Finally, going back to the time domain signal and substituting back in for K, we have our equation for the inverse filter.

$$\tilde{s}(t) = \frac{\omega_1 \ln\left(\frac{\omega_2}{\omega_1}\right)}{T} e^{-t\ln\left(\frac{\omega_2}{\omega_1}\right)/T} \cos\left[\frac{\omega_1 T}{\ln\left(\frac{\omega_2}{\omega_1}\right)}\left(e^{-t\ln\left(\frac{\omega_2}{\omega_1}\right)/T} - 1\right)\right] \tag{2.31}$$

This derivation is general and thus can be applied to other types of chirps, so long as the chirp's frequency evolution is monotonic. In general the inverse filter will consist two parts: an amplitude envelope containing the frequency evolution of the forward chirp but reversed in time $f(-t)$, and a time-reversed version of the chirp itself as the oscillatory part, $e^{j\phi_s(-t)}$.

2.3.1 Analysis of the Generation of Impulses

The nonlinear system response to the chirp $s(t)$ can be broken up into two stages:

- The generation of harmonics of the input signal $s_i(t)$
- The convolution of each of the harmonics with the corresponding harmonic IR $h_i(t)$

Each of these harmonics convolved with an IR are then mixed together to form the output $r(t)$. A diagram illustrating the relationship of these elements is shown in Figure 2–11. Mathematically, we can write:

$$r = \sum_{i=1}^{N} s_i * h_i \tag{2.32}$$

The task then is to disentangle the impulse responses $h_i(t)$ from the input signal. If an inverse filter $\tilde{s}(t)$ can be found that satisfies Eq. (2.8), *the impulse responses*

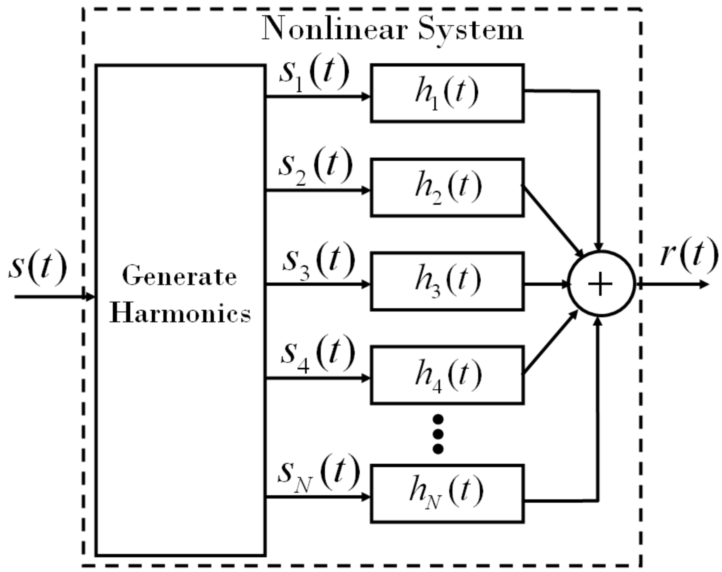


Figure 2–11: The creation of the response $r(t)$ of a nonlinear system to a chirp $s(t)$.

$h_i(t)$ can be extracted by convolving it with the response $r(t)$.

$$\begin{aligned}(\tilde{s} * r) &= \tilde{s} * \left(\sum_{i=1}^{N}(s_i * h_i)\right) \\ &= \sum_{i=1}^{N}(\tilde{s} * s_i) * h_i \\ &= \sum_{i=1}^{N}\delta(t+\Delta t_i) * h_i \\ &= h(t)\end{aligned} \tag{2.33}$$

Where $\delta(t+\Delta t_i)$ are pure unit impulses, delayed by a time Δt_i. The end result is a single time-domain signal $h(t)$ which contains the impulse responses $h_i(t)$ spaced in time and packed into impulse responses. Each impulse response occurs at a precise time delay, with the higher order responses occurring earlier than the lower ones.

At this point in the discussion, nonlinear convolution still has some mysteries to reveal. Why do the higher order responses pack into impulses at precise time delays? How does the inverse filter deconvolve the higher order responses? Farina [3] [4] did not provide the details of how the method generated these higher order responses and oddly, no precise explanation of exactly why nonlinear convolution works could be found in the literature. These questions are addressed in the following section.

2.4 Convolution

The equivalence between convolution in the time domain and multiplication in the frequency domain (see Eq (1.2)) means that the convolution of the system's nonlinear response $r(t)$ with the inverse filter $\tilde{s}(t)$ can be analyzed in two different but equivalent ways. Each view highlights some important features of the MIR sequence and about why Farina's method works. In this section, some empirical and qualitative reasoning will first be presented in both the time and frequency domains, followed by a more formal quantitative explanation of how Farina's method produces the MIR output.

2.4.1 Frequency Domain

When the spectra of $s(t)$ and $\tilde{s}(t)$, denoted $S(f)$ and $\tilde{S}(f)$ are multiplied, the result should have a magnitude response which is flat within the bounds of the chirp, since the spectrum of an impulse is flat. As mentioned in Section 2.2, a chirp with an exponential frequency evolution has more energy in the lower

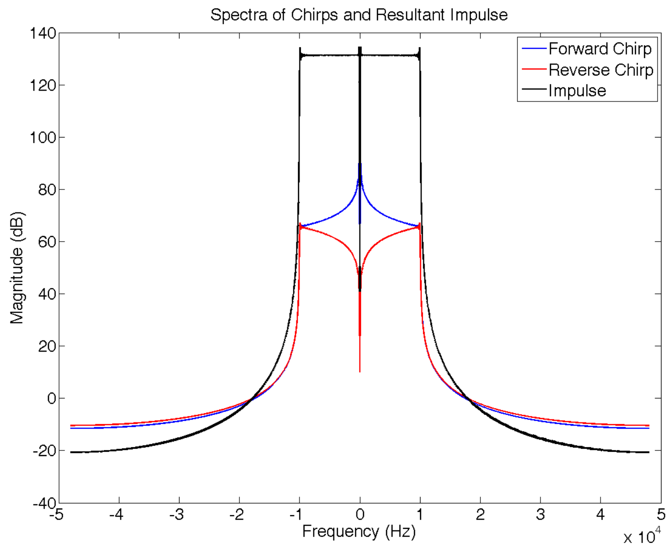


Figure 2–12: Spectra of forward and reverse chirps, and resultant impulse. Chirps sweep from 100 Hz to 10 000 Hz, and thus the impulse has a flat spectrum through these frequencies.

frequencies. Thus if a flat spectrum is desired in the convolved result, then $\tilde{s}(t)$ must have proportionally less energy in the lower frequencies and more in the higher frequencies. This is the role that the amplitude envelope plays in Eq. (2.31). The magnitude spectra of $S(f)$ and $\tilde{S}(f)$, as well as the spectrum of the resultant band-limited impulse from their convolution are shown in Figure 2–12. From this Figure, it is clear that the product of magnitude spectra of $s(t)$ and $\tilde{s}(t)$ will give the spectra of the impulse. This however, is not definitive proof that the magnitude spectrum of the impulse in Figure 2–12 actually yields an impulse in the time domain. Two different complex spectra can share the same magnitude spectrum, but differ in their phase spectrum, yielding a different time-domain signal upon inverse FFTing [8]. For example, a linear chirp can also have a flat spectrum, but is clearly very different from an impulse. Time domain arguments in the following section will show how impulses result when performing the nonlinear convolution.

The effect of convolving $\tilde{s}(t)$ with a harmonic generated by a nonlinear system is much the same as the case of convolving $s(t)$ and $\tilde{s}(t)$, with a few minor differences. This is best illustrated by example. Consider a nonlinear system whose input/output relation is a 3rd order Chebyshev Polynomial:

$$y(x) = T_3(x) = 4x^3 - 3x \tag{2.34}$$

The Chebyshev polynomials $T_n(x)$ have a very useful property that when a cosine function is applied to $T_n(x)$, one gets purely the nth harmonic of the cosine:

$$T_n(cos(x)) = cos(nx) \tag{2.35}$$

Thus if the exponential chirp $s(t)$ has a bandwidth of $[f_1, f_2]$, the response $r(t)$ has a bandwidth of $[3f_1, 3f_2]$. Convolving the response $r(t)$ with the inverse filter yields a spectra as shown in Figure 2–13. The spectrum of $(r * \tilde{s})(t)$ is flat in the region where both functions have significant energy, but sharply drops off outside the bandwidth of $\tilde{s}(t)$. Thus as the order of the nonlinearity increases, the bandwidth of the harmonic IR decreases. It is bounded above by the maximum frequency of $\tilde{s}(t)$ and below by nf_1. Novak [1] dealt with this problem by extending the bandwidth of $\tilde{s}(t)$ out to nf_2, where n is the number of harmonics to be extracted, thus yielding a much broader bandwidth in the resulting impulse. This is discussed further in the following chapter. However, even without this modification, the result of the convolution still packs into a sharp impulse in the time domain.

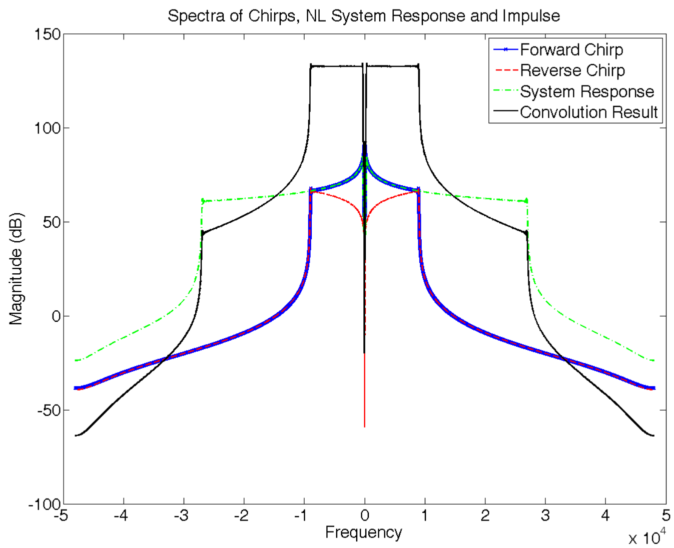


Figure 2–13: Spectra of forward and reverse chirps, the response of the nonlinear system of Eq.(2.34), and resultant impulse. The system response's bandwidth is increased to 3 times the bandwidth of $s(t)$. The impulse spectrum is only flat in the region where both $r(t)$ and $\tilde{s}(t)$ have significant energy.

2.4.2 Time Domain

To illustrate how these higher order responses pack into MIR upon deconvolution, we will proceed by stepping through a number of plots at different times during the convolution. The example nonlinear system will be an even mix of the first three Chebyshev polynomials $T_n(x)$:

$$\begin{aligned} T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \end{aligned} \tag{2.36}$$

Thus the nonlinear system response can be written as in Eq.(2.37). It will produce equal strength 1^{st}, 2^{nd}, and 3^{rd} harmonics of the input.

$$r(t) = \frac{1}{3}\left(T_1(s(t)) + T_2(s(t)) + T_3(s(t))\right) \tag{2.37}$$

For illustration purposes, $s(t)$ evolves exponentially from 1 Hz to 5 Hz in this example in order to show the convolution process clearly. Because of the band-limited nature of the chirps, the resultant impulses are not ideal, and a lot of 'pre-ringing' and 'post-ringing' artifacts are evident surrounding the impulse. When using this method with audio-bandwidth chirps, these extraneous effects become very small compared with the magnitude of the chirps, which are typically orders of magnitude larger.

Figure 2–14 shows the beginning of the convolution between $\tilde{s}(t)$ and $r(t)$. $\tilde{s}(t)$ is time-reversed for the convolution as in Eq. (1.1). Since both $s(t)$ and $\tilde{s}(t)$ are originally causal, they do not exist for $t < 0$ and $\tilde{s}(t)$ does not need to be shifted towards $-\infty$. $\tilde{s}(t)$ is then progressively moved to the right, and at each step, the product of the functions is integrated.

The important spots in the convolution will be illustrated in the following Figures, namely those points in time that produce an impulse. To see where the impulses occur, we can look at the MIR result in Figure 2–15. If we observe graphically the convolution at the locations of the peaks, it starts to become apparent why the impulses occur. The convolution plots at the three marked points are shown in Figures 2–16, 2–17, and 2–18. In each Figure, a full view of both functions is provided, followed by a zoomed in image of the *interaction*

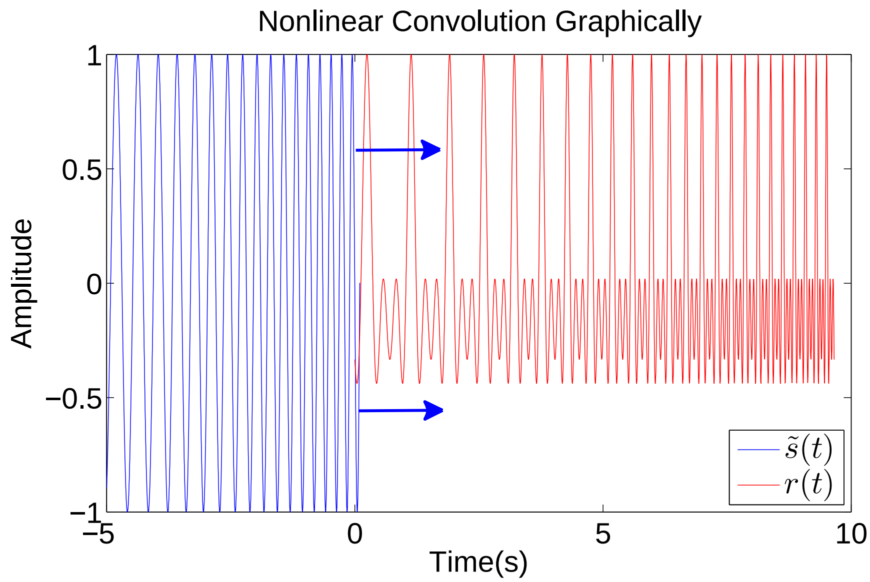


Figure 2–14: The beginning of the the convolution between $\tilde{s}(t)$ and $r(t)$. $\tilde{s}(t)$ is time-reversed and shifted until there is no overlap between the functions. Blue arrows indicate the direction of shifting of $\tilde{s}(t)$.

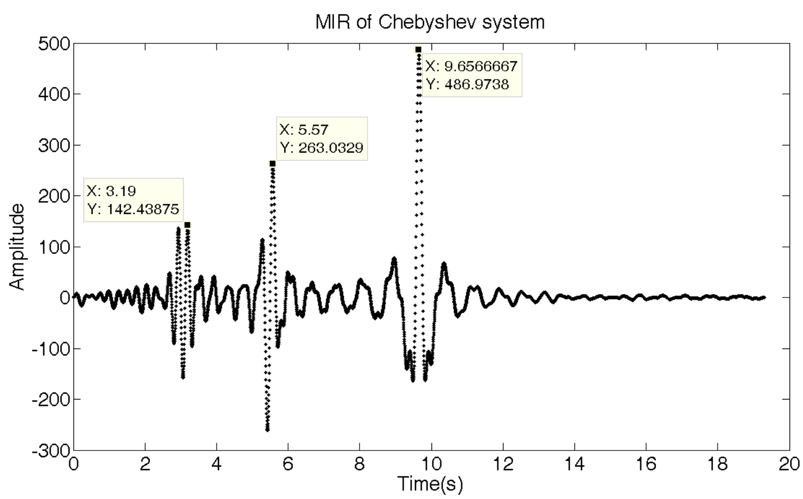
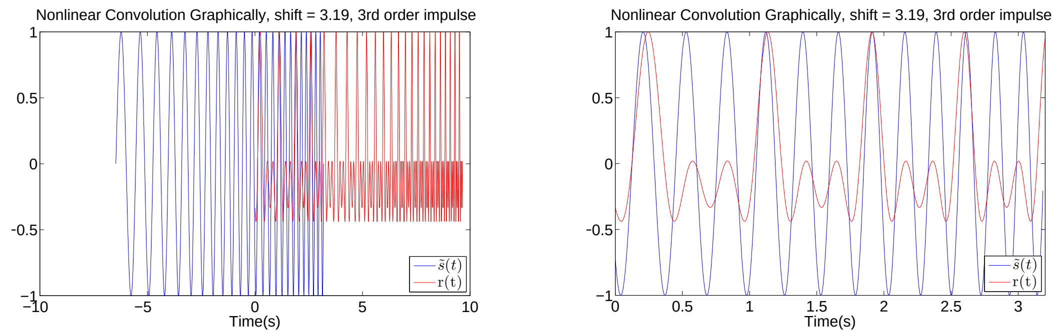


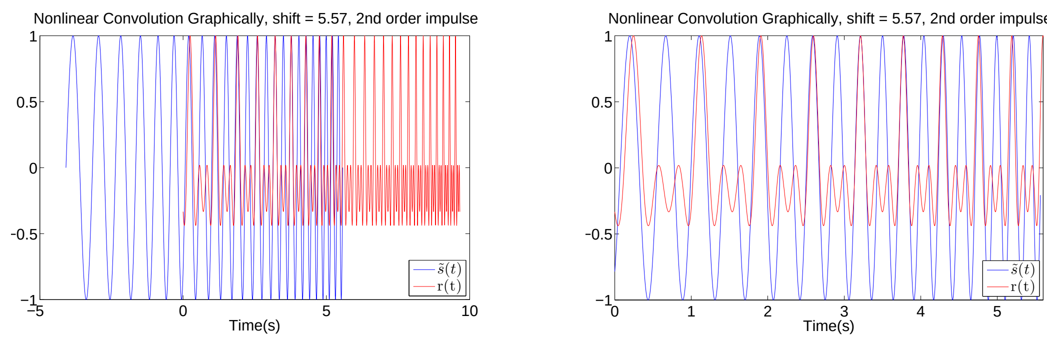
Figure 2–15: MIR of the system in Eq. (2.37).

region, which shall be defined as the region where the two functions overlap during the convolution. The parts of the function in the interaction region at a given time contribute to the convolution at that time.



(a) (b)

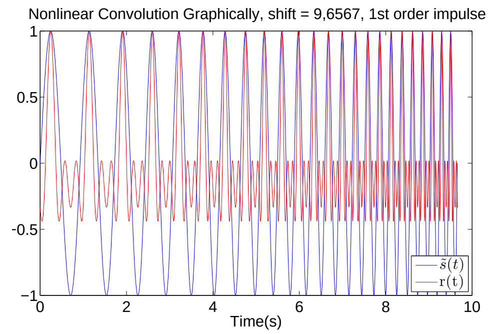
Figure 2–16: Graphical convolution at the time of the 3^{rd} order impulse. (a) Full view of the convolution. (b) Interaction Region of the two functions.



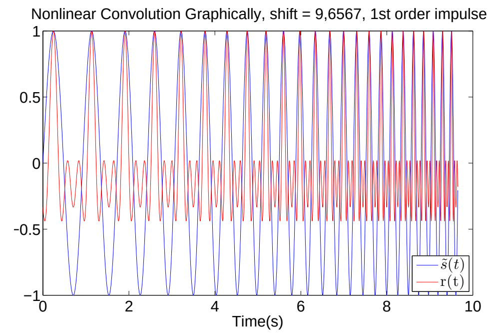
(a) (b)

Figure 2–17: Graphical convolution at the time of the 2^{nd} order impulse. (a) Full view of the convolution. (b) Interaction Region of the two functions.

The plot in Figure 2–16 shows the large peaks of the response lining up with every 3^{rd} peak of $\tilde{s}(t)$. At this instant the interaction region consists of approximately the last 3 seconds of $\tilde{s}(t)$ and the first 3 seconds of $r(t)$. Because



(a)



(b)

Figure 2–18: Graphical convolution at the time of the 1^{st} order impulse. (a) Full view of the convolution. (b) Interaction Region of the two functions.

exponential chirps and hence $\tilde{s}(t)$ spend equal time in each octave, the frequency range spanned by $\tilde{s}(t)$ in the interaction region is the same as that of the 3rd harmonic of the response at the time of the impulse. Furthermore note that from the point of view of the convolution, both functions are of increasing frequency from left to right. Thus when these functions are multiplied and integrated, a large value in the integral at that time results. The case is similar for all other impulse responses as shown in Figures 2–17 and 2–18. When the frequency range spanned by $\tilde{s}(t)$ and a harmonic of $r(t)$ in the interaction region is the same, a large value results in the integral. Figure 2–19 shows how the frequency content of $r(t)$ and $\tilde{s}(t)$ evolve in the interaction region over the course of the convolution. As indicated in the Figure by the arrows, there are times along the convolution where both $\tilde{s}(t)$ and a harmonic of $r(t)$ have the exact same frequency content in the interaction region. Thus each resulting band-limited harmonic impulse response contains the frequencies common to both signals in the convolution at the time of the impulse.

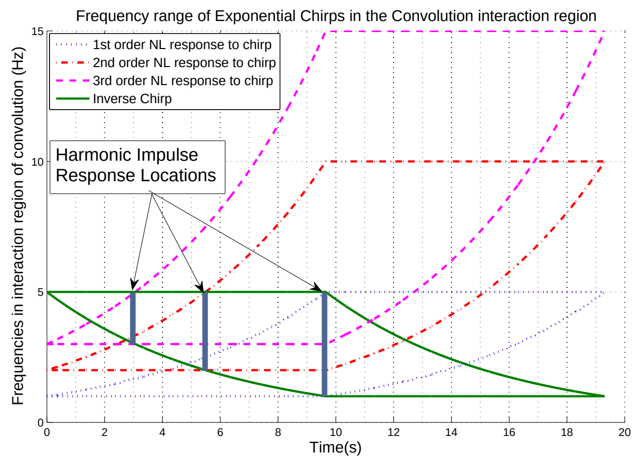


Figure 2–19: Frequency content of exponential chirps in the interaction region. The content of a harmonic of $r(t)$ or $\tilde{s}(t)$ in the region is bounded above and below by the curves of the same colour. The impulses occur when the inverse chirp $\tilde{s}(t)$ has identical frequency content in the interaction region of the convolution as a harmonic of the nonlinear response.

For other types of chirps, this is not the case. Figure 2–20 shows how the frequency content of $r(t)$ and $\tilde{s}(t)$ evolve in the interaction region when a linear chirp excitation is used. When both $\tilde{s}(t)$ and a harmonic of $r(t)$ have the same maximum frequency in the interaction region, their minimum frequencies do not line up at the same instant, and thus the response does not 'pack' into a sharp impulse as is apparent in the spectrograms of Figures 2–7 and 2–9. This is indicated in Figure 2–20 by the non-vertical lines showing the higher order responses 'smearing' in time.

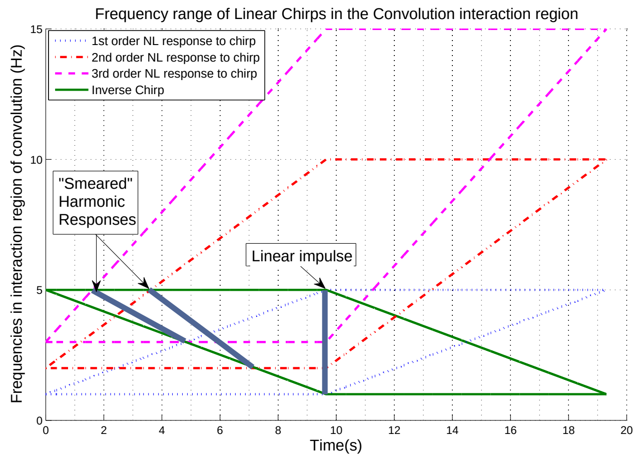


Figure 2–20: Frequency content of linear chirps in the interaction region. The content of $\tilde{s}(t)$ or a harmonic of $r(t)$ in the region is bounded above and below by the curves of the same colour. The higher order responses smear out in time because $\tilde{s}(t)$ and a given harmonic of $r(t)$ never share the exact same frequency content at any instant.

Because $s(t)$ and $\tilde{s}(t)$ are deterministic, and the nonlinear system response produces integer harmonics of the input, the arrival times of the impulses in the MIR are easily calculated a priori. To find the amount of time the n^{th} harmonic IR precedes the linear one (which occurs at time T), we look at when the instantaneous frequency of $s(t)$ is n times the current frequency [3]. In other words, we solve the following equation for Δt_n:

$$n\frac{d\phi_s(t)}{dt} = \frac{d\phi_s(t+\Delta t_n)}{dt} \tag{2.38}$$

Where $\phi_s(t)$ is the phase of $s(t)$. For the exponential chirp, the result is:

$$\Delta t_n = \frac{T\ln(n)}{\ln\left(\frac{\omega_2}{\omega_1}\right)} \tag{2.39}$$

Eq.(2.39) gives the advance time that the n^{th} harmonic appears ahead of the linear impulse at time T. Thus in the MIR output, the highest order IR appears first at time $T - \Delta t_N$ and are followed by the next lower one at time $T - \Delta t_{N-1}$, and so on until the linear IR ($n = 1$) at time T.

Eq.(2.38) also provides insight as to why the exponential chirp sets itself apart from the other types of chirp for nonlinear convolution. The critical outcome is that when Eq. (2.38) is solved for Δt_n, the result is *independent* of both the current frequency $\omega(t)$ and on t itself. When the expressions for the frequency of the linear (Eq. (2.2)) and cosine (Eq. (2.4)) chirps are used in Eq. (2.38) and solved for Δt_n, we see that the times at which the instantaneous frequency is n times the current frequency are dependent on $\omega(t)$ and t. For linear chirps the result is

$$\Delta t_n = n - 1 \frac{\omega_{lin}(t)}{(\omega_2 - \omega_1)/2} \tag{2.40}$$

and for cosine chirps

$$\Delta t_n = -t + \frac{T}{\pi} \arccos \left[\frac{2}{\omega_1 - \omega_2} \left(n\omega_{cos}(t) + \frac{\omega_1 - \omega_2}{2} - \omega_1 \right) \right] \tag{2.41}$$

This explains why when using linear and cosine chirps the higher order responses from nonlinear convolution do not pack into impulses but instead are smeared out over time (See Figures 2–7 and 2–9). Both the linear and cosine chirps do however yield an impulse for the linear response (n=1) , and solving the above equations with $n = 1$ yields a constant $\Delta t_1 = 0$ as expected.

Eq.(2.39) implies that the impulses are spaced unevenly in time, with the higher order impulse responses being spaced closer together. This can be problematic if the system is highly reverberant; the higher order impulse responses can overlap in time with each other. The impulse responses can however be spaced further apart by making the length of the chirp T longer. If one knows the length of the longest impulse response (or the T_{60} [32]), then the length of the chirp required to extract the MIR without overlap is given by using Eq. (2.39) for two different values of n and subtracting to give the time delay between two adjacent HIRs. This time lag must be greater than the length of the longest IR (or longest T_{60}) in the system:

$$T > \frac{T_{longest} \ln\left(\frac{\omega_2}{\omega_1}\right)}{\ln\left(\frac{n+1}{n}\right)} \tag{2.42}$$

To conclude, nonlinear convolution is a relatively simple method for the characterization of nonlinear systems. The method takes advantage of the equal time per octave property of exponential chirps, which allows the response of a nonlinear system to be deconvolved into multiple impulse responses, which provide more detailed information about the nonlinear system than a simple THD or ID measurement. These impulse responses can then be used in a nonlinear model for the system under study. This is the subject of the next chapter.

Chapter 3
Re-Synthesis of Nonlinear Systems

The previous chapter has explained the method of nonlinear convolution and how it can be used to extract MIR, which describe which harmonics are generated by a nonlinear system, their relative strengths, and how each harmonic of the input is filtered. In the analysis/synthesis approach, the extraction of the MIR is the *analysis* phase of the procedure, and the MIR are the pieces of information, or the 'parameters' necessary to the *synthesis* phase.

This chapter focuses on a model for the software simulation of nonlinear systems. The MIR obtained from the nonlinear convolution method are used as parameters in this model. The goals of this chapter are as follows: to describe the re-synthesis model used and some closely related models, to describe modifications to the nonlinear convolution method necessary for accurately re-synthesizing nonlinear systems, and to show the results of simulating some synthetically created nonlinear systems with this model.

The model employed here (Section 3.4) is based on the polynomial Hammerstein and MISO models, which are briefly discussed in Sections 3.2 and 3.3. Before the synthesis models are discussed, some modifications to the signals used in the nonlinear convolution method are necessary to improve the final result. This is detailed in Section 3.1 below. Finally Section 3.5 presents results of synthesizing some simple synthetic nonlinear systems.

3.1 Modifications to the Nonlinear Convolution Method

When Farina [3][4] first presented the method for nonlinear convolution, his intent was to use only the rightmost HIR as a linear impulse response with any nonlinear distortion artifacts 'pushed out' of the response. The higher order HIR were regarded as artifacts and not used. If these HIR are to be used in synthesis, some modifications to the excitation signals must be made.

3.1.1 Synchronization of Harmonics of $s(t)$

The first modification is described in [1], and involves the synchronization of the phases at harmonics of the excitation signal. That is, it is desired that $s(t)$ have the same value and sign of the slope when its instantaneous frequency is n times the start frequency of the chirp f_1. A diagram showing the desired result for $s(t)$ being an exponential chirp is given in Figure 3–1. First, an expression is found

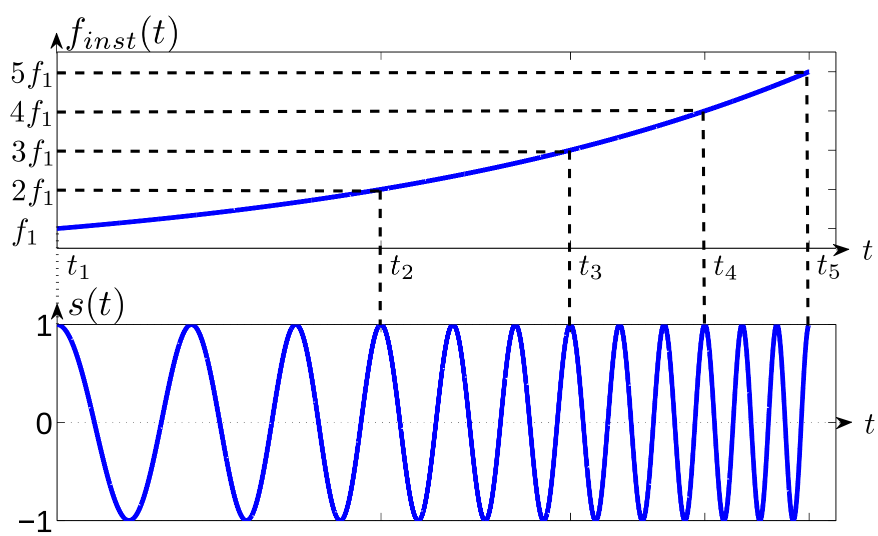


Figure 3–1: Illustration of synchronized harmonics of the exponential chirp. The value of the chirp when $f_{inst} = nf_1$ is the same for integers n.

for the times at which the instantaneous frequency equals Nf_1. The equation for the instantaneous frequency (Eq.(2.6)) is equated to Nf_1 and solved for $t = t_N$:

$$t_N = K \ln N \tag{3.1}$$

where $K \equiv \frac{T}{\ln \frac{\omega_2}{\omega_1}}$ as before. When the above is substituted into the equation for the phase of the exponential chirp (see Eq. (2.1)), we obtain

$$\phi(t_N) = 2\pi K f_1 (N - 1) \tag{3.2}$$

Now if $\phi(t_N)$ were an integer multiple of 2π, then $s(t_N) = 1$ for every N. This condition is thus satisfied if we modify K such that

$$K \equiv \frac{1}{f_1} round\left(\frac{f_1 \hat{T}}{\ln \frac{\omega_2}{\omega_1}}\right) \tag{3.3}$$

Where $\hat{T}$ is an approximative time length of the chirp $s(t)$ used for the design [1]. This modification of K is manifested in the length of the chirp T, such that

$$T = \frac{\ln \frac{\omega_2}{\omega_1}}{f_1} round\left(\frac{f_1 \hat{T}}{\ln \frac{\omega_2}{\omega_1}}\right) \tag{3.4}$$

Finally using Eq.(3.3) in Eq.(2.1) we have the modified chirp:

$$s(t) = \cos\left[2\pi f_1 K \left(e^{\frac{t}{K} - 1}\right)\right] \tag{3.5}$$

Thus by slightly modifying the length of $s(t)$, we can be assured that the phase of the chirp is equal to 2π at multiples of the start frequency f_1. This is important for obtaining large values for the impulses in the MIR output of nonlinear convolution. If the phases of the harmonics were not synchronized, artifacts such as pre and post-ringing of the impulses would be larger relative to the main peaks in the impulses, and thus a less ideal impulse with smaller signal-to-noise ratio would result.

3.1.2 Extension of the Inverse System

When testing systems that were only weakly nonlinear, Farina could allow the bandwidth of the excitation to be from the subsonic region to near the Nyquist frequency [33]. However for nonlinear systems the bandwidth of the excitation must be limited because the presence of an n^{th} order nonlinearity increases the bandwidth of the system response n times. For an excitation of bandwidth $[f_1, f_2]$, the n^{th} harmonic response of a nonlinear system will have a bandwidth $[nf_1, nf_2]$. The inverse system used in the nonlinear convolution method has the same bandwidth as the excitation, and when nonlinear convolution is performed the bandwidth of each of the HIR contains those frequencies common to both $\tilde{s}(t)$ and the n^{th} harmonic response. Therefore all HIR have an upper bandwidth limit f_2. An example of the spectra of the chirp, system response, inverse system, and resultant impulse can be found in Figure 2–13.

Novak [1] addressed this problem by allowing the bandwidth of the inverse system $\tilde{s}(t)$ to extend out to Nf_2, where N is the number of harmonics to be considered. If the amplitude envelope is left unchanged then the nth HIR will have a bandwidth of $[nf_1, min(nf_2, Nf_2)]$. This is done by extending the length of the inverse chirp, such that it starts at a frequency of Nf_2 and ends at f_1. The length of the chirp T is changed to

$$T_{\tilde{s}} = \frac{K}{f_1} \ln\left(\frac{Nf_2}{f_1}\right) \tag{3.6}$$

With K unchanged from Eq.(3.3). Due to the presence of K above, $\tilde{s}$ undergoes a similar phase alignment to $s(t)$ as described above in Section 3.1.2. Finally keeping the amplitude envelope the same, $T_{\tilde{s}}$ can be substituted into the argument of the cosine of Eq.(2.31) to arrive at the equation for the modified inverse system:

$$\tilde{s}(t) = \frac{\omega_1 \ln\left(\frac{\omega_2}{\omega_1}\right)}{T} e^{-t\ln\left(\frac{\omega_2}{\omega_1}\right)/T} \cos\left[\frac{\omega_1 T_{\tilde{s}}}{\ln\left(\frac{\omega_2}{\omega_1}\right)}\left(e^{-t\ln\left(\frac{\omega_2}{\omega_1}\right)/T_{\tilde{s}}} - 1\right)\right] \tag{3.7}$$

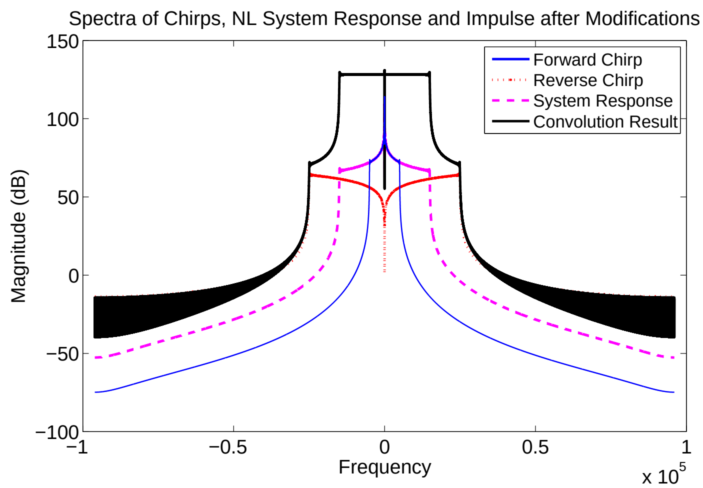


Figure 3–2: Magnitude spectra of forward and extended reverse Chirps, the response of the nonlinear system of Eq.(2.34), and resultant impulse. $\tilde{s}(t)$'s bandwidth (1-25000 Hz) is increased to 5 times that of $s(t)$(1-5000 Hz). The nonlinearity increases the response's bandwidth to 3 times the bandwidth of $s(t)$.

A figure showing the magnitude spectra of a chirp, the extended inverse system, the response of a chirp to the nonlinearity described in Eq. (2.34), and the result of convolving the response of this system with the inverse system is shown in Figure 3–2. This figure presents the same concepts as Figure 2–13, except Figure 3–2 shows the signal spectra with the modifications above implemented. In this example $\tilde{s}(t)$ is extended to 5 times the bandwidth of $s(t)$. The resultant impulse now has a larger bandwidth, extending in this case out to the bandwidth of $r(t)$.

3.2 Polynomial Hammerstein Model

The Polynomial Hammerstein model [37] is a method for the simulation of nonlinear systems. The so-called simple Hammerstein model consists of a zero memory polynomial operator followed by a linear dynamics element such an impulse response. A zero-memory or static non-linear system modifies the signal by some instantaneous process. This means that the output of the system at some time t depends solely on the input at that same time and not on past or future inputs. An example of a simple Hammerstein model is shown in Figure 3–3. The

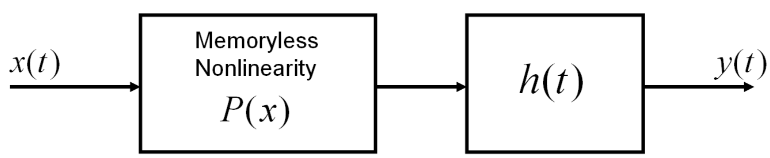


Figure 3–3: A simple Hammerstein model with general memoryless nonlinearity and linear part with impulse response $h(t)$.

Hammerstein representation of a nonlinear system is advantageous because the nonlinear parts of the system are separated from the linear part, simplifying the conceptualization of a nonlinear system. The idea of the simple Hammerstein model can be expanded into a parallel branched structure containing multiple different nonlinearities and linear elements, thus expanding the model's flexibility. This idea is illustrated in Figure 3–4, where we restrict the nonlinearity to be a polynomial of order N. Any polynomial nonlinearity of the form $a_1x + a_2x^2 +$

$a_3x^3 + \ldots$ can be represented by multiplying the respective linear impulse response $h_i(t)$ by a_i.

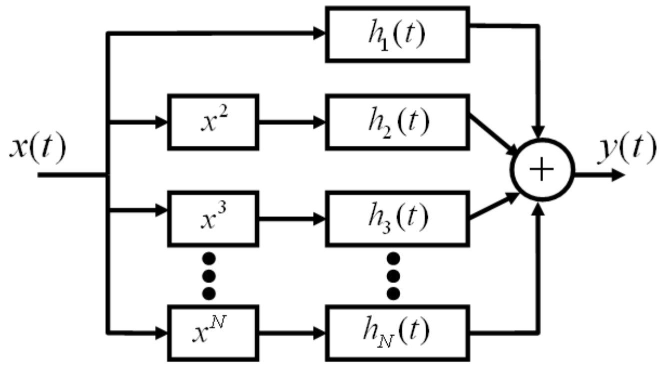


Figure 3–4: Multiply branched polynomial Hammerstein model.

This model can be thought of as an expansion to the classical waveshaping paradigms initally proposed in [34] and [35], which processes a signal by specifying an input/output relationship of a system as a polynomial function. The difference in the Polynomial Hammerstein model is that each of the branches has its own independent filter, which could be used to control the transients or colourize different powers of the input independently.

3.3 MISO Model

Bendat and Piersol [36] proposed a statistically-based model similar in structure to the Polynomial Hammerstein model of Section 3.2 called the Multiple Input Single Output (MISO) Model. The MISO model was originally developed for the identification of systems and is primarily used under the assumption that the input signals are statistical in nature such as Gaussian white noise. The goal was the identification of the linear filters $h_i(t)$ and polynomial coefficients of the nonlinearities from the input and output signals using multiple cross-correlations and autocorrelation techniques [37] [36] [38].

The MISO model can also serve as a framework for re-synthesis if the impulse responses and nonlinear functions are known. In this case, the MISO model takes up the same form as the polynomial Hammerstein model shown in Figure 3–4, but each of the nonlinear functions can be *any* function of the input signal, rather than a simple power x^n. That is, the nth nonlinear function can be written as $P_n[x(t)]$, where P_n depends solely on the input x at the instant t. Thus the Polynomial Hammerstein model is a specific case of the MISO model.

3.4 Re-Synthesis Model

The synthesis model developed in this work for the re-synthesis of nonlinear systems has structure similar to both the Hammerstein and MISO models. The impulse responses $h_i(t)$ are known, derived from the nonlinear convolution method. The nonlinear functions $P[x(t)]$ however have yet to be identified.

The approach taken in this re-synthesis method is called a 'blind identification', because no assumptions are made about the nonlinear system under study. Thus the nonlinear system to be re-synthesized is treated as a black box having one input and one output. In this case the nonlinearities could simply be powers of $x(t)$. However these nonlinear elements would not be compatible with the MIR derived from the nonlinear convolution method, since the $h_i(t)$ are responses to a *pure harmonic* of the input signal. For harmonic signals like the excitation $s(t)$, the simple power-laws x^n are related to harmonics of the input signal by

$$\cos^n \theta = \begin{cases} \frac{2}{2^n} \sum_{k=0}^{\frac{n-1}{2}} \binom{n}{k} \cos((n-2k)\theta) & n \text{ odd} \\ \frac{2}{2^n} \sum_{k=0}^{\frac{n}{2}-1} \binom{n}{k} \cos((n-2k)\theta) + \frac{1}{2^n}\binom{n}{\frac{n}{2}} & n \text{ even} \end{cases} \tag{3.8}$$

This implies that applying a power-law nonlinearity would produce multiple harmonics instead of just one. For example, Eq.(3.8) with $n = 5$ implies the signal contains harmonics of 1^{st}, 3^{rd}, and 5^{th} order of differing strengths:

$$\cos^5 \theta = \frac{10cos\theta + 5cos3\theta + cos5\theta}{16} \tag{3.9}$$

3.4.1 Change of basis of the HIR

There are two options to solve this problem. One option is to perform a change of basis on the set of HIR, as done in [1]. Eq.(3.8) gives expressions for $\cos^n(\theta)$ as linear combinations of $\cos(n\theta)$. It is possible to write out Eq.(3.8) for each n and create a matrix equation:

$$\begin{pmatrix} 1 \\ \cos(\theta) \\ \cos^2(\theta) \\ \cos^3(\theta) \\ \vdots \end{pmatrix} = A \begin{pmatrix} 1 \\ \cos(\theta) \\ \cos(2\theta) \\ \cos(3\theta) \\ \vdots \end{pmatrix} \tag{3.10}$$

Where A is the matrix of coefficients given by Eq.(3.8). We then create a vector of the HIR $h_i(t)$ obtained by nonlinear convolution, denoted H. The goal is to find a set of impulse responses G that are compatible with power-law nonlinearities. It can be shown that

$$H = (A^T)G \tag{3.11}$$

The equation can then be inverted to get an expression for the impulse responses G to be used in the re-synthesis model:

$$G = (A^T)^{-1}H \tag{3.12}$$

Thus by a change of basis the set of impulse responses H obtained through nonlinear convolution can be changed into the set G for use with power-law nonlinearities.

3.4.2 Use of Chebyshev Polynomials as nonlinear elements

Instead of performing a complicated change of basis, which involves a matrix inversion, it would be advantageous if the nonlinearities could instead be chosen such that they are compatible directly with the HIR $h_i(t)$. That is, we seek a nonlinearity such that when it is applied to the input, a single pure harmonic of the input results. The Chebyshev polynomials $T_n(x)$ have the useful property that when they are applied to a cosine function, a pure harmonic results:

$$T_n(\cos\theta) = \cos(n\theta) \tag{3.13}$$

Chebyshev polynomials are most simply defined recursively:

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_n(x) &= 2xT_n(x) - T_{n-1}(x) \end{aligned} \tag{3.14}$$

Similar to the process of creating the matrix A in Section 3.4.1, a matrix, denoted here as C can be created by taking the coefficients of the Chebyshev polynomials such that

$$\begin{pmatrix} T_0(x) & T_1(x) & T_2(x) & \dots \end{pmatrix} = \begin{pmatrix} 1 & x & x^2 & \dots \end{pmatrix} C \tag{3.15}$$

Perhaps unsurprisingly, it turns out that this matrix C is related to the matrix A by

$$C = (A^T)^{-1} \tag{3.16}$$

This makes sense if we consider that in Section 3.4.1 the change of basis was done to express the set of impulse responses in terms of a basis of powers of x, where the Chebyshev polynomials are used to express the nonlinearities in the same basis as the $h_i(t)$. It is also even more apparent when we make the substitution $x = cos\theta$ in Eq. (3.8) and also utilize Eq. (3.13). Thus Eq. (3.8) can be written as [39]

$$x^n = \sum_{k=0}^{n} a_{n,k} T_k(x) \tag{3.17}$$

Where $a_{n,k}$ are the constants defined in Eq. (3.8). Thus Eq. (3.17) is simply the inverse of Eq. (3.15) as is confirmed by Eq. (3.16).

Since the $T_n(x)$'s are easily defined, this is the method chosen in this work. Therefore the HIR are the exact same ones extracted via the nonlinear convolution method, and the nonlinearities chosen for each branch of the synthesis model are the Chebyshev polynomials. A block diagram is shown in Figure 3–5.

3.5 Implementing and Testing the Algorithm

In order to demonstrate the entire analysis and synthesis method, two artificially created nonlinear systems were analyzed by the nonlinear convolution

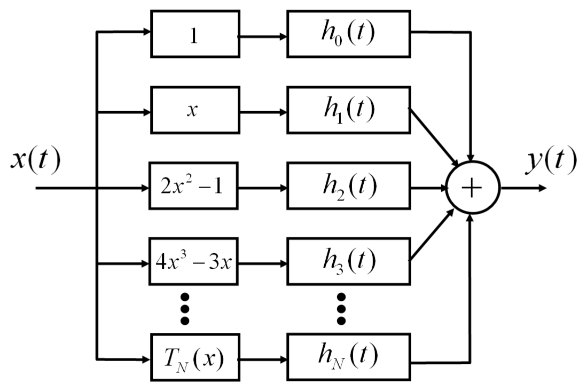


Figure 3–5: Synthesis model with Chebyshev polynomials.

method and then re-synthesized using the Hammerstein/MISO model described in the previous section. Both the NSUT and simulated system will be compared by sending an identical input to both systems and comparing the outputs in the time and frequency domains. The first system will be a system that generates a square wave out of a sine wave, and the second system is a simple power-law nonlinearity with memory. The nonlinear convolution and re-synthesis algorithms were developed and tested in the MATLAB programming environment.

3.5.1 Fast Convolution

To perform the convolutions necessary in both the analysis and synthesis stages, a fast convolution algorithm was developed making use of the efficiency of the FFT [9]. The excitation and system response signals have a length on the order of 10^6 samples each. Convolving two of these sequences in direct form (see Eq. (1.1)) is extremely costly, $O(n^2)$. The FFT algorithm employed here uses the overlap-add approach [10] and extends it to segment both functions into chunks for FFT-ing, creating a double overlap-add algorithm for fast convolution.

The computational savings are tremendous ($O(n \log_2 n)$), and the convolution of $r(t)$ and $\tilde{s}(t)$ can be done quickly and efficiently. This convolution algorithm is used in both the nonlinear convolution method and in the re-synthesis model for convolving the signals with the HIR.

3.5.2 Square Wave Generator

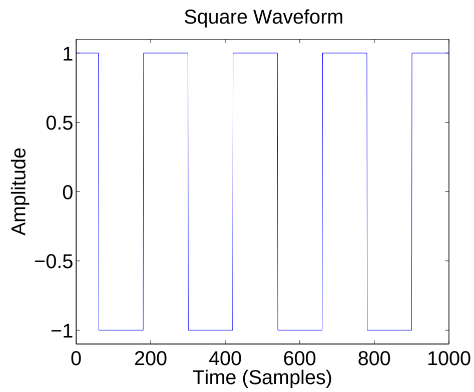


Figure 3–6: Square waveform.

The square waveform is common in signal processing, and is illustrated in Figure 3–6. An even square wave of amplitude 1 and frequency f_o in continuous time is given by

$$x_{f_o,square}(t) = \mathrm{sgn}(\cos(2\pi f_o t)) \tag{3.18}$$

Where sgn((x)) is +1 when $x > 0$, and -1 when $x < 0$. A more useful definition of the square waveform is the Fourier Series definition:

$$x_{f_o,square}(t) = \frac{4}{\pi} \sum_{n=0}^{\infty} (-1)^n \frac{\cos(2\pi(2n+1) f_o t)}{2n+1} \tag{3.19}$$

An even square wave is composed of the sum of odd harmonics of a cosine scaled inversely by the harmonic number. Therefore in order to generate n harmonic components of the input signal, the highest multiple of the fundamental that needs to be generated is $2n + 1$. Because the square wave can be written as a sum of harmonics of a cosine function, a square wave of some frequency f_o can be thought of as the output of a nonlinear system to a cosine input with the same f_o. The nonlinear system would ideally generate harmonics of the same strength as those in the Fourier Series definition given in Eq.(3.19). A simple block diagram illustrating the concept of this system is given in Figure 3–7.

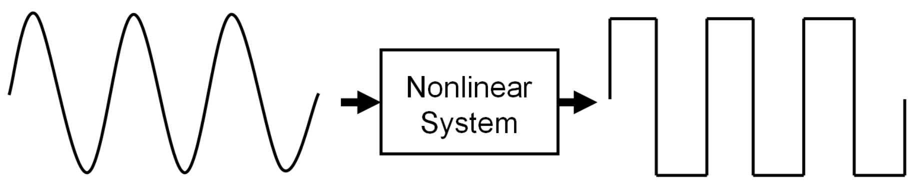


Figure 3–7: Concept of using a nonlinear system to turn a sine wave into a different waveform. In this case a square wave is being generated. The actual result will be a band-limited square wave and will not have such sharp edges.

This function is of particular interest in the case of re-synthesis because the cosine term in Eq. (3.19) allows for us to take advantage of the pure harmonic-generating property of the $T_n(x)$ (see Eq. (2.35)) to generate harmonics of the excitation $s(t)$.

This system was chosen for a number of reasons. First, it is very easy to calculate the amplitude of the n^{th} generated harmonic through Eq.(3.19). Secondly, the only signals involved in the re-synthesis stages are sinusoids and pure impulses. Thus this example serves as a simple introduction to the re-synthesis scheme in practice.

In discrete time, it is not possible to have an arbitrary number of harmonics due to aliasing, thus the notion of a band-limited version of the square wave is such that sum in Eq. (3.19) goes to the the highest harmonic that does not cause aliasing. Because only a finite number of harmonics are used, the band-limited version is not a perfect square wave but exhibits oscillations along the flat sections of the square wave. This is the nonlinear system that was investigated in this section.

The nonlinear response $r(t)$ was artificially created for this case and consists of the excitation $s(t)$ with 5 scaled harmonics components as in $x_{f_o,square}(t)$.[1] Thus the response can be written as

$$r(t) = \frac{4}{\pi}\sum_{n=0}^{5}(-1)^n \frac{s((2n+1)t)}{2n+1} \tag{3.20}$$

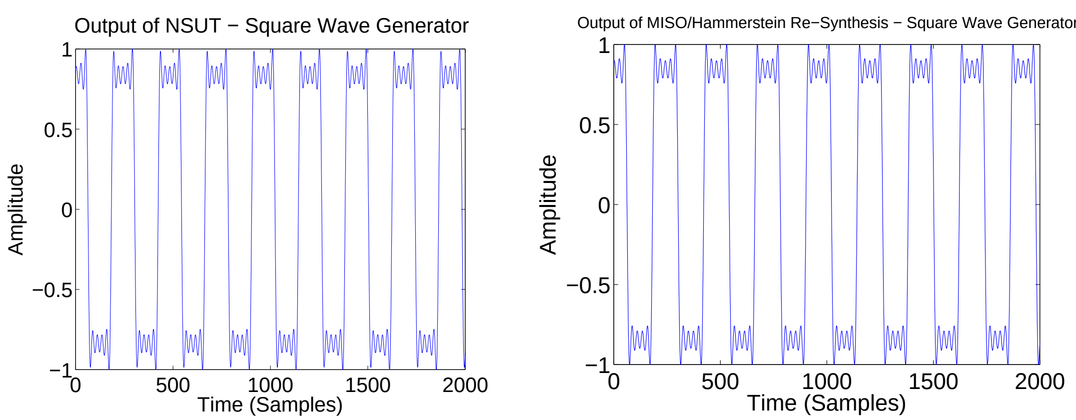
Where $s(t)$ is the chirp defined in Eq. 2.1.

Convolution of the response $r(t)$ with $s(t)$ should produce an MIR output consisting of pure impulses. Each HIR's strength is proportional to the inverse of its harmonic number. By convolving the original input $s(t)$ with each of the HIR as described in the re-synthesis method of Section 3.4, we expect simply scaled versions of the input sinusoid. Upon summation the result should give a band-limited square wave, exactly as Eq.(3.20) describes.

[1] Therefore, up to the 11^{th} harmonic of $s(t)$ was generated.

The parameters of the first test were as follows: $T = 10$ s, $f_1 = 10$ Hz, $f_2 = 9000$ Hz, $f_s = 96$ kHz with the actual length of the chirp adjusted to synchronize the harmonics as explained in Section 3.1.1.

The MIR were extracted by the nonlinear convolution method and appropriately windowed to get separate HIR which were then used as the filters $h_i(t)$ as in Figure 3–5. For a test input $x(t)$, a full-amplitude cosine wave of frequency 400 Hz was used. This was chosen again because of the nature of the nonlinear system and its use of harmonics of cosine functions. $x(t)$ was fed both to the re-synthesized model and the NSUT as given in Eq. (3.20), with $s(t)$ replaced by $x(t)$. Figure 3–8 shows a comparison of a portion of the output of the actual and simulated systems.



(a) (b)

Figure 3–8: Comparison of output waveforms for the band-limited square wave generator. (a) the NSUT , and (b) the re-synthesized version.

A look at the spectra of both outputs in Figure 3–9 show close qualitative agreement. The amplitude of each harmonic of the re-synthesis is very close to the respective harmonics of the NSUT (< 0.1 dB). A comparison of the spectral phase is also given in Figure 3–10. To give a general quantification of the closeness of

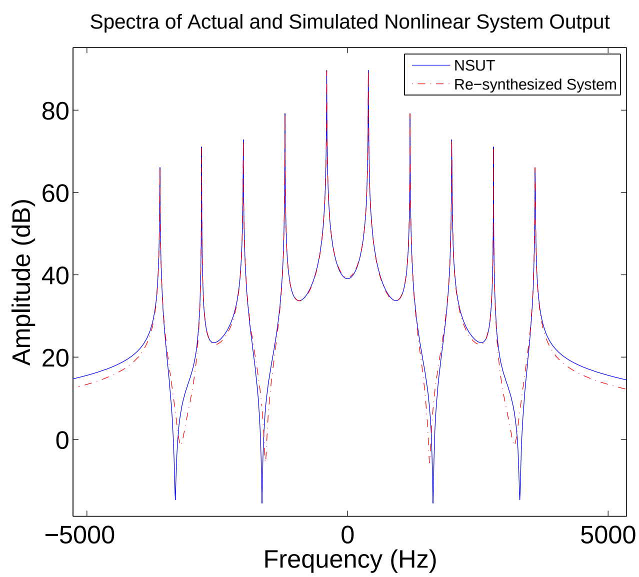


Figure 3–9: Comparison of the magnitude spectra of both waveforms.

the output of the MISO/Hammerstein model to the NSUT, we use the spectral flatness (SF) as described in Section 1.4, which gives a good measure of the precision of the re-synthesis model:

$$SF = 10 \log \left(\frac{FFT(|\Theta|)}{FFT(|\hat{\Theta}|)} \right) \tag{3.21}$$

Where Θ and $\hat{\Theta}$ are the NSUT and re-synthesized system outputs, respectively. The spectral flatness of the square wave re-synthesis is shown in Figure 3–11, evaluated at harmonics $n \cdot 400$ Hz of the input signal. Overall the re-synthesis performed extremely well. One can see a few locations (harmonics $n = 4, 6, 8, 10$) where the re-synthesis significantly underestimates the magnitude of the harmonic.

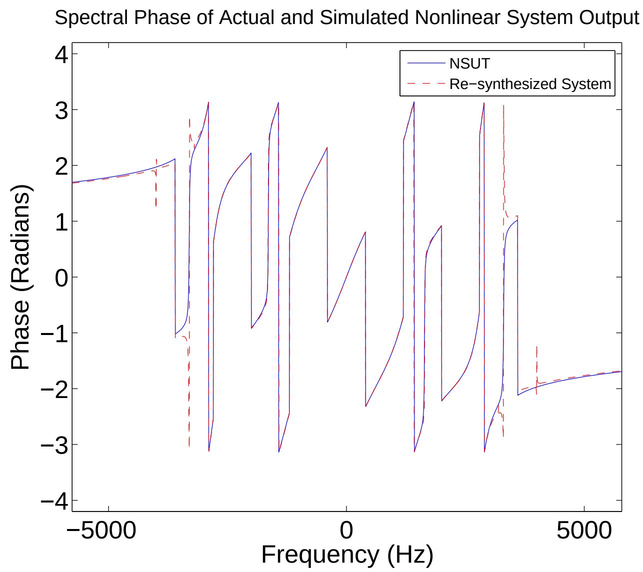


Figure 3–10: Comparison of the spectral phase of both waveforms.

However, the square wave only consists of *odd* harmonics and thus the inaccuracies lie in the 'valleys' of the spectrum, which are far less perceptually important than the accuracies of the peaks which are in fact accurate to 0.1 dB in this case.

This example illustrates an interesting consideration. Here a more complex waveform was constructed out of simple harmonics of a sinusoid of some frequency. While Fourier analysis is over 200 years old, it is the means by which these harmonics were generated that is interesting. By using the Chebyshev polynomials to generate harmonics of the input signal, one can then interpret any harmonic signal with overtones of the fundamental as being generated by a nonlinear system with an input consisting of just a single frequency. The generation of complex signals by the combination of a few simple signals was very advantageous for musical synthesis in the 1970's when computing costs were of the utmost

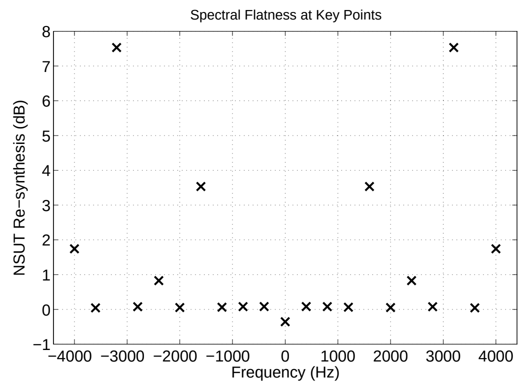


Figure 3–11: Spectral flatness at relevant points on the spectrum for the square wave generator.

importance. By using only a few sinusoidal oscillators and performing simple operations on them, complex spectra could be produced with little computational cost. Some examples include AM and FM synthesis, and simple waveshaping via lookup tables [32]. In the context of this work, the Chebyshev polynomials $T_n(x)$ generate harmonics of the input, and the strength of a purely impulsive HIR convolved with one of the $T_n(x)$'s act as a gain control for the n^{th} harmonic of the input. Of course this method in general offers greater flexibility by being able to specify an impulse response rather than a simple gain control to each harmonic, but in the context of constructing a more complicated waveform from a sine wave and its harmonics, purely impulsive HIRs of varying strengths are sufficient.

With this in mind, the analysis/synthesis method described in this work could also have applicability as a type of function generator. For the analysis phase, if a chirp of fundamental frequency f_1 to f_2 can be generated whose waveform is the intended function to model, then the nonlinear convolution procedure would work in the same way as for any nonlinear system: by convolution of this signal with a sinusoidal chirp, extracting the HIR, and using them in the MISO model.

Alternatively, if the Fourier coefficients a_n, b_n can be extracted for a period of the intended function, then the MISO method could be used in a slightly different manner to re-create the function. The re-synthesis method is slightly modified to include two input signals, which are a sine and cosine function. Both signals would have a frequency equal to the fundamental frequency of the function to generate. The $T_n(x)$ nonlinearities are then applied to the both inputs separately, thus creating $2n$ branches for the MISO model. The HIR in the MISO model can then be created from the a_n, b_n by creating pure impulses scaled by the appropriate Fourier coefficient to encode the strength of that particular harmonic. A block diagram of this setup is shown in Figure 3–12. This method of Fourier synthesis is advantageous because only two sinusoidal signals are required to recreate a complex waveform, rather than the $2n$ oscillators required to traditionally model a function up to harmonic number n. Furthermore, the MISO method also offers the flexibility to filter the harmonics of the input independently.

3.5.3 power-law Nonlinearity with Memory

The next nonlinear system under test consisted of two nonlinear branches with fourth and fifth-order nonlinearities, each followed by a different linear filter. A block diagram representation is shown in Figure 3–13.

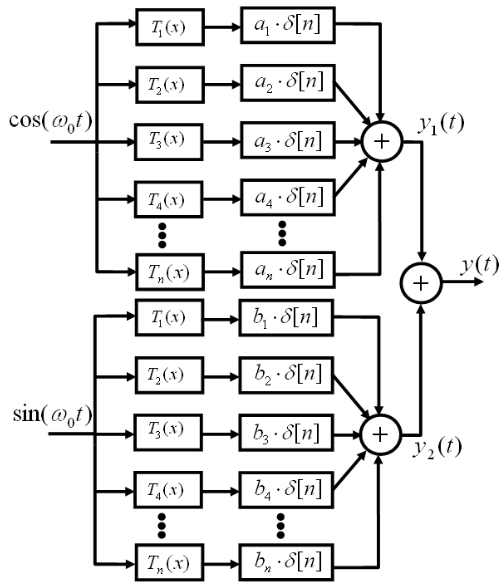


Figure 3–12: Block diagram of the MISO method used as a function generator. The a_n's and b_n's are the Fourier Coefficients and $\delta[n]$ is the Kronecker function.

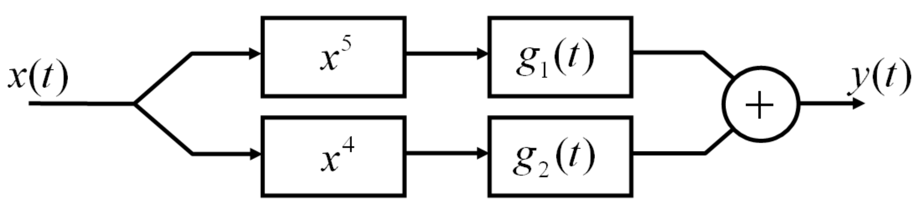


Figure 3–13: Block diagram of the NSUT: branched power-law nonlinearity with independent filters for memory effects.

The filters $g_i(t)$ were impulse responses taken from a free online IR repository[2] and chosen for their short length and character.

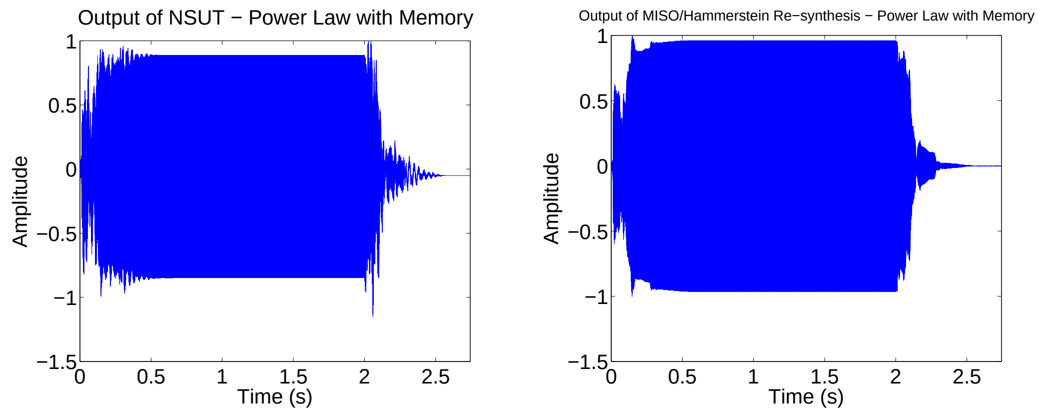
4^{th} and 5^{th} order nonlinearities generate different harmonics of the input as implied by Eq. (3.8). The 5^{th} order term generates 1^{st}, 3^{rd}, and 5^{th} order harmonics, while the quartic generates the 2^{nd} and 4^{th} and also introduces a DC component.[3] Thus the odd order harmonics will have a different reverberant characteristic than the even ones since they are filtered by different impulse responses $g_1(t)$ and $g_2(t)$ in separate branches of the model.

The parameters of this test were as follows: $T = 20$ s, $f_1 = 10$ Hz, $f_2 = 9000$ Hz,$f_s = 96$ kHz. A particularly long excitation chirp had to be used because the NSUT is highly reverberant. Therefore to avoid overlap of the transients of one HIR with another, the length of the chirp must be sufficiently long as discussed on page 47.

To test this system, a full amplitude sine wave of frequency $f_0 = 1000$ Hz and duration 2 s was used as a test signal in both the NSUT and re-synthesized system. Figures 3–14 and 3–15 show a comparison of the time-domain outputs of each system.

[2] www.irlibrary.org

[3] The DC component is easily filtered out of the nonlinear response to the excitation by subtracting the mean value of the response. Since the re-synthesis method does not detect zero order components, and has no audible effect on the output in this context, it is filtered out in the system's nonlinear response to both the chirp and input signal.



(a) (b)

Figure 3–14: Comparison of output envelopes for the power-law nonlinear system with memory. (a) the NSUT, and (b) the re-synthesized version.

Figure 3–15: Comparison of output waveforms for the power-law nonlinear system with memory.

It appears from Figure 3–14 that the method has some difficulty representing the transients of the system. This is likely a consequence of only considering a finite number of harmonics of the input. The steady state characteristic however is well modeled as indicated in Figure 3–15. This can also be seen by looking at the spectra and phase of a portion of the steady-state part of the output in Figures 3–16 and 3–17.

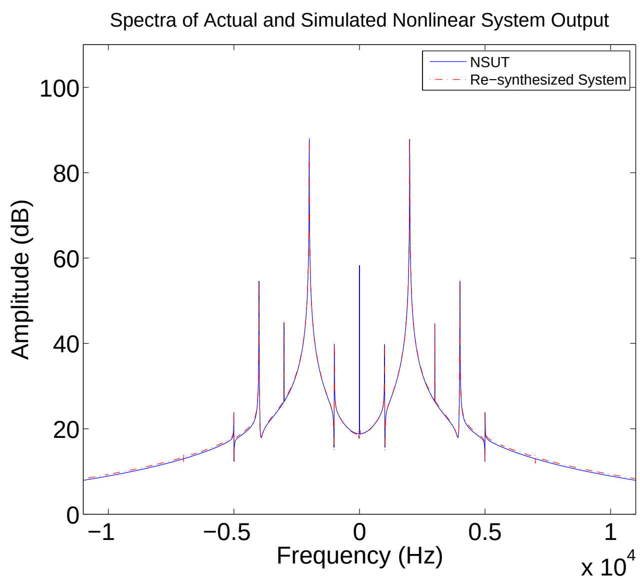


Figure 3–16: Comparison of the magnitude spectra of both waveforms.

The spectral flatness of the steady state portion of the output is shown in Figure 3–18. The spectral peaks of the re-synthesis agree with expected values to within 0.3 dB, for the first 4 harmonics, and within 0.5 dB for the fifth[4] . The

[4] There is a prominent DC component in the NSUT from the quartic nonlinearity that is not modeled. Since this only affects the mean value of the signal and is inaudible, it is considered insignificant.

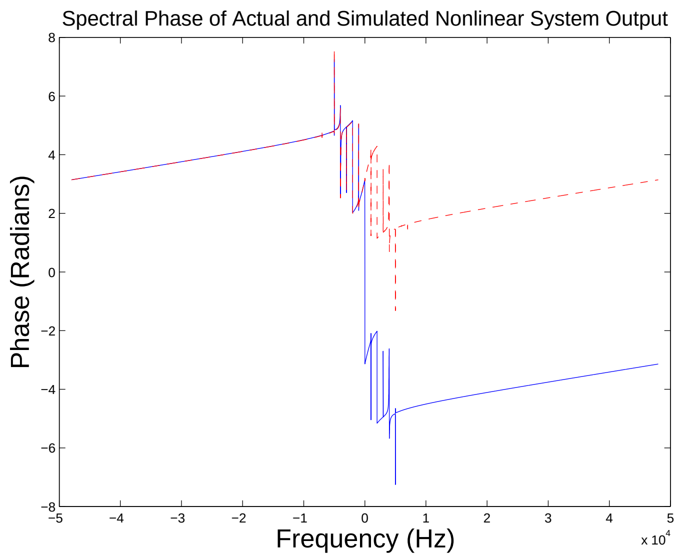


Figure 3–17: Comparison of the unwrapped spectral phase of both waveforms.

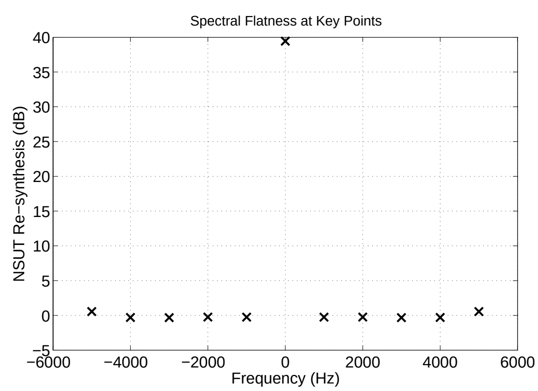


Figure 3–18: Spectral flatness at relevant points on the spectrum for the power-law nonlinearity with memory.

phase plot also matches closely, however it appears as though at around zero frequency the re-synthesis does not account for a decrease in phase delay, and thus all frequencies > 0 have a greater spectral phase shift than the NSUT. This is likely because the DC component was not modeled in the re-synthesis, and thus the spectral phase is different at 0.

3.5.4 Problems with Chebyshev Polynomials

While the idea of Chebyshev polynomials as the nonlinear elements in the synthesis model is conceptually simpler than performing a linear transformation on the $h_i(t)$ (Section 3.4.1), it presents a particular problem. The $T_n(x)$'s only produce a pure harmonic of x if it is a *full-amplitude* sinusoidal signal. Otherwise, applying T_n to a signal of amplitude < 1 will produce a harmonic at order n but additional harmonics at order $n - 2k$, k an integer in $[0\ n/2]$. This poses a fundamental problem if the input signal is to be arbitrary. Even if the signal can be decomposed into a sum of sinusoids, none of them could possibly be of full-amplitude. To test the severity of this issue, a simple clipping nonlinearity was created and is shown in Figure 3–19. This system does not affect samples less than α, and clips any samples greater than α. As the input/output relationship is an odd function, this system will produce only odd harmonics of the input. It was analyzed and then re-synthesized with Chebyshev polynomials as the nonlinear elements in the MISO model, followed by the $h_i(t)$ as linear filters following each nonlinearity. The $h_i(t)$ are the HIR produced by the nonlinear convolution method. The same re-synthesis procedure was also performed using simple power-law nonlinearities with a corresponding linear transformation to

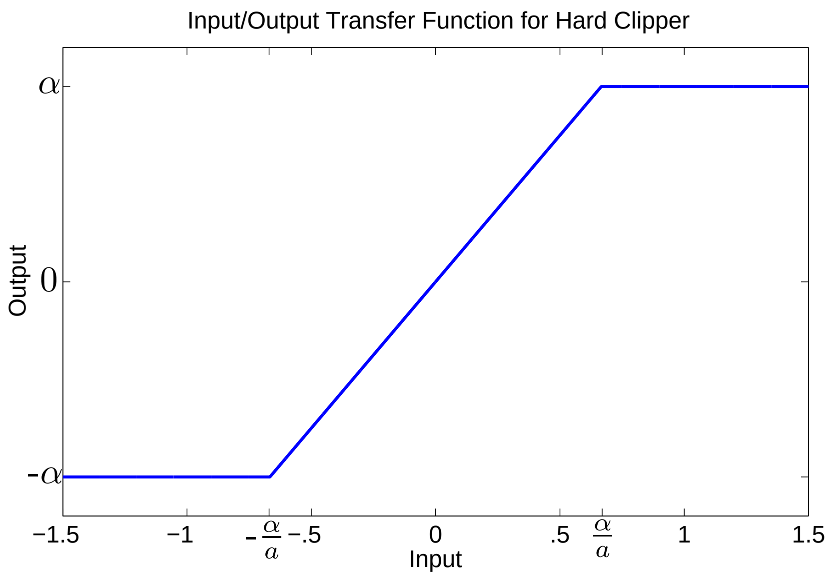


Figure 3–19: Input/Output relationship for the hard clipping nonlinearity. Any input values $> \alpha/a$ are clipped to a value of α.

the HIR for the MISO method as described in Section 3.4.1. Sinusoidal inputs of differing amplitude and character were used to investigate the severity of the amplitude problem. Temporal and spectral comparisons of the outputs of both synthesis methods along with the actual output of the hard clipper are shown below for different sinusoidal inputs.

For these tests, 10 harmonics were modeled, with a sweep from $1 - 9000$ Hz @ 192 kHz. Because only a finite number of harmonics can be used in the resynthesis, the bandwidth of the output will be limited. A compromise has to always be made in this analysis/synthesis method between the number of harmonics to model and the bandwidth of the chirp used. Too small of a chirp bandwidth will not capture the high frequencies in the linear and low-order responses; too few harmonics will not capture the sound of the nonlinear effect correctly.

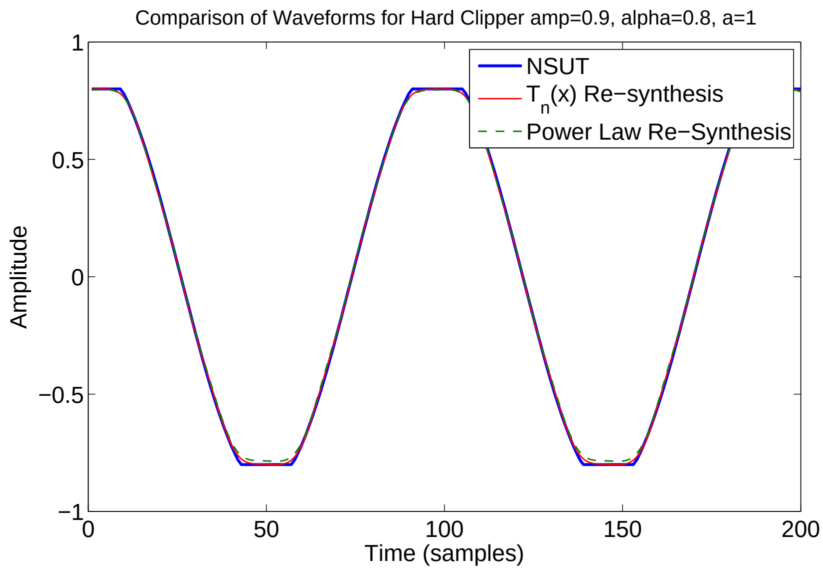


Figure 3–20: Waveform comparison of the outputs of the NSUT, the Chebyshev re-synthesis, and the power-law re-synthesis. The input sine amplitude is 0.9, clipping value is 0.8, and 10 harmonics were used in the re-synthesis.

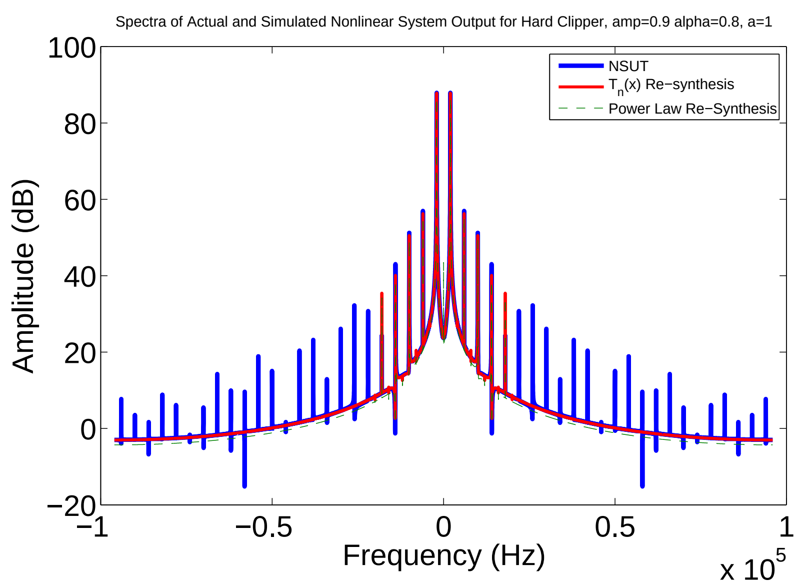


Figure 3–21: Spectral comparison of the outputs of the NSUT, the Chebyshev re-synthesis, and the power-law re-synthesis. The input sine amplitude is 0.9, clipping value is 0.8.

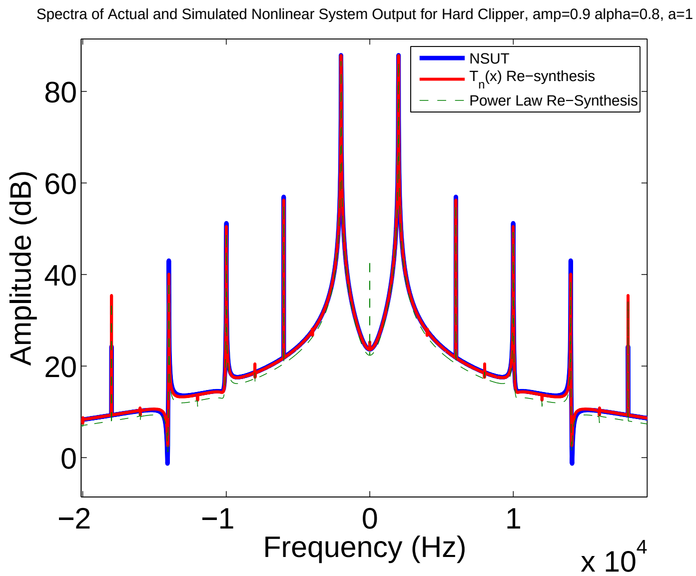


Figure 3–22: Spectral comparison of the outputs of the NSUT, the Chebyshev re-synthesis, and the power-law re-synthesis, showing only the locations where harmonics were modeled. The input sine amplitude is 0.9, clipping value is 0.8.

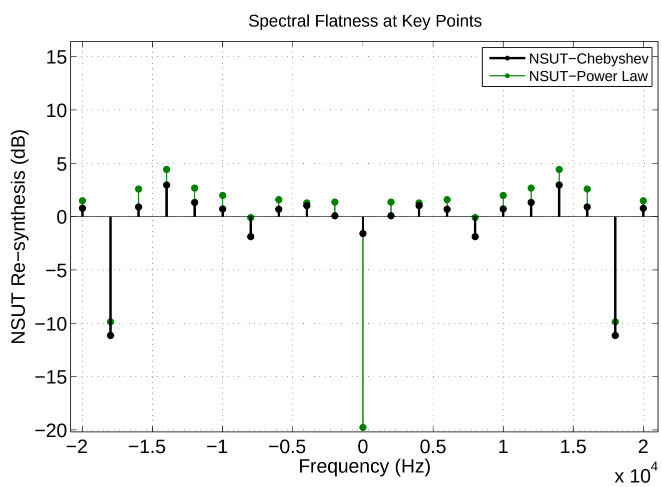


Figure 3–23: Spectral Flatness for the hard clipper spectrum in Figure 3–21, showing the spectral flatness of each synthesis method compared with the NSUT output. The re-synthesis only models harmonics up to order 10 and only these values of the spectral flatness are shown.

The first two tests have the hard clipper threshold at $\alpha = 0.8$. The first input used is a simple sinusoid of amplitude 0.9 producing a slightly clipped output waveform. Temporal and spectral comparisons are shown in Figures 3–20, 3–21, 3–22 and 3–23. It is apparent that both re-synthesis schemes perform well and that the re-synthesis outputs are nearly identical to each other.

Analysis of the spectral flatness in Figure 3–23 shows that the re-synthesis for both methods performed quite well, with all harmonics modeled (except the 9^{th}) being within 5 dB for both synthesis methods.

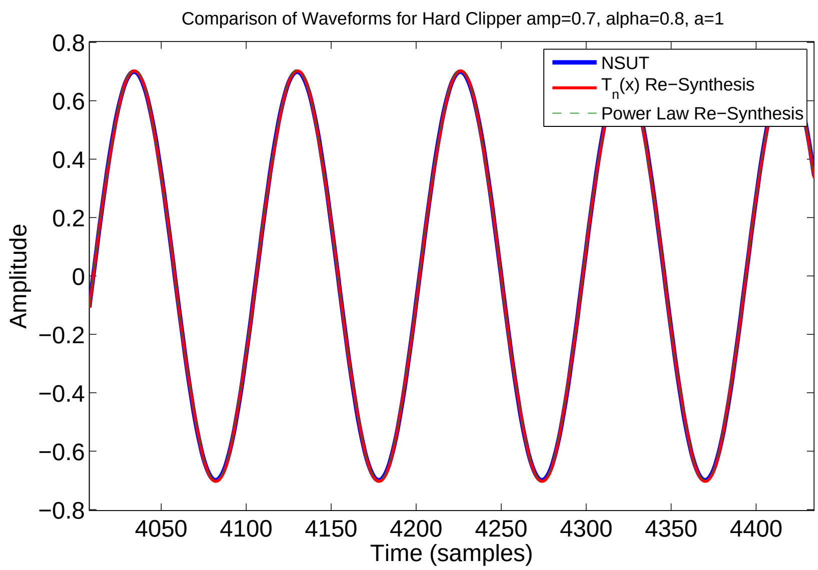


Figure 3–24: Waveform comparison of the outputs of the NSUT, the Chebyshev re-synthesis, and the power-law re-synthesis. The input sine amplitude is 0.7, clipping value is 0.8, and thus the output should pass unaffected.

The same clipper is used in the next comparison. The input is again a sinusoid but of amplitude 0.7 - less than the clipping threshold $\alpha = 0.8$. Thus we expect the output waveform to remain unaffected by the nonlinear system

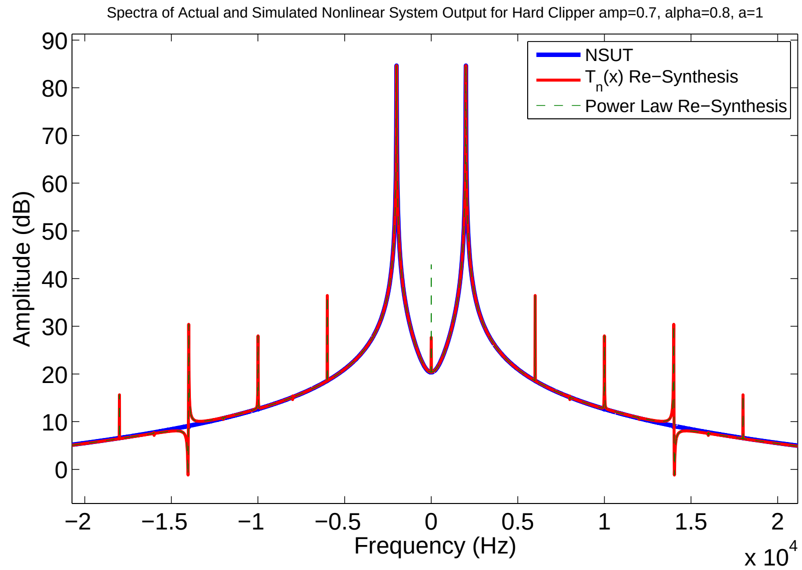


Figure 3–25: Spectral comparison of the outputs of the NSUT, the Chebyshev re-synthesis, and the power-law re-synthesis. The input sine amplitude is 0.7, clipping value is 0.8.

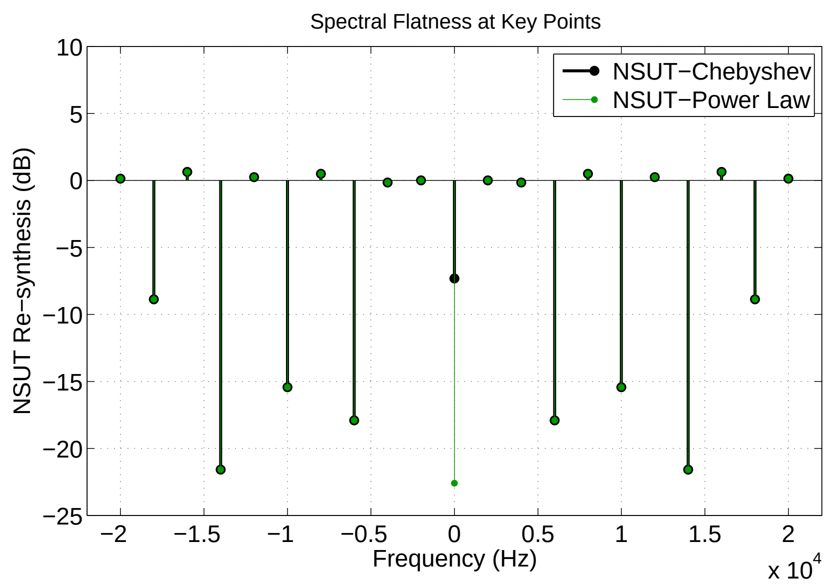


Figure 3–26: Spectral Flatness for the hard clipper spectrum in Figure 3–25, showing the spectral flatness of each synthesis method compared with the NSUT output. The re-synthesis only models harmonics up to order 10.

since the system only clips samples at 0.8 and greater. Comparisons are shown in Figures 3–24, 3–25, and 3–26.

As expected neither the NSUT output nor the either model's output clip the waveform. Both re-synthesis methods perform identically. When looking at the spectra in Figure 3–25, it is interesting to note that higher harmonics still appear to be present in the output. However they are approximately 55 dB lower than the power of the fundamental. Because the model is modeling a nonlinear system, there exist HIRs of order > 1 which have some meaningful content in them in order to model harmonics when the input is actually clipped as in the first clipping example. As such, it is expected that there would be some frequency contribution from the higher harmonics in the output. Fortunately, these higher harmonics are very weak compared to the fundamental in this case and the time domain outputs in Figure 3–24 appear indistinguishable for this case.

The 2^{nd} clipping system identified had a different threshold for clipping at $\alpha = 0.55$. The input signal used was also more complicated than a simple sine wave. The input signal used was an AM signal consisting of two sinusoids at frequencies 125 Hz and 2000 Hz. This is equivalent to two inputs at frequencies 2125 Hz and 1875 Hz. This input was chosen to produce a more dense spectral response with the clipper and to illustrate how the system produces intermodulation distortion. In addition, 20 harmonics were modeled in this case, however the excitation chirp bandwidth was cut in half to 4500 Hz. The nonlinear system tested is merely the same clipper as in the previous two cases, with a modified value of α, and the AM signal is the input signal to the nonlinear system which is used here to compare the NSUT with the two methods of re-synthesis. For clarity, the method is characterizing the clipping system only, and using the AM signal as the input to this nonlinear system.

Temporal and spectral comparisons of the re-synthesis methods to the NSUT output are given in Figures 3–27, 3–28 and 3–29. The spectral flatness is plotted for multiples of 125 Hz, since the intermodulation distortion will produce harmonics at *some* multiples of 125 Hz. Strong components are produced at the two input frequencies 125 Hz and 2000 Hz as well as intermodulation products at sums and differences of the two. Again both synthesis methods perform essentially identically and appear to produce the intermodulation products fairly well. It is apparent that the method has some difficulty accurately representing higher harmonics however, and the spectral flatness in Figures 3–29 shows increasing disagreement as the frequency increases. There is general agreement at frequencies < 10 kHz usually within 5dB, with the accuracy degrading as frequency increases. However these higher order components are roughly 60-70 dB lower than the highest peaks of the output and are not as important perceptually compared to lower harmonics.

These tests show that the amplitude not a significant source of error when using simple inputs. Furthermore, these tests also confirm that the two methods described in Sections 3.4.1 and 3.4.2 are indeed equivalent. This is to be expected, as the matrix of coefficients C of Chebyshev polynomials is related to the matrix A used for the linear transformation of the HIR by Eq. (3.16). The Chebyshev method uses the $T_n(x)$ as nonlinearities to isolate specific harmonics of the input,

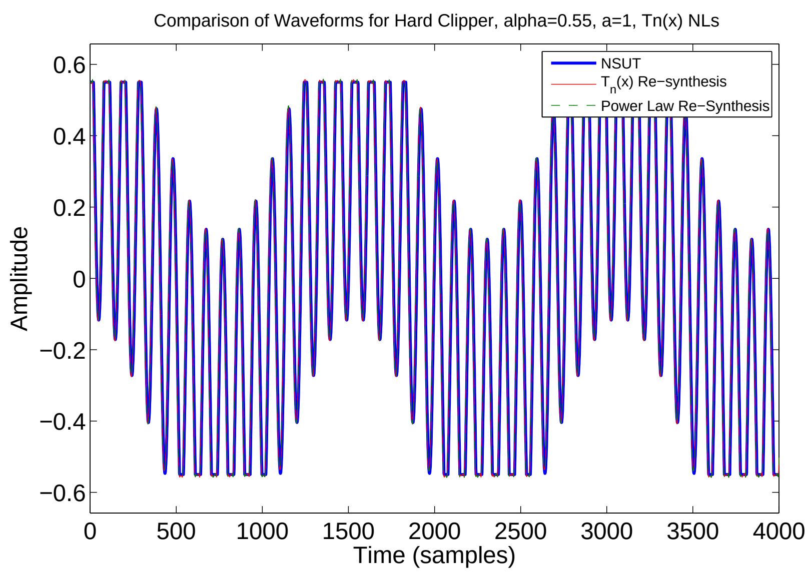


Figure 3–27: Waveform comparison of the outputs of the NSUT, the Chebyshev re-synthesis, and the power-law re-synthesis. The input is an AM signal consisting of 2 sine waves of frequency $f_1 = 125$ Hz, $f_2 = 2000$ Hz. The clipping value is 0.55. 20 harmonics were considered in the model.

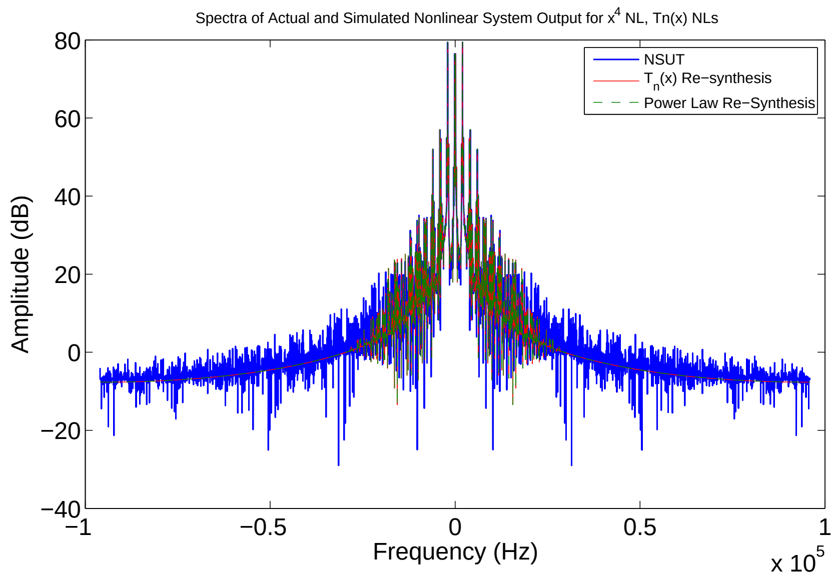


Figure 3–28: Spectral comparison of the outputs of the NSUT, the Chebyshev re-synthesis, and the power-law re-synthesis. The input is an AM signal consisting of 2 sine waves of frequency $f_1 = 125$ Hz, $f_2 = 2000$ Hz. The clipping value is 0.55.

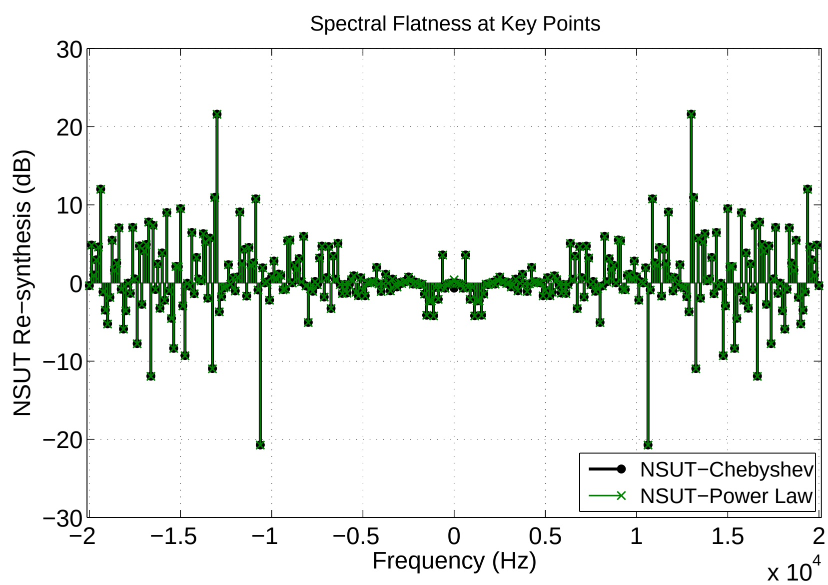


Figure 3–29: Spectral Flatness for the hard clipper spectrum in Figure 3–28, showing the spectral flatness of each synthesis method compared with the NSUT output. The re-synthesis models harmonics up to order 20.

and the MIR control the gain of each harmonic. The power-law method uses simple x^n nonlinearities with each nonlinearity generating multiple harmonics of differing strengths. The linear transformation of the HIR then combines the HIR in such a way that it takes into account the strength of the different harmonics generated by applying each x^n.

Concerns of the $T_n(x)$ nonlinear elements producing harmonics other than the n^{th} (i.e. harmonics $n - 2k$) is not an issue. While the reason why may not be obvious from the Chebyshev method, we can take advantage of the fact that the method is equivalent to the method using power-law nonlinearities and a linear transformation. When a power-law nonlinearity is applied to a signal $x(t)$, we get simply $x^n(t)$. If $x(t)$ is instead accompanied by an envelope $a(t)$, where $0 < a < 1$ and $|x| < 1$, then applying an n^{th} order nonlinearity to $a(t)x(t)$ yields $a^n(t)x^n(t)$. Thus the output scales as the input to the power of the power-law applied. For example, if the input was a scaled cosine $a(t)\cos[\theta(t)]$, then the harmonics produced by applying a power-law nonlinearity (see Eq. (3.8)) would all simply be scaled by $a^n(t)$. The re-synthesis would then continue as before and yield simply a scaled output.

While certainly the application of the $T_n(x)$ polynomials does not give the desired harmonic content after application to a signal of amplitude < 1, the equivalency of both methods described in Sections 3.4.1 and 3.4.2 means that upon summation of all the $T_n(x)$'s after applying the HIR's, this effect disappears.

Since it is equivalent to use either method in terms of the end result, re-synthesis using the $T_n(x)$ as the nonlinear elements offers an advantage in that no linear transformation of the HIR is necessary; they are instead used directly as the filters following the nonlinearities. This method is also more intuitive and conceptually simpler than the power-law method.

To conclude, this chapter has discussed two modifications to the signals used in the nonlinear convolution method: the synchronization of the phase of the excitation with its instantaneous frequency, and the bandwidth extension of the inverse system (Section 3.1). Brief descriptions of some nonlinear synthesis methods have been given in Sections 3.2 and 3.3, followed by a discussion of how the MIR extracted by the nonlinear convolution can be used directly in a re-synthesis method for nonlinear systems, or used after a suitable linear transformation of the HIR is performed (Section 3.4). Two artificially created nonlinear systems were then analyzed and re-synthesized using Chebyshev polynomials as nonlinear elements. The input used was a simple full-amplitude sinusoid, and the performance of the method is discussed in Section 3.5. Results in general showed good agreement with the NSUT. Finally, a comparison of the two re-synthesis methods was performed using a hard clipping nonlinearity in Section 3.5.4. The input signals used were simple sinusoids and combinations thereof with amplitudes less than 1 to test the effects of the input amplitude on the re-synthesis. Results showed that both methods were equivalent and provided good agreement with the intended output waveform. Having simple inputs of amplitude less than 1 did not adversely affect the re-synthesis.

The use of complicated or musically interesting input signals to these systems was avoided in this chapter in order to clearly show how the re-synthesis model emulates the NSUT. The next chapter puts the synthesis method to the test by emulating some real nonlinear systems. A guitar distortion pedal, practice amplifier with on-board distortion, and home-made distortion pedal are analyzed and tested with both a pure signal as well as a musically interesting input signal. Because of the wide bandwidth of real musical inputs signals, further modifications to the re-synthesis method are required and will be discussed.

Chapter 4
Testing Nonlinear Systems

Chapter 2 introduced the concept of nonlinear convolution and explained how the use of exponential chirps could be used as input to a nonlinear system to extract information about the system. By de-convolving the system's response with the input, multiple impulse responses (MIR) could be extracted which describe the strength and transient nature of each harmonic response to the input by the nonlinear system. Chapter 3 described a re-synthesis method for emulating the sound of the particular nonlinear system under study. The MIR were used in a Hammerstein/MISO framework as linear filters, preceded by Chebyshev Polynomials which were used to create harmonics of the input signal. Some simple 'synthetic' nonlinear systems were then tested to ensure the model was functioning properly. Finally this chapter tests the analysis/synthesis method in a musical context with real nonlinear systems. The nonlinear systems to be tested are in the realm of guitar distortion pedals. While the method is applicable to any other harmonic-generating nonlinear system, guitar effect and amplifier emulation is the primary goal of this research.

Three nonlinear effects were tested. The first was a Roland Micro Cube Practice Amplifier. The second was a DigiTech RP-6 Multi-Effects Pedal, and the third is a home-made effects pedal created by a fellow student in the Music Technology Department at McGill University.[1]

4.1 Further Modifications to the MISO Method

The inputs used to the nonlinear systems studied in Chapter 2 produced fairly accurate re-creations through the MISO method. Upon initial study with real nonlinear systems however, it was observed that in particular low amplitude input signals were not being re-synthesized as closely as with the synthetic nonlinear systems. In general, the re-synthesis was lacking in high-harmonic content. This was due to the issues described in Section 3.5.4. When a low-amplitude input or component of amplitude a (where $a < 1$) is passed through an k^{th} order nonlinearity such as $T_k(x)$, the amplitude of this component is reduced to a^k. For high-order polynomials, this greatly suppresses weaker components in the input. Two solutions have been implemented to attempt to counteract this effect: envelope following and pre-emphasis. Each are explained in the following sections.

4.1.1 Envelope Following

Envelope following is a method of tracking the amplitude variations of a signal to be used for the control of some parameters, or to produce another signal that resembles purely these variations. An example of an audio signal and its corresponding envelope is shown in Figure 4–1.

[1] Pedal designed and built by Avrum Hollinger - http://idmil.org/people/avrum_hollinger.

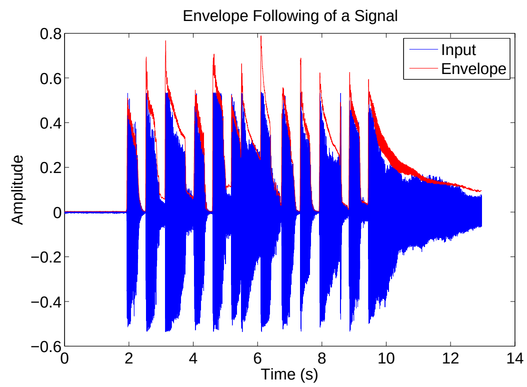


Figure 4–1: Example of envelope following of an audio signal.

There are a few basic methods to perform envelope following. One class of envelope tracking algorithms employ the Hilbert Transform [8]. From a time-domain standpoint, the Hilbert Transform applied to a signal $x(t)$, denoted as $H(x(t))$ causes a 90° phase shift on $x(t)$. One can use the analytic signal $z(t) = x(t) + iH(x(t))$ to generate an approximation of the envelope for $x(t)$. There are two common approaches from here to generate an envelope. The first approach rectifies the real and imaginary parts of the analytic signal individually, and then sums the result:

$$Env_{abs} = |Re(z)| + |Im(z)| \tag{4.1}$$

Because the imaginary part is phase shifted by 90°, the summing together of the real and imaginary parts has a low-pass filtering effect, and the rectification ensures that the envelope is always positive. The second approach is similar, except that the rectification is intrinsically done while calculating the energy of the real and imaginary parts of the analytic signal:

$$Env_{energy} = \sqrt{Re(z)^2 + Im(z)^2} \tag{4.2}$$

The result is a positive-definite signal that approximates the envelope. In either of these methods however the envelope varies far too quickly, almost at the same rate as the audio signal. Thus the result is low-pass filtered to give an envelope that varies slowly in comparison to the frequency of the signal.

The other class of envelope follower uses a sliding window on the signal and calculates an average quantity based on samples in the window. A window of some length is chosen and the signal's RMS or mean absolute value is calculated for the signal samples in that window. The window then slides forward by one sample, and the calculation is repeated. Because the value of the calculated quantity at a given point is based on an averaging of the signal around it, the envelope is inherently smoothed relative to the audio signal. Low pass filtering however is also necessary after this to provide a smoother measure for the envelope. Care must be taken to choose the appropriate window length to achieve a balance between responsiveness of the envelope function to sharp attacks and decays, and smoothness of envelope function.

After much experimentation with the four methods, the use of a sliding window which calculates the energy in the current window was found to be the best in terms of amplitude tracking and smoothness after low pass filtering:

$$Env_{sliding}[n] = \sqrt{\frac{1}{w}\sum_{k=n}^{n+w-1} x[k]^2} \tag{4.3}$$

Where the window length is given by w. Low pass filtering was performed with a biquad filter with a cutoff frequency of 40 Hz. This cutoff was chosen as a compromise between the responsiveness of the filter and the smoothness of the envelope, during signal transients.

The idea of envelope following in the MISO method is to boost low-amplitude signals in order to preserve their harmonic content after the $T_n(x)$ are applied. The envelope is first extracted and scaled accordingly. The mathematical inverse of the envelope is then applied to the signal to (ideally) make the signal of uniform amplitude. Division by 0 is controlled by a thresholding value. This signal is then sent as the input to the MISO method. Following the application of the nonlinearities and filtering of the branches, the envelope is re-applied to the output signal in order to restore the dynamic character of the signal.

Envelope following did improve the harmonic content when testing real nonlinear systems. However it was apparent that further compensation of the high harmonics was necessary.

4.1.2 Pre-Emphasis

A second solution to compensate for the lack of high harmonics is to anticipate the losses that will occur when the input passes through the nonlinearities and 'pre-emphasize' the high harmonics with a high-pass filter.

The filter was carefully tuned to boost the high-frequencies sufficiently without compromising the sound of the lower harmonics. A biquad treble shelving filter [5][40] was used to accomplish this and its frequency and phase response are shown in Figure 4–2.

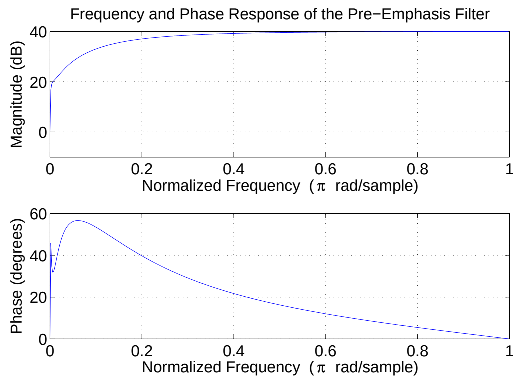


Figure 4–2: Frequency and phase response for the pre-emphasis treble-shelving filter.

At the end of the signal chain, a 'de-emphasis' filter is placed to 'undo' what was done with the pre-emphasis filter. This filter has the same form as the shelving filter but is instead the treble cutting shelving filter.

4.1.3 Revised Synthesis

With the two modifications described above, the MISO re-synthesis method can be summarized with the block diagram in Figure 4–3. The pre-emphasis filter is placed immediately at the input. The envelope following algorithm then generates an envelope for the input. The signal input is then divided by this amplitude envelope to boost low-amplitude signals before the the polynomial nonlinearities. The remainder of the MISO method continues as before, and just before the output the envelope is placed back onto the re-synthesized signal for dynamics, and the de-emphasis filter just precedes the output to restore the balance of high and low frequencies.

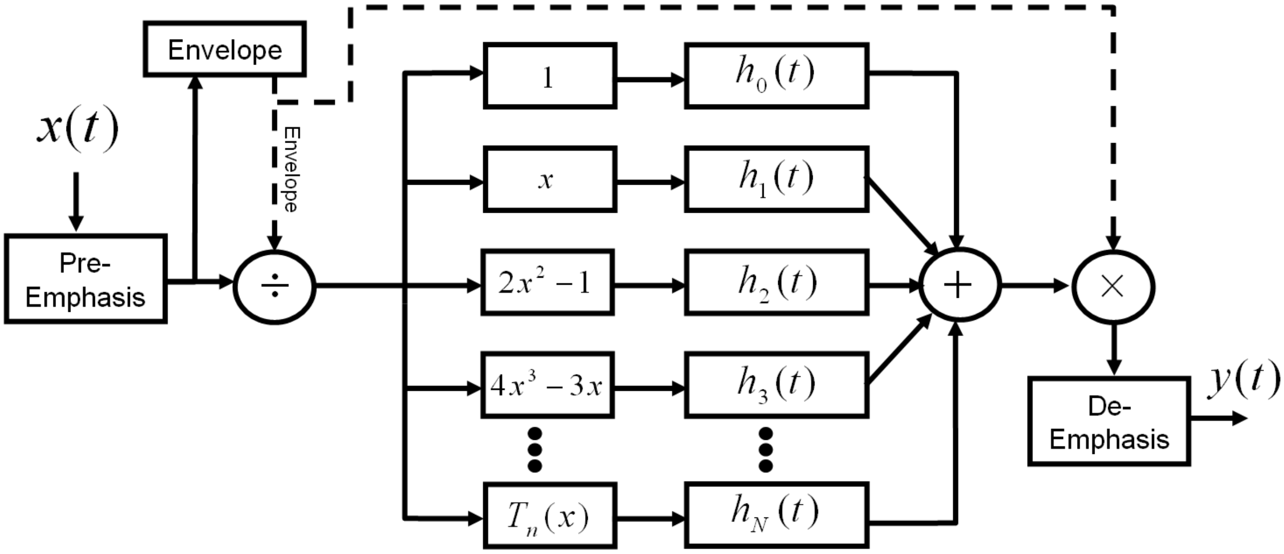


Figure 4–3: Revised MISO method, with envelope tracking and pre-emphasis filter.

With these modifications, the analysis/synthesis method is ready to be tested with real nonlinear devices. Results of these tests are summarized in the next section.

4.2 Tests of Real Nonlinear Systems

Three distortion effects were used to test the analysis/synthesis method laid out in this work in a musical context: a Roland Micro Cube Practice Amplifier, a DigiTech RP-6 Multi-Effects Pedal, and finally a home-made fuzz effect pedal. After the analysis of each system, the system was fed a sinusoidal input, and then a musical guitar pluck to be used to compare the system's output with the model's. Results are presented in the following sections.

4.2.1 Roland Micro Cube practice amp

The Roland Micro Cube practice amp is a small 2 Watt Amplifier and speaker combo intended as a portable/practice amplifier for guitarists.[2] It has a built-in DSP and offers 6 built-in effects, and 7 simulated guitar amp models. The amplifier and its controls are shown in Figure 4–4.

(a) (b)

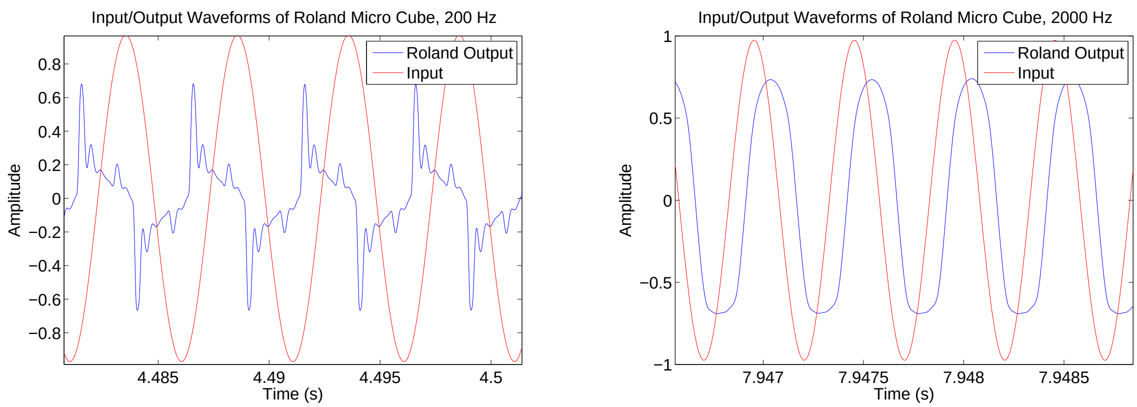
Figure 4–4: Roland Micro Cube practice amp. a. amp unit, b. control panel.

For this test, the amplifier was set to 'Classic Stack' with the following parameters: gain - 2, volume - 7, tone - 8. The Classic Stack setting is a model of a Marshall JMP1987, a popular amplifier in classic and hard rock [41]. These settings produce pleasant subtle overdriven sound without being too 'harsh' sounding or strong in the higher harmonics. The power THD was measured for these settings and found to be 13% at 1000 Hz with harmonics up to 20 kHz.

[2] http://www.roland.com/products/en/Micro-CUBE/ - last accessed 2011 02 04.

Since the re-synthesis method can only model a finite number of harmonics, this was reasoned to be a good starting point for testing the method with a real nonlinear device.

To give a sense of how the Roland modifies the input signal, Figure 4–5 shows the two outputs of the system to two different sine wave inputs of different frequencies.



(a) (b)

Figure 4–5: Input/Output waveforms for the Roland Amplifier at a. 200 Hz, b. 2000 Hz.

An exponential chirp was first sent through the system and the extraction of the MIR was performed. The parameters of the sweep were the same ones used in the hard clipping nonlinearity in Section 3.5.4: $f_1 = 10\ Hz$, $f_2 = 9000\ Hz$, $T = 10\ s$, $f_s = 192\ kHz$. Figure 4–6 shows the MIR output. It is clear that the Roland produces many strong odd-order harmonics, and weak even-order harmonics. The MISO re-synthesis was then performed with two different input signals: a simple sine wave of amplitude 0.5 at 1000 Hz, and a guitar pluck, note

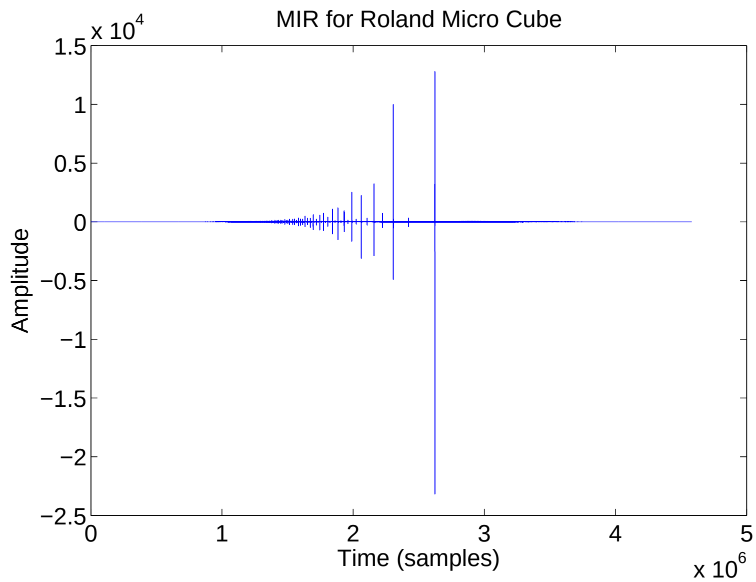


Figure 4–6: MIR output of the nonlinear convolution method for the Roland Micro Cube.

C2. These two inputs were used as the signals for comparison of the nonlinear system to its model for each nonlinear system under investigation in this work. In the re-synthesis 15 harmonics were modeled. While harmonics > 10 alias during the sine sweep, this only occurs when the chirp is near its end. In addition, these harmonics are very weak compared to the lower order ones and therefore their inclusion is not considered problematic.

Temporal and spectral comparisons of the MISO output with the actual system output are shown in Figures 4–7, 4–8, and 4–9 for the sine wave input, and in Figures 4–10, 4–11, and 4–12 for the guitar pluck input.

Sinusoidal Input

Since the input was a pure sinusoid, it was convenient to make a THD measurement for the re-synthesis. The power THD at 1000 Hz for the MISO

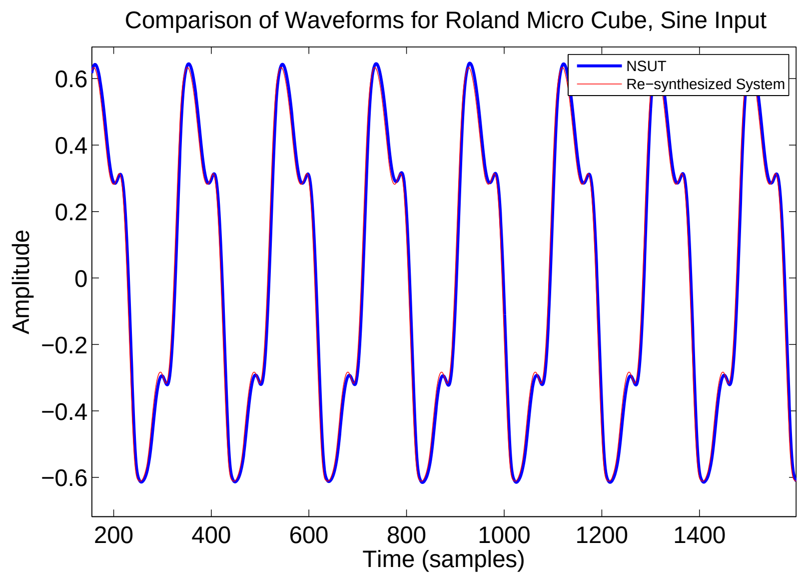


Figure 4–7: Waveform comparison for the Roland Micro Cube with a sinusoidal input of frequency 1000 Hz.

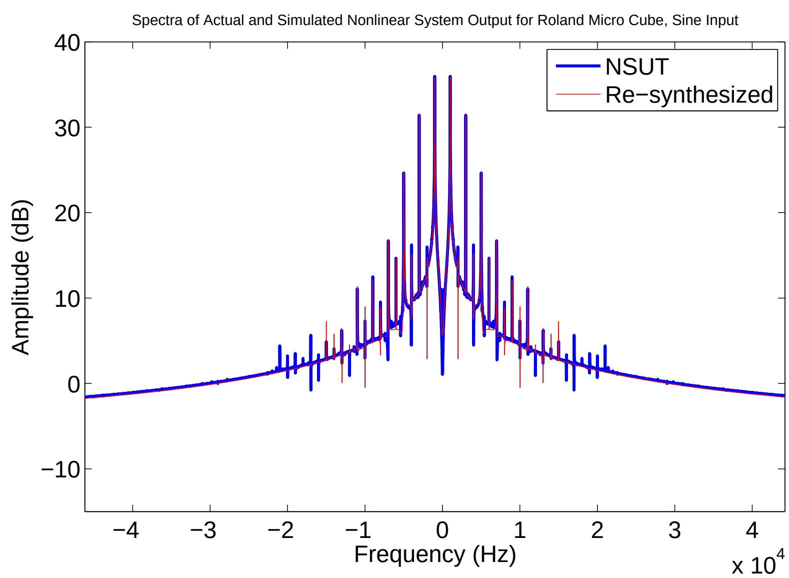


Figure 4–8: Spectral comparison for the Roland Micro Cube with a sinusoidal input of frequency 1000 Hz.

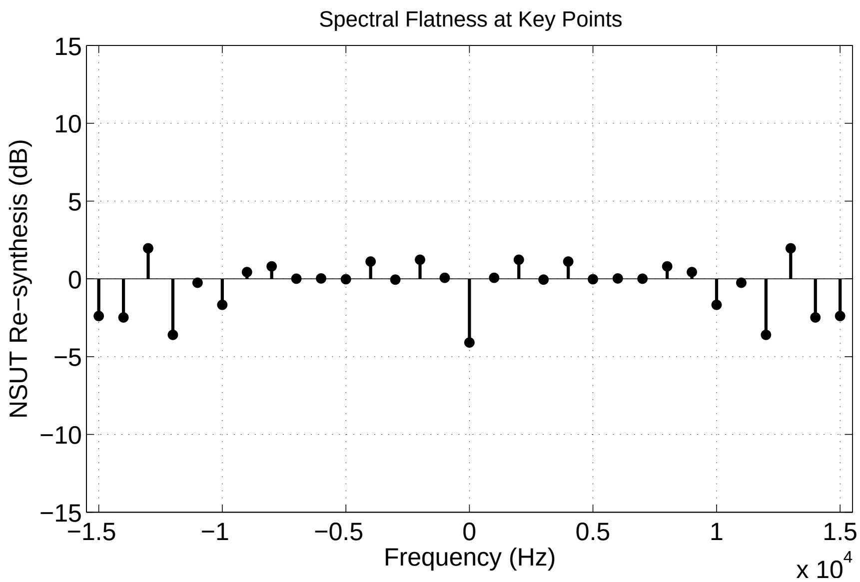


Figure 4–9: Spectral Flatness for Roland Micro Cube with a sinusoidal input of frequency 1000 Hz.

re-synthesis was 13.6%, while the THD of the actual system was measured to be 13%.

From Figure 4–9, it is clear that the re-synthesis has performed well. Lower order odd harmonics were modeled very accurately ($<$ 0.1dB), with the error increasing to $<$ 2dB past 10 kHz. The weak even harmonics are not represented as well, but the general trend of high accuracy at low frequencies with decreasing accuracy as frequency increases is apparent. Perceptually the two sound nearly identical, with the actual system output having slightly more high frequencies, which agrees with the analysis. It is worth noting that the harmonics $n = 11$ to 15 are reasonably modeled and within 3dB of the system's output. This is comforting despite the fact that these harmonics alias at the higher frequencies of the exponential chirp.

Musical Input

The system and the MISO were then fed with a simple guitar pluck at note C2 ($f_0 = 130.81\ Hz$). Comparisons with the actual system output are given in Figures 4–10, 4–11, and 4–12. Spectra were calculated on a section of the signal 1 second into the signal. From the temporal plot in Figure 4–10, it is clear that the method has some difficulty representing this waveform. Figures 4–11 and 4–12 show that the method does not perform as well with a more complex input signal. Most of the harmonics considered in the model are within 5 dB of the NSUT output, somewhat greater than with a sinusoidal input. Perceptually, the emulation output does have a significant overdriven character, but seems to lack the gain of the actual effect, which is in agreement with the spectral plots.

It appears as though despite the method's accuracy in simulating the system's behaviour in response to simple inputs, the method has difficulty when presented with a musical input signal. This could be due to the wide bandwidth and/or rich harmonic content of a musically interesting signal. This is disccused further in Section 4.3.

4.2.2 DigiTech RP-6

The DigiTech RP-6 (Figure 4–13) is a multi-effects guitar pedal intended for guitar. It includes two different distortion types, a noise gate, 3 band-equalizer, compressor, speaker cabinet modeling, 8 modulation effects, dual delay, a reverb unit, and expression pedal for continuous control of some parameters. Despite being a digital pedal, employing an ADC and DAC, the distortion module is analog and placed before the ADC in the pedal [42].

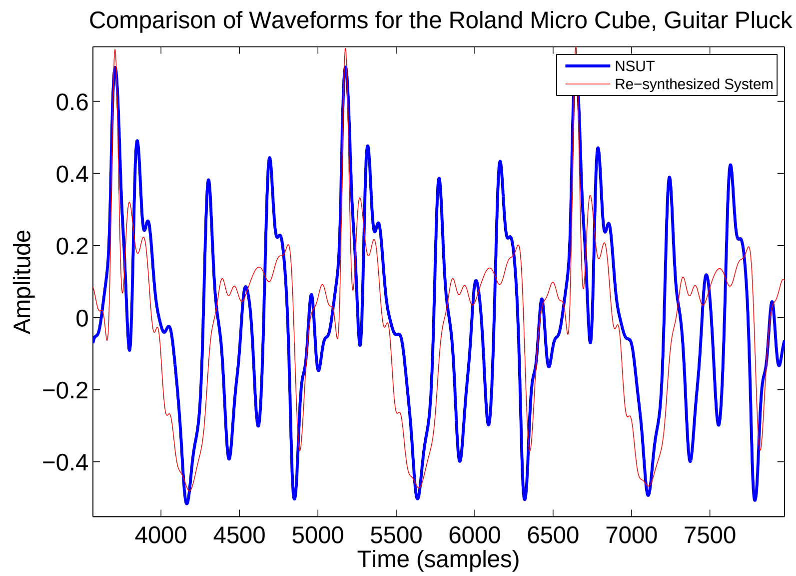


Figure 4–10: Waveform comparison for the Roland Micro Cube with a musical signal as input.

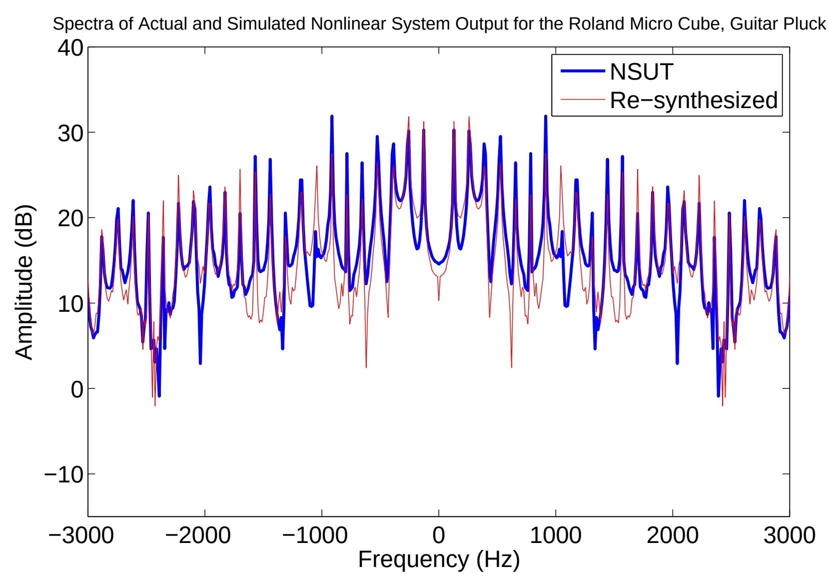


Figure 4–11: Spectral comparison for the Roland Micro Cube with a musical signal as input. The relevant part of the spectrum is shown only.

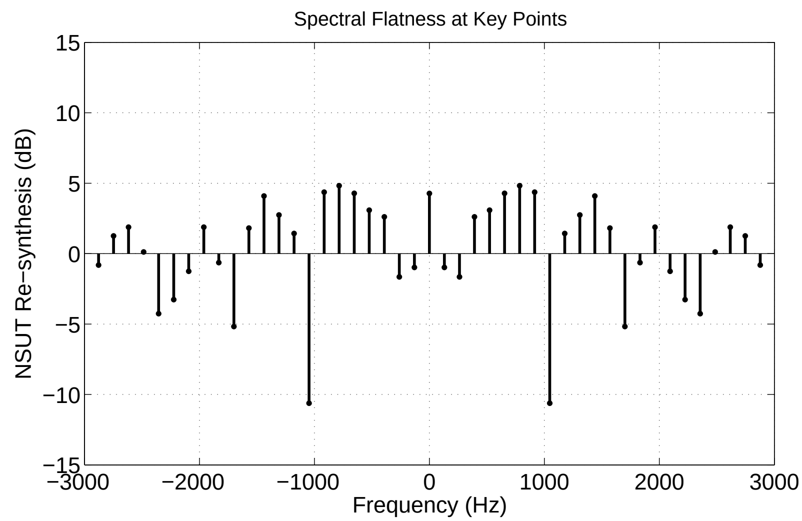
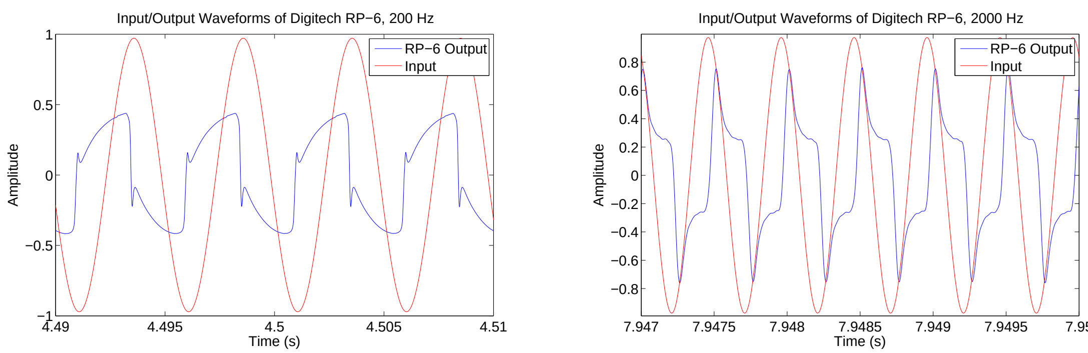


Figure 4–12: Spectral Flatness for the Roland Micro Cube with a musical signal as input.

Figure 4–13: DigiTech RP-6 multi-effects pedal.

The analysis of the RP6 was performed with the following parameters: The distortion type was set to ‘grunge’, which is an over-the-top high gain distortion [42], with a gain setting of 0.6, and the noise gate activated to prevent the characteristic hum from the unit (and most non-gated overdrive pedals and amplifiers) from sounding when there was a very small input signal. All other parameters and effects were de-activated to purely look at the distortion effect. Figure 4–14 shows in/out relationship for two sine waves of different frequency. A power THD measurement was made at 1000 Hz and found to be 70% with these settings. Thus this is a more extreme nonlinearity to model compared with the Roland Micro Cube, with a THD measurement of 13%.



(a) (b)

Figure 4–14: Input/Output waveforms for the RP6 at a. 200 Hz, b. 2000 Hz.

Nonlinear convolution was performed by sending an exponential chirp through the system and recordings its response. The parameters of the sweep were the same as when testing the Roland Micro Cube (Section 4.2.1) and hard clipping nonlinearity (Section 3.5.4). Figure 4–15 shows the MIR output. The RP-6 appears to exhibit the same behaviour as the Roland Micro Cube: many strong odd-order harmonics are produced, and weak even-order harmonics are also present. The MISO re-synthesis was then compared to the system's behaviour with

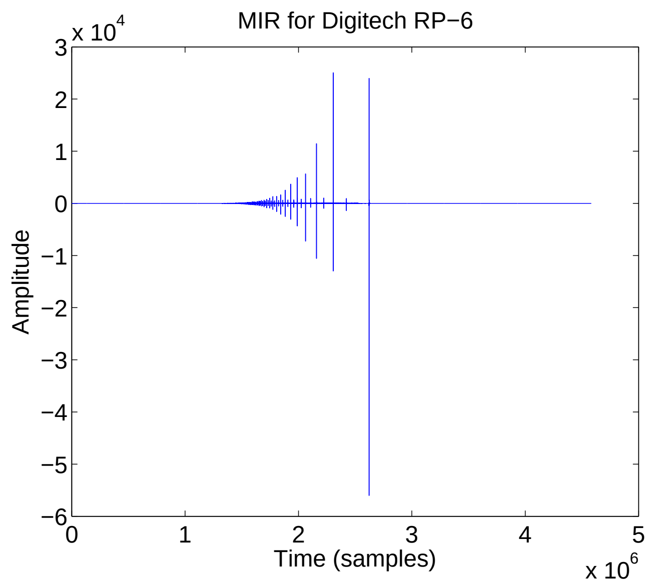


Figure 4–15: MIR output of the nonlinear convolution method for the DigiTech RP6.

the same two input signals as in Section 4.2.1: a simple sine wave of amplitude 0.5 at 1000 Hz, and a guitar pluck, note C2.

Temporal and spectral comparisons of the MISO output with the actual system output are shown in Figures 4–16, 4–17, and 4–18 for the sine wave input, and in Figures 4–19, 4–20, and 4–21 for the guitar pluck input.

Sinusoidal Input

A power THD measurement at 1000 Hz for the MISO output gives 76%, 6% higher than the system's output, despite not modeling all of the harmonics. It was observed that although not all of the harmonics are represented in the MISO output, the odd-order harmonics are slightly stronger than the RP6 ouput, and the

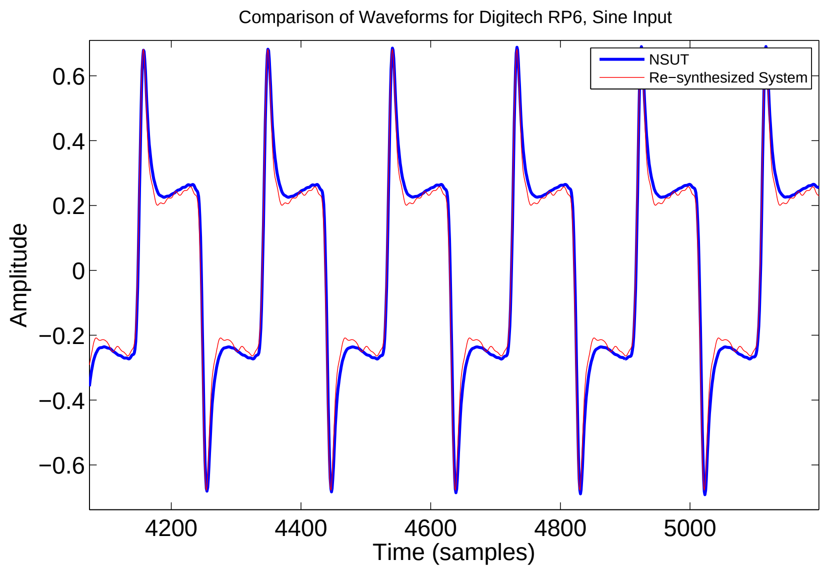


Figure 4–16: Waveform comparison for the RP6 with a sinusoidal input of frequency 1000 Hz.

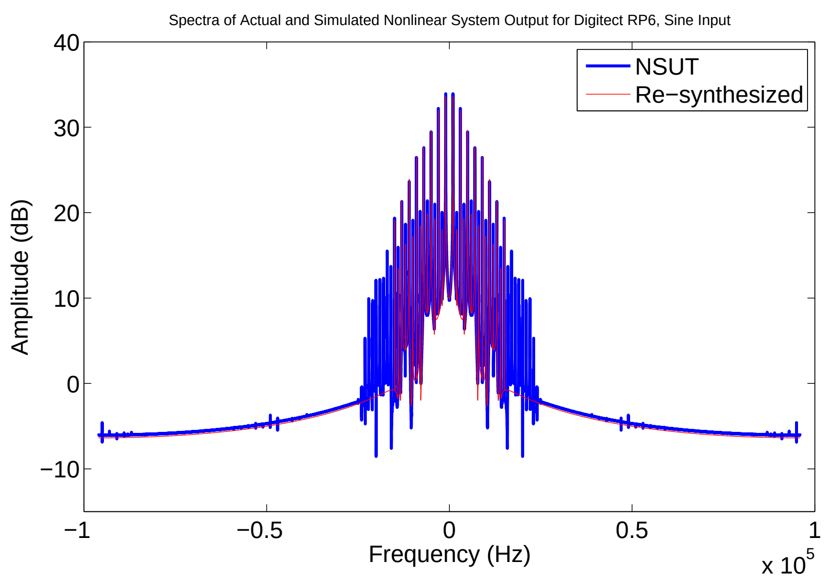


Figure 4–17: Spectral comparison for the RP6 with a sinusoidal input of frequency 1000 Hz.

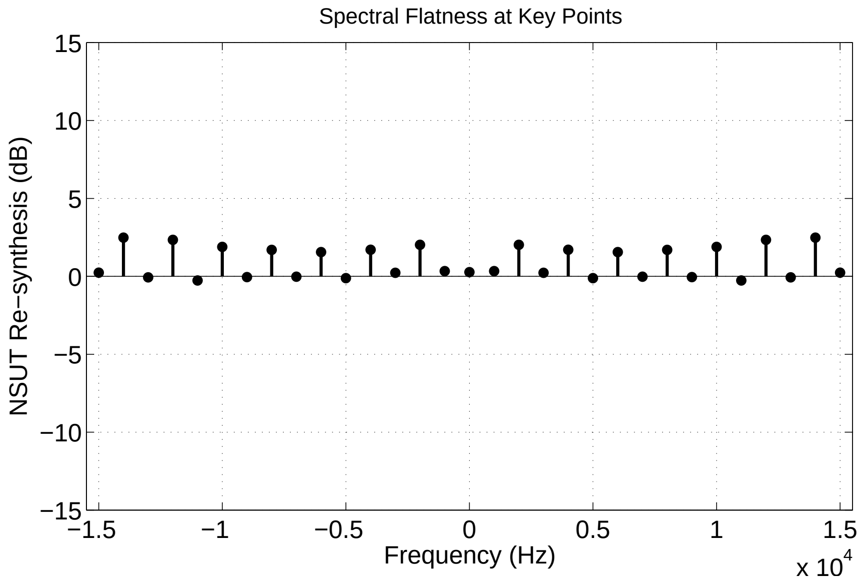


Figure 4–18: Spectral Flatness for RP6 with a sinusoidal input of frequency 1000 Hz.

even-order ones are slightly weaker than the RP6 output. This is the cause for the difference in harmonic distortion in this case.

Figure 4–18 shows that the prevalent odd harmonics of the nonlinearity are well modeled. Their accuracy, shown in Figure 4–18 show agreement to within 0.3dB. The weak even harmonics are not as well modeled, but are still fairly accurate. The re-synthesis has the even-order harmonics on average 2dB lower than the system's output. Once again the two sound nearly identical, with the actual system output having slightly more high frequencies, which agrees with the analysis. It is also encouraging again to see that harmonics higher than $n = 10$ are represented as well as the lower harmonics, despite the aliasing issue mentioned in the previous section. Thus for high order harmonics that are relatively weak, aliasing of the harmonic response when a system is excited with a chirp is not devastating to their re-synthesis in the MISO method.

Musical Input

The system and the MISO model were then fed with a simple guitar pluck at note C2 ($f_0 = 130.81\ Hz$) as with the Roland. Comparisons with the actual system output are given in Figures 4–19, 4–20, and 4–21. The temporal plot in Figure 4–19 shows not much resemblance to the NSUT's output. From the spectral plots of Figures 4–20 and 4–21 we see that the model represents the harmonics 1 to 15 fairly accurately, with the accuracy ususally within 5 dB for each harmonic modeled. While the re-synthesis performed well for the harmonics considered, the accuracy of the re-synthesis when musical inputs are used seems to suffer some when compared to when a sinusoidal input was used. As with the Roland, the model's output to the pluck does have an overdriven character, but it is lacking in high harmonic content and overall gain. Results are discussed further in Section 4.3.

4.2.3 Home-Made Distortion Pedal

The final nonlinear system tested in this framework is a 'Fuzz' distortion pedal designed and built by Avrum Hollinger, a fellow student in the Music Technology department at McGill University.[3] A schematic of the pedal is shown in Figure 4–23. The pedal uses pre-gain, comparator, and gain stages (op amps $U1A$, $U1B$, and $U1D$ respectively) to produce a highly clipped waveform that is

[3] Avrum Hollinger - http://idmil.org/people/avrum_hollinger.

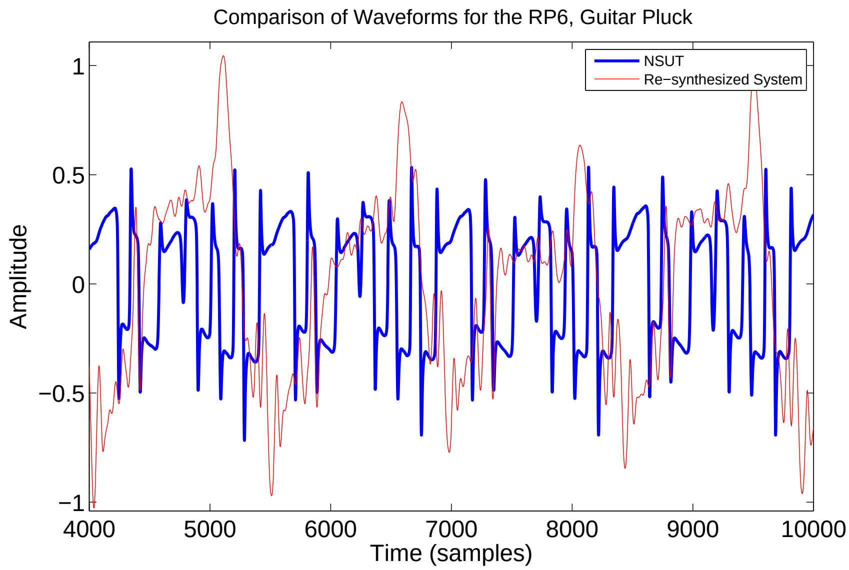


Figure 4–19: Waveform comparison for the RP6 with a musical signal as input.

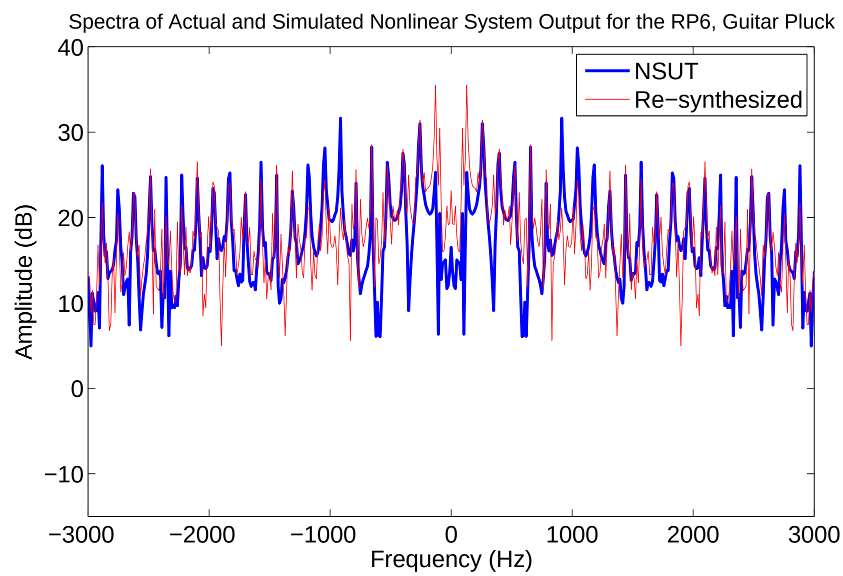


Figure 4–20: Spectral comparison for the RP6 with a musical signal as input. The relevant part of the spectrum is shown only.

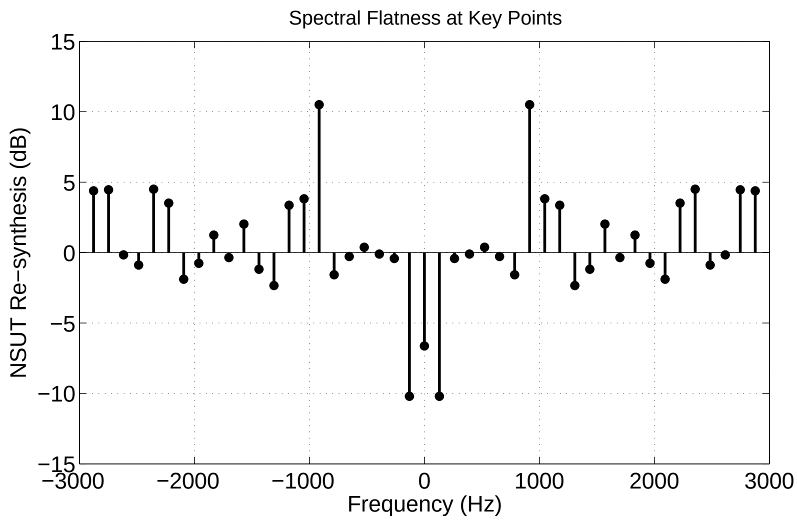


Figure 4–21: Spectral Flatness for the RP6 with a musical signal as input.

very rich in harmonic distortion [43]. The three knobs on the pedal control the pre-gain ($R4$), output volume ($R11$), and the amount of mix between the original and distorted signal ($R9$).

Figure 4–22: Home-made fuzz pedal.

For this test, the pedal was experimented with and adjusted to give a pleasant highly overdriven sound with a midrange mix of the clean and overdriven signal.

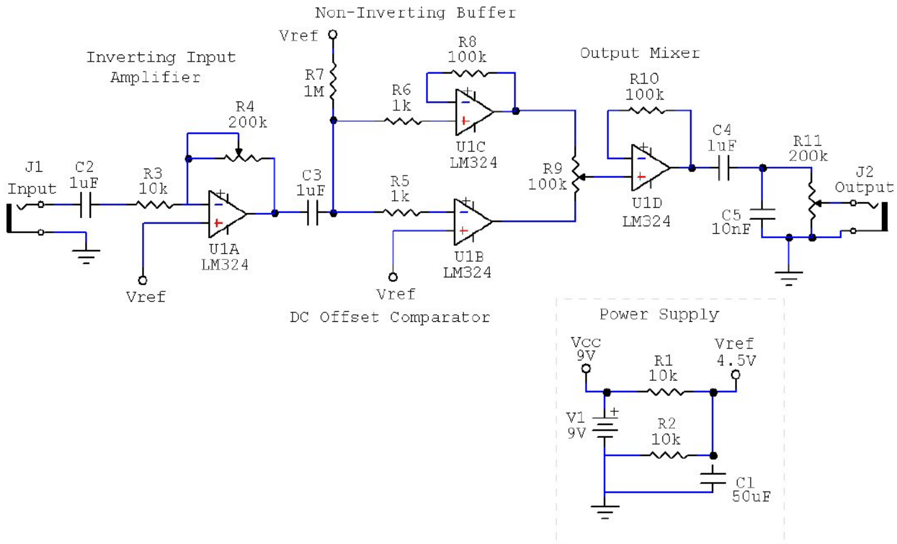


Figure 4–23: Fuzz pedal schematic.

Figure 4–24 shows the input/output relationship for two sine waves of different frequency. The pedal appears to have a slight DC offset and a sharp clipping character. A power THD measurement was made at 1000 Hz and found to be 23% with these settings.

Nonlinear convolution was performed in the usual way to extract the MIR. The parameters of the sweep are again the same as those for the testing of the other real systems. Figure 4–25 shows the MIR output. It is apparent that the fuzz pedal generates both even and odd-order harmonics of appreciable strength. The MISO re-synthesis was then tested with the same sine wave and guitar pluck as before. Parameters of the MISO model are the same as those used in the testing of the Roland and RP6. Temporal and spectral comparisons of the MISO output

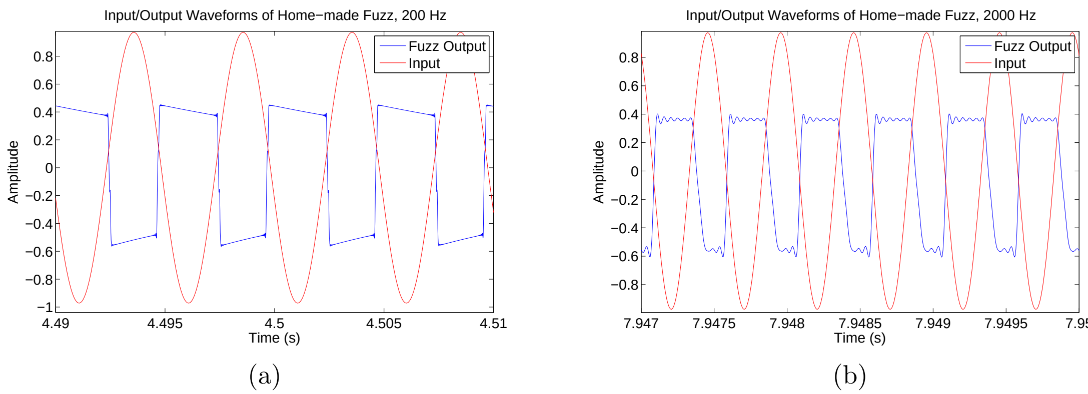


Figure 4–24: Input/Output waveforms for the home-made fuzz at a. 200 Hz, b. 2000 Hz.

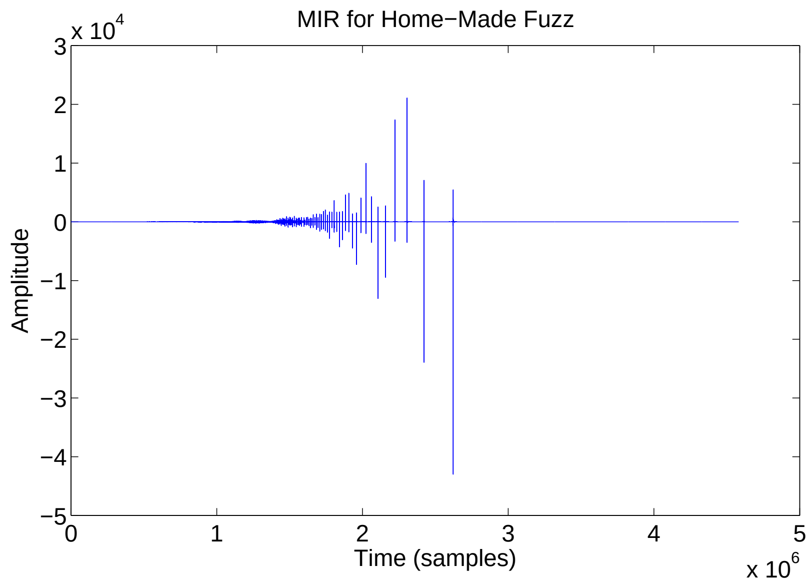


Figure 4–25: MIR output of the nonlinear convolution method for the home made fuzz pedal.

with the actual system output are shown in Figures 4–26, 4–27, and 4–28 for the sine wave input, and in Figures 4–29, 4–30, and 4–31 for the guitar pluck input.

Sinusoidal Input

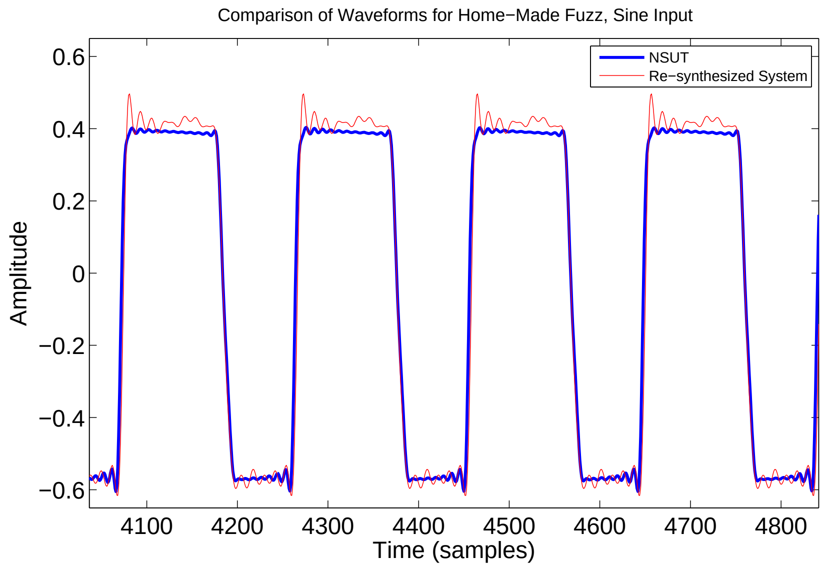


Figure 4–26: Waveform comparison for the home made fuzz with a sinusoidal input of frequency 1000 Hz.

A power THD measurement at 1000 Hz for the MISO output gives 23%, which agrees with the system output's THD. This does not mean necessarily that the model works perfectly in this case, but rather states that the re-synthesis and the system output produce an equal amount of harmonic content, but does not however specify where that harmonic content is. The spectral flatness in Figure 4–28 gives a clearer picture of this. The re-synthesis of the fuzz pedal is on par with the other systems tested. Both odd and even harmonics of the re-synthesis are within 3dB of the system output. What is interesting to note are the signs of the points in Figure 4–28. Previous spectral flatness plots have usually stayed

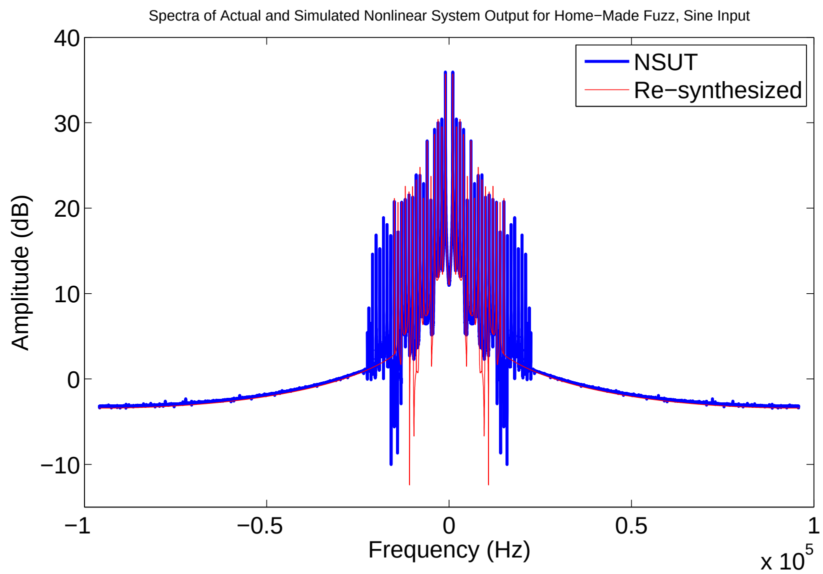


Figure 4–27: Spectral comparison for the home made fuzz with a sinusoidal input of frequency 1000 Hz.

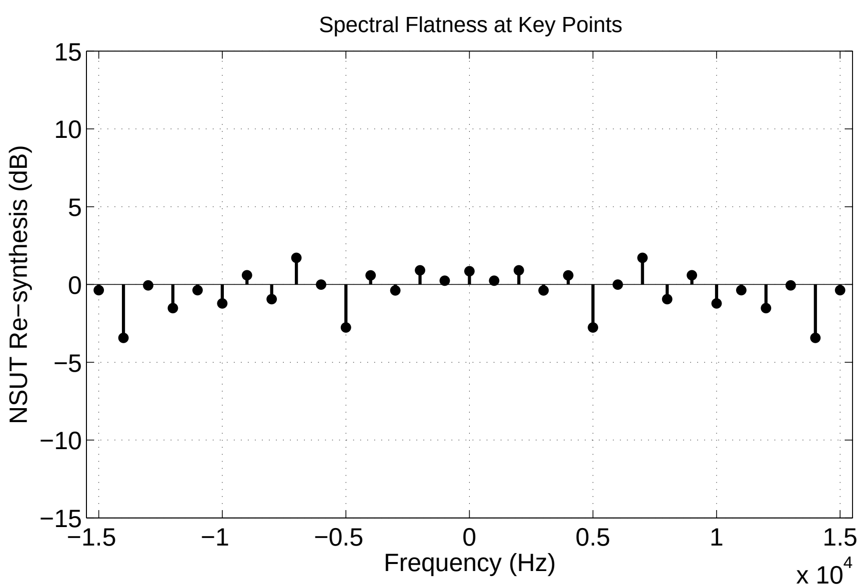


Figure 4–28: Spectral Flatness for the home made fuzz with a sinusoidal input of frequency 1000 Hz.

on one side of zero, but here there are a comparable number of points both above and below zero, implying that some harmonics were too strong and others too weak. There does not appear to be a pattern (such as the odd harmonics being too strong and the even ones being too weak) as was the case with the Roland and RP6.

Harmonics $n = 11$ to 15 are again well modeled and of comparable accuracy to the lower order harmonics despite their potential for aliasing. Perceptually, the system output and re-synthesis sound equivalent to one another with the re-synthesis lacking in the highest harmonic content, which is to be expected.

Musical Input

The fuzz pedal and the MISO model were then fed with a guitar pluck at note C2 ($f_0 = 130.81\ Hz$) as with the other effects. Comparisons with the actual system output are given in Figures 4–29, 4–30, and 4–31. The temporal plot in Figure 4–29 again shows that the model does not simulate the temporal behaviour of the system well. From the spectral plots of Figures 4–30 and 4–31 it is clear that the model is much less accurate with a musical input than with a sinusoidal input. Harmonics 1 to 15 of the model are in agreement to within 10 dB. The timbral qualities again are lacking in high harmonic content and gain as with the re-synthesis of the Roland and RP6 with a guitar pluck as input. Results are discussed further in Section 4.3.

4.3 Discussion

Sound examples for the results presented in this chapter can be found at `http://mt.music.mcgill.ca/~collicuttm/Sound_Examples/index.html`.

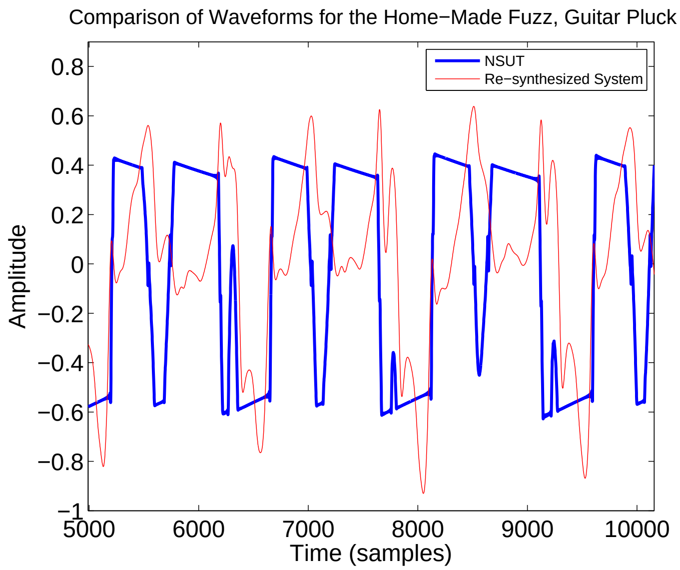


Figure 4–29: Waveform comparison for the home made fuzz with a musical signal as input.

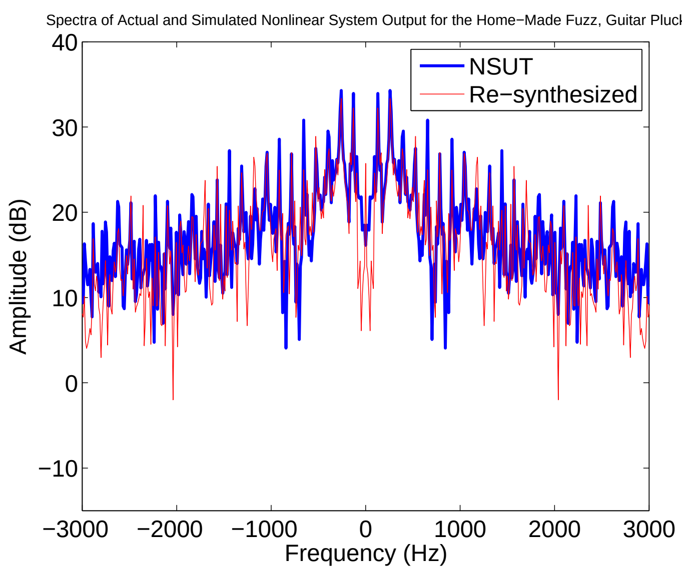


Figure 4–30: Spectral comparison for the home-made fuzz with a musical signal as input. The relevant part of the spectrum is shown only.

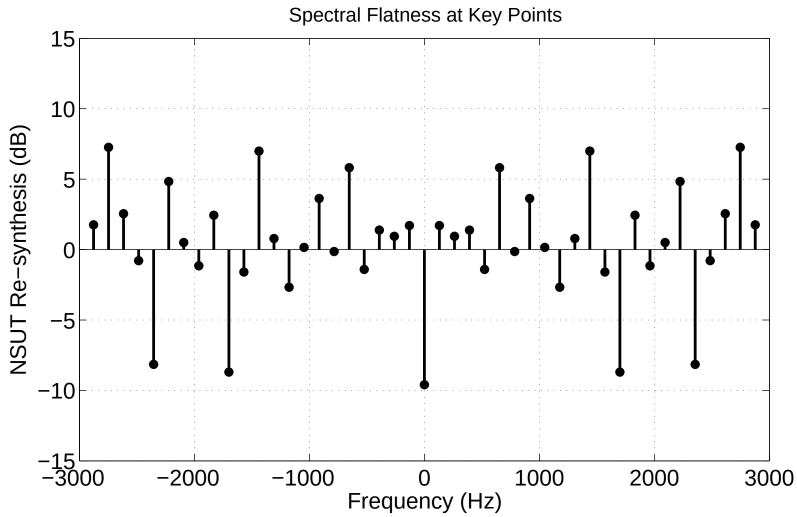


Figure 4–31: Spectral Flatness for the home-made fuzz with a musical signal as input.

The three nonlinear guitar effects were analyzed with Farina's method and MIR were extracted for each system. These MIR were used in the MISO/Hammerstein re-synthesis model to emulate the behaviour of each of the systems. The accuracy of each model was measured by sending both the NSUT and its model a sine wave of amplitude 0.5, and a musical guitar pluck.

The model performed generally well for a sinusoidal input with these systems. Strong harmonics were often modeled to within 1 dB of what the NSUT produced. Weaker harmonics produced by the NSUT were not as accurately modeled, but still within 3 dB of the NSUT's harmonics. Perceptually the model's response to a sinusoidal input was very close to its NSUT fed with the same signal. The model's output seemed to be missing the highest harmonics compared to the NSUT which was expected, since only a finite number of harmonics can be modeled in this framework.

The model did not perform as accurately with a musical input signal. Spectral flatness plots show general agreement to only within 10 dB at the harmonic locations of the plucked note for the harmonics considered.

Each of the effects tested produced harmonics up to roughly 20 kHz regardless of the fundamental frequency of the input signal. Because a low note (C2) was used as the input to the systems under test and the model, a very large number of harmonics are generated in the nonlinear systems just from the fundamental of input. Thus many harmonics must be considered in the model in order to produce the same timbre as the NSUT output for low notes. Because of the trade-off between chirp bandwidth and number of harmonics to model, it becomes impractical at some point to model many harmonics, due to the highly band-limited nature of each HIR that will result. The lack of bandwidth of the harmonics is most important in the linear and low order harmonics, which are typically the strongest. Thus perceptually, the re-synthesis sounds lacking in harmonic content and overall gain.

Higher harmonics ($n > 15$) of the output were not recreated as well. Because the input is a wide-bandwidth signal (containing overtones of the fundamental of the note C2), the nonlinear system also generates harmonics of these overtones as well as the fundamental. Therefore we would expect harmonics higher than $n = 15$ in the output of the NSUT and model. Indeed this is the case, and the NSUT output contains spectral content all the way up to about 20 kHz. Because the input to the model is the same as that of the NSUT, we would also expect the output of the model to contain spectral content up to 20 kHz and even beyond. What was observed was that the spectral content of the model's output dropped off very quickly relative to the NSUT output as the frequency increased, suggesting that the model did not generate harmonics of weaker overtones of the input signal sufficiently. This could be due to the issue of low-amplitude signals being suppressed by the polynomial nonlinearities in the model, as discussed in Section 4.1, and shows that potentially the countermeasures described in Sections 4.1.1 and 4.1.2 were insufficient in addressing this amplitude issue. Thus perceptually, the model's output in response to the guitar pluck does have a significant overdriven character, but seems to lack the gain of the actual effect, which is in agreement with the spectral plots.

Chapter 5
Conclusions and Future Work

The goal of this thesis was to develop, test, and validate a largely automatic analysis/re-synthesis method for harmonic nonlinear systems. The method extracts meaningful parameters in the analysis that characterize a NSUT, and then use these parameters in a generic synthesis model to emulate the response of the NSUT to any input signal as closely as possible. While the focus of this work is on nonlinear audio systems, the methods employed here are intended to be valid across many scientific disciplines where nonlinear systems may arise.

Farina [3] developed a method using sinusoidal sweeps to extract the impulse response of a linear system while simultaneously 'pushing out' any nonlinear responses of the system. Because of this feature, he soon realized that this method could also be used to extract *multiple* impulse responses which provided information about the strength and character of harmonics generated by nonlinear systems as well. By sending a exponentially varying sinusoidal chirp to the system and measuring the response, a convolution of the response with the time-reversed chirp would produce multiple impulse responses separated in time, with each impulse response providing information about a harmonic generated by the nonlinear system. Farina did not provide a precise explanation as to how the method worked to push out these artifacts. One major goal in this work was to provide a more complete mathematical explanation of how nonlinear convolution works as well as illustrate the nonlinear convolution process qualitatively in order to provide a sense of intuition with the method.

A re-synthesis model was developed based on the Multiple Input Single Output and polynomial Hammerstein models. It uses Chebyshev polynomials in a branched structure as the nonlinear elements to generate harmonics of the input signal. Following each Chebyshev polynomial is a linear filter whose impulse response is one of the multiple impulse responses output by the nonlinear convolution method. The branches of the re-synthesis are then combined to form a re-synthesis of the NSUT's output for a given input.

The method was first tested with a few 'synthetic' nonlinear systems created using polynomials or lookup tables as the nonlinear element. Inputs consisted of sinusoids or simple combinations thereof and were sent to both the actual nonlinear system and model and were compared both temporally and spectrally. Results showed that the analysis/synthesis method could accurately reproduce the response of these nonlinear systems with simple inputs.

The analysis/synthesis method was then put to the test with real nonlinear systems. Three guitar distortion effects were tested: a Roland Micro Cube Practice Amplifier with built-in effects, a DigiTech RP6 multi-effects guitar pedal, and a home-made fuzz pedal. It was observed with all of these systems that the re-synthesis was not reproducing higher harmonics to the same degree as with the synthetic nonlinear systems, even with the same input used. To compensate for this apparent weakness in the high harmonics, a pre-emphasis filter was implemented to boost the high-frequency content of the input before being processed by the nonlinearities. In addition an envelope-tracking algorithm was implemented. The envelope was used to boost low-amplitude segments of the input before the applied nonlinearities. This was done because the application of any polynomial to a signal of amplitude less than 1 inherently reduces its amplitude. This effect becomes more drastic as the order of the nonlinearity increases. The amplitude envelope of the signal is then restored after the application of the nonlinearities and filters and the branches summed followed by a 'de-emphasis' filter to counteract the pre-emphasis filter at the beginning of the signal chain. Both compensation measures improved the presence of high harmonics in the re-synthesis, particularly when the systems were tested with simple sinusoids.

The re-synthesis model of the three nonlinear systems was then evaluated using a musical guitar pluck as the input. The signal was sent to both the NSUT and the synthesis model and compared. Results showed that the re-synthesis modeled the nonlinear response of the fundamental of the input reasonably well. It was observed however that in general the outputs lack higher harmonic content compared to the system output. For these distortion modules, the re-synthesis does have a pleasant overdriven, distorted quality, however it sounds as if the traditional 'gain' setting on an amplifier of distortion device is too low. This is thought to be due to the issue of low-amplitude signals or weaker overtones of the input being suppressed by the polynomial nonlinearities in the model, which is inherent in using polynomials. The countermeasures to this issue of envelope following and pre-emphasis therefore require further refinement. The compromise that must be made between the number of harmonics to model and the bandwidth of the chirp also affects high harmonic content in the output. Fewer harmonics considered means that fast transients and high harmonic content cannot be represented, and a low bandwidth chirp means that frequencies higher than f_2 will not be re-synthesized in the linear portion of the output. The analysis could benefit from an upsampling scheme which allows for the expansion of the bandwidth of the chirps while performing the convolution in the analysis. Finding a more elegant solution to capturing the high harmonic content produced by nonlinear systems such as the ones investigated is left as future work.

This analysis/synthesis system was developed in the MATLAB programming environment and hence all computations with input signals were not performed in real-time. Based on this study, the implementation of the re-synthesis method for real-time processing seems feasible and computationally tractable. The re-synthesis involves passing the input into a number of different branches for separate processing on each one, and a convolution operation for each branch. While this sounds like a high-bandwidth scheme, the entire input signal does not need to be stored at once, and convolution techniques involving the FFT can be employed efficiently in real-time [12]. Thus the employment of this synthesis method for real-time use is also proposed as future work.

This thesis focused on the analysis and re-synthesis of nonlinear audio systems, and in particular guitar overdrive and distortion effects. The method however is applicable to other nonlinear systems that may arise in other areas of engineering and science. So long as the system has a single input and single output, and produces harmonic distortion, the method can be applied. Because the analysis method takes a 'black box' approach to the system's characterization and makes very few assumptions, little has to be known about the nonlinear system under study before conducting the analysis. This is advantageous as it allows non-specialists to be able to analyze how a signal is affected by the nonlinear system. This can be of particular interest to musicians interested in emulating audio effects in software. Having a framework that can automatically characterize, modify, and emulate new nonlinear systems or effects could allow for someone with little technical knowledge to be able to create their own amplifier or effect emulations, or even create truly unique new effects.

References

[1] Antonin Novak. *Identification of Nonlinear Systems in Acoustics*. PhD thesis, Czech Technical University in Prague, 2008.

[2] Tom Ritchey. Analysis and Synthesis: On Scientific Method - Based on a Study by Bernhard Riemann. *Systems Research*, 8(4):21–41, 1991.

[3] Angelo Farina. Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique. In *108th AES Convention*, Paris, France, 2000.

[4] Angelo Farina, Alberto Bellini, and Enrico Armelloni. Non-linear convolution: A new approach for the auralization of distorting systems. In *110th AES Convention*, Amsterdam, The Netherlands, 2001.

[5] Julius O. Smith. *Introduction to Digital Filters with Audio Applications*. `http://ccrma.stanford.edu/~jos/filters/`, accessed (2011-02-17). online book.

[6] Julius S. Bendat. *Nonlinear System Techniques and Applications*. John Wiley & Sons, Inc., New York, NY, USA, 1998.

[7] Brian Amberg. Convolution of spiky function with box, MATLAB code on Wikimedia Commons. MATLAB code on Wikimedia Commons, 2010. `http://en.wikipedia.org/wiki/File:Convolution_of_spiky_function_with_box.gif`, Accessed 2011-02-16.

[8] Alan V. Oppenheim, Ronald W. Schafer, and John R. Buck. *Discrete-Time Signal Processing (3rd ed.)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2009.

[9] J.W. Cooley and J.W. Tukey. An Algorithm for the Machine Calculation of Complex Fourier Series. *Mathematics of Computation*, 19(90):297–301, 1965.

[10] Thomas G. Stockham, Jr. High-Speed Convolution and Correlation. In *AFIPS '66 (Spring): Proceedings of the April 26-28, 1966, Spring joint computer conference*, pages 229–233, New York, NY, USA, 1966. ACM.

[11] J.-S. Soo and K.K. Pang. Multidelay block frequency domain adaptive filter. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 38(2):373 –376, feb 1990.

[12] William G. Gardner. Efficient Convolution without Input-Output Delay. *J. Audio Eng. Soc*, 43(3):127–136, 1995.

[13] John Chowning. The Synthesis of Complex Audio Spectra by means of Frequency Modulation. *J. Audio Eng. Soc*, 21(7):526–534, 1973.

[14] Marc Le Brun. Digital waveshaping synthesis. *J. Audio Eng. Soc*, 27(4):250–266, 1979.

[15] Khalil K. Hassan. *Nonlinear Systems*. Prentice Hall, New Jersey, 2002.

[16] Brian Santo. Volume Cranked Up in Amp Debate. *Electronic Engineering Times*, 817, October 1991.

[17] Eric Barbour. The Cool Sound of Tubes. *IEEE Spectrum*, 35(8):24–35, 1998.

[18] Native Instruments. Guitar Rig Control. http://www.native-instruments.com/#/en/products/guitar/guitar-rig-kontrol, accessed 2011-02-16.

[19] Universal Audio Inc. Universal Audio. http://www.uaudio.com, accessed 2011-02-16.

[20] David Griesinger. Beyond MLS - Occupied Hall Measurement with FFT Techniques. In *101st AES Convention*, Los Angeles, California, 11 1996.

[21] M.R. Schroeder. Integrated-Impulse Method measuring Sound Decay without using Impulses. *Journal of the Acoustical Society of America*, 66(2):497–500, 1979.

[22] Bill Waslo. The IMP goes MLS. *Speaker Builder Magazine*, 6:130–136, 1993.

[23] Richard C. Heyser. Acoustical Measurements by Time Delay Spectrometry. *J. Audio Eng. Soc*, 15(4):370–382, 1967.

[24] Cecil R. Cable. Time Delay Spectrometry is Alive and Well. In *60th AES Convention*, Los Angeles, California, (5) 1978.

[25] Walt Kester. Understand SINAD, ENOB, SNR, THD, THD+N, and SFDR so you don't get lost in the noise floor. Tutorial MT-003, Analog Devices, 2003. `http://www.analog.com/static/imported-files/tutorials/MT-003.pdf`.

[26] William Press, Saul Teukolsky, William Vetterling, and Brian Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, UK, 2nd edition, 1992.

[27] Patrick Flandrin. *Time-frequency/time-scale analysis*, volume 10 of *Wavelet Analysis and its Applications*. Academic Press Inc., San Diego, CA, 1999.

[28] Athanasios Papoulis. *Signal Analysis*. McGraw-Hill, New York, 1977.

[29] Bob York. Stationary phase method. ECE 201C Antenna Theory Lecture Notes, 2003. http://my.ece.ucsb.edu/bobsclass/201C/Handouts/StationaryPhase.pdf, accessed 2011-02-16.

[30] E.T. Copson. *Asymptotic Expansions*. Cambridge University Press, Cambridge, 1965.

[31] Leon Cohen. *Time Frequency Analysis : Theory and Applications (Prentice-Hall Signal Processing)*. Prentice Hall PTR, December 1995.

[32] Charles Dodge and Thomas A. Jerse. *Computer Music: Synthesis, Composition, and Performance*. Schirmer Books, 2nd edition, 1997.

[33] Angelo Farina. Advancements in Impulse Response Measurements by Sine Sweeps. In *122nd AES Convention*, Vienna, Austria, 2007.

[34] Marc Le Brun. Digital Waveshaping Synthesis. *J. Audio Eng. Soc*, 27(4):250–266, 1979.

[35] Daniel Arfib. Digital synthesis of complex spectra by means of multiplication of nonlinear distorted sine waves. *J. Audio Eng. Soc*, 27(10):757–768, 1979.

[36] Julius S. Bendat and Allan G. Piersol. *Random Data: Analysis and Measurement Procedures*. John Wiley & Sons, Inc., New York, NY, USA, 1990.

[37] S. A. Billings and S. Y. Fakhouri. Non-linear system identification using the Hammerstein model. *International Journal of Systems Science*, 10(5):567–578, 1979.

[38] M. Boutayeb, M. Darouach, H. Rafaralahy, and G. Krzakala. A new technique for identification of MISO Hammerstein model. In *American Control Conference*, pages 1991–1992, San Francisco, California, 1993.

[39] L. Fox and I.B. Parker. *Chebyshev Polynomials in Numerical Analysis*. Oxford University Press, London, 1967.

[40] Robert Bristow-Johnson. Cookbook formulae for audio eq biquad filter coefficients. `http://www.musicdsp.org/files/Audio-EQ-Cookbook.txt`. accessed 2011-02-17.

[41] Roland Corporation. *Roland Micro Cube Owner's Manual*, 2003. `http://lib.roland.co.jp/support/en/manuals/res/1811450/M-CUBE_e1.pdf`.

[42] Harman Music Group Inc. *Digitech RP6 Owner's Manual*, 1996. `ftp://ftp.digitech.com/pub/PDFs/Discontinued/Manuals/RP6manual.pdf`.

[43] Avrum Hollinger. Distortion for electric guitar. February 2003.