High-Voltage Wideband Switching Amplifier for Capacitive Loads

by

John C. Olsson

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

M.Eng in Electrical Engineering

at the

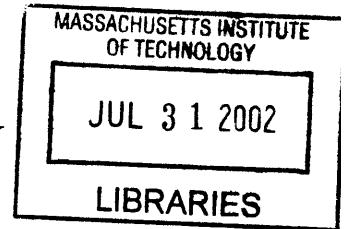## MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2001

(September 2001)

Author . . . .
Department of Electrical Engineering and Computer Science
August 30, 2001

Certified by. . . . .
Barry Vercoe
Professor of Media Arts & Sciences
Thesis Supervisor

Certified by. . . . .
F.JosephPompei
Underling PhD Candidate
Thesis Supervisor

Accepted by . . . . . .
Arthur C.Smith
Chairman, Department Committee on Graduate Students

High-Voltage Wideband Switching Amplifier for Capacitive Loads

by

John C. Olsson

Submitted to the Department of Electrical Engineering and Computer Science
on August 31, 2001, in partial fulfillment of the
requirements for the degree of
M.Eng in Electrical Engineering

Abstract

Why is it that arbitrarily driving imaginary loads has always required lots of power? In this thesis, a highly efficient switching amplifier class is developed that is capable of delivering energy to, as well as taking energy from, a capacitive load in a finely controllable, dissipationless manner. Several control schemes were investigated, and a simple version of the amplifier was then built and tested using both synchronous and asynchronous controllers. The amplifier proved to be capable of driving high voltage, high frequency signals across a capacitive transducer with extremely low total power consumption and very low distortion.

Thesis Supervisor: Barry Vercoe
Title: Professor of Media Arts & Sciences

Thesis Supervisor: F.JosephPompei
Title: Underling PhD Candidate

Acknowledgments

I'd like to acknowledge the following businesses for their reliable, daily support of my efforts towards this thesis:

- Dunkin Donuts

- Sweet Touch Cafe

- Magnificent Muffin & Bagel Shoppe

- MIT Coffeehouse

- Phil Collins

- Au Bon Pain

- Tosci's

Contents

List of Figures

Chapter 1

Introduction

Figure 1-1: The Beast

The fundamental picture to visualize is Figure 1-1. We have a capacitor consisting of two conductors oriented such that when charged, an electric field is constrained between them, along with its associated energy. The "voltage" across the plates is purely a state variable describing the amount of charge that has been accumulated into that geometry, and is given by $V = Q/C = \frac{1}{C} \int_{-\infty}^{t} i(t)dt$. Although the electric field does the mechanical work which we are interested in, one typically controls the voltage that corresponds to [1] to the electric field and can be used as a measure and

---

[1] In some situations this is a first order approximation, as the geometry itself can change as a

control signal in electric circuitry.

The capacitor itself is an energy storage device; energy stored in the electric field varies with voltage to the second power $U_E = \frac{1}{2}CV^2$. Aside from negligible wire resistance during charge transfer, *there is no inherent power dissipation in the capacitor.* Our goal, therefore, is to control the voltage by integrating charge into and out of the capacitor, while ideally losing zero energy to heat in the process. From the historical introduction to capacitive drives, this goal can be viewed as, in no fewer words, quite formidable.

An undamped LC tank circuit approximates the lossless energy transfer just described, but with the shortcoming that it only operates at the single frequency $\omega_o = (LC)^{\frac{1}{2}}$.

We are faced with a unique challenge when trying to achieve resonant efficiency over a broad bandwidth. Looking at a snapshot of a resonant circuit, the energy stored in the capacitor's electric field ($\frac{1}{2}CV^2$) and the inductor's magnetic flux ($\frac{1}{2}LI^2$) has no concept of time. It is through the coupling of these energies, via the circuit wiring, that the state variables influence each other in a fixed, temporal manner, hence ω_o. If we control, in the time domain, this "coupling" of energy, then we control the time-evolution of the system. Thus, we have achieved bandwidth.

This we do by switching.

---

function of voltage and therefore the capacitance changes too

Chapter 2

Historical Approaches

Capacitive loads are reactive – they have no real component of their impedance ($\mathcal{R}\{Z\} = 0$). In theory[1], a pure capacitance shouldn't require any energy to be dissipated when driving signals across them. Unfortunately the means to do this have not been well developed; instead researchers and consumers alike have needlessly resorted to burning large amounts of power.

One of the greatest things to come out of the 1950's were class-B push-pull transistor amplifiers. With this output stage, driving an output signal $V_o(t) = A\sin(\omega t)$ across a purely resistive load R draws positive and negative current $V_o(t)/R$ across half of the rail-to-rail supply voltage ($2V_{cc}$) at any given time. In terms of power,

$$P_{out} = \frac{1}{\pi}\int_0^\pi V_o^2(t)/R \, d(\omega t) = \frac{A^2}{2R}$$

$$P_{amp} = \frac{1}{\pi}\int_0^\pi (V_{cc} - A\sin(\omega t)\frac{A}{R}\sin(\omega t)d(\omega t) = \frac{2V_{CC}A}{R\pi} - \frac{A^2}{2R}$$

$$\text{efficiency} = \frac{P_{out}}{P_{amp} + P_{out}} = \frac{A^2/2R}{2V_{cc}A/R\pi} = \frac{A\pi}{4V_{cc}}$$

This is what one expects from a load that dissipates power. Notice that the optimal efficiency is bounded by $\pi/4 \approx 79\%$. For a capacitive load $Z = 1/(Cs)$ the situation is different. The current is 90° out of phase ($I_o(t) = (A/R)\omega\cos(\omega t)$), and we get

---

[1] a theory naive of thermodynamics

12

$$P_{out} = \frac{1}{2\pi} \int_0^{2\pi} \left(A \sin(\omega t) \right) \left(AC\omega \cos(\omega t) \right) d(\omega t) = 0$$

$$P_{amp} = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} (V_{cc} - A \sin(\omega t)) \left(AC\omega \cos(\omega t) \right) d(\omega t) = \frac{2V_{cc}AC\omega}{\pi} \qquad (2.1)$$

Power consumption looks similar to the resistive load but now is dependent on ω!. This is because the energy $U_e = \frac{1}{2}CV^2$ is transferred to and from C each half-period. Therefore $power = \frac{energy}{time} = 2CA^2 f$, which is the minimum power in Equation 2.1 for a given A.

A manifesto entitled *Driving Capacitive Loads* [Ape] prepared by APEX CORP. delineates the usage of their power op-amp product line for driving such reactances. They offer the same formula written for the optimal case where the supply voltage is the peak output voltage:

$$P_{opt,C} = \frac{4A^2\omega C}{2\pi} \qquad (2.2)$$

The problem is made obvious by the following table for a $200V$ amplitude sinewave on $C = 10nF$:

Frequency	Power
100Hz	160mW
5kHz	8W
10kHz	16W
50kHz	80W
80kHz	128W!!!

2.0.1 Resonant Topologies

Fortunately a lot of capacitive trandsducer applications require operation either at a single frequency or over a very narrow bandwidth (i.e. using short pulses of ultrasound to measure distance). This allows for a resonant LRC "tank" to be built into the load. At the resonance frequency, energy sloshes around between the two reactive

13

components, alleviating the workload of the amplifier. These are known as "high-Q" circuits, where 'Q' is defined as $Q \equiv \frac{\text{center frequency}}{\text{bandwidth}}$ which also takes the form $Q \equiv \frac{2\pi(\text{peak energy stored})}{\text{energy dissipated per cycle}}$ [KSV91].

The power savings are dramatic, but it is at the tradeoff of an extremely non-flat response. In some cases this is acceptable; a lot of the time it absolutely is not (such as any audio applications).

2.0.2 Problems with Prior Methods

1. Increasing losses with increasing frequency. Equation 2.2 shows that there is no upper bound on power use versus frequency. Power consumption is also directly proportional to the load capacitance. Both of these statements are unfortunate realities considering that the load is dissipating none of this power itself.

2. The power used in driving the capacitance appears as heat in the amplifier. Where one finds heat, one usually finds a large, heavy heat sink, vents, and probably a noisy, dusty fan. The whole package isn't very attractive.

3. Power cord must reach wall outlet.

4. There must be that wall.

Chapter 3

Overview

In this section we discuss a class of amplifiers that operate by switching different sources of constant current into a load capacitor. These are constant-slope integrating amplifiers, whose specific control and physical implementation are covered in the following chapters.

3.1   Converter Waveforms
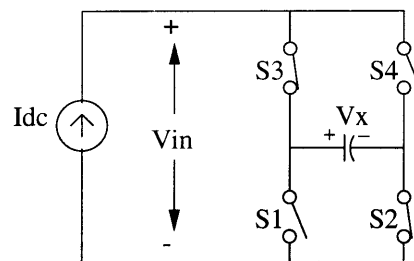
Figure 3-1: Simplified Converter Topology

The basic switching topology is illustrated in Figure 3-1. For now we assume that the current source is constant and an H-bridge of switches connects it in various arrangements to the load. The switches are grouped in pairs: S_1 and S_2 are the "low-side switches" while S_3 and S_4 are "high-side". One of the switches in each pair is always on at any time, which is designated by using the odd-numbered switch as a

boolean variable. Thus, $S_1 = 1$ means S_1 is ON and S_2 is OFF. This allows for the following modes:

Mode	Switch Positions	Final Value
Charge Up	$S_2 S_3$	$V_o + \frac{I_{dc}}{C}\Delta t$
Charge Down	$S_1 S_4$	$V_o - \frac{I_{dc}}{C}\Delta t$
Neutral	$S_1 S_3 \ OR \ S_2 S_4$	V_o

Note that during "Charge Up" energy is being transferred to the load, and during "Charge Down" energy is being transferred back into the current source; the neutral position "uncouples" the source from the load, transferring no net energy. **This conservation is the primary goal of the switching amplifier** [1].

Another useful way to rewrite this, which will come in handy later on, is to view the converter output as the sum of two saturated controllers passed through an integrator. It's also a good time to disregard the scale factor $\frac{I_{dc}}{C}$. The controller $u_{lo}(t) \in \{-1, 0\}$ and $u_{hi}(t) \in \{0, 1\}$, and this is illustrated in Figure 3-2.
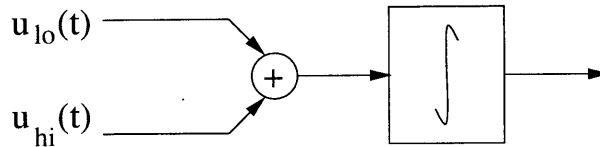
Figure 3-2: Saturated Control Representation of Converter

Theoretically these controls $u_{lo}(t)$ and $u_{hi}(t)$ can be applied in any arbitrary manner to achieve some purpose (the purpose will be revealed in Section 3.2). Practically, however, the decisions of when to throw the switches are determined by a finite, limited amount of knowledge about the state of the system, the future state of the system, and physically what the system is capable of doing. The governing of these switches also reflects tradeoffs between different measures of system performance.

---

[1] Of course we aren't violating any laws of thermodynamics; on a much larger time scale, energy is being dissipated in switching and parasitic elements. Then again, the word "converter" implies efficiency strictly below 100%.

In many cases the system isn't endowed with any degree of prescience and is left to follow simple rules determining its output. A common instantiation is that of a **relay**, in which the output is determined by a single input that is compared to specified threshold values. The input is usually some function of the system error and its derivatives. Several relay transfer functions are illustrated in Figure 3-3.

generic two-valued relay   relay with hysteresis (time delay)   relay with dead zone
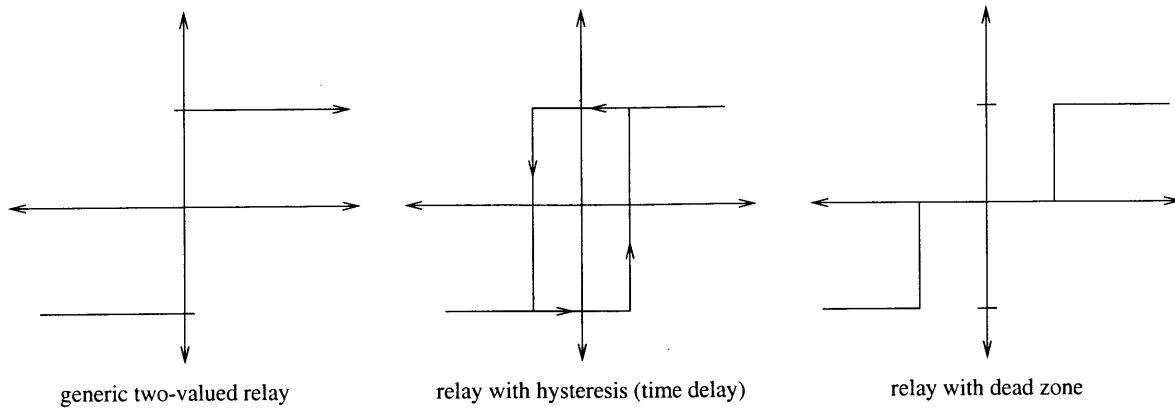
Figure 3-3: Different Relay Transfer Characteristics

Note: Historically the advent of the relay as a fundamental engineering building block came in the beginning of the 20th century. Relays would appear as saturated actuators that were typically used to achieve some mechanical action. An explosion of nonlinear control theory involving relay mechanics came after WWII, in which funding from the military went into improving the stabilization of jet airplanes.

3.1.1 Ramp Response

To get an idea of the type of waveforms produced by this converter, the output shown in Figure 3-4 is what to expect with a simple feedback error amplifier. The converter is attempting to track a signal of slope a by either holding a constant output or ramping with slope A. The output is switched when it's absolute error exceeds a bound $|y(t) - x(t)| > \alpha$, which is the relay configuration just described. The error sensing, control logic and switching transients are effective after a finite time delay τ. The switching points are derived from simple geometry:

17

$$t_1 = \frac{\alpha + a\tau}{A - a}$$

$$t_2 = t_1 + \frac{\alpha}{A - a}$$

$$t_3 = t_2 + \tau$$

$$t_4 = t_3 + \frac{\alpha + (A - a)\tau}{a}$$

$$t_5 = t_4 + \frac{\alpha}{a}$$
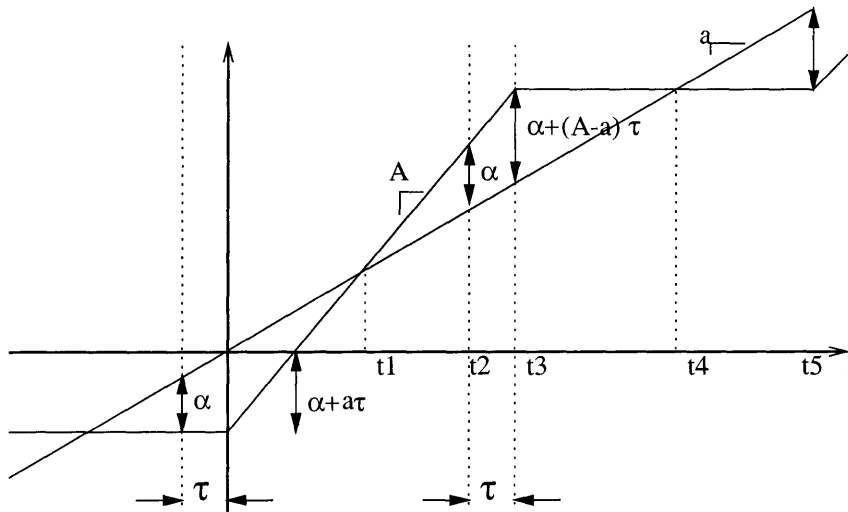
Figure 3-4: Tracking a Constant Ramp

There are a few things to observe from this exercise:

- The local-average slope of the output $A\frac{t_3 - t_0}{t_5 - t_0}t$ is the same as the input slope a.

- The converter cannot track any ramp faster than it's slope A. If it does, the error grows unbounded as long as that input persists. This is the mechanism that will lead to *overload distortion* described later.

- The error of the system doesn't converge to zero, but oscillates as an asymmetric triangle wave around zero. The peak amplitude of the error is $\alpha = \left| \frac{d(input)}{dt} - \frac{d(output)}{dt} \right| \tau$. Immediately we can see that switching delay, which is a

18

finite quantity resulting from the physical implementation, seems to impose a lower bound on the error signal. This will remain true in most cases, but isn't always necessarily counterproductive.

- Momentarily ignoring the additional affect of switching delay τ, the error period is

$$T = \frac{2\alpha(A - \alpha) + 2\alpha(a)}{a(A - a)} = \frac{2\alpha A}{a(A - a)} \rightarrow f_{osc} = \frac{a(A - a)}{2\alpha A}$$

- From the last two items we begin to see the inherent link between error energy and control activity. We can see how the state evolution $\dot{x} = slope \approx \frac{\Delta signal}{\Delta time}$ relates signal to time through a system of differential equations; as the controller activity is reduced by allowing more time between transitions, the direct consequence is more error growth during that period. It sounds basic, but it is fundamental to the entire controller design so there is no harm in becoming intimately acquainted with its nature. Through describing function analysis in Section 6.2 the stability versus amplitude of these oscillations can be visualized in the frequency domain as a limit cycle from nonlinear feedback.

3.2 The Goal, In So Many Words

The brain of a 4-year old learns how to actuate the muscles up and down the arm and hand to trace a crayon along a predetermined path, using visual feedback to close the loop. A dozen years later that boy traces city roads with a three thousand pound vehicle. He wants to come close to the target path by making continuous, minute control adjustments. While the crayon example may judge performance by how close his line comes to the input, passengers in his car may take into consideration the first or second derivatives of his error as well, especially if they are trying to trace something with a crayon.

In the amplifier business, the designer is faced again with the age-old, kindergarten nemesis of tracing. However we must limit the possible trajectories to those whose output derivatives take on a finite number of fixed values. In other words, we can

ramp up, ramp down, or hold steady (this is illustrated later in Figure 6-1).

3.3 So, Mathematically, What is This Thing?

The output of the converter is a continuous signal of piecewise-linear segments. These segments can be theoretically any length [2], but can only remain constant or change up or down at a fixed rate. As just described, this simple device allows for several different mathematical analyses, each with its associated assumptions, performance limitations, and ease of translation into a controller. In brief, the system can be described as follows:

- A nonlinear analog feedback circuit. One might use describing function analysis to approximate gain and phase shift for sinusoidal inputs, in which the 3-slope integration is modeled as a relay function connected to an integrator. Loop delay is the fundamental quantity that controls the tradeoff between performance and controller action, as well as instability.

- A fixed-frequency 3-level delta-modulator. This is typically a special case of a differential pulse-width modulator (DPCM) which has an extremely reduced-level quantizer (in this case, three levels). The quantizer-integrator combination acts as a linear predictor for the input signal based on the output. This is subtracted from the input and the differences are quantized, thus improving dynamic range over a straight 1.6-bit quantizer.

- A fixed-frequency, oversampled noise-shaping (ONS) A/D converter, also known as a sigma-delta converter. The idea is to increase input-signal adjacent-sample correlation by prepending the deltamodulator with an integrator, and then differentiating the output. A side-effect of this is that the in-band component of quantization noise is subtracted from the input, thereby "shaping" the noise by moving it to the out-of-band frequencies.

---

[2]In fact there are finite minimum time constraints which play are large role in the converter operation, as we will see later

- A linear system with a constrained class of controller inputs, namely three saturated input states with a minimum time between changing. This is a type of problem called "bang-bang control"; the optimal controller minimizes a specified performance measure, usually by applying the extremum of a constrained control signal. This will be discussed thoroughly later.

After mathematically specifying the intended input classes of signals, each of these analytical approaches will be investigated. The output performance resulting from applying each of the mathematical models will be presented, and their tradeoffs will be discussed thoroughly in the conclusion.

3.4 Input Signal Information

Anything we *know* about the input signal – bounds, shape, periodicity – can significantly bring us closer to our goal of reproducing it. Quantitative measures of how well we achieve this goal are now discussed.

3.4.1 Sinusoid Inputs

Initially we are interested in step and ramp inputs to qualify performance aspects of this converter. Our ultimate goal, thankfully, is to reproduce something far more practical – music. The analysis of reproduced audio fidelity is an entirely separate and complicated field, encompassing not only mathematical but also physiological sciences, and colored by the specious doctrines of the mock-scientific audiophile world[3]. For the purposes of this thesis, therefore, we shall investigate the reproduction of *sine waves*. Not only are their properties cleanly, mathematically defined, but measures such as SNR and THD thereby derived provide much insight into the reproduction of wider-spectrum signals. With no further delay, our input is

$$r(t) = R_o \sin(w_c t)$$

---

[3]see *The Absolute Sound* publication

Because the amplifier topology is inherently an integrator, the immediate properties of interest are derivatives:

$$\frac{dr(t)}{dt} = R_o w_c \cos(w_c t), \qquad \frac{dr(t)}{dt}\Big|_{max} = R_o w_c \qquad (3.1)$$

A fundamental law of switched constant-slope integrating converters is that the product of output frequency and amplitude is bounded by the abilities of the converter. In this sense, the "ability" is the amount of charge that can be delivered in a finite time to the integrating capacitor, namely $\frac{I_{dc}}{C}\Delta t$.

When the input changes faster than the amplifier can integrate, over a finite period of time, i.e. $\frac{dr(r)}{dt} > \frac{I}{C}$, we say the amplifier is *saturated*. In the next section we will discuss these consequences as they relate to performance. For the time being assume that, under normal conditions, occasional saturation is acceptible.

Saturation on the derivative of a sinewave results in distortion. Figure 3-5 shows sine waves $v(t) = A\sin(t)$ with amplitude $A \in [1..2]$ and $|\frac{dv(t)}{dt}| <= 1$. THD measurements (see Section 3.5.1) are shown on the plots for $v(t)$. A plot of THD vs. amplitude is given in Figure 3-6.

It is easy to generate first-order filters that anticipate near-future inputs, i.e. for $r(t) = \sin(2\pi f_c t)$, a filter that would predict $\sin(2\pi f_c(t + T_s))$, $(T_s << f_c^{-1})$. This becomes especially useful with controllers that base decisions on trying to optimize predicted future performance. For short distances ahead, we can just use a linear approximation $r(t + T_s) \approx \tilde{r}(t) = k[r(t) + \dot{r}(t)T_s]$. The approximation error of the corresponding filter $H_1(s) = k(1 + sT_s)^4$ can be found from:

$$H_{err1}(s) = H_1(s) - e^{sT_s} = k(1 + sT_s)^2 - \left(1 + sT_s + \frac{(sT_s)^2}{2!} + \frac{(sT_s)^3}{3!} + \cdots\right) \qquad (3.2)$$

For a sine wave or band-limited signal at frequency f_c, we can keep the amplitude $|H_1(s)||_{f_c} = 1$ by setting $k = \frac{1}{\sqrt{1+(2\pi f_c T_s)^2}}$. (Note $k \approx 1 - \frac{1}{2}(2\pi f_c T_s)^2$, which matches

---

[4]Of course there are poles much higher than the frequencies of interest that keep the transfer function proper.
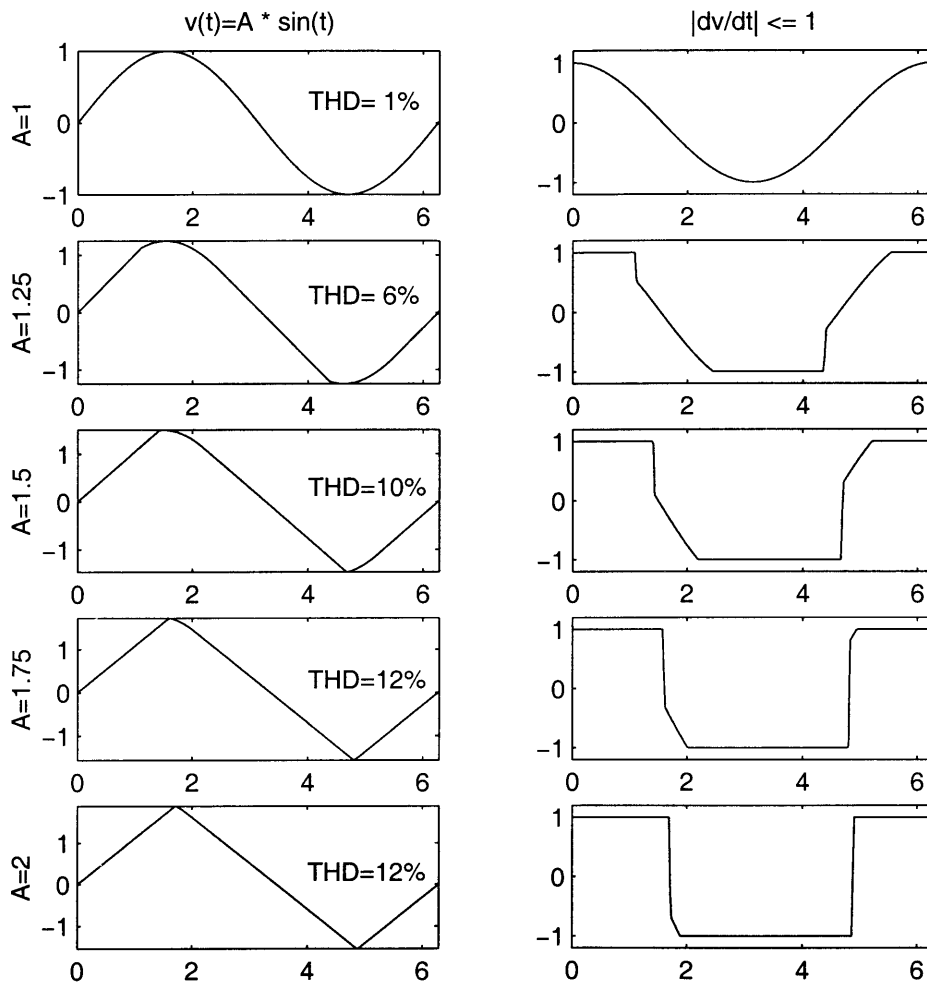
Figure 3-5: Distortion on Slope-Overloaded Sine Waves
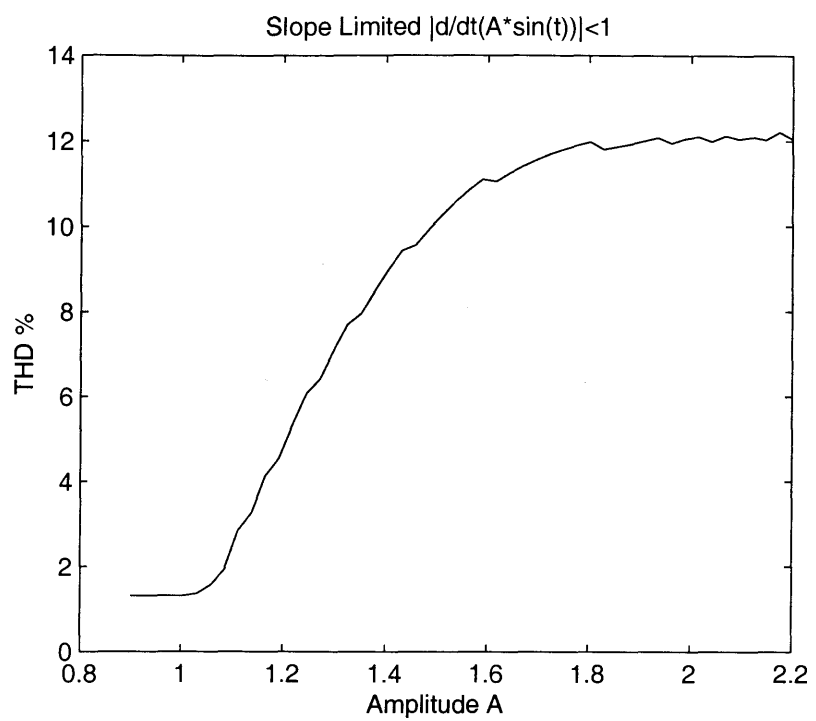
Figure 3-6: Slope Overload Distortion vs. Amplitude

the third order term of Equation 3.2.) Obviously from the expansion of e^{sT_s}, the error grows quickly as f_cT_s gets larger.

3.5   Output Performance Measures

Without some repeatable method of mathematically quantifying the performance of the system, everybody can be a winner [5]. Historically the most popular characteristic of an audio amplifier is the Total Harmonic Distortion (THD) measure. This is particularly applicable to "linear" amplifiers which have their inherent nonlinearities minimized by using frequency-dependent feedback loop gain. In switching converters, however, THD isn't always the best choice of a performance figure. For example, one-bit modulators are theoretically linear, and therefore the generation of harmonics of the input signal is avoided. However, the mechanism of quantization injects noise into the circuit, which may have its own coloring (i.e. pattern noise). Therefore the Signal-to-Noise Ratio (SNR) is an extremely useful performance measure.

3.5.1   THD

$$\text{THD} \equiv \sqrt{\frac{\sum_{i=2}^{\infty} V_{i,rms}^2}{V_{1,rms}^2}} \qquad \text{where} \qquad V_{i,rms}^2 = \frac{1}{2\pi} \int_0^{2\pi} V_X^2(t) sin(i\omega t) dt$$

In other words, it is the ratio of power in the harmonics to the power in the fundamental component. Sometimes the noise power n^2 is factored in to yield:

$$\text{THD+N} \equiv \sqrt{\frac{\sum_{i=2}^{\infty} V_{i,rms}^2 + n^2}{V_{1,rms}^2} + n^2}$$

When the switching frequency is a multiple of the fundamental frequency, we definitely get harmonics. Though purely academic, this is illustrated in Figure 3-7.

As soon as a discrete-time modulator isn't at an integral multiple of the input sinusoid, the combination of two fixed frequencies leads to modulations around the

---

[5]see Audiophile magazine

4 segments, 11.88% THD

6 segments, 4.06% THD

8 segments, 11.88% THD

8 segments, 28.69% THD

8 segments, 11.88% THD

10 segments, 6.90% THD

10 segments, 18.08% THD

10 segments, 18.79% THD
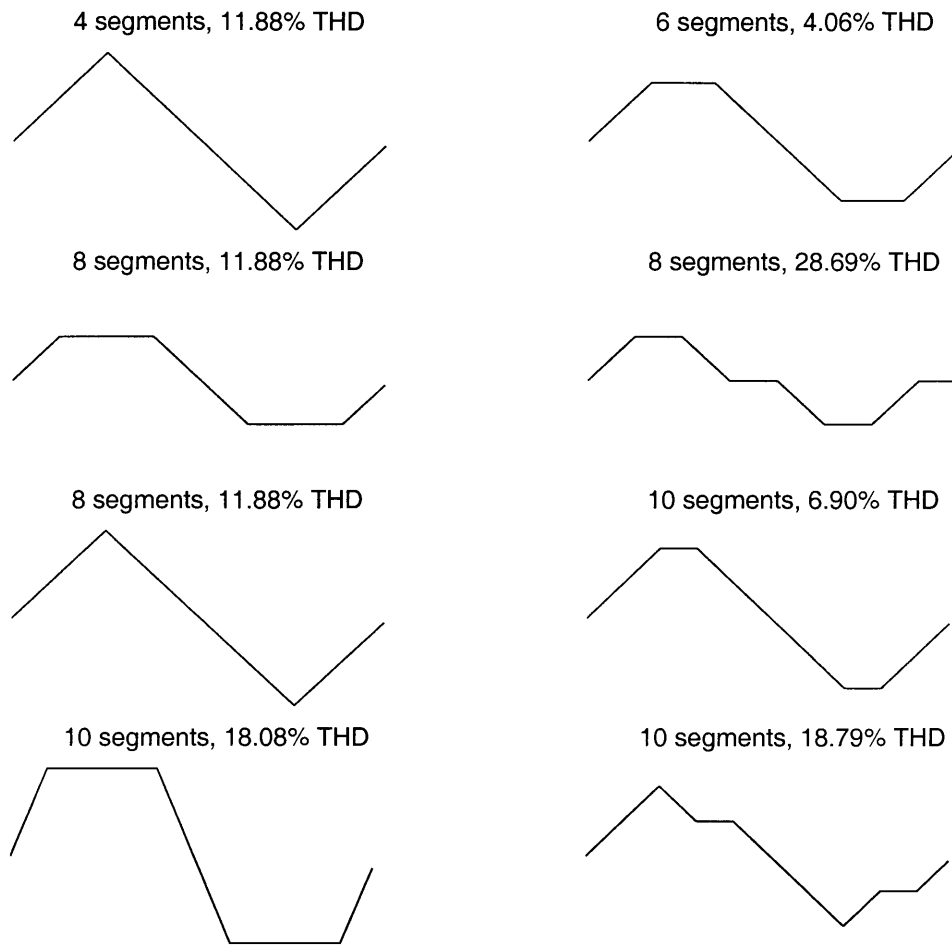
Figure 3-7: THD for N-Segment Piecewise Linear Sine Waves

center frequency. This is discussed in Section 4.5, and illustrated briefly here in Figure 3-8. As you can see from the figure, the modulating frequencies aren't at multiples of the fundamental frequency, suggesting a straight THD evaluation may not be the best measure of performance. At the same time, the THD+N measurement dismisses some of the information about the output waveforms by grouping all nonharmonic energy into one lump noise power n^2.
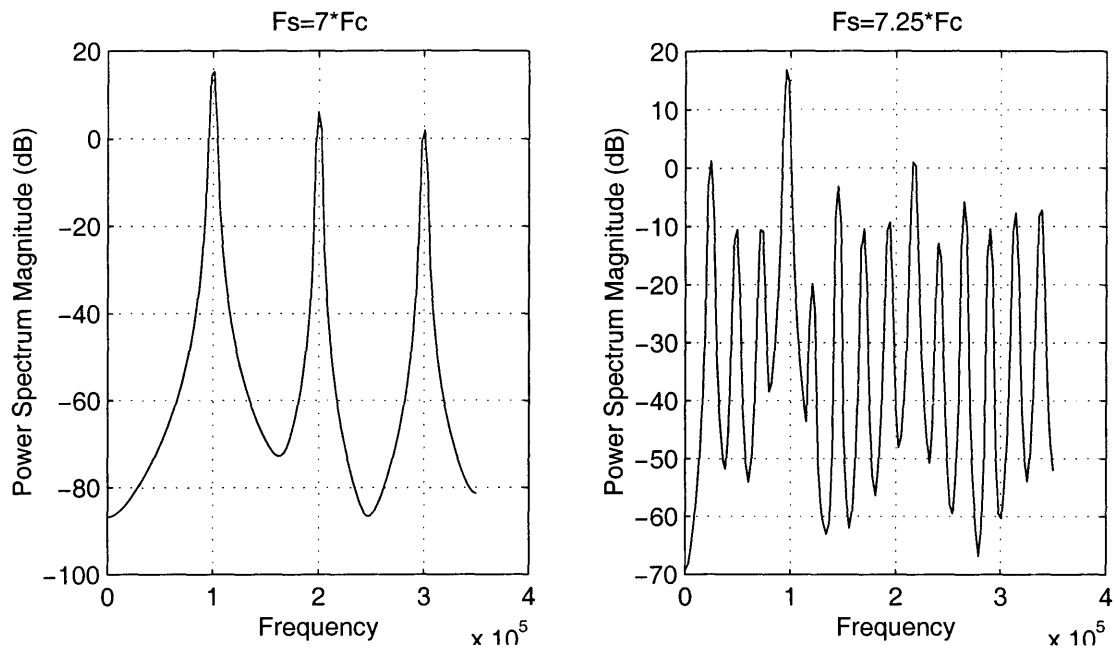
Figure 3-8: Integral and Nonintegral Oversampling of 100kHz

Throughout this thesis, therefore, we will not try to "sum up" the time or frequency domain of an amplifier output with a single numeric value. Instead, the frequency power spectrums of the output, as well as time-domain snapshots, will be analyzed to show a more complete picture of what the amplifier is doing. This way we can tell, for example, which harmonics are contributing the most towards harmonic distortion, or what the spectral shape of the baseband noise is.

Chapter 4

Discrete Time

CD players, communication systems, satellite video broadcasts – they all transmit analog signals as digital values sampled at fixed intervals of time. The channels used to process these signals – internet packets, compact disc optics, etc. – are inherently discrete-time entities. A switching amplifier connected directly to its load, on the other hand, has no implicitly-defined discrete time nature associated with it. Motivation to use DT methods for controlling the switching amplifier comes from the great wealth and simplicity of existing DT mathematics. Although we intuitively compromise something by doing this, it will be shown that for certain conditions digital control is a very practical approach.

The device used in all of these systems to achieve this representation is an analog-to-digital converter (ADC). In this chapter, the switching converter is treated as a high-speed ADC that converts an input discrete-time signal $\{x[n]\}$ to a sequence of quantized output values $\{y[n]\}$. The error between these, $e[n] = x[n] - y[n]$, is our primary concern; we measure the performance of the converter by the signal-to-noise ratio $SNR = \sigma_x^2/\sigma_y^2$.

4.1 Slope Quantization

The topology of our converter accommodates any number of fixed-current sources to be connected positively or negatively into the load capacitance. For the purposes of

this thesis, one current source is used, i.e. $C\,dV_c/dt \in \{-I_1, 0, I_1\}$. This situation is similar to 1-bit D/A conversion, which ensures linearity of the output waveform. The "midtread" 3 level quantizer is shown in Figure 4-1, which is equivalent to 1.6 bits of resolution. It has been shown in oversampling Sigma-Delta ADCs that using "tri-level" quantization rather than the usual 1-bit method improves noise on the output by up to 15dB [PBSA87].
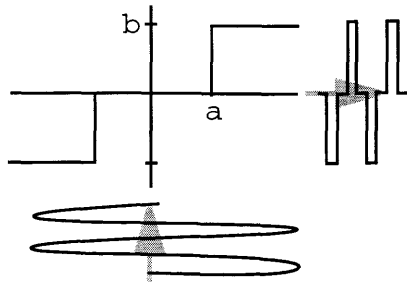
Figure 4-1: 3-level Quantizer

Quantization noise for L levels of quantization is defined [Jay84] as:

$$\sigma_q^2 = \sum_{k=1}^{L} \int_{x_k}^{x_{k+1}} (x - y_k)^2 p_x(x) dx$$

where x_k is the boundary of the kth quantization interval, y_k is the quantizer output for that interval, and $p_x(x)$ is the probability that the input falls in the interval. Inputs to the quantizer that occur outside of the intervals $(x_k, x_{k+1}), k \in [0, L]$ are said to "overload" the quantizer. The enery of the error signal due to overloaded quantizations is called "overload noise", whereas the typical quantization noise is called "granularity noise". The latter noise is typically more high-pass in nature, while overload noise is usually more low-pass, and therefore perceptually less annoying [Jay84].

In this thesis we deal with a 3-level midtread quantizer feeding into an integrator. When this is controlled at discrete time intervals, the output is therefore uniformly quantized as well. The number of effective output levels depends on the ratio of

29

switcing frequency f_s to input frequency f_o according to

$$L = \left\lceil \frac{T_o/2}{T_s} \right\rceil = \left\lceil \frac{f_s}{2f_o} \right\rceil \qquad \rightarrow \qquad B(\text{bits}) = \log_2 \lceil (f_s/2f_o) \rceil$$

We need to choose the quantization parameters $\{a, b\}$ of Figure 4-1 to optimize the quantization noise that appears on the output. The quantization noise of a signal $v[n]$ is defined by $q[n] = v[n] - Q\{v[n]\}$; the quantizer performance is described by $\epsilon \ q = \sigma_q^2/\sigma_x^2$. From the system we describable later, given in Figure 4-7, we find the appearance of quantization noise at the output

$$
\begin{aligned}
y[n] &= y[n-1] + \hat{e}[n-1] \\
&= y[n-1] + (e[n-1] + q[n-1]) \\
&= y[n-1] + (x[n-1] - y[n-1] + q[n-1]) \\
&= x[n-1] + q[n-1]
\end{aligned}
\tag{4.1}
$$

$$\tag{4.2}$$

Therefore the quantization noise sequence $q[n]$ appears directly on the output. The optimization of this error is discussed in Section 4.3.

4.2 Pulse Code Modulation

A large number digital systems we are familiar with use pulse-code modulation as a vehicle for representing analog signals with digital streams. We designate the sequence $x[n] = x_c(nT)$ with its ideal reconstruction as $x(t) = \sum_{-\infty}^{\infty} x[n] h_T(t - nT)$ where $h_T(t) = \text{sinc}(t\pi/T)$. For straightforward PCM, the SNR of the output is simply the quantization noise of the converter, which for a 3-bit converter would be pretty gruesome. Luckily this is not our case; we are quantizing current into a capacitor and our output is its voltage. That is, we are quantizing the differences in the input stream.

4.3 Differential PCM

Traditionally DPCM is a technique used to exploit information we have about the input signal to improve SNR of an analog-to-digital conversion using the same number of quantization intervals. Figure 4-2 shows the basic DPCM topology, in which $H_p(z)$ is used to predict the next value of $x[n]$, call it $\hat{x}[n]$, based on the previous outputs $\{y[n-1], y[n-2], \ldots\}$.
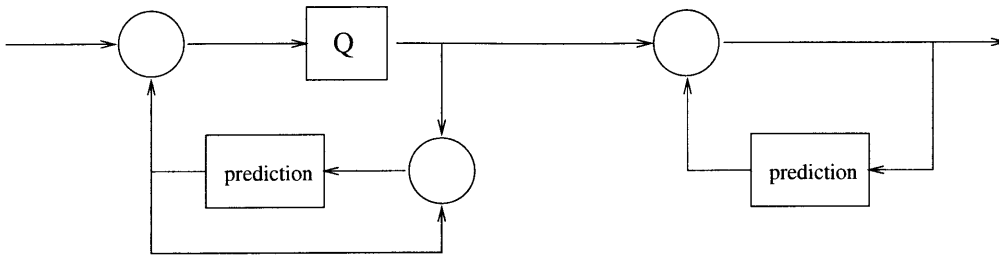
Figure 4-2: DPCM Algorithm for ADC

We call the prediction error $d[n] \equiv x[n] - \hat{x}[n]$; the quantity defined as $G_p \equiv \sigma_x^2/\sigma_d^2$ is called *prediction gain* which shows the SNR improvement in DPCM over PCM according to $SNR|_{DPCM} = SNR|_{PCM} + 10\log(G_p)$. The prediction error is bounded by the inverse of γ_x^2, which is a measure of "spectral flatness"[1] of the input. Therefore, for a white noise input, $\gamma_{white}^2 = 1$, and the SNR of DPCM is the same as regular PCM. [Jay84].

Our prediction filter is $h[z] = z^{-1}$ – a delay. For highly oversampled signals this make sense, as we dont expect the input to have changed that much during each interval. More importantly, however, is that the reconstruction of the DPCM pulses uses the same prediction filter (H_{pred2} in Figure 4-2). Because a capacitor fed sequence of current pulses has a DT transfer function $z^{-1}/(1 - z^{-1})$, it inherently outputs a voltage based on incrementing the previous output.

The input to the quantizer is $e[n] = x[n] - x[n-1]$. For a sinusoid input $x[n] =$

---

[1]spectral flatness is defined as $\gamma_x^2 = \dfrac{\exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln S_{xx}(e^{jw})dw\right]}{\frac{1}{2\pi}\int_{-\pi}^{\pi}S_{xx}(e^{jw})dw}$ in [Jay84]

$x_c(nT_s) = A\sin(2\pi f_o nT_s)$, we are quantizing the differences

$$
\begin{aligned}
e[n] &= A\sin(2\pi f_o nT_s) - A\sin(2\pi f_o(n-1)T_s) \\
&= A\sin(2\pi f_o nT_s) - A\left[\sin(2\pi f_o nT_s)\cos(2\pi f_o T_s) - \sin(2\pi f_o T_s)\cos(2\pi f_o nT_s)\right] \\
&\approx 2\pi f_o T_s \cos(2\pi f_o nT_s)
\end{aligned}
\tag{4.3}
$$

The factor $f_o T_s$ in Equation 4.3 just derived demonstrates how a slower f_o or faster sampling frequency (lower T_s) lowers the energy at the quantizer input, and therefore the quantization SNR that appears on the output. This is because f_s/f_o determines how closely correlated the adjacent input samples are, which is the same thing as saying that the prediction filter z^{-1} becomes more accurate.

From Equation 4.2 we find that the slope quantization noise appears directly at the output, so we can optimize the 3-level quantizer simply for the input defined above [2].

Generally our sampling frequency is never an exact multiple of the input frequency. We can assume that the phase at which we sample an input sinusoid is uniformly distributed across $\theta \in [-\pi, \pi]$. For positive values of S, $Prob(S > \sin(\theta)) = 2\sin^{-1}(S)/\pi$. Taking a derivative of the CDF we find the distribution shown in Figure 4-3.

$$
f_{sine}(s) = \frac{d}{ds}P_{s>=\sin\theta}(s) = \begin{cases} \frac{2}{\pi}\frac{1}{\sqrt{1-s^2}} & |s| < 1 \\ 0 & \text{else} \end{cases}
$$

We hope to optimize the quantizer paramaters a, b in Figure 4-1 by minimizing

$$
\sigma_q^2 = 4f_o T_s \left(\int_0^a \frac{s^2\,ds}{\sqrt{1-s^2}} + \int_a^{2\pi f_o T_s} \frac{(s-b)^2\,ds}{\sqrt{1-s^2}} \right)
$$

(Note that this optimization assumes Equation 4.2 holds. For oversampling circuits that filter the noise spectrum, either by low-passing or noise-shaping the out-

---

[2] A similar procedure can be followed for signals other than sinusoids. I'm just particularly fond of sinusoids.
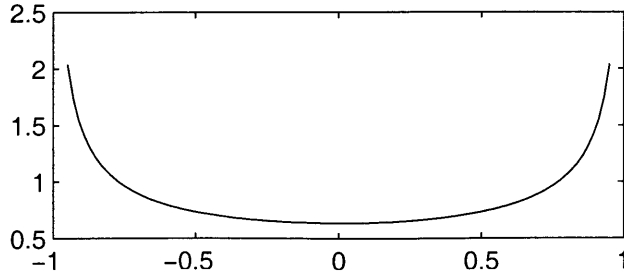
Figure 4-3: Distribution of $\sin(x)$, x uniform $\in [-\pi, \pi]$

put, the SNR-optimal $\{a_{opt}, b_{opt}\}$ will depend on the noise transfer function in effect. [Jay84])

If we begin to include more bandwidth than a single frequency, fixed-amplitude signal, the quantizer input probability density begins to flatten out. This is shown in Figure 4-4.
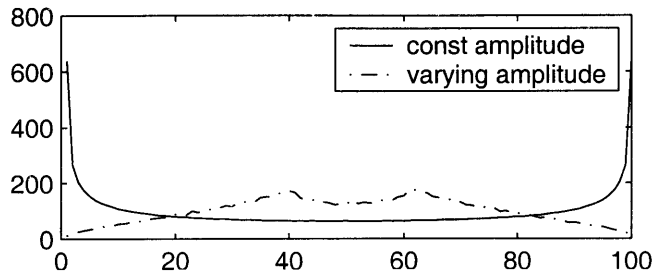
Figure 4-4: Quantizer Input Distribution with Varying Amplitude Sinewaves

As the distribution of inputs to the quantizer begins to flatten out, we can jump at the chance to optimize a quantizer for a uniform distribution. That is, for inputs to the quantizer over the convenient interval $[-1, 1]$, minimize

$$
\begin{aligned}
\sigma_q^2 &= \int_0^a x^2 \, dx + \int_a^1 (x - b)^2 \, dx \\
&= \left. \frac{x^3}{3} \right|_0^a + \left. \left(\frac{x^3}{3} - x^2 b + b^2 x \right) \right|_a^1
\end{aligned}
$$

33

$$= \frac{1}{3} + (a^2 - 1)b + b^2(1 - a)$$

Which boils down to $a = \frac{1}{3}$, and $b = \frac{2}{3}$. For an interval of $[-A, A]$, scale a and b by A. Now before showing the DT feedback loops, it's probably a good time to point out the caveats of not living entirely in a discrete time world.

4.4   DT Processing of CT Signals

In this chapter we talk about the system input $x[n]$, output $y[n]$, and things such as error $e[n] = x[n] - y[n]$. However, the converter output is a CT signal – the error energy $E \neq \cdots + e^2[n] + e^2[n + 1] + \cdots$, rather it's $E = \int |x(t) - y(t)|^2 dt$. Inbetween the DT points, the output ramps at a constant slope. This is illustrated in Figure 4-5.

Figure 4-5: DT and CT View of Converter Output

The piecewise-linear connectivity of the output samples is called a *first order hold*, and has the following transfer function:

$$h(t) = \begin{cases} 1 - |t| & |t| < T_s \\ 0 & else \end{cases}$$

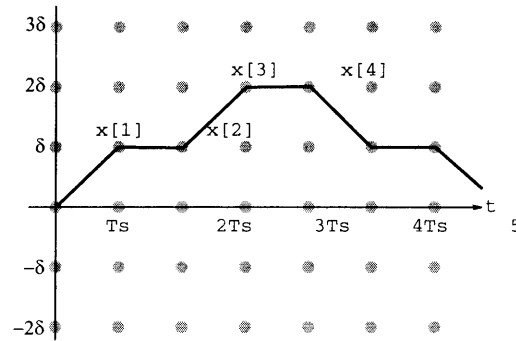The CT output can then be expressed as:

$$y(t) = \sum y[k]h(t - nT_s) = \left(\sum y[k]\delta(t - nT_s) \right) * h(t)$$

34

Transform!

$$
\begin{aligned}
Y(j\omega) &= \int_{-\infty}^{\infty} \sum y[k]\delta(t - kT_s)e^{-j\omega t}dt \cdot H(j\omega) \\
&= \sum y[k]e^{-j\omega kT_s} \cdot H(j\omega) \\
&= Y(e^{-j\Omega})\Big|_{\Omega=\omega T_s} \cdot H(j\omega)
\end{aligned}
$$

As for $H(j\omega)$:

$$
\begin{aligned}
H(j\omega) &= \int_{-\infty}^{\infty} h(t)e^{-j\omega t}dt \\
&= \int_{-T_s}^{0} (1 + t/T_2)e^{-j\omega t}dt + \int_{0}^{T_s} (1 - t/T_s)e^{-j\omega t}dt \\
&= \int_{0}^{T_s} (1 - t/T_s)(e^{-j\omega t} + e^{j\omega t})dt \\
&= \int_{0}^{T_s} (1 - t/T_s)2\cos(\omega t)dt \\
&= \frac{2(\cos(\omega T_s) - 1)}{\omega^2 T_s}
\end{aligned}
$$

It should come as no surprise that

$$
\lim_{\omega \to 0} \frac{2(\cos(\omega T_s) - 1)}{\omega^2 T_s} = \lim_{\omega \to 0} \frac{2\cos(\omega T_s)T_s^2}{2T_s} = T_s
$$

So we can see that for slow signals ($\omega T_s < 1$) the continuous-time output spectrum is simply a scaled version of the DT spectrum that we are mathematically optimizing. For higher frequencies, this spectrum will get a little colored by the filter $H(j\omega)$, but typically the signal frequencies we are concerned with aren't affected by this filtering.

4.5   Modulations

One of the "nasties" that results from creating a fixed-frequency sinewave using a fixed-frequency switching speed is a modulation of the output frequency. This is a byproduct similar to pattern noise, which is a non-white appearance of quantization

noise that frustrates many DAC's. We'll analyze the situation of tracking a frequency $f_c = T_c^{-1}$ using switches operated at $f_s = T_s^{-1}$.

For the signal $\sin(2\pi f_c t)$, the periods recur at $t = nT_s, n = 1, 2, \ldots$. The first period of the output sinewave is the integral of n_1 fixed-length outputs, such that $n_1 T_s < T_c < (n_1 + 1)T_s$. The output waveform starts it's next period early by $\delta T = \delta T = T_c - n_1 T_s$ where $n_1 = \lfloor T_c/T_s \rfloor$. Each subsequent output period is $n_1 T_s$ long an additional δT short until $k \delta T >= T_c$. The periods begin at:

$$
n_1 = \left\lfloor \frac{T_c}{T_s} \right\rfloor, \; n_2 = \left\lfloor \frac{2T_c}{T_s} \right\rfloor, \; \ldots, \; n_k = \left\lfloor \frac{kT_c}{T_s} \right\rfloor
$$

At n_k there have been k periods output by the converter, of which $\lfloor k\delta/T_c \rfloor$ were $n_1 T_s$ long, and $k - \lfloor k\delta/T_c \rfloor$ were $(n_1 + 1)T_s$ long. This is illustrated in Figure 4-6.
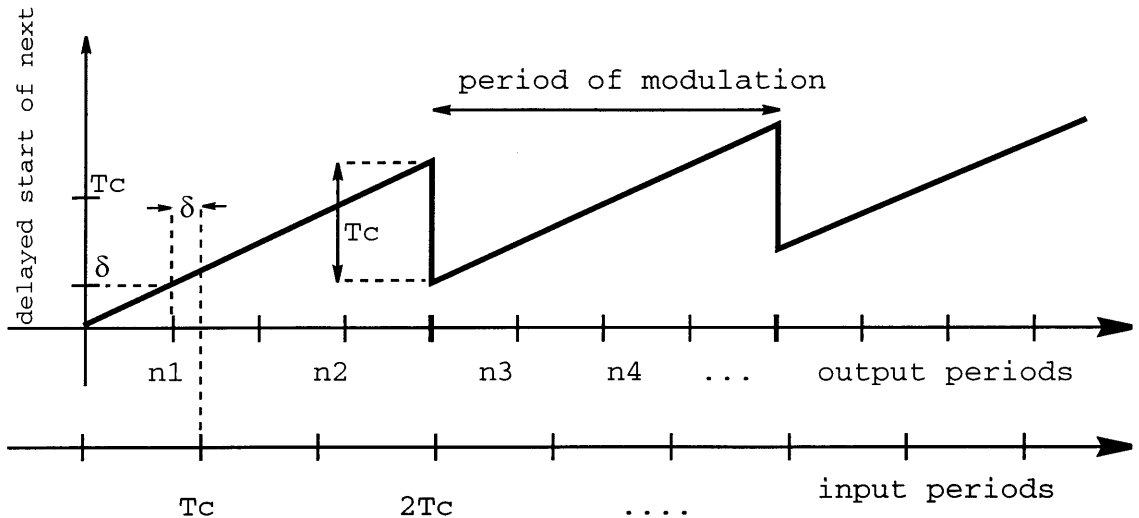

Figure 4-6: Modulation of Carrier on DT Output

4.6 First Order Loop

The term *Deltamodulation* is typicaly used to describe 1-bit DPCM. We're using a little over 1.5 bits, but the idea is the same – oversample by as large a ratio as

practical, and filter out the upper frequencies. The DPCM loop of Figure 4-2 is redrawn with the prediction filter $H_1(z) = z^{-1}$ in Figure 4-7.
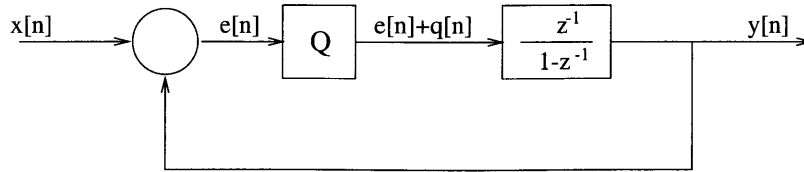
Figure 4-7: First Order Deltamodulator

The transfer functions from the input sequence $x[n]$ and the quantization noise sequence $q[n]$ to the output $y[n]$ are given by

$$\frac{Y_1(z)}{X(z)} = \frac{Y_1(z)}{Q(z)} = \frac{\frac{z^{-1}}{1-z^{-1}}}{1 + \frac{z^{-1}}{1-z^{-1}}} = z^{-1}$$

Fortunately noise injected at the output of the integrator due to nonidealities in its implementation has the familiar transfer function $1 - z^{-1}$. We all recognize this in continuous time as

$$H(e^{jw}) = (1 - z^{-1})|_{z=e^{jw}} = e^{-jw/2} \left(e^{jw/2} - e^{-jw/2} \right) = 2je^{-jw/2}\sin(w/2)$$

Yes, a zero at $z = 1$ filters noise out from the power density spectrum around the lower frequencies according to $|H(e^{jw})|^2 = 4\sin^2(w/2)$ [OS99].

4.7 Second Order Loop

The problem with the first-order loop is that quantization noise is still unfiltered, and with low oversampling ratios (see Equation 4.3) it will make for a shitty amplifier. The second-order loop in Figure 4-8 achieves the transfer function

$$\frac{Y_2(z)}{Q(z)} = 1 - z^{-1}$$

37

This system is a bona-fide member of the oversampled noise-shaping (ONS) ADC family.

The transfer function from the input to output is a two-cycle delay $y_2[n] = x[n-2]$. More interestingly, the transfer function from noise introduced by integration error at the output is filtered by $\frac{Y_2(z)}{Q_{int}} = (1 - z^{-1})^2$. This is a second-order noise shaping filter. Keep in mind that noise-shaping filters do not eliminate noise energy; in fact, they add to it, but they move it out of the spectrum of interest, which in our case means it is moved into the upper frequencies.
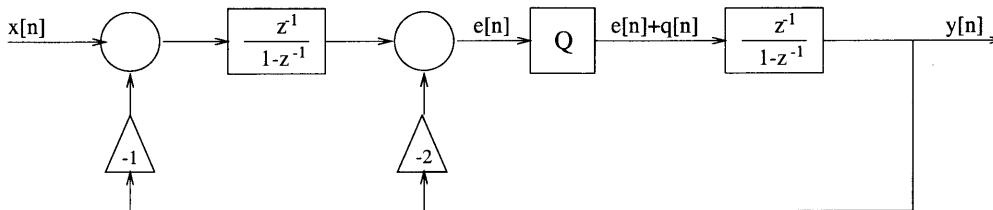
Figure 4-8: Second Order Deltamodulator, with Noiseshaping

Higher and higher orders of noiseshaping become double-edged swords. In typical A/D converters, for example, any 3rd-order or higher noise-shaping modulator is inherently unstable [Rab99]. In all cases, the converter must be able to generate the output sequence and filter out the noise energy in subsequent steps. Remember that the output is slope-limited, which leads maximum product of output frequency with amplitude as given in Equation 3.1. Putting too much noise energy into the upper frequencies can overload the converter and worsen the filtered SNR instead of improve it.

4.8 DT Modulator Simulations

A C++ program was written to directly simulate the block diagrams given for the first and second-order deltamodulators. The DT system in Figure 4-7 was simulated for varying input amplitudes and OSRs.

Figures 4-10 and 4-9 show the output spectrum of the converter as it varies with input amplitude. The oversampling ratios of 20 and 10 on a 1MHz switching frequency would correspond to input sinusoids of 25kHz and 50kHz respectively. In both figures we can see only even harmonics, corresponding to the even rate-limiting effects of the slope quantization. For the OSR of 10 (Figure 4-9), the THD is approximately 10% throughout the gain range except at the input amplitudes 2.3 and 4.5. This is just due to coincidental alignment of the piecewise-linear segments of the output wave with the input waveform. When we move to an OSR of 20 (Figure 4-10) the THD drops to 8.67%.
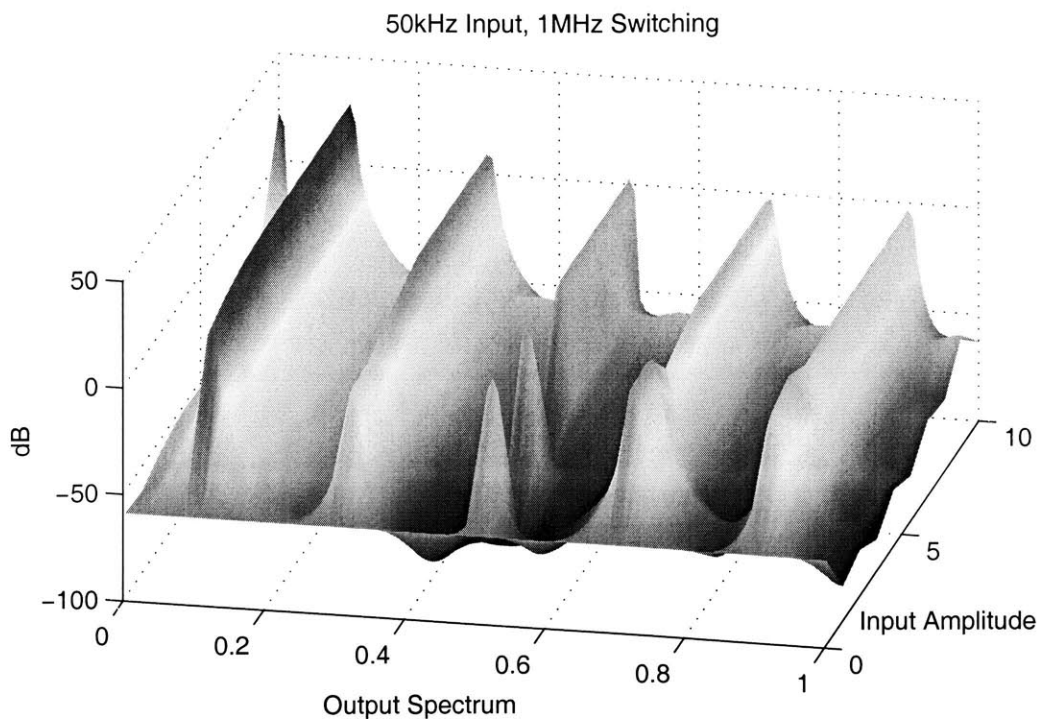
Figure 4-9: Output Spectrum vs. Amplitude, OSR=10, 1st Order

Keeping with our arbitrary switching frequency of 1MHz, the spectrum of the output as the input is swept across the frequency range $f \in [25kHz, 75kHz]$ is shown in Figure 4-11.

The second order loop was simulated over the same ranges of input amplitudes

25kHz Input, 1MHz Switching
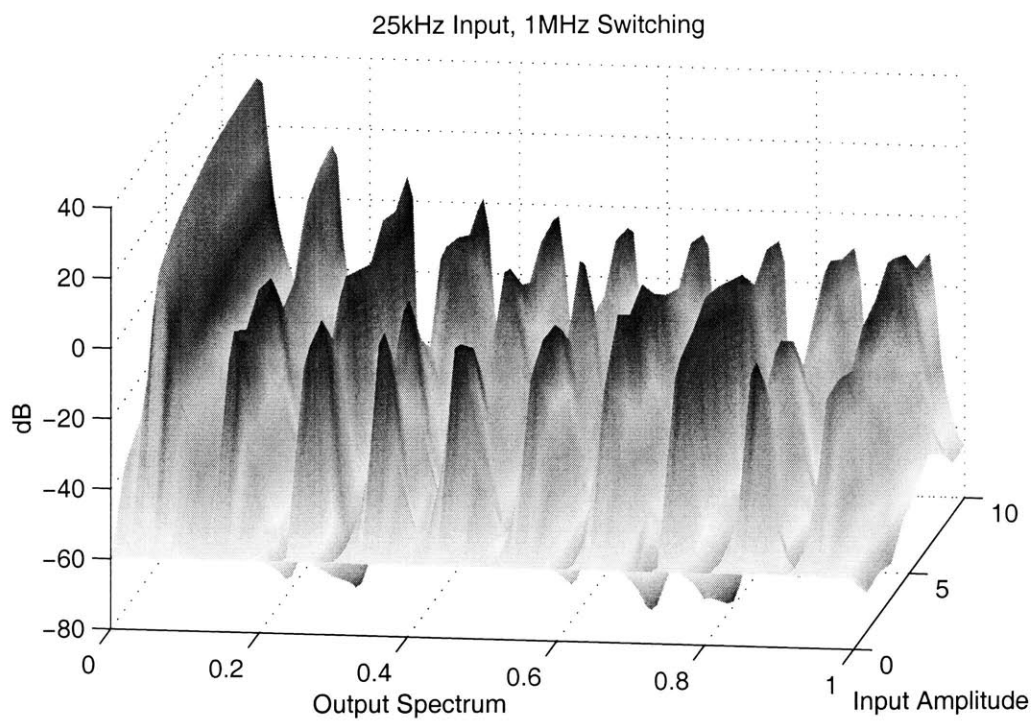
Figure 4-10: Output Spectrum vs. Amplitude, OSR=20, 1st Order
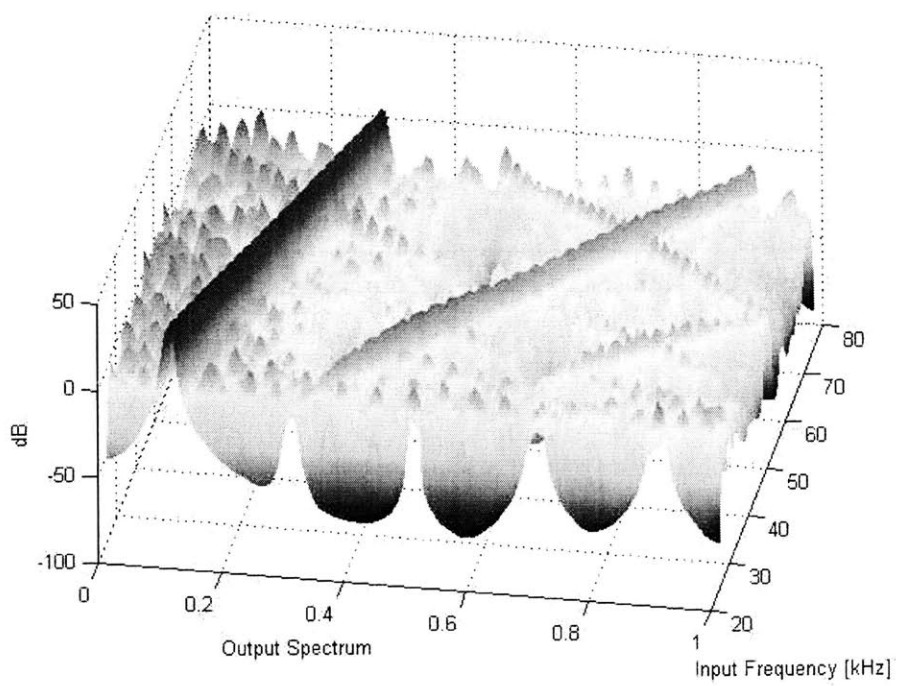
Figure 4-11: Output Spectrum vs. Frequency, First Order

and frequencies, shown in Figures 4-12 - 4-15.

Figure 4-12 shows how the second order filter works. High-frequency oscillations occur on the output that average out to closer approximations of the input. The error from this occurs as out-of-band noise in the upper spectrum, thus achieving the "noiseshaping" goal. The second harmonic is about $33dB$ below the fundamental, unlike the first-order converter which was only down by $20dB$.
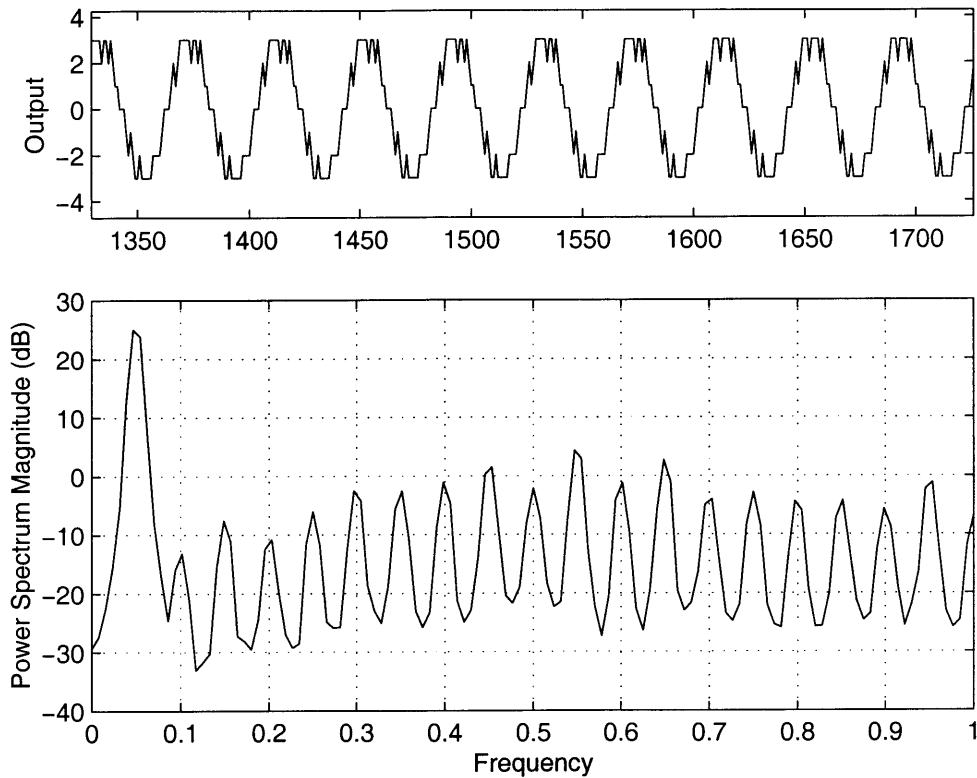
Figure 4-12: Output Waveform & Spectrum, OSR=20, 2nd Order

The output for an OSR of 20 (Figure 4-14) has a THD of about 7%, while an OSR of 10 (Figure 4-13) has an average THD of 18%. This is clearly unacceptible for any application requiring low distortion. One possible reason is that the high-frequency corrections made by the noiseshaping function move too much energy into the upper frequencies, thereby saturating the slope quantizer.

Finally, a plot of the output spectrums for a constant input amplitude while the

42
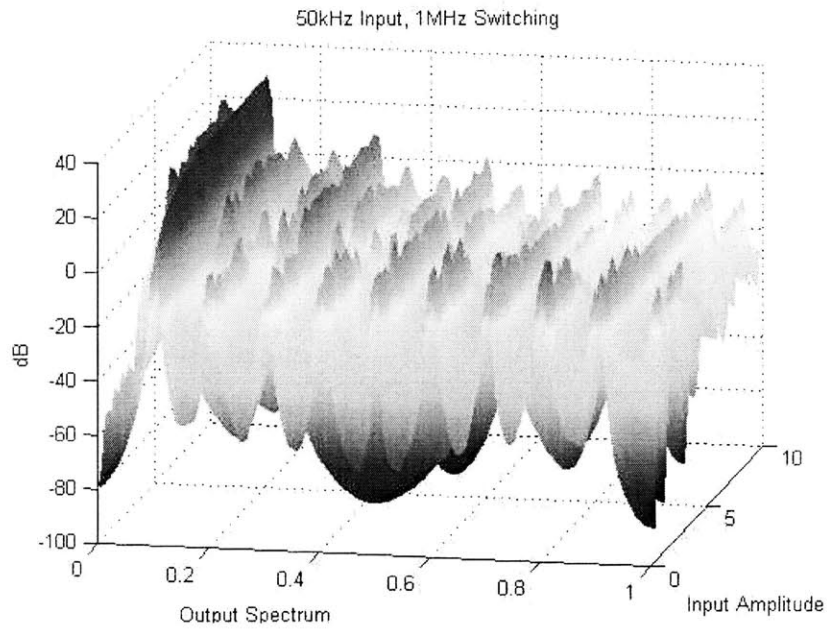
Figure 4-13: Output Spectrum vs. Amplitude, OSR=10, 2nd Order
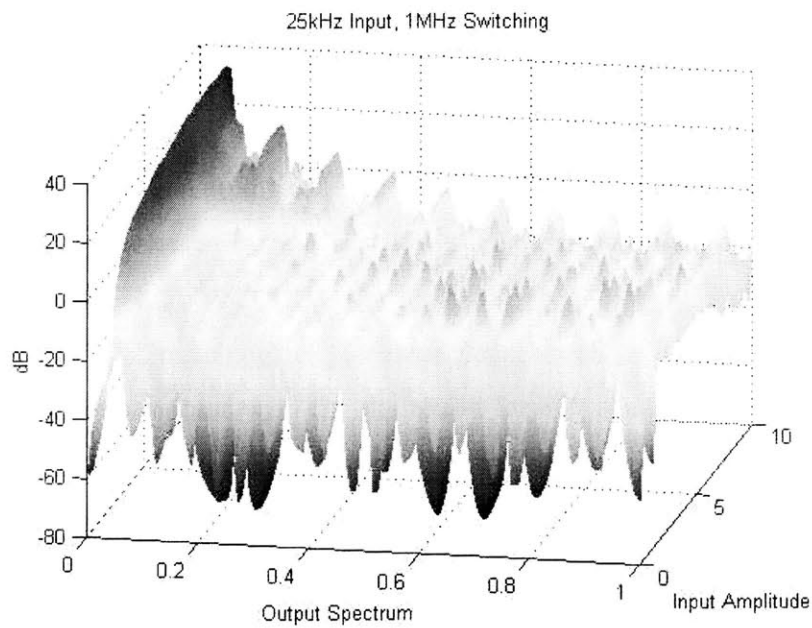
Figure 4-14: Output Spectrum vs. Amplitude, OSR=20, 2nd Order

frequencies are swept from 25kHz (OSR=40) to 75kHz (OSR=13). One can generally observe that the broadband noise is altogether too high in relation to the fundamental.
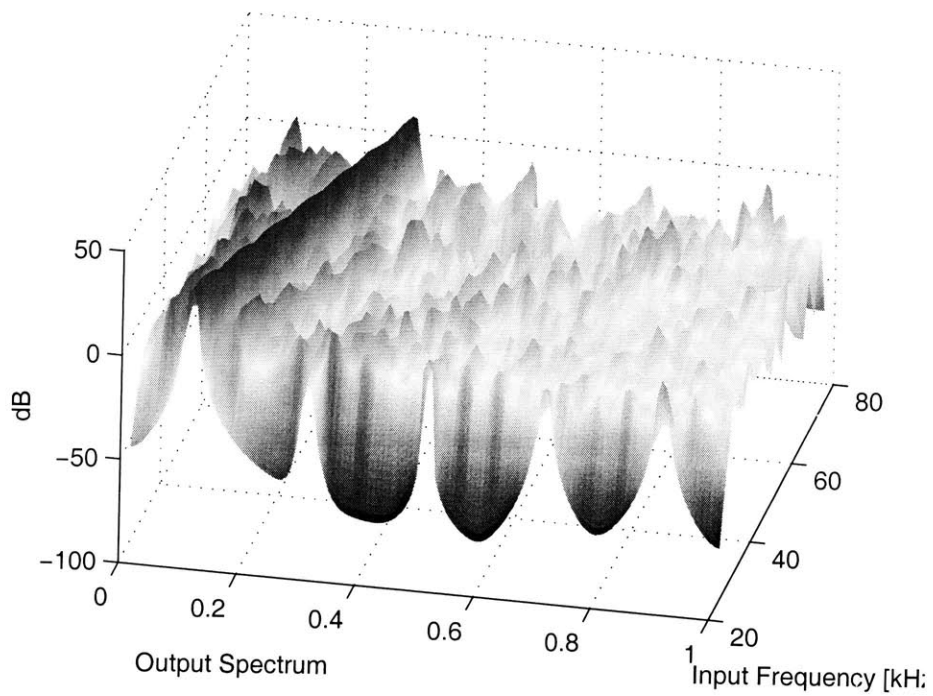
Figure 4-15: Output Spectrum vs. Frequency, 2nd Order

Chapter 5

Pulse-Width Modulation

When someone says "switching amplifier", the first thing we think of is pulse-width modulation (PWM). This means we're on T_{on}, off T_{off}; our control is the average value of the input $d(t) = \int_{t-T}^{t} input(t)dt = T_{on}/(T_{on} + T_{off})$. Because the feedback loop controls averaged quantities, switching speeds are assumed to be about an order of magnitude faster than the input signal.

It is not entirely realistic to expect this averaging approach to be the best control scheme for high-frequency input signals. For example, National's LM2622 1.3MHz switching computer power supply controller recommends the feedback pole to be set around 3-8kHz [Nat]. The PWM algorithm described in this section is used to create sinewaves at oversampling ratios of 100, 40, 20 and 13.3. The integrated PWM output, a filtered version of it, and it's power spectrum for each oversampling ratio is shown in Figure 5-1.

There are many ways to control piecewise constant-slope integration to achieve a local time-average consistant with the definition for $d(t)$ above. These are shown below along with a *dynamic range*$= d_{max}/d_{min}$ measure for each scheme if we reasonably impose a T_{min} and T_{max} on the system:

Figure 5-1: PWM-Generated Sinewaves for Different OSRs

fixed variable	T_{on}	T_{off}	$T = T_{on} + T_{off}$	dynamic range
$frequency$	dT	$(1-d)T$	T	$\frac{T_{max}-T_{min}}{T_{min}}$
$ontime$	T_{on}	$\frac{1-d}{d}T_{on}$	$d^{-1}T_{on}$	$\frac{T_{max}+T_{on}}{T_{min}+T_{on}}$
$offtime$	$\frac{d}{1-d}T_{off}$	T_{off}	$(1-d)^{-1}T_{off}$	$\frac{T_{max}+T_{on}}{T_{min}+T_{on}}$

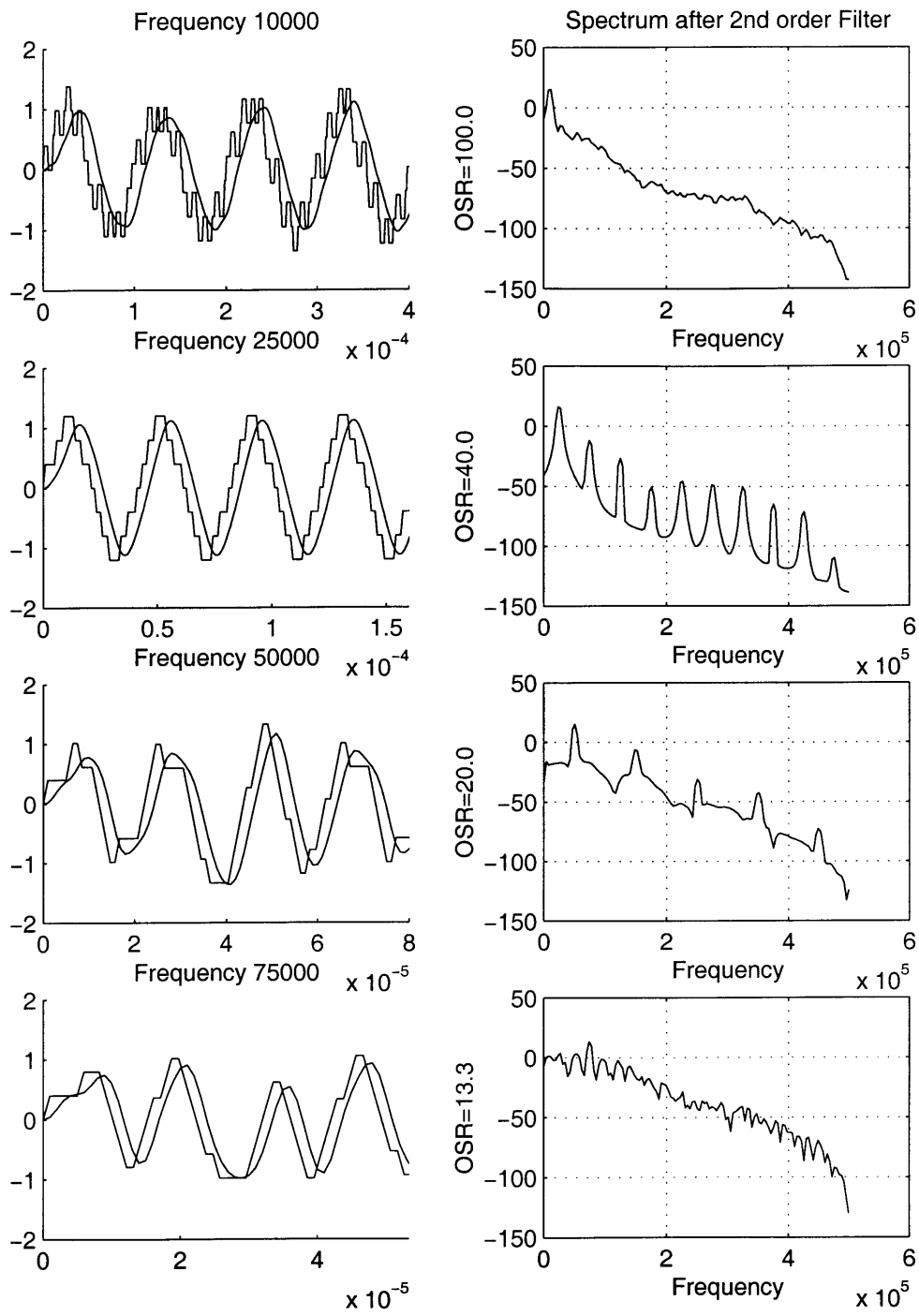The minimum switch position time T_{min} cannot be made infinitely small because the switches have their own state and dynamics. The maximum allowable cycle time T_{max} cannot be infinitely long or the time-averageing notion of the $d(t)$ begins to break down. The problem with the above on/off-time methods is that T_{max} and T_{min} reduce the dynamic range. A more efficient scheme can be defined as follows:

$$T_{on} = \begin{cases} T_{min} & d <= \frac{1}{2} \\ \frac{d}{1-d}T_{min} & d > \frac{1}{2} \end{cases}$$

$$T_{off} = \begin{cases} \frac{1-d}{d}T_{min} & d <= \frac{1}{2} \\ T_{min} & d > \frac{1}{2} \end{cases}$$

(5.1)

The dynamic range of this algorithm is T_{max}/T_{min}, which makes better use of the constraints than the previous methods.

5.1 PWM Implementation

Say $d < \frac{1}{2}$, so that the converter is ramping positive current into the output for $T_{on} = T_{min}$. The converter then holds [zero slope] for $T_{off} = \frac{1-d}{d}T_{min}$. The duration that the current is diverted around the output capacitance (T_{off}) is shown in Figure 5-2 for duty cycles $d \in [0.1, 0.5]$.

Fortunately this curve looks a lot like an exponential decay, which allows for a very slick method of implementing this control scheme. Just as fixed-frequency PWM controllers determine on time $T_{on} = d \cdot T$ with the intersection of a sawtooth waveform and the d signal (see Figure 5-3 left side), the max-DR algorithm above
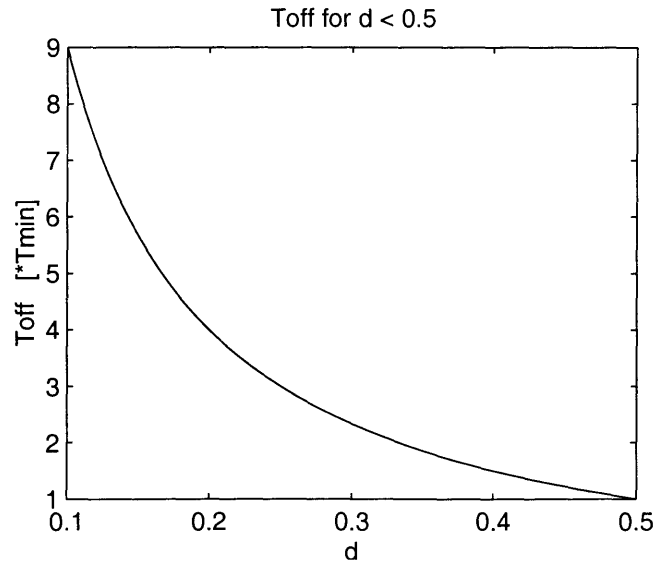
Figure 5-2: Off-time for $0 < d < \frac{1}{2}$

can be implemented with the intersection of an exponential decay and d (Figure 5-3, right).
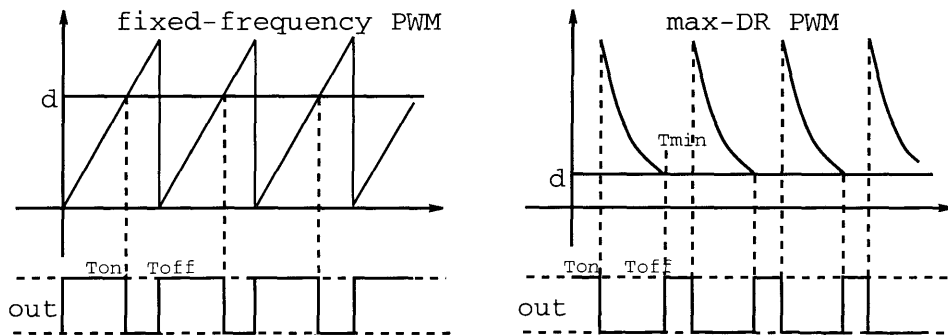
Figure 5-3: Implementation of two PWM Schemes

By finding the intersection of $Ae^{-t/\tau} = d + C$, for constants $\{\tau, C\}$, we get the approximation for the off-time

$$T_{off} = \frac{1-d}{d} T_{on} \approx \tau[\log(A) - \log(d + C)] \qquad (5.2)$$

An easy way to determine the constants is to let $C \equiv 0$ and find the remaining two degrees of freedom by solving for $\{d_1, T_{off,1}\}$ and $\{d_2, T_{off,2}\}$ from Equation 5.1. Thus,

$$\log(A) = \frac{T_{off,2} \log(d_1 + C) - T_{off,1} \log(d_2 + C)}{T_{off,2} - T_{off,1}}$$

For $d > \frac{1}{2}$, the equations for T_{on} are the same with the substitution $d \to (1-d)$. We can improve this approximation of $d/(1-d)$ by iteratively optimizing C for minimum error:

$$C = \arg\min_c \left[\int_{d_{min}}^{\frac{1}{2}} \left(T_{min} \frac{1-d}{d} - \tau \log\left(\frac{A}{d+C}\right) \right)^2 \right]$$

For $d_{min} = 0.2$ and $T_{min} = 1\mu s$, the values for $\{A, \tau, C\}$ used in the simulations are $\{0.6542, 1.67\mu s, -0.14\}$. The approximation is shown in Figure 5-4.
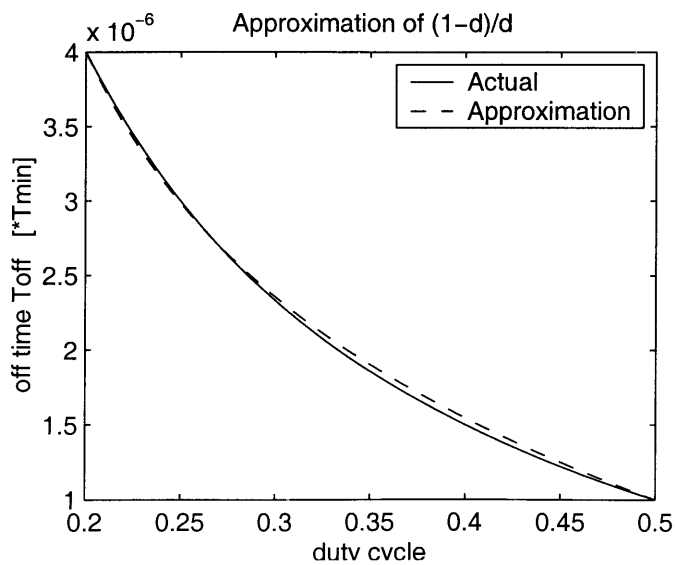
Figure 5-4: Approximation error for $(1-d)/d$

5.2 PWM Simulations

The algorithm for PCM signals was simulated on a computer for varying amplitudes and input frequencies. The system in simulation imposes a minimum duty cycle ratio of $D_{min} = 0.2$, which corresponds to a minimum switch on/off-time $T_{min} = 1/mus$. Figure 5-5 shows in the upper plot the output for a 50kHz signal synthesized using the actual timings given by Equation 5.1, and on the lower plot the timings from the approximation in Equation 5.2. The "mistakes" from the approximation appear as the infrequent bumps shown in the lower plot.
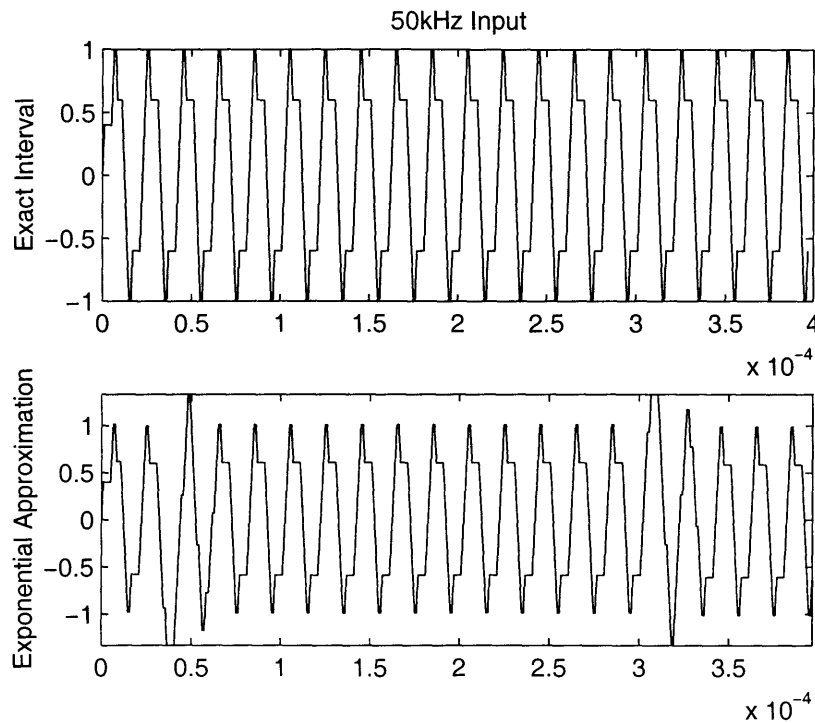
Figure 5-5: Output From Actual vs. Approximate Duty Cycle Algorithms

The output spectrums over the range of 25kHz - 75kHz for a fixed amplitude is shown in Figure 5-6. One noticeable characteristic is that the harmonics aren't as prominent as they frequency sweeps of the DPCM algorithm (compare with Figure 4-11).

In Figure 5-7 we hold the frequency fixed at 25kHz while sweeping input amplitude
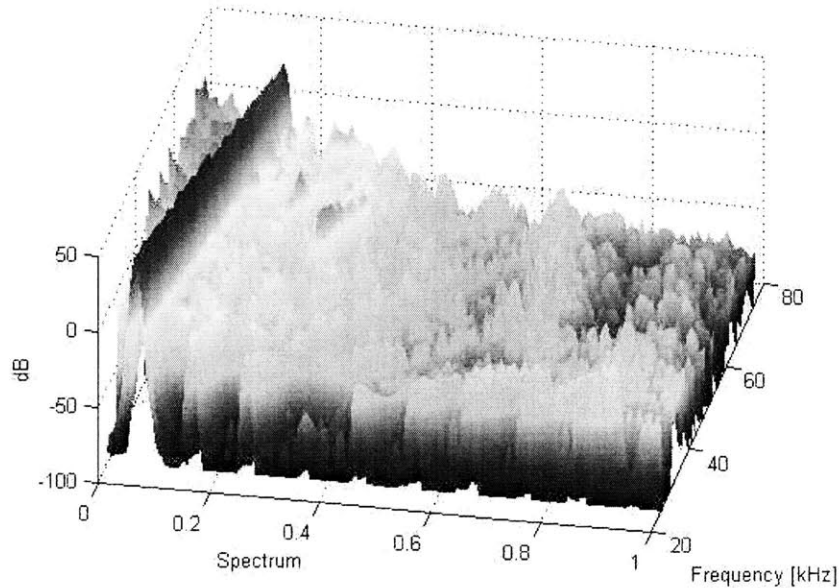
Figure 5-6: PWM Output Power Spectrum vs. 25kHz-75kHz

and we plot the spectrums vs. output amplitude. The algorithm works really well at larger oversampling ratios; the second harmonic is on average $30dB$ down from the first, with a THD of around 3%. In Figure 5-8 the 50kHz input causes a dramatic rise in THD above a certain amplitude. This is because the current source becomes saturated and the output starts to look like a triangle wave.

5.2.1 SPICE Simulation

The PWM algorithm described above was simulated in SPICE using the model shown in Figure 5-9. A 50kHz, 8V sine wave is recreated by generating a sequency of PWM pulses from the absolute value of the input and integrating them positively or negatively depending on the sign of the input. For the first half of the period, R_1 and C_1 decay from 10V until they cross the input value, which is signalled by DOWN COMPARATOR. At this point C_1 is reset while C_2 in parallel with R_2 begin to decay from zero volts upwards towards towards 10 until they cross the input. Thus, the PWM signal is generated.
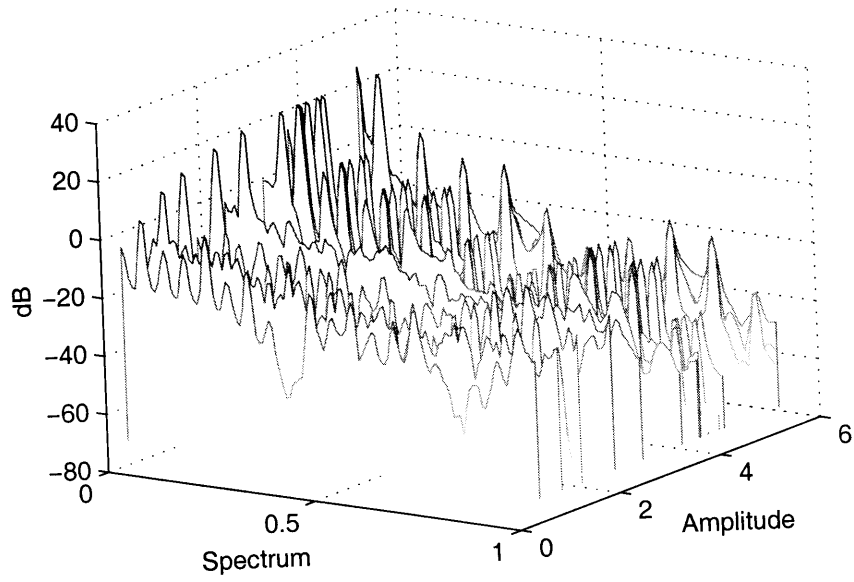
Figure 5-7: PWM Output Power Spectrum vs. Amplitude @25kHz
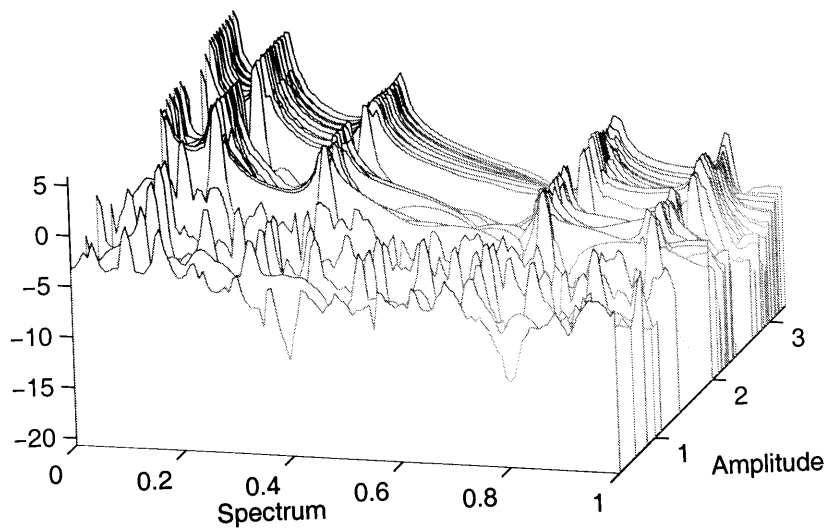
Figure 5-8: PWM Output Power Spectrum vs. Amplitude @50kHz
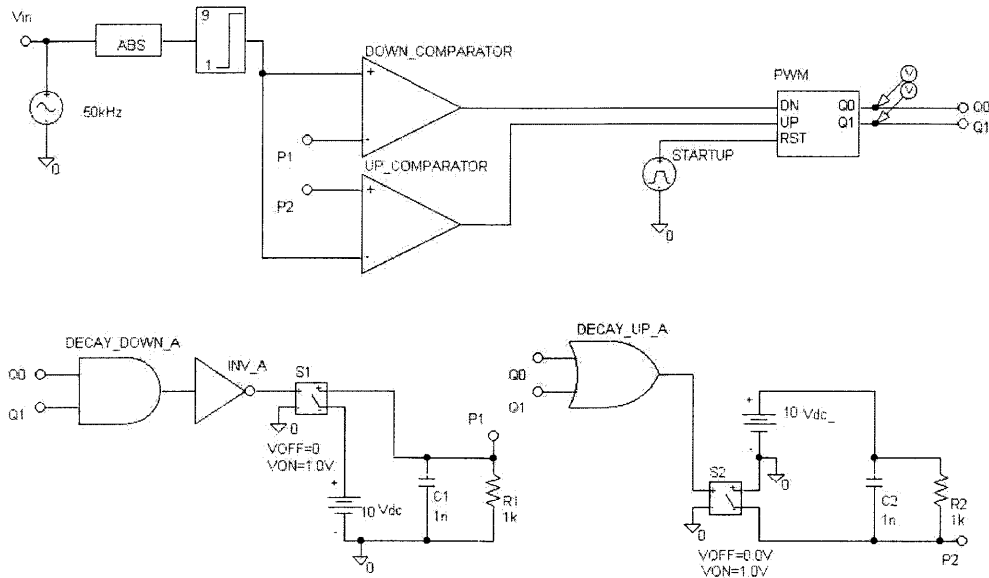
Figure 5-9: SPICE Exponential-Decay PWM Implementation

The control logic is shown in Figure 5-10. The state variable Q_0 determines which half of the duty cycle is being determined, and Q_1 is an auxiliary variable used to hazard-correct the wait condition from the reset action under asynchronous operation. The signals UP and DOWN are generated when the respective decaying RC pair intersects the input level. The resulting state equations are

$$Q_0 = Q_0 \bar{Q}_1 + \bar{Q}_1 UP + Q_0 UP + Q_0 \overline{DOWN}$$

$$Q_1 = Q_0 Q_1 + Q_0 \overline{DOWN}$$

The OrCAD Spice Student Edition reached it's circuit maximum complexity limit, so the output voltage was integrated in the waveform viewing window as shown in Figure 5-11.
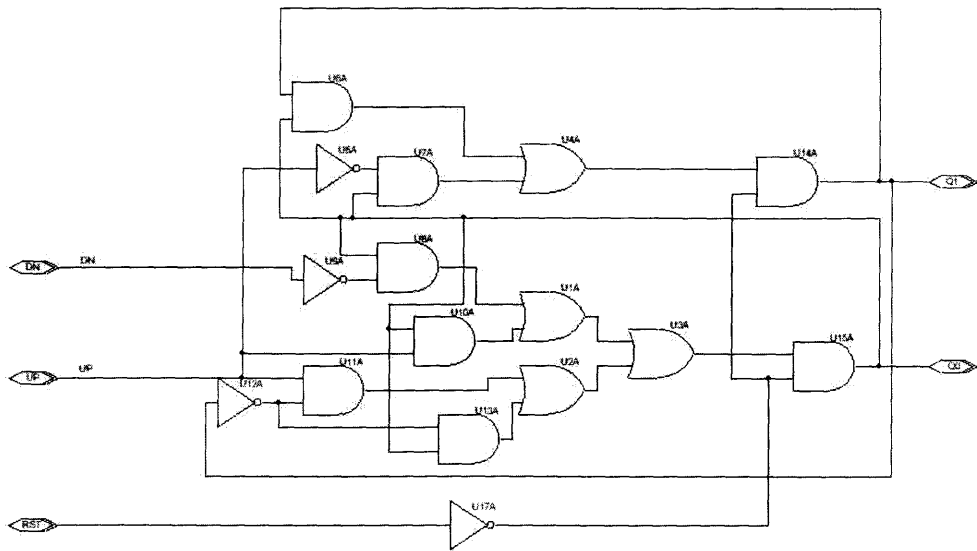
Figure 5-10: SPICE PWM Control Logic
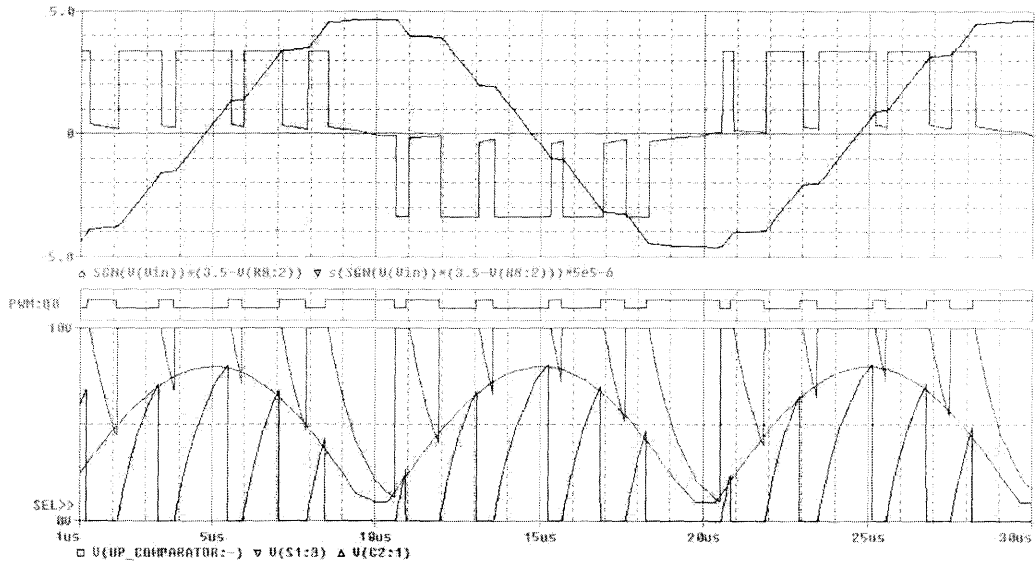
Figure 5-11: SPICE Simulation Output

54

Chapter 6

Continuous-Time Math

Operating our integrating switching converter as a discrete-time device may simplify the math, but it imposes unnecessary constraints upon the system – constraints that get in the way of achieving optimal performance. This becomes especially problematic as the output frequency gets closer to the switching frequency; we've seen the DPCM and PWM algorithms suffer from unacceptible noise and distortion as the OSR was increased enough. In order to derive a control scheme that performs at the device's inherent limitations, we must view *when* and *what* we do with the switches as a sequence of decisions. Subjecting ourself to certain limitations of the switches, and keeping in mind measures of output performance, we find an optimal control problem sitting on our doorstep.

The name of the game is "optimization"; however, we cannot simply look to the well-developed theory of optimal linear feedback systems. Our system ramps steadily up, down or holds its mark. Fortunately, these characteristics are very similar to systems with saturating controls, such as airplanes and space shuttles. Most notably during the second world war, the enviable advantage of stable aircraft led to enthusiastic development of saturating controller theory. Names such as LaSalle and Pontryagin, among many others, began to appear in the journals in the 50's and 60's, along with the related techniques of the maximum principle and dynamic programming [Rya82].

Before attempting a rigorous approach, it is important to understand the simple

problem in the simple manner illustrated in Figure 6-1. We are standing at time t_0. The input is $r(0)$ and moving with velocity $\dot{r}(t)$. The output is $c(0)$ and currently integrating a constant so that $\dot{c}(0) = u$, where $u(t_0) \in \{-1, 0, 1\}$. We have to stay at any switching position for T_{min}, we are trying to track the input, i.e. minimize $|e(t)| = |r(t) - c(t)|$. We have to choose what the next control $u(t_1)$ will be to do this, as well as when we will switch to it. We then hold this new control until $t_2 > t_1 + T_{min}$ and choose the new $u(t_2)$, and so on.
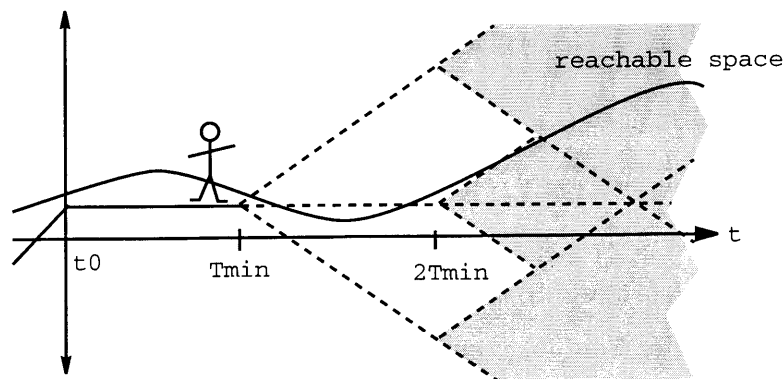
Figure 6-1: CT Decision Sequence

6.1   Optimal Control

We have on our hands an optimal control problem, which has the following essential elements [AF66]

- a dynamical system (i.e. $\dot{\mathbf{x}}(t) = f[\mathbf{x}(t), u, t]$)

- a desired output r(t)

- an allowable class of controls: $u \in \Omega$, $\Omega = -1, 0, 1$

- a performance functional

Each of these components and the associated approximations are quickly discussed.

6.1.1 Dynamical System

The system stores energy in two components - current I_l in the source inductance and voltage V_X of the output capacitance - which comprise the system's physical state. The relations can be derived from Figure 7-4

$$\frac{dv_C}{dt} = \frac{u}{C} i_L$$

$$\frac{di_L}{dt} = \frac{-1}{L}[v_D + uV_c + i_L R_s]$$

for loop voltage drop v_D and switch resistance R_s. The system equations can be expressing in matrix form, along with the hope $L \to \infty$

$$\lim_{L\to\infty} \dot{\mathbf{x}} = \lim_{L\to\infty} \begin{bmatrix} \dot{v}_C \\ \dot{i}_L \end{bmatrix} = \lim_{L\to\infty} \begin{bmatrix} 0 & u/C \\ -u/L & -R/L \end{bmatrix} \begin{bmatrix} v_C \\ i_L \end{bmatrix} + \begin{bmatrix} 0 \\ -v_D/L \end{bmatrix} = \begin{bmatrix} u/C \, v_C \\ 0 \end{bmatrix}$$

Setting $C \equiv 1$ we get the simplified state dynamics

$$\dot{x} = f[x(t), u(t), t] = u$$

Another set of devices that carry information accrued from the past are the switches. The MOSFETs are field-operated devices; a finite amount of charge required to setup the fields must enter and leave the MOSFETs through the gates at a finite rate determined by the driving circuitry, terminal voltages, and device parasitics. A simple way to sum up this process in a useful quantity is to measure the time after a switch is turned off or on and to impose a minimum time T_{min} between transitions. As the switching frequency increases (T_{min} decreases) in order to track the input signal more rapidly, the switch driving circuits become a significant source of energy dissipation. T_{min} becomes a key parameter that controls the tradeoff between energy dissipation and tracking performance.

6.1.2 Performance Functional

One of the most important parts of any optimization problem – which may sound obvious – is to figure out what to optimize. The classical audio performance measure is low output THD. Unfortunately it is difficult to relate the effect of a single decision to a harmonic analysis that involves thousands of data points. In this case it is most reasonable to minimize the error energy $\int_{-\infty}^{\infty}(r(t) - x(t))^2 dt$, which minimizes SNR for a fixed output signal energy.

The switching constraint can be taken into consideration as additional *cost*, which simplifies the decision logic significantly. Each decision to switch to a new control $u(t_o)$ must be held for a finite period of time, during which

$$r(t) - x(t) = (r(t_o) - x(t_o)) + (\dot{r}(t_o) - u(t_o))t + \frac{1}{2}\ddot{r}(t_o)t^2 + \cdots$$

Thus, the factor $\dot{r}(t_o) - u(t_o)$ appears as a component of the error signal. The amount that the slope error contributes is related to switching speeds; we use the scaling factor λ in the cost function $J(u)$ as follows

$$J(u) = \int_{t_0}^{t_1} (r(t) - x(t))^2 + \lambda(\dot{r}(t) - u)dt \tag{6.1}$$

6.1.3 Derivation of Switching Function

We are now in a position to employ Pontryagin's minimum principle[1], which determines the conditions necessary for a control to be optimal if one exists. First, the Hamiltonian is defined as

$$H = L + <\mathbf{p}, \mathbf{f}> = (x - r)^2 + \lambda(\dot{r} - u)^2 + pu$$

where L is the integrand of the cost function and the adoint variable p must satisfy the differential equation $\dot{p}(t) = -\partial H/\partial x = 2(x - r)$. We'll define $e \equiv r - x$; the optimal control problem can then be found by

---

[1] excellent references for the Minimum Principle are [Rya82, AF66, FL68, Ber95]

$$
\begin{aligned}
u_{opt} &= arg\ min_{u \in \Omega}\ H \\
&= arg\ min_{u \in \Omega}\ [(x-r)^2 + \lambda(\dot{(r)} - u)^2 + 2u(x-r)t] \\
&= arg\ min_{u \in \Omega}\ [e^2 + (\dot{r}^2 - 2\dot{r}u + u^2)\lambda + 2uet] \\
&= arg\ min_{u \in \Omega}\ [-2u(\dot{r}\lambda - et)] \\
&= -sgn(et - \dot{r}\lambda)
\end{aligned}
$$

This is intuitively what we would expect — for a constant input ($\dot{r}(t) = 0$) and positive error we expect $u = -1$, and vice versa. For a changing input, if $e(t) > 0$ and $\dot{r}(t) > 0$ then we expect the error to only become more positive and also choose $u = -1$. However, if $r(t) > 0$ and $\dot{r}(t) < 0$ then we may want either $u = 1$ or $u = -1$ depending on how quickly the error about to reverse direction. The λ parameter must be chosen to optimize the tradeoff between minimum switching times and error energy.

6.1.4 Impact of Switching Function

The controller of this system tracks the function

$$
S(t) = (r - x) + \lambda \dot{r} \tag{6.2}
$$

and determines which control $u \in -1, 0, 1$ to apply by passing $S(t)$ through the relay function shown in Figure 4-1. The upshot of $\lambda \neq 0$ can be reinterpreted as tracking the signal $r'(t) = r(t) + \lambda r(t)$. This has interesting consequences which we will illustrate with the following examples.

Ramp Input $r(t) = R_o + \dot{r}t$: The effect of a nonzero λ adds only an offset to the ramp, resulting in situation already discussed in the overview (Figure 3-4). That is,

$$
r'(t) = R_o + (\lambda + t)\dot{r}
$$

Sine Input $r(t) = \sin(\omega_c t)$: We are now effectively tracking an input scaled and phase-shifted slightly due to $\lambda \neq 0$:

$$r'(t) = r(t) + \lambda \dot{r}(t) = \sqrt{1 + (\lambda \omega_c)^2} \sin\left(\omega_c t + \cos^{-1}\left(\frac{1}{\sqrt{1 + (\lambda \omega_c)^2}}\right)\right)$$

The effect of this is that we are a little more anticipatory of the input by generating an error signal between the current output and the future input. With a single-frequency input, all this does is circumvent the closed loop delay of the input signal, which in many cases would be irrelevant. However, this "predicting" nature comes into play when bandwidth is involved, as shown in the last example.

Modulated Signal $r(t) = A(t) \sin(\omega_c t)$. : We carry out the same type of expansion:

$$
\begin{aligned}
r'(t) &= A(t) \sin(\omega_c t) + \lambda(\dot{A}(t) \sin(\omega_c t) + A(t)\omega_c \cos(\omega_c t) \\
&= (A(t) + \lambda \dot{A}(t)) \sin(\omega_c t) + \lambda A(t)\omega_c \cos(\omega_c t)
\end{aligned}
$$

Here we notice more than the time delay shown from the single frequency input. The switching function is anticipating the input signal and trying to minimize the cost function defined in Equation 6.1.

6.1.5 Implementation

The switching curve derived in the previous section can be implemented by creating the signal of Equation 6.2 and using a comparator to determine when it crosses zero. An opamp with the transfer function $H(s) = RCs$ pulls of a very convincing $\lambda = RC$, and can be used to add in information about derivative of the input.

The "midtread" nature of the converter is added by using two comparators at $S(t) > +\alpha$ and $S(t) < -\alpha$. The error function for a 50kHz input with $alpha = 0.3$ is shown in Figure 6-2 on the right. The left hand side of the figure shows the trajectory of the state and change of input.

A plot of spectrums is shown in Figure 6-3 for $\lambda \in \frac{1}{2}..5 * 10^{-1}$, but it must be kept

60
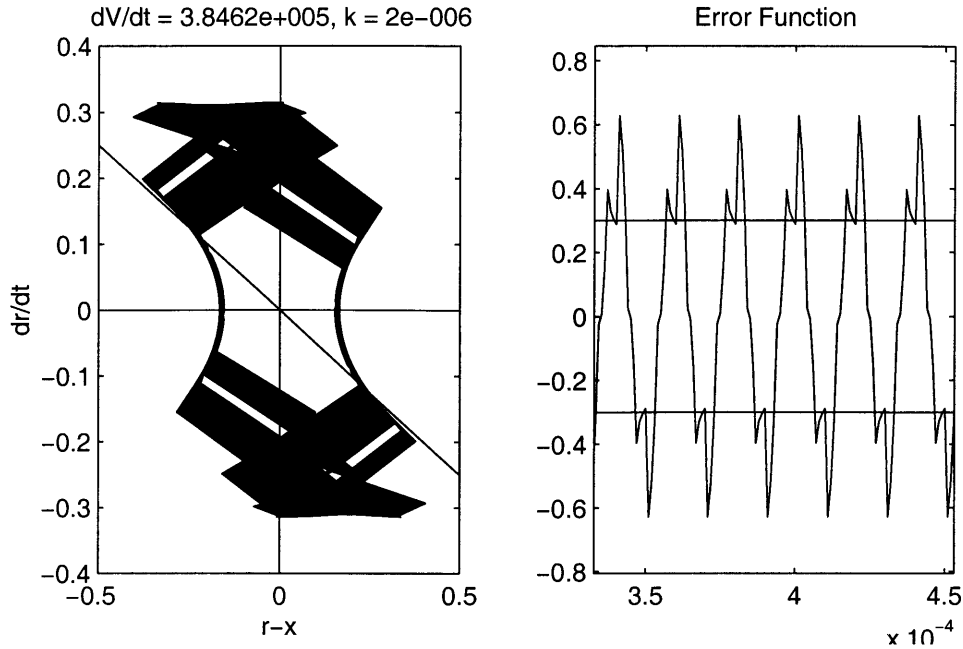
Figure 6-2: 50kHz Switching Function Error, $\lambda = 2e - 6, \alpha = 0.3$

in mind that the spectrum doesn't tell the whole picture and this must be looked at together with Figure 6-5 to judge where the optimal switching curve lies. For example, when the output begins to form a triangle wave from saturating the error function (large λ), the third harmonic drops out of the spectrum but the error energy ramps upwards.

The energy of the error signal can be seen in relation to λ in Figure 6-5. This control method is very close to "sliding control"[SL91] in which a multiple-order system is controlled with a single sliding variable $s \equiv (1 + \lambda \frac{d}{dt})e(t)$. However, in this case the error energy from using a control of this nature is larger, as shown in Figures 6-4 - 6-5 as it varies with λ. Sliding control incorporates the discontinuous state derivative u into the switching function through the $\dot{e}(t) = \dot{r}(t) + u(t)$. Discontinuities in the switching function incline the controller more towards a "chattering" response. In Figure 6-5 the error for $s = e(t) + \dot{e}(t)$ jumps for $\lambda = 2.2 * 10^{-6}$ as the discontinuity results in noncontrolling chatter.

The control has effectively become a feedback system with a nonlinear element

61
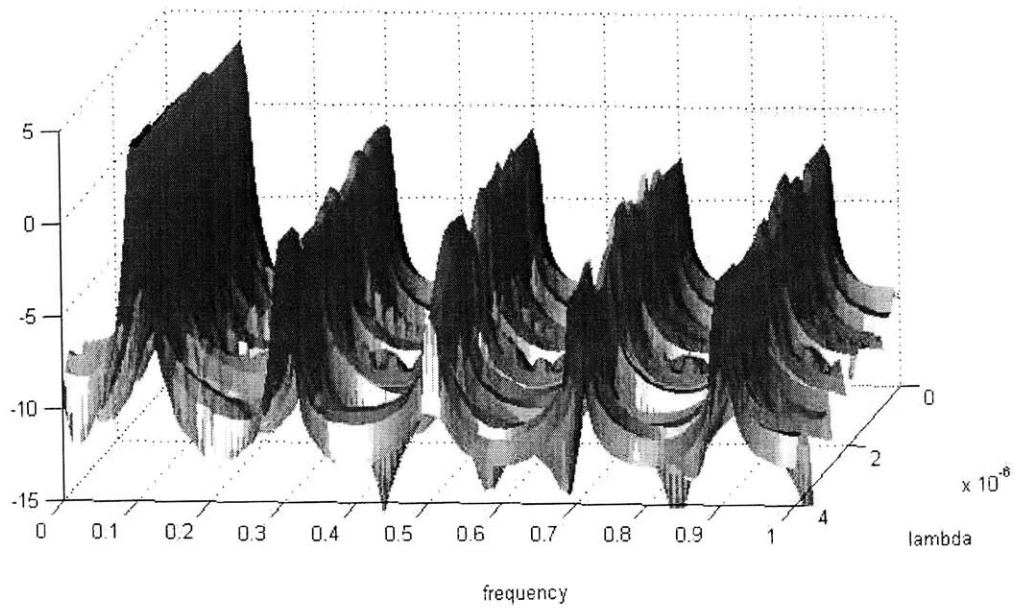
Figure 6-3: Output Spectrum vs. λ
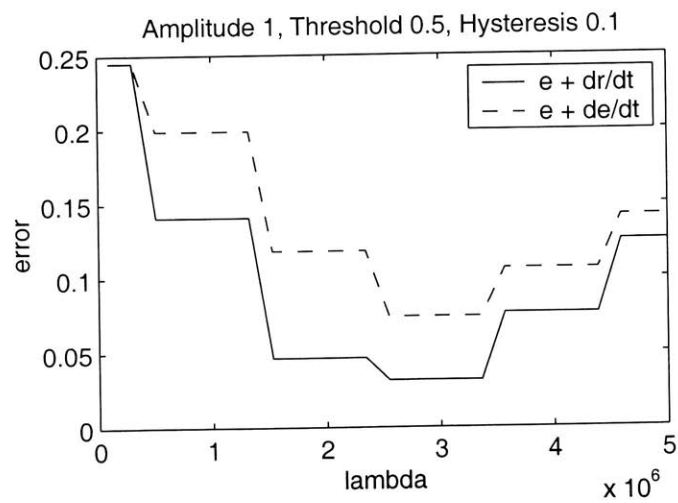
Figure 6-4: Error Improvement over Sliding Mode
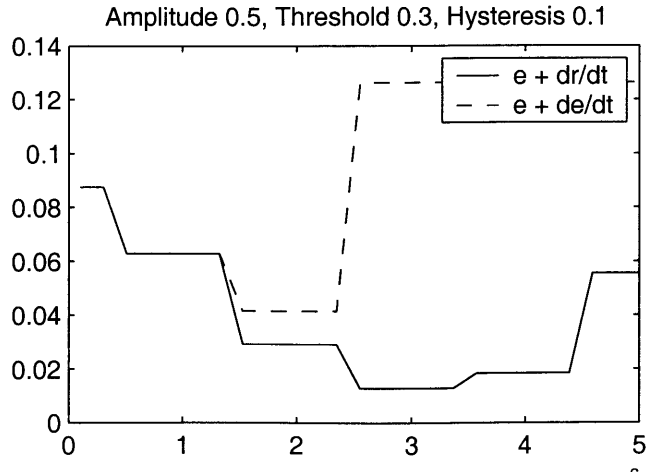
Figure 6-5: Error Improvement over Sliding Mode

determining the switching behavior. Looking at fast-changing input signals over a single period, the situation closely resembles that of Figure 3-4 where the switching function allows for a hysteresis margin of $\pm \alpha$ around the transitions points, and the switching action takes place a finite amount of time τ later. The time delay and hysteresis creates a limit cycle condition in which the integrator input can oscillate at fixed frequency while trying to track an input. Conversely, we can effectively guarantee a minimum switching period by adding some hysteresis to the switching thresholds if the inherent time-delay is by itself insufficiently long. An elegant method for analyzing these conditions is using describing functions.

6.2 Describing Function Analysis

An informative tool for analyzing continuous-time nonlinear systems is the describing function. This is particularly applicable to systems with nonlinearities such as saturation, dead-band regions and hysteresis; our midtread 3-level quantizer has such a transfer function (see Figure 4-1), defined by a and b.

The describing function looks at the gain and phase relationship of the fundamental component output by the nonlinear loop element, which, for zero-offset inputs,

usually depends on input amplitude as well as frequency. We call this function G_D and define it for odd functions as [SL91]

$$G_D(E, \omega) \frac{1}{\pi E} \int_0^{2*\pi} y(t) \sin(t) dt$$

where $y(t)$ is the output of the nonlinearity for an input of $E \sin(t)$. For the 3-bit quantizer, this integral is shown in Figure 6-6. There is obviously no phase shift (assuming an ideal implementation of this transfer function) and the gain of the fundamental is determined by:
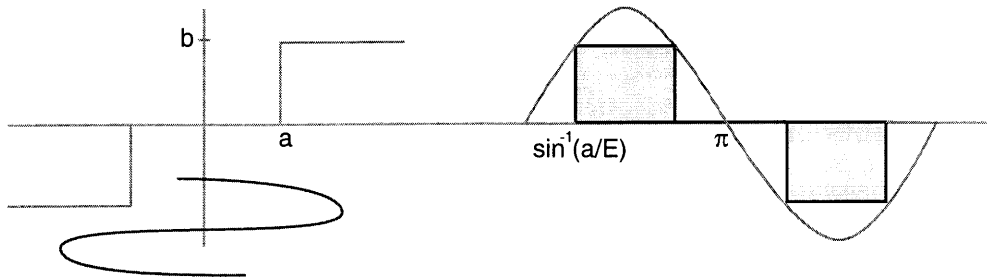
Figure 6-6: Describing Function Integral

$$
\begin{aligned}
|G_D(E, w)| &= \frac{1}{\pi E} \int_0^{2*\pi} y(t) \sin(t) dt \\
&= \frac{1}{4\pi E} \int_0^{\pi/2} y(t) \sin(t) dt \\
&= \frac{b}{4\pi E} \int_{\sin^{-1}(a/E)}^{\pi/2} \sin(t) dt \qquad \text{for } a < E \\
&= \frac{b}{4\pi E} \left[-\cos(t) \big|_{\sin^{-1}(a/E)}^{\pi/2} \right] \qquad \text{for } a < E \\
&= \begin{cases} \frac{b\sqrt{1-(a/E)^2}}{4\pi E} & a < E \\ 0 & else \end{cases}
\end{aligned}
$$

We notice $|G_D(E, \omega)|_{E=a} = \lim_{E \to \infty} |G_D(E, \omega) = 0$, but for $E \in (a, \infty) \Rightarrow G_D(E, \omega) \neq 0$. Setting $\frac{\partial G_d}{\partial E}(E, \omega) = 0$ we find that the peak gain of the output

fundamental component for an input sine wave is $\frac{b}{8\pi a}$, which occurs at $E = \sqrt{2}a$.

This nonlinearity is rewritten as the describing function in the loop pictured in Figure 6-7. All of the linear elements of the loop are grouped into $H_{lin}(s) = \frac{ke^{-sT}}{s}$, so the loop transfer function $L(s) = G_D(E, \omega)H_{lin}(j\omega)$. We are interested in the stability of this approximate system (G_D doesn't take into consideration any of the harmonics generated by the nonlinearity, which is usually safe enough for an approximate stability analysis). The system classicaly goes unstable for $L(s) = -1$, which means we want to find [Coo94]
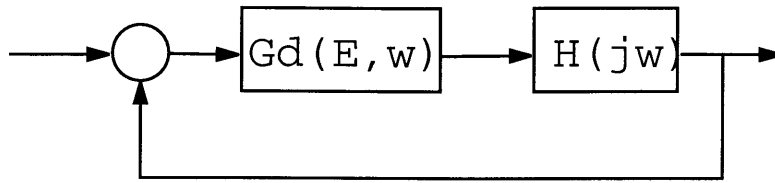
Figure 6-7: Feedback System with Linear and Nonlinear Elements

$$H_{lin}(s) = \frac{-1}{G_D(E, \omega)} \tag{6.3}$$

It's easy to find this solution on the gain-phase diagram shown in Figure 6-8.

The gain-phase plot of $-1/G_D(E)$ starts and ends at infinite magnitude, falling to $8\pi a/b$ at one point. Therefore the intersection defined by Equation 6.3, if it occurs, happens at two value's of E. The derivative of $G_D(E)$ with respect to E determines whether the limit cycle itself will be stable. In this case it is true for the $G_D(D)$ decreasing with increasing E.
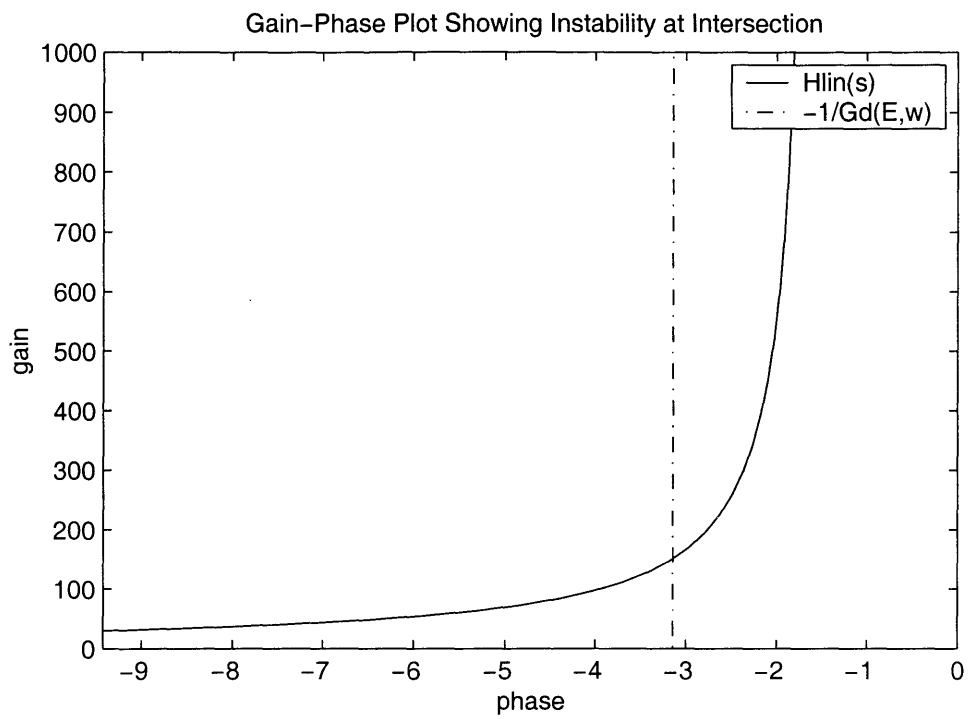
Figure 6-8: Gain-Phase Plot Showing $L(s) = -1$

Chapter 7

Implementation
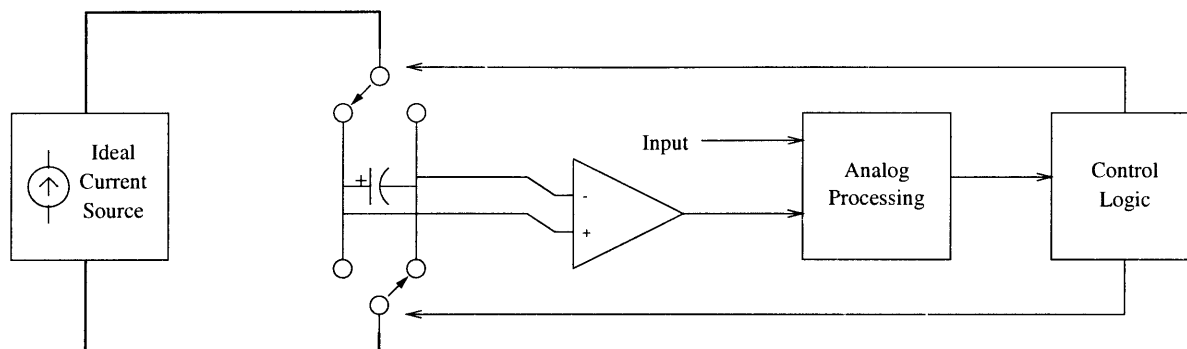
Figure 7-1: H-Bridge Diagram

The physical topology of the amplifier is a current-switching H-bridge, as illustrated in Figure 7-1. Charge from a current supply is either integrated positively or negatively onto the output capacitor, or bypasses it entirely. The output is continuously measured and combined with the input voltage to generate an error signal that is fed into the controller. Most of the filtering and algebraic manipulation is done via analog processing to reduce the propagation delay of the digital controller.

There are several methods of realizing each block; generally performance can be improved with more complicated schemes. For the purposes of this thesis, a rudimentary arrangement was constructed and optimized because the basic principles are the same but the effects of the optimizations are more apparent. For example,

no current source will be "ideal" and therefore the control system must be designed to adapt to current source variations. However, if a more complicated current source is used, it might not be as obvious from the measurements which feedback loop is responsible for the remaining error.

The current H-bridge system consists of the following components:

- Current Source

- H-Bridge

- Analog Processing

- Control Logic

7.1   Current Source

The current source was made from either an inductor or the magnetizing inductance in a transformer, as in a "flyback converter". The output impedance of the source can be calculated for an initial current I_o in an inductor L used to move the load C_x through its full swing $\pm V_{max}$. After delivering energy ΔU to the load, the current source has dropped to $I_{new} = I_o - \Delta I$ according to:

$$
\begin{aligned}
\frac{1}{2}LI_{new}^2 &= \frac{1}{2}LI_o^2 - \Delta U, \text{ where } \Delta U = \frac{1}{2}CV_{max}^2 \\
I_{new} &= \sqrt{I_o^2 - \frac{\Delta U}{L}} \\
\Delta I &= I_o\left(\sqrt{1 + \frac{\Delta U}{LI_o^2}} - 1\right) \\
Z &\equiv \frac{\Delta V}{\Delta I} = \frac{\Delta U}{LI_o} = \frac{LI_o}{CV_{max}}
\end{aligned}
\tag{7.1}
$$

We can see from Equation 7.1 an inductor is a good approximation of an ideal current source if $LI \gg CV$, which is a state variable analgous to momentum. A realistic load of **40nF** driven to **±100V** at **50kHz** would require a minimum DC

68

V_{peak}	I_{peak}	L	C	U_{elec}	U_{mag}	ΔU percent	θ
50	1	1mH	1nF	1.25uJ	500uJ	.25%	2.87
100	1	1mH	1nF	5uJ	500uJ	1%	5.74
150	1	1mH	1nF	11.3uJ	500uJ	2.25%	8.63
200	1	1mH	1nF	20uJ	500uJ	4%	11.5
100	2	1mH	1nF	5uJ	2mJ	.25%	2.87
150	2	1mH	1nF	11.3uJ	2mJ	.56%	4.3
200	2	1mH	1nF	20uJ	2mJ	1%	5.74
100	2	1mH	10nF	50uJ	2mJ	2.5%	9.1
150	2	1mH	10nF	113uJ	2mJ	5.63%	13.72
200	2	1mH	10nF	200uJ	2mJ	10%	18.43
100	2	.5mH	10nF	50uJ	1mJ	5%	12.9
150	2	.5mH	10nF	113uJ	1mJ	11.25%	19.6
200	2	.5mH	10nF	200uJ	1mJ	20%	26.57

Table 7.1: Changes in Source vs. Output Level

current of **1.25A**, to avoid overload distortion, and an inductance of $L \gg 3.18\mu H$, which is very reasonable. The only problem with having an low output impedance is that over a single switching period the assumption of a constant I_{dc} breaks down and must be taken into account by the controlling algorithm. Some algorithms, such as the PWM scheme outlined earlier, rely on a constant source and do not inherently compensate for such variations.

Table 7.1 shows the percentage change in the peak inductor current for different levels and loads, where U_{elec} and U_{mag} are peak stored energy levels in the capacitor and inductor. This can also be thought of the differential equation $\ddot{V_c} = -\frac{1}{LC}V_c$, where the column for θ is the max angle (in degrees) in the (V_c, I_l)-plane through which the state trajectories rotate, which is given by

$$theta = \sin^{-1}\left(\frac{V_{peak}}{I_{peak}}\sqrt{C/L}\right)$$

As it can be seen from Table 7.1 and Figure 7-2 and 7-3, for increasingly higher output signals into larger loads, the ratio of energy transferred to and from the capacitor as compared to the energy maintained in the inductor rises. If the controller is assuming a fixed current source input, $I_{dc}(t) = I_o$, then at the peak of an output sinewave the current source has dropped by a factor of $\cos(\theta)$. In the ideal case,
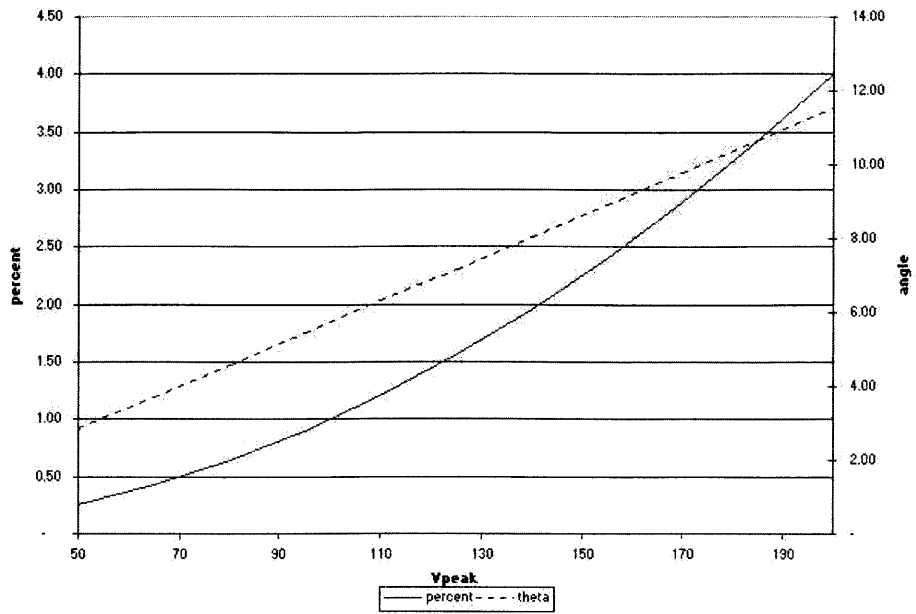
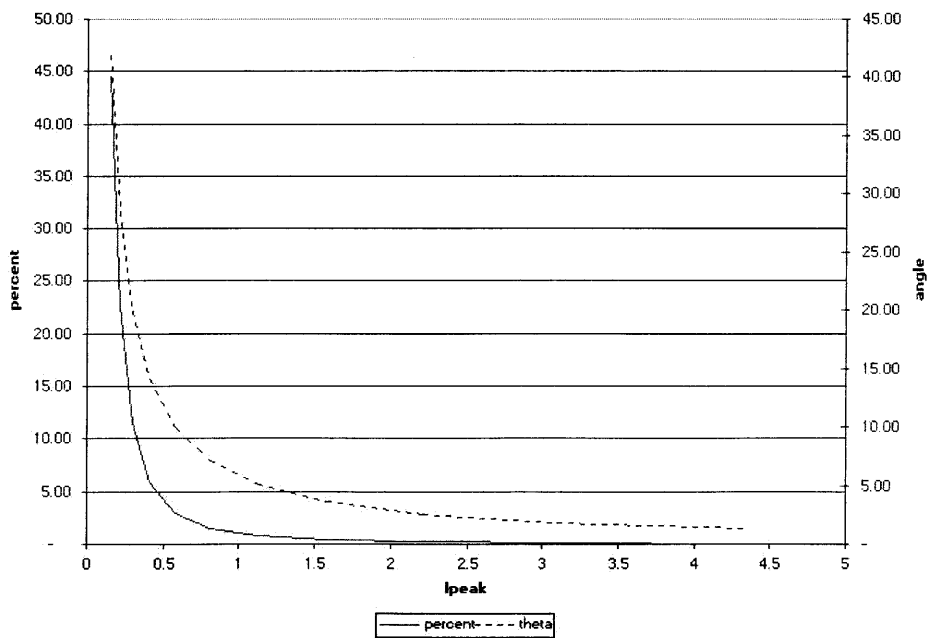Figure 7-2: Variations in Source vs. Output Level

Figure 7-3: Variations in Source vs. Source Level

$L \to \infty$ which implies $\omega_o = (LC)^{-\frac{1}{2}} \to 0$. If we start from zero output,

$$V_{ideal}(t) = \lim_{\omega_o \to 0} A\sin(\omega_o t) = \lim_{\omega_o \to 0} \frac{I_o \sin(\omega_o t)}{C\omega_o} = \frac{I_o}{C}t$$

so the difference between an ideal inductor and a real one is

$$V_{real}(t) - V_{ideal}(t) = \frac{I_o}{C}\left(\frac{\sin(\omega_o t)}{\omega_o} - t\right) = \frac{I_o}{C}(t - \frac{\omega_o^2 t^3}{3!} + \frac{\omega_o^4 t^5}{5!} + \cdots) \qquad (7.2)$$

This is really the difference between approximating the converter as a perfect integrator or acknowledging the whole dynamic system involving both capacitor and inductor states and differentials.

The discrete time implementations of the control loop for this amplifier approximate the inductor current by periodic sampling. If the converter is connected to the load, then the error grows according to the series 7.2. If we assume the switching patterns are uncorrelated with the input (reasonable for most inputs), we can model this error mechanism as a source of white noise, $\sigma_{I_{dc}}^2 \approx \frac{I_o^2}{C^2 f_s^2}$, injected into the output. Note that it is inversely related to switching frequency squared.

7.1.1 Inductor Current Source

An inductor in a "buck converter" configuration works well as a simple current source, as shown in Figure 7-4. Unlike a buck converter, however, the ideal operation of the converter would never require an "on" period, since the inductor is 90° out of phase with the load, and thus all energy delivered to it is returned within the same cycle.

The buck converter operation that takes place is purely in terms of supplying a constant DC voltage to the V_{BE} and R_{DS} components of the diodes and switches. To eliminate unnecessary current fluctuations, the DC-current control loop operates simply by dumping a fixed amount of energy into the inductor whenever it's current (or energy) falls below a minimum value. This control loop will reach a steady state cycle. We can determine this duty cycle D from Figure 7-4 by setting the average
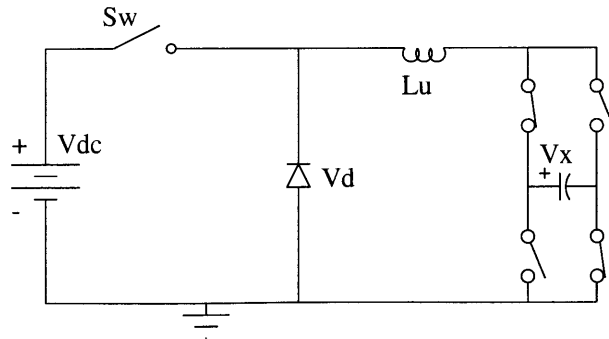
Figure 7-4: Inductor Current Source

voltage across the inductor to zero:

$$
\begin{aligned}
V_L &= (1-D)[A\sin(\omega_o t) - V_D - V_{SW}] + D[A\sin(\omega_o t) - V_{SW} + V_{DC}] \\
&= DV_{DC} - (1-D)V_D - V_{SW} + A\sin(\omega_o t) \\
&= DV_{DC} - (1-D)V_D - V_{SW} \qquad \text{if } \omega_o >> (LC)^{-\frac{1}{2}} \\
&= 0
\end{aligned}
$$

$$
D = \frac{V_{SW} + V_D}{V_{DC} + V_D} \tag{7.3}
$$

The power consumed by the converter can be given as $P = DI_{DC}V_{DC}$. Note that this doesn't take into consideration dissipation of the H-bridge MOSFET drivers, which also is a significant component of the total power. The MOSFET voltage drop $I_{DC}R_{DS}$ is proportional to R_{DS}.

The circuit constructed used a 1mH inductor which has a maximum current rating of about $1A$ and a maximum resistance $R_L < 1\Omega$. Figure 8-9 in the Results section shows the actual waveforms captured.

7.1.2 Transformer Current Source

Another viable option for a current source is the magnetizing flux of a transformer, which we link through a secondary winding. The advantage of this is that current source can be "charged up" much faster than the inductor by a factor N, where N is the turns ratio. The battery will see a much lower inductance on the primary, and can transfer energy in a quick high-current pulse while the secondary coil can source current to the load through a much higher inductance.

Although the load capacitor doesn't remove any net energy from the source, energy stored in the transformer is eventually dissipated through wire resistances and semiconductor voltage drops. Energy is added back into the transformer through a secondary coil with a much lower inductance. Therefore, a voltage pulse on the secondary can transfer a lot of energy quickly into the core, which will then decay again at a slow rate relative to the switching frequency until the next pulse. During this "recharge" [1] period, the primary current path must be turned off.
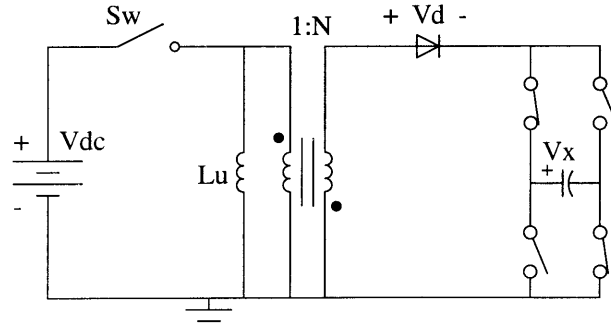
Figure 7-5: Transformer Re-Energizing

In a circuit with magnetic storage, conductors around the flux must form a closed path at all times to avoid dangerous voltage spikes. Therefore it is not practical to rely on controller/driver timing to turn off the primary path at the instant the secondary is energized for recharge. Instead, we keep the primary path unidirectional with a diode, and apply a secondary charging voltage large enough to force the primary path

---

[1] perhaps a more appropriate term would be "reflux"

to turn off. This results in a very simple constraint on the turns ratio. Looking at Figure 7-5, when S_w closes $V_d = -N \cdot V_{dc} \pm V_x$ which leads to the condition

$$V_{dc} \cdot N > |V_{x,max}| \tag{7.4}$$

While the DC voltage is connected during T_{on}, the current in the transformer primary ramps up at a constant slope, creating a net change ΔI_p:

$$\int_t^{t+T_{on}} v(t)dt = L(I_p(t + T_{on}) - I_p(t)) \rightarrow \Delta I_p = \frac{V_{dc}T_{on}}{L} \tag{7.5}$$

More important is the increase in stored energy during the charging period, which we can equate with the power dissipation during converter operation:

$$
\begin{aligned}
\Delta U_{mag} = \frac{1}{2}L((I_p + \Delta I_p)^2 - I_p^2) &\approx LI_p\Delta I_p \\
&= P_{diss}T_{off} \\
&= \left(\frac{I_p^2}{N^2}(2R_{DSon} + R_{sense}) + 2\frac{I_p}{N}V_{diode} \right) T_{off}
\end{aligned}
$$

By substiting in Equation 7.5 for ΔI_p we now have a relationship between run and recharge periods in a single-inductor converter implementation. Letting $R_{diss} = 2R_{DSon} + R_{sense}$, that is

$$V_{dc}N^2T_{on} = (I_pR_{diss} + NV_{diode})T_{off} \tag{7.6}$$

Realistic values for a transformer source are given in Table 7.1.2. Notice that a higher source voltage and lower dissipating resistance both reduce the T_{on}/T_{off} ratio favorably. At the expense of high peak primary currents, a larger N will also lead to smaller charing on-time.

Since the output is held constant while the current-sourcing inductor is being re-energized, converter performance clearly improves with the ratio T_{off}/T_{on}. From Equation 7.6 we can increase the percent of the time the current source is available by changing V_{dc}, I_p or N. (Note that magnetizing inductance L is not a factor.)

Vdc	$I_{secondary}$	N	I_p	R_{diss}	P_{diss}	T_{on}/T_{off}	V_{max}
12	1	10	10	1	1	.83%	120
15	1	10	10	1	1	.66%	150
24	1	10	10	1	1	.41%	240
36	1	10	10	1	1	.28%	360
15	1	8	8	1	1	.83%	120
24	1	5	5	1	1	.83%	120
36	1	3.33	3.33	1	1	.83%	120
12	1.5	10	15	1	2.25	1.23%	120
12	2.5	10	25	1	6.25	2.04%	120
12	5	10	50	1	25	4%	120

Minimizing N according to Equation 7.4, we get

$$V_{max}T_{on} = (I_p R_{diss}/N + V_{diode})T_{off}$$

Or more interestingly, if we let $V_{drop} = I_p R_{diss}/N + V_{diode}$

$$\frac{V_{drop}}{V_{max}} = \frac{T_{on}}{T_{off}}$$

The transformer used in the test circuit consisted of a 22-turn primary and 200-turn secondary coil (N=8.8) around an E-core. The inductances measured in the lab are $L_p = 12.5\mu H$, $L_s = 9.7mH$.

7.1.3 Further Possibilities

We can sidestep the refreshing cycle problems from the previous section by, for example, maintaining two DC current sources and using one when the other is being re-energized. This extension upon the converter framework discussed in this thesis has not been implemented as it is trivial and bothersome to implement while leaving the theory unaffected. However, some implementation possibilities are interesting enough to merit description:

- Two identical current sources that alternate between use and energizing periods.

- A primary magnetizing inductance of a transformer for DC current, with smaller inductor used to source a brief pulse of current while the transformer recharges.

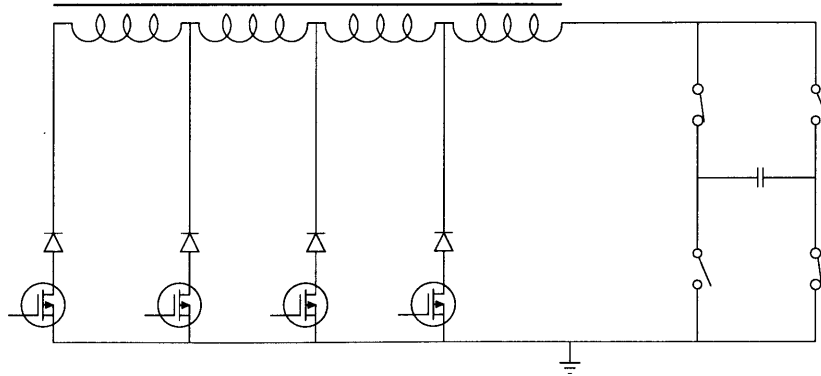- An array of differently valued current-sources. (See Figure 7-6).

Figure 7-6: Multiple-Slope Implementation

The final item is a veritable doorway to a superset of integrating converters with multiple-slope capabilities. For example, the state functions $dr/dt = \alpha_i, i \in 1, 2 \ldots n$ become available.
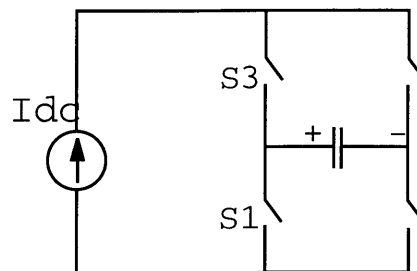
7.2 Current H-Bridge

Figure 7-7: Simplified Current H-Bridge

An H-bridge is used to connect a load across a fixed-polarity source in either direction, as in Figure 7-7. The output impedance of the source is irrelevant; most

Switches	$S_1 S_3$	$S_1 S_2 S_3$	$S_2 S_3$	$S_2 S_3 S_4$	$S_2 S_4$	$S_1 S_1 S_4$	$S_1 S_4$	$S_1 S_3 S_4$
V_{Idc}	0	0	V_c	0	0	$-V_c$	$-V_c$	$-V_c$

Table 7.2: Overlapping Switch Evolutions and Corresponding Load Connections

often a voltage source is used; in this thesis, current is similarly routed through a load in either direction. Current can also be bypassed around the load.

One of the potential dangers of voltage H-bridges is the possibility of "shoot-through", where the simultaneous switching of the low and high power devices can overlap enough to short the voltage to ground. Because the current H-bridge is a dual of the voltage H-bridge and therefore always requires a path to ground. The switching arrangement must be able to guarantee a minimum overlap so that the sourcing inductance won't open-circuit, resulting in harmful voltage spikes. This is illustrated in Figure 7-8.
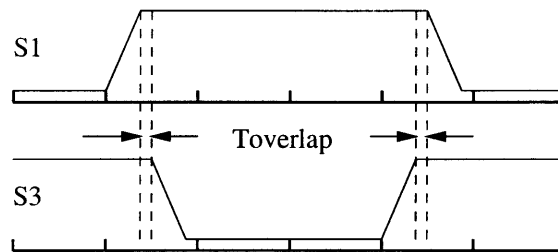

Figure 7-8: Overlapping Switch Transitions

At the same time that we always want a path to ground to avoid voltage spikes, as with water, the current will flow through the "path of least resistance", which in this case is better called the "path of most entropy". That is, when three switches are closed, the path of forward biased diodes transfers the least energy from the source by applying the more negative voltage across the source. This is illustrated in Figure 7-9.

From Figure 7-9 and Table 7.2 it is clear that the voltage waveforms aren't symmetric around zero. We will assume a constant switching frequency $f_s = T_s^{-1}$ and allow the switches to overlap a duration T_{ov}, and define the ratio $\delta = T_s/T_{ov}$. The normalized input current to the capacitor $u \in \{-1, 0, +1\}$ is integrated each period. For example, alternating between a positive current during one period ($u = 1$) and
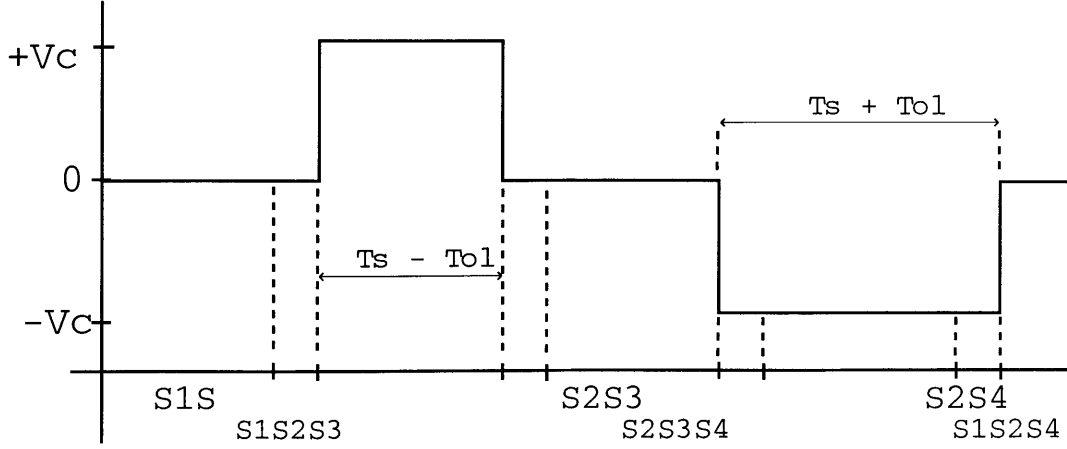
Figure 7-9: Waveform across Current Source with Positive Load

zero current ($u = 0$) over the next period, the capacitor sees zero current during both transitions (T_{ov}) and the average voltage increase is $\frac{1}{2}(u_0 + u_1) = \frac{1}{2}(1 - \delta)$. This means the transition $u_0 \to u_1$ as well as $u_1 \to u_0$ comes out $\frac{1}{2}\delta$ below the control input average.

To find the effective control values $\tilde{u}_i$, we create a table of transitions:

$$u_0 \to u_1 := 1 - \delta \qquad u_0 \to u_0 := 0 \qquad u_{-1} \to u_0 := -1 - \delta$$

$$u_1 \to u_1 := 2 \qquad u_1 \to u_0 := 1 - \delta \qquad u_1 \to u_{-1} := -2\delta$$

$$u_{-1} \to u_1 := -2\delta \qquad u_{-1} \to u_0 := -1 - \delta \qquad u_{-1} \to u_{-1} := -2$$

If we assign the transition probabilities $P(u_i \to u_j) = p_{ij}$, then the effective control input $\tilde{u}_1$ can be found by looking at the transitions to and from u_1:

$$
\left.\begin{array}{ll} u_1 & 0 \\ u_0 & -\delta/2 \\ u_{-1} & -\delta \end{array}\right\} \underline{\qquad u_1 \qquad} \left\{\begin{array}{ll} u_1 & 0 \\ u_0 & -\delta/2 \\ u_{-1} & -\delta \end{array}\right.
$$

We find effective control $\tilde{u}_{1,V>0}$ for a positive load voltage by taking the expectation:

78

$$\tilde{u}_{1,V>0} = \mathrm{E}[(u_i \to u_1) + (u_1) + (u_1 \to u_j)]$$
$$= 1 + (0)p_{11} + (-\delta/2)p_{01} + (-\delta)p_{-11}$$
$$+ (0)p_{11} + (-\delta/2)p_{10} + (-\delta)p_{1-1}$$

The probabilities p_{ij} will depend on the switching curves and converter input signal; in general they cannot be assumed to be independent, i.e. $P(u_i \to u_k) \neq P(u_j \to u_k)$ for $i \neq j$. However, it's reasonable to assume $p_{1,j} = p_{-1,-j}$, which means the signal is symmetric around zero. The reader will now be left with the following:

$$\tilde{u}_{1,V>0} = 1 - \frac{\delta}{2}(p_{01} + p_{10} + 4p_{1-1})$$
$$\tilde{u}_{0,V>0} = 0 - \delta(p_{10} + p_{01})$$
$$\tilde{u}_{-1,V>0} = -1 - \frac{\delta}{2}(p_{0-1} + p_{-10} + 4p_{1-1})$$

In short, if $\delta > 0$ (finite switching overlap) then the integrated voltage output for a positive load will be a little less negative than expected, i.e. $\tilde{u}_{i,V>0} < u_i$. For $V < 0$, the situation is just the opposite.

7.2.1   Node Voltage Waveforms

In order to guarantee a finite overlap, while at the same time not allowing two paths for the current for too long (ideally $\delta << 1$), we must understand the voltage waveforms across the switches. They are illustrated in Figure 7-10.

With the node voltages as defined in Figure 7-11, we can focus on the $\{S_1, S_2\}$ loop. The simplified device law for the diodes yields $V_D <= 0$. In this thesis we use exclusively power MOSFETs for the current switches (this choice will be discussed
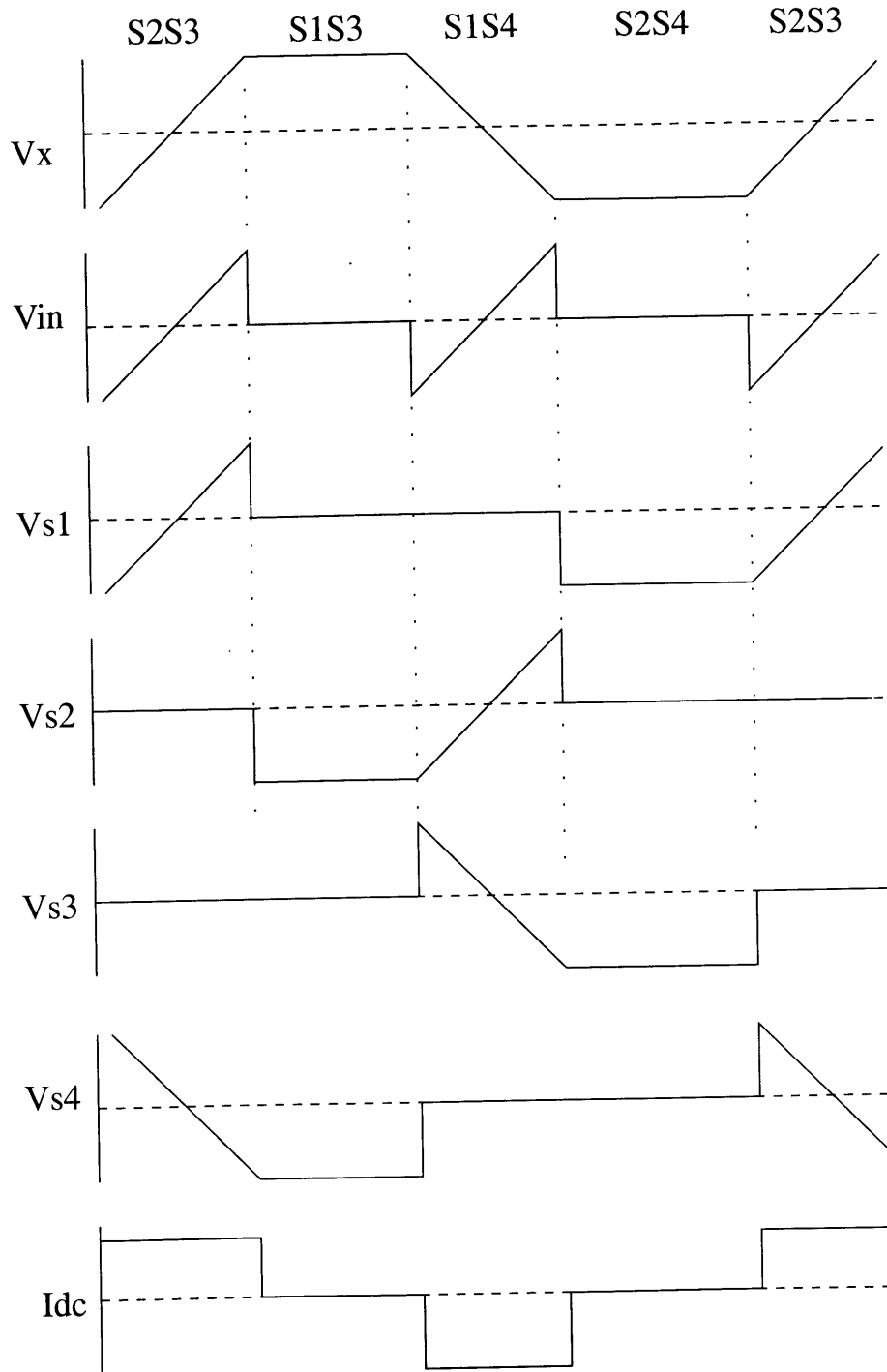
Figure 7-10: Node Voltages

shortly), which typically have a vertical geometry with the substrate tied to the source. This acts as a reverse body diode, yielding $V_S >= 0$. KVL can be written as:

$$V_{S1} + V_{D1} = V_{S2} + V_{D2} + V_X$$

The condition that at least S_1 or S_2 is on at any time can be imposed by the constraint:

$$(V_{S1} + V_{D1})(V_{S2} + V_{D2}) = 0$$

Therefore, if $V_X > 0$ and one switch is closed:

$$S_1 \text{ open}: \quad V_{S1} + V_{D1} = V_X > 0 \quad \rightarrow V_{S1} = V_X$$
$$S_2 \text{ open}: \quad V_{S2} + V_{D2} = -V_X < 0 \rightarrow V_{D2} = -V_X$$
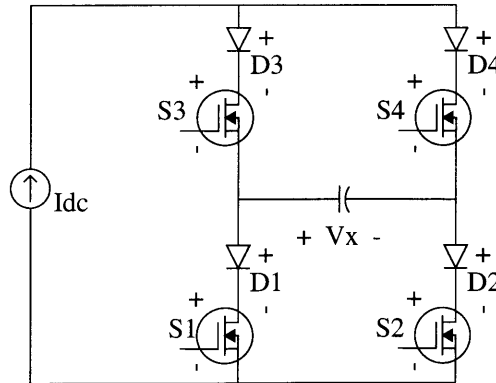

Figure 7-11: Node Definitions

When the output capacitor voltage is nonzero, the switches $S_1 \ldots S_4$ have either $|V_X|$ or zero volts across them, while the diodes support either $-|V_X|$ or zero volts. First off this makes it easy to specify the voltage stresses that the transistors must

be capable of sustaining. More importantly for this discussion, however, is that the drain-to-source voltage of a MOSFET about to turn on determines a large part of the turn-on waveform.

7.2.2 Transistor Selection

MOSFETs were chosen because they're easy to drive. The two simple specifications to meet are drain-to-source breakdown voltage (V_{DSS}) and maximum drain current, which are determined by the output requirements. For a load C driven up to f_{max} at a peak voltage of $\pm V_{max}$,

$$I_{D,max} > 2\pi f_{max} C V_{max}$$

$$V_{DSS} > \begin{cases} V_{max} & \text{inductor source} \\ N \cdot V_{DC} & \text{transformer source} \end{cases}$$

The primary consideration in choosing the MOSFETS is the compromise between switching drive power consumption and R_{DS} voltage drops that absorb energy from the current source (see Equations 7.3, 7.6). Datasheets usually supply a Q_{gate} specification as the amount of charge necessary to sufficiently turn on the device without having to resort to parasitic capacitance calculations [Intb]. Transistor current capacity has a logical inverse correlation with $R_{DS,on}$ while it is roughly proportional to Q_{gate}. Figure 7-12 shows a loose inverse relationship beween $R_{DS,on}$ and Q_{gate} themselves, suggesting an opportunity to optimize performance with careful selection.

For the arbitrarily chosen specifications

V_{max}	200V
C	10nF
f_{max}	75kHz

then $I_{max} = 1A$. It reasonable to make the DC power used by the transistors on the order of the power consumed by the diodes of the H-bridge. Thus we'll try for $I_{max} R_{DS} \approx \frac{1}{2} V \rightarrow R_{DS} < \frac{1}{2}$. A cheap MOSFET with the following parameters was chosen for S_1 and S_2
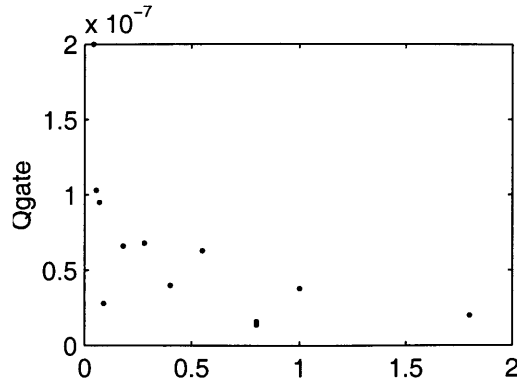
Figure 7-12: Q_{gate} vs. $R_{DS,on}$ for Assorted MOSFETS

IRF644

V_{DSS}	250V
$R_{DS,on}$	0.28Ω
I_D	14A
Q_{gate}	46nC
C_{iss}	1300pF
C_{oss}	330pF

For S_3 and S_4 IRL640's are preferred because they have a logic-level gate drive. These switches are floating and their turn-on charge is supplied through a pulse-coupling transformer (PE-64973) made by Pulse Engineering. The transformer is a 1:1 ratio pair of bifilar windings wrapped seven times around a torroid core and optimized for 500kHz operation. The windings have an inductance of $100\mu H$ and a maximum leakage inductance of $180nH$. This schematic is shown in Figure 7-13. The logic level allows for complete turn-on at the lower drive voltage $V_{GS} = \pm 5V$ from the isolated gate drive at the expense of an additional $20nC$ of necessary gate charge.

Switches S_1 and S_2 are grounded N-channel MOSFETS that are driven by a TC4469 quad FET driver by Telcomm Semi. The power consumed by the driver can be approximated by $P_{tot} = P_l + P_t + P + q$, where the last two quantities are transition and quiescent power [Tel]. The power consumed by driving the gate capacitance of the MOSFET can be approximated by $P_l = f_o C V_{DC}^2$. At a switching frequency of
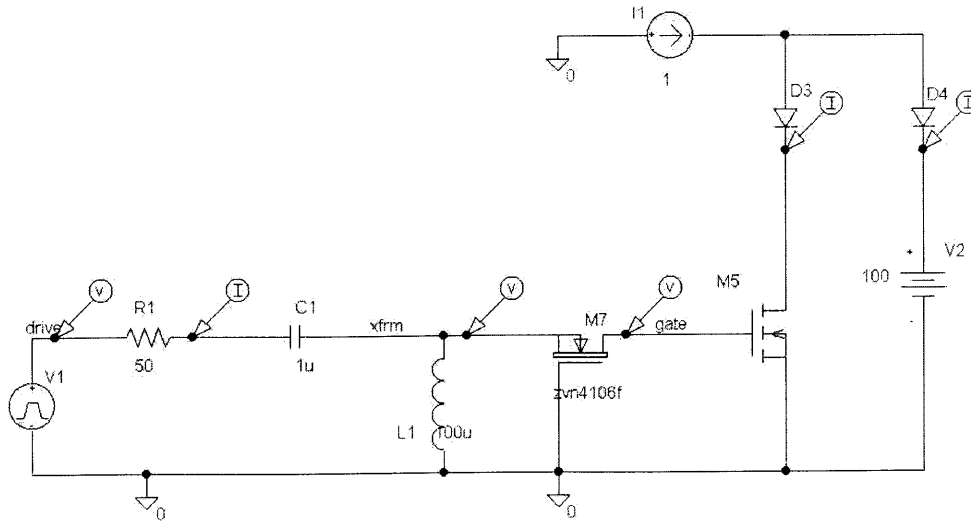
Figure 7-13: High-side Transistor Drive

$f_s = 500kHz$ this comes to under half a Watt.

Rudimentary estimates show that the TC4469, which can source up to $250mA$ per load, can turn on in about $150ns$. A more thorough investigation will illuminate the sources of switching delays and will justify the control decisions to have more of the switching action occur between S_1 and S_2.

7.2.3 Switching Delays

Between the time when the error signal crosses it's threshold level, or when the quantizer detects a new output level, and when the MOSFETs have actually achieved redirection of current through or around the load, a period of time has passed. This is a loop delay; it is consequential, inevitable, and important. It is loop delay that makes a control algorithm "chatter" instead of asymptotically approach zero error. It is loop delay that adds phase to our otherwise single-pole transfer function, thus imposing a limit on the loop gain and crossover frequency. In other words, it is delay that keeps the converter in the real world.

The switching delays we concern ourselves with now are between the control signal

given by the FPGA and it's change to the load current. The FPGA drives high-impedance logic inputs of the TC4469, which happens essentially instantly. The TC4469 driver then begins to source or sink a limited amount of current from the gate of a MOSFET. The TC4469 itself guarantees a minimum propagation delay of under $100ns$, after which it's outputs either rise or fall within $50ns$. [Tel].

Switches S_1 and S_2 are driven by one of the four outputs of the TC4469, which can source up to $250mA$. This can be seen in Figure 7-14. First this current charges up the MOSFETs gate-to-source capacitance until it reaches the threshold voltage and begins to conduct drain current. Before the gate voltage rises any further, the drain voltage lowers by discharging the gate-to-drain capacitance in a typical "Miller effect". At this point the MOSFET is on and a charge estimated by the "gate charge" of the transistor datasheet has been delivered [Intc]. After this point the gate charges until it meets the driver output volatage. The important apsect is that the turn-on time depends partly on the off drain-source voltage $V_{DS,off}$, which depends on the load state V_{OUT}. We can estimate that the gate charge ($Q_G = 68nC$ for a IRF644) delivered at 250mA will turn on the transistor in about $250ns$.

The circuit for the high-side MOSFETs S_3 and S_4 is more complicated because of the pulse-coupling transformer. This circuit was simulated in Pspice and the results are shown in Figure 7-15. Not only are the high switches slower, but their maximum turn-on volage from the floating drive is half of the TC4469's output. This is why we use logic-level transistors for the upper switches and let the lower switches do most of the work.

The auxiliary MOSFET,shown in Figure 7-13, is chosen to be a real lightweight – the VN0300L, a 60V transistor with turn-on/off times quoted to be under 30ns. The transistor can pass up to $\frac{1}{2}A$ to the power switch, and will leak no more than $100nA$ across the gate-body region. This leakage is important because it determines along with the input capacitance of the switch (1800pF for the IRL640) how long we can expect the switch to stay on ($T_{max} \approx 90ms$). Additional gate-source capacitors can be added for longer hold times. However, once the maximum on-time has passed, the logic cannot assume there is a closed loop for the current source and must perform
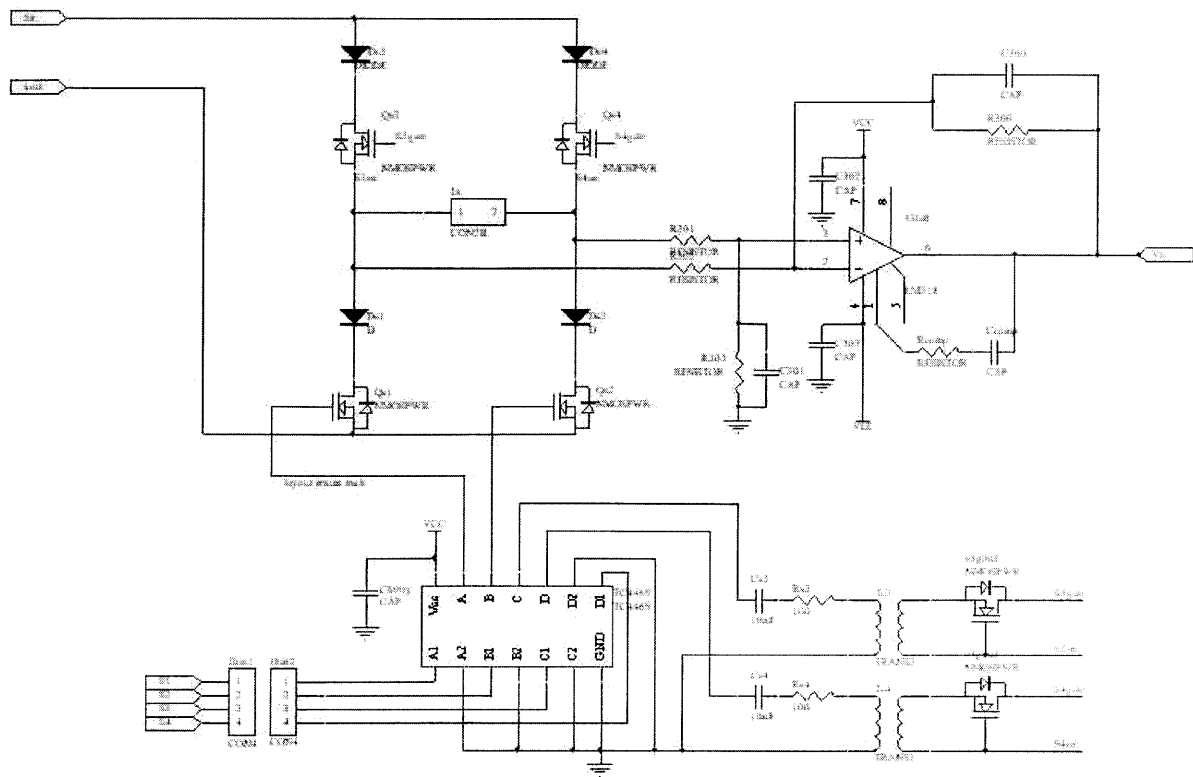
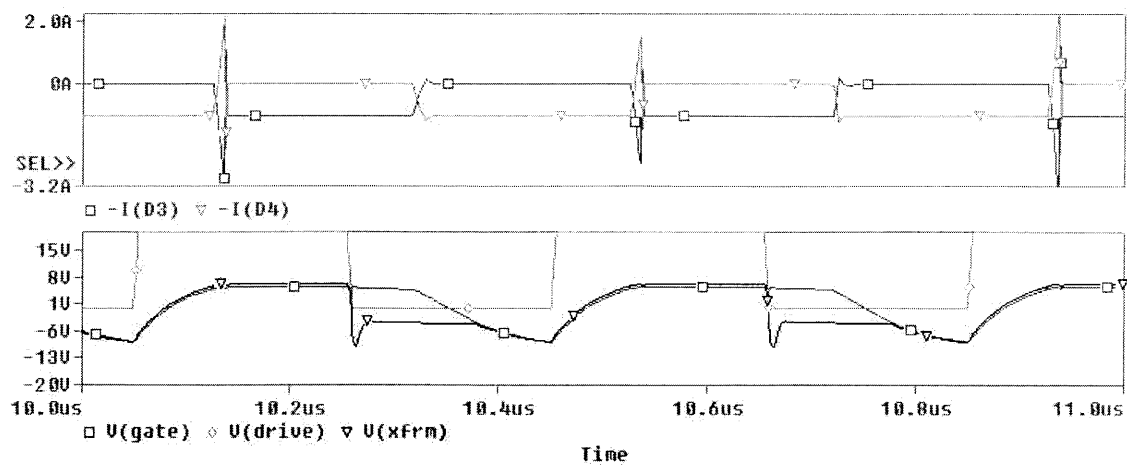Figure 7-14: H-Bridge Drive and Sense Circuitry

Figure 7-15: Simulation of Hi-Side MOSFET Drive

some operation to refresh the switch states.

For the minimum turn-on time one can expect around 75ns with a 10 ohm resistor [Inta]. In the demo circuit built for this thesis, a current-limiting resistor of 47 ohms was chosen to meet the output current limits of the TC4469; subsequently the simulation of Figure 7-13 is captured by Figure 7-15 and shows turn-on times under $200ns$.

7.2.4 Current Sense Amplifier

Although we aspire towards an ideal current source with a constant DC value, inevitable dissipation in the circuit slowly saps energy from the current source. There are a few reasons to have an accurate measure of the current at any given moment. First, when the current falls below a threshold the converter "kicks" it back up (this was discussed in Section 7.1.2).

The more important reason to have an accurate measure of the inductor current is because it's a state variable of the system. Over any switching interval we try to assume it is constant by meeting the the condition $LI >> CV$ from Section 7.1. The value $I(t)$ affects the loop gain whether the current H-bridge is modeled as a CT integrator $H(s) = \frac{I_o}{Cs}$ or a slope quantizer as shown in Figure 4-2 where $H(z) = \frac{(I/C)T_s z^{-1}}{1-z^{-1}}$.

A $\frac{1}{2}$-Watt, 1% carbon film resistor was used as the current sense element; it's main requirements are low inductance so that it doesn't slow s witching speeds of S_1 or S_2, and a sufficient power rating which should be very small. The threshold-sensing comparator is an LM393, which has an input common-mode range down to ground. Therefore the threshold voltage can be as necessary to minimize power consumption, but shouldn't be too low that circuit noise will affect it's accuracy. As a compromise, a threshold value of $V_{thres} = 0.6V$ was chosen, which leads to $R_{sense} = V_{thresh}/I_{DC}$. For a 0.6A minimum current, we select a 1Ω resistor rated at $\frac{1}{2}$Watt.

An RC-filter smooths out any switching noise on the current-sense measurement that might falsely trip the threshold comparator. This first order filter is simply $H_{lp}(s) = \frac{1}{1+s/p}$ The current ramps down from I_o with an average slope $k \equiv V_L/L$,

where V_L is the DC voltage presented by the load. Thus starting each period after the current source has been charged, we find $I(t) = I_o(1 - kt)u(t)$. Now we determine tracking error

$$I_{error}(s) = I_o \left(\frac{1}{s} - \frac{k}{s^2} \right) (1 - H_{lp}(s)) = I_o \left(\frac{s-k}{s^2} \right) \frac{s}{s+p} = I_o \left(\frac{s-k}{s(s+p)} \right)$$

Taking inverse Laplace shows $I_{error}(t) = I_o \left((1 - \frac{1}{p})e^{-pt} - \frac{1}{p} \right) u(t)$, but it's just as simple to find the steady-state tracking error by

$$I_{SS\ error} = \lim_{s \to 0} s\, I_{error}(s) = \frac{k}{p}$$

Since the resistor has a tolerance of 1%, it is reasonable to require the tracking error to be as accurate, thus $p = (RC)^{-1} = 100k = 100V_L/L$. A typical value is $RC = 1\mu s$. The results of different RC's is shown in Figure 7-16.

7.3 Load

The final stage of processing often occurs in the transducer itself. For example, many electrostatic transducers are resonant structures in which a memebrane of foil is stretched over a pocket of air. The mass of the foil and the compressibility of the air can be modeled as in Figure 7-17.

The energy in the electric field is $U_e = \frac{1}{2}CV^2 = \frac{\epsilon AV^2}{d}$, so $F = \frac{dU_e}{dd} = -\frac{\epsilon AV^2}{2d^2}$. Typically the transducers are operated with a large bias voltage V_B, so $V^2 = (V_B + v)^2 \approx V_B^2 + 2V_B v$. If we equate force with current and velocity with voltage, we can model the foil mass as a capacitance M. The air in the pocket acts as a spring with coefficient k, which we model as an inductance k^{-1}. The overall transfer function from signal voltage input v to foil velocity output is approximately given by

$$H_{load}(s) = \frac{velocity\ u(s)}{voltage\ v(s)} = \left(\frac{2V_B A}{2d^2} \right) \frac{s}{k + Ms^2}$$

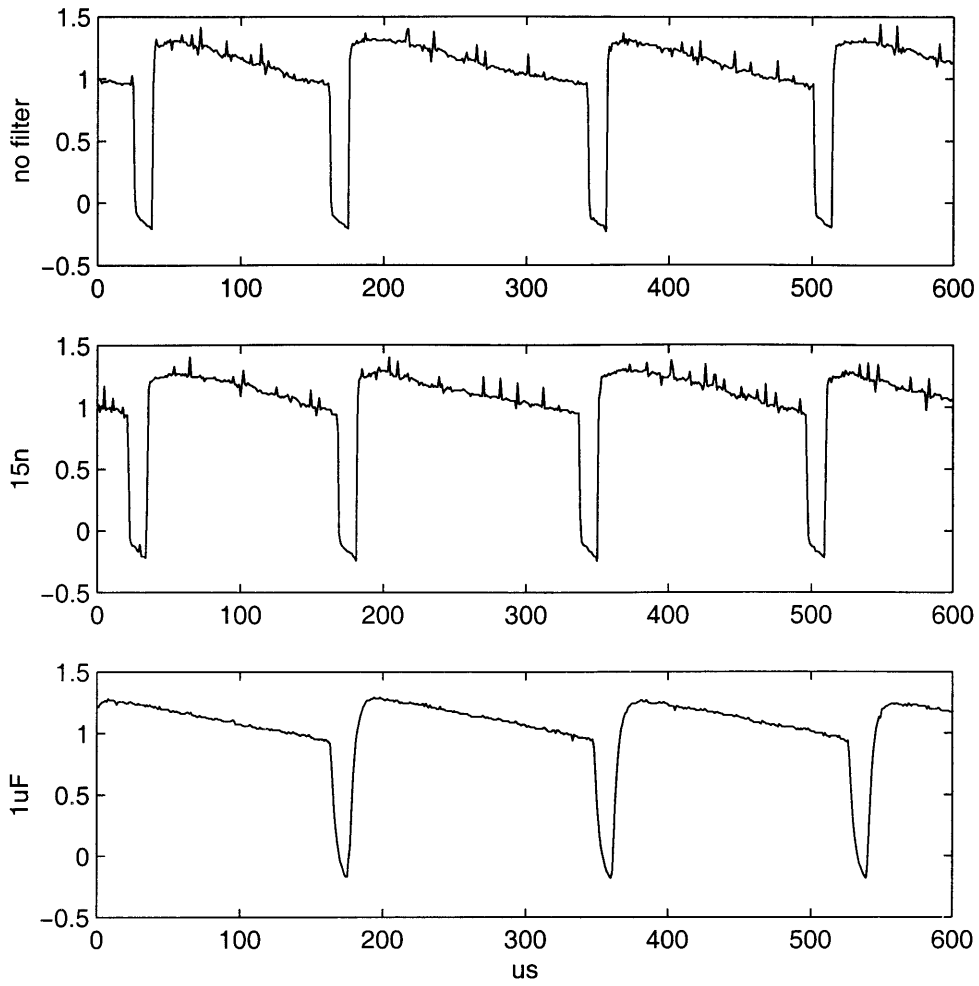The main point is there usually are physical mechanisms in transducers that filter

88

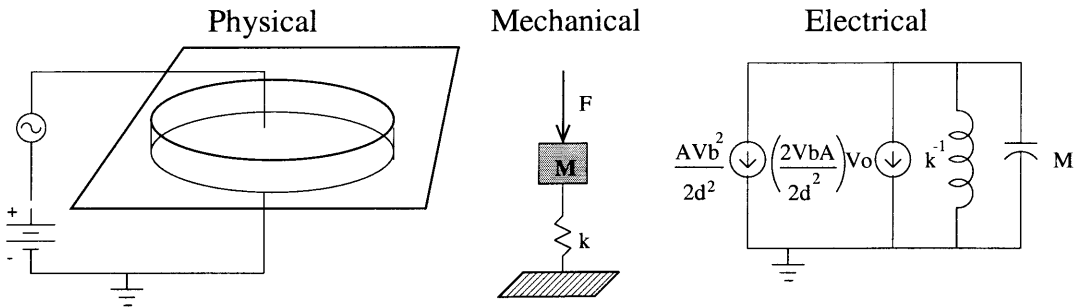Figure 7-16: Filtering of Current Sense Output

Figure 7-17: Simple Model of Electrostatic Transducer

89

out some out-of-band noise generated by the switching actions. This isn't always the case, but when it is it should be considered as part of the system transfer function. (For the measured output of an electrostatic ultrasound transducer, see Figure 8-11.)

7.4 Control Logic

A Xilinx FPGA Spartan XCS10 was used to implement both clocked and asynchronous control schemes. The chip runs on a $5V$ power supply, while the converter operates on bipolar $\pm 12V$ rails. All connections into the FPGA are comparator outputs; we use exclusively LF393 dual comparators which have open collector outputs. The LF393 collectors drive loads connected to the FPGA's 5V power regulator, whic are clamped through a diode and resistor to ground so that no FPGA inputs are pulled negative. This is shown in Figure 7-18.

An unimpressive method of implementing finite-overlap switch transitions is also shown in Figure 7-18. The logic outputs from the FPGA are filtered through a low-pass RC-network to the TC4469 driver inputs for turning a switch off. For turn-on, the FPGA output bypasses the RC filter through a diode. The filter properties are variable, and therefore arbitrary-overlapping switch transitions are possible. This arrangement depends on input capacitance of the TC4469 and therefore should not be replicated with great enthusiasm.

For both CT and DT controllers, the FPGA keeps the current source in business by waiting for the current to drop below a minimum threshold and then outputting a fixed-time pulse to the transistor which charges back up the current source. This digital logic is shown in Figure 7-19.

7.4.1 Synchronous [DT] Controller

The heart of the H-bridge algorithm is shown clearly in the flowchart of Figure 7-20. The premise of this control scheme has to do with the H-bridge impelmentation given in Figure 7-13. The upper switches, S_3 and S_4, are activated through an isolated gate drive using a pulse-coupling transformer. This is inevitably slower than connecting
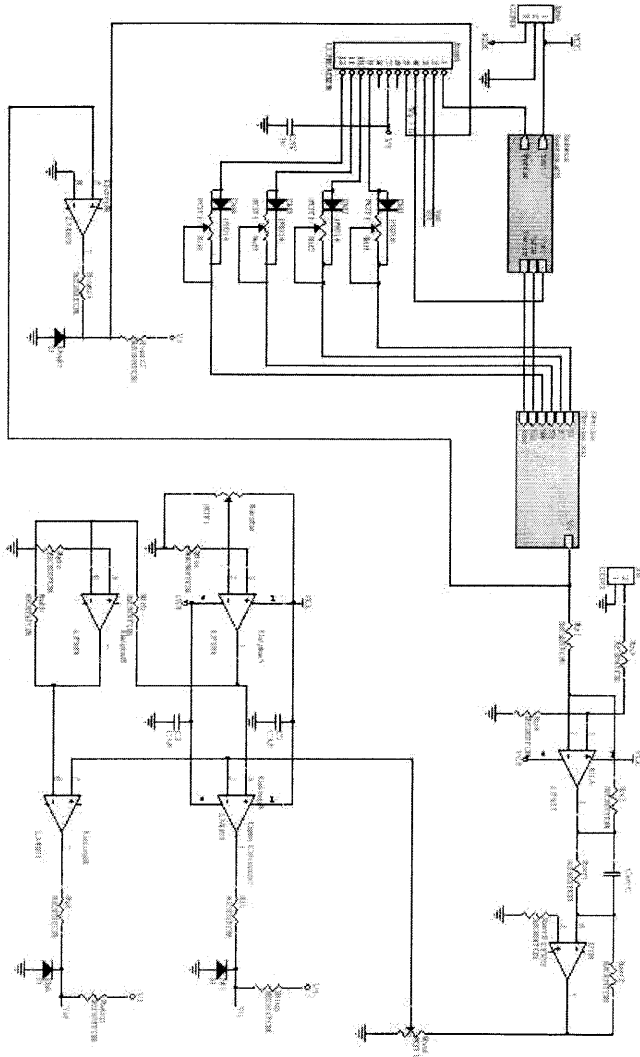
90

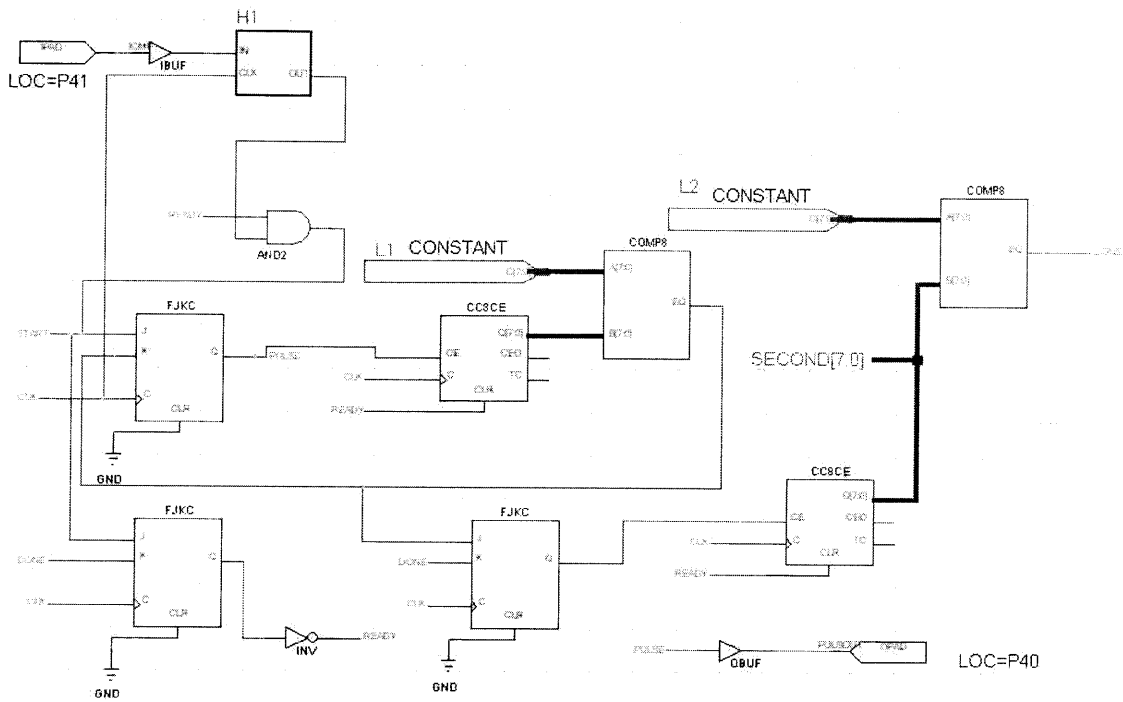Figure 7-18: FPGA interface with analog loop components.

Figure 7-19: Current Source Control Logic

the gates directly to the driver output pins. Therefore, when possible, the lower switches S_1 and S_2 do more of the work.
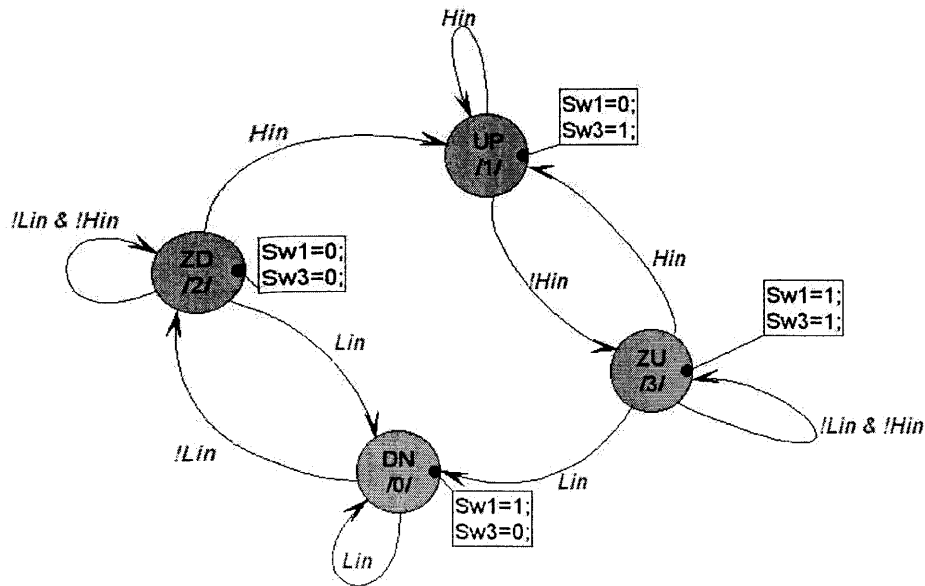

Figure 7-20: Flowchart of Switch States

As seen in Figure 7-7, turning on $S_2 + S_3$ charges the load positively and $S_1 + S_4$ charges it negatively. However, there are two choices for diverting the current around the load, namely $S_1 + S_3$ or $S_2 + S_4$. Because we are switching above the signal bandwidth, most of our outputs will either increase or decrease monotonically over many switching intervals, resulting in either a staircase-upwards or staircase-downwards output voltage. Either staircase can be generated by holding one of the high-side switches constand and toggling the lower two back and forth. That is, the direction of integration is selected by the slower S_3 and S_4, and the rate of change is controlled by the faster S_1 and S_2. Figure 7-20 implements this. The flowchart is connected to the analog components with the stunning simplicity depicted in Figure 7-21.

Another hack of the DT implementation is designed to keep the gain of the quantizer at unity. This is important for the DT algorithms. The discrete time equation for the integrator is $y[n] = y[n-1] + \frac{I_{dc}(t)T_s}{C}u[n]$. For an analog implementation with
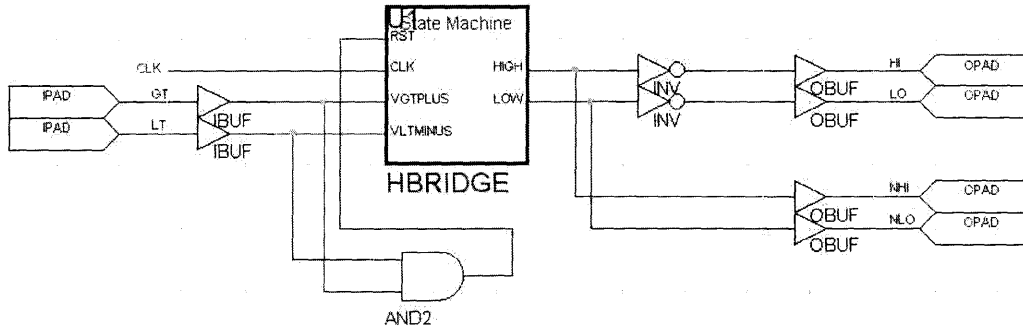
Figure 7-21: Connection of State Control to Analog Board

current $I_{DC}(t)$ and output capacitor C_L, this leads to an I/O transfer function

$$\frac{Y(z)}{X(z)} = \frac{\frac{I_{dc}(t)T_s z^{-1}}{C\alpha(1-z^{-1})}}{1 + \frac{I_{dc}(t)T_s k z^{-1}}{C\alpha(1-z^{-1})}} = \frac{IT_s z^{-1}}{C\alpha(1 - z^{-1}) + T_s I_{dc}(t)k z^{-1}}$$

We see that unless $\frac{T_s I_{dc}(t)k}{C\alpha} = 1$ we get more than a simple delay filtering the signal. Unfortunately none of our current sources are ideal, but they vary slowly in time, i.e. $I_{dc}(t) \neq const$. A cheap hack is implemented in the FPGA controller which keeps $\alpha \propto I_{dc}(t)$. A 4-bit DAC generates the threshold α to the comparators which quantize the error signal. When the current source is given the refresh pulse, the input to the α-DAC is reset to it's full value. It linearly decreases this threshold value in unison with the decaying current until the next refresh. The FPGA logic is shown in Figure 7-22.

7.4.2 Asynchronous [CT] Controller

It must be understood that there is really no such thing as a completely synchronous contol system – there will always be a natural world with no concept of the system's local oscillator. In clocked controllers, the real world is synchronized with input registers, and the beauty of the clocked state machine is that it consolidates all of the asynchronous race conditions into setup and hold requirements of the state flip-
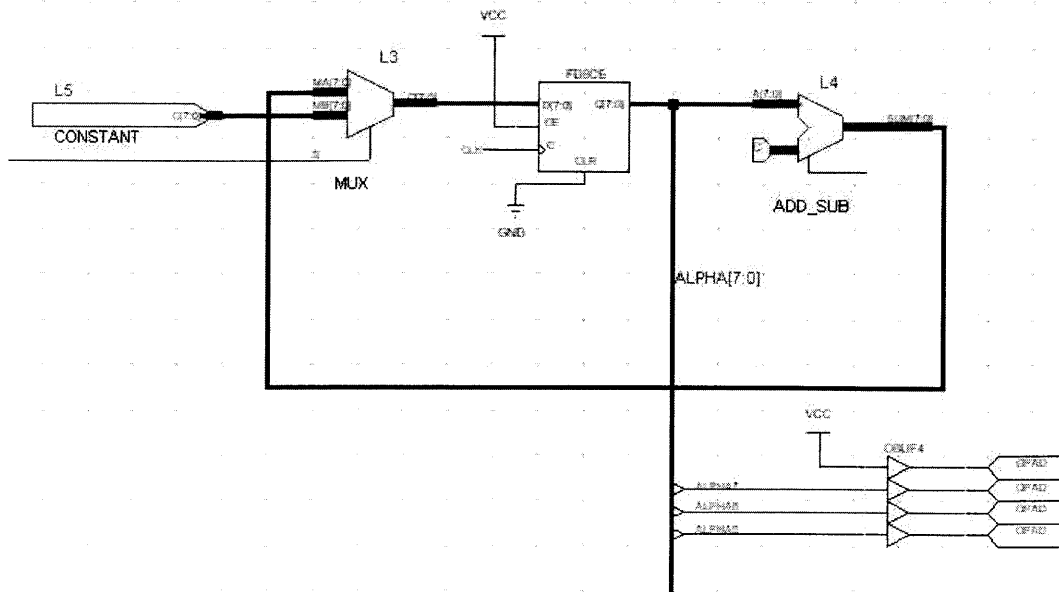
Figure 7-22: Compensation for Current Source Decay

flop registers [Tre92]. We determine a maximum propagation delay path, add a little margin, and clock away, right?

Let me remind you that we aren't making a adding machine, we're trying to amplify an analog wave with as high accuracy as possible. The discrete-time section demonstrated the use of a second-order DPCM loop with one order of noise shaping for the quantization noise. And for high-frequency signals, the switching speeds of transistors along with logic delay results in marginally effective oversampling ratios. With clocked control, we lose the information about how far into a clock period the signal transitions occur. This form of "rounding" transition times to the nearest clock period will show up as modulations on the output signal (see Section 4.5).

Fortunately we have two inputs to the switching logic which have one essential property – they will never be concurrent. The inputs are comparisons $H_{in} = V_{out} > \alpha$ and $L_{in} = V_{out} < -\alpha$. Unless $\alpha = 0$ these two conditions cannot change at the same time. All we have to worry about is state transition race conditions, which can be avoided with the use of wait and delay structures. Furthermore, the lower

and upper switch pairs are always in complementary states, meaning that the entire output state is 2 bits, which we take as Moore-type output structures to avoid output race conditions [Tre92].

A Karnaugh map can be used to ensure the following: [Tre92]

- Each delay or wait transitions to an adjacent state.

- All in-arrows to a branch-wait state must cover the wait condition.

- All state equations must be hazard corrected so at least one product term remains active.

$Q_0Q_1 \setminus H_{in}L_{in}$	00	01	11	10
ZD	0	0	X	1
UP	1	1	X	1
ZU	1	0	X	1
DN	0	0	X	1

$$Q_0 = H_{in} + \overline{L_{in}}Q_0 + Q_0\overline{Q_1}$$

$$Q_1 = L_{in} + \overline{L_{in}H_{in}}Q_0 + Q_0Q_1\overline{H_{in}}$$

This is implemented in Figure 7-23.

7.5 Layout

The analog circuitry is laid out as shown in Figure 7-24. The basic rules of PCB layout were [hopefully] adhered to; some of the important points are pointed out here:

- The main current loop involving the sourcing inductance, the h-bridge switches and the sensing resistor are laid out as close to each other as possible, not only to minimize resistive loss but also to minimize parasitic inductances from adding spikes to the switching waveforms.
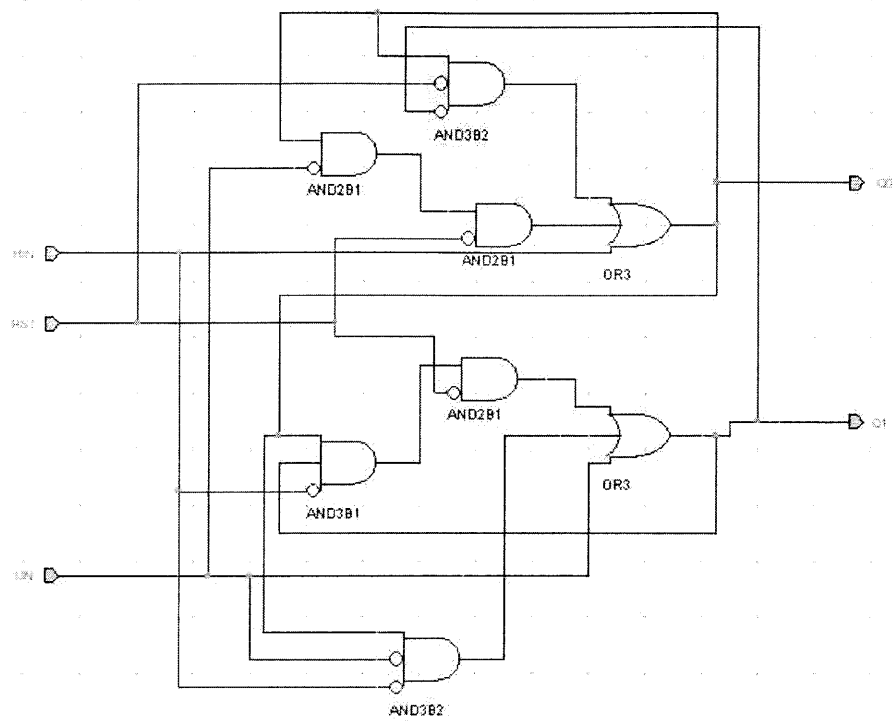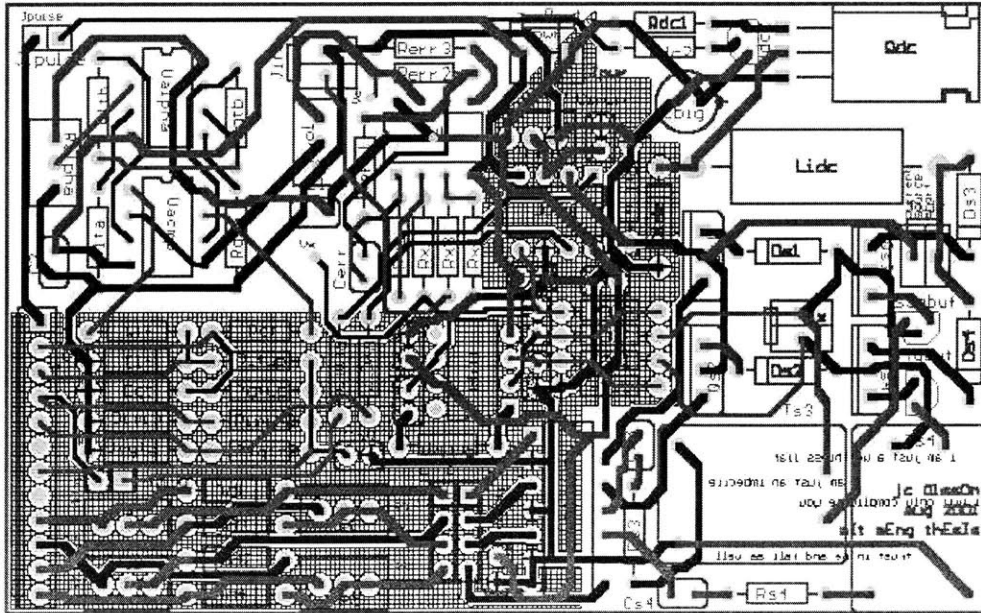
96

Figure 7-23: State Logic for Asynchronous Controller

Figure 7-24: PC-Board Layout

- The gate drives from the TC4469 to MOSFETS S_1 and S_2 as well as the pulse-coupling transformers are laid out close together including the ground return paths to the chip and its bypassing capacitor.

- All of the digital waveforms on the circuit are placed over a ground-planed area to minimize return-path inductance.

Chapter 8

Measurements

Two switching converters were built and tested, one using a DT implementation of Figure 4-7, and one based on the asynchronous controller of Figure 7-23. Test signals used were sinewaves between 25kHz and 75kHz. The feeback amplifier has a gain of $A = 27k/2.2k = 12.3$, so an input signal of $10V_{pp}$ corresponds to an output of $123V_{pp}$. Most of the measurements shown were taken using a LabView data acquisition [1].

8.0.1 DT System Results

Figure 8-1 shows an unimpressive spectrum and extremely rough ouptut waveform for the discrete time implementation. (The gaps in the output waveform ocurring at $180\mu s$ intervals are due to the current-source recharging, and thus shouldn't be blamed on the control scheme as well).

The DT system used a transformer current source whose waveforms are shown by the current-sense resistor in Figure 8-2. The current source is relatively constant for about $180\mu s$, and then is refreshed for $12\mu s$. Unless a type of standby current-source is available to stand in while the energy is restored to the primary source, these gaps add in low-frequency noise and are generally unacceptible.

---

[1] The recording and processing was done by a monster piece of LabView code written by fj Pompei
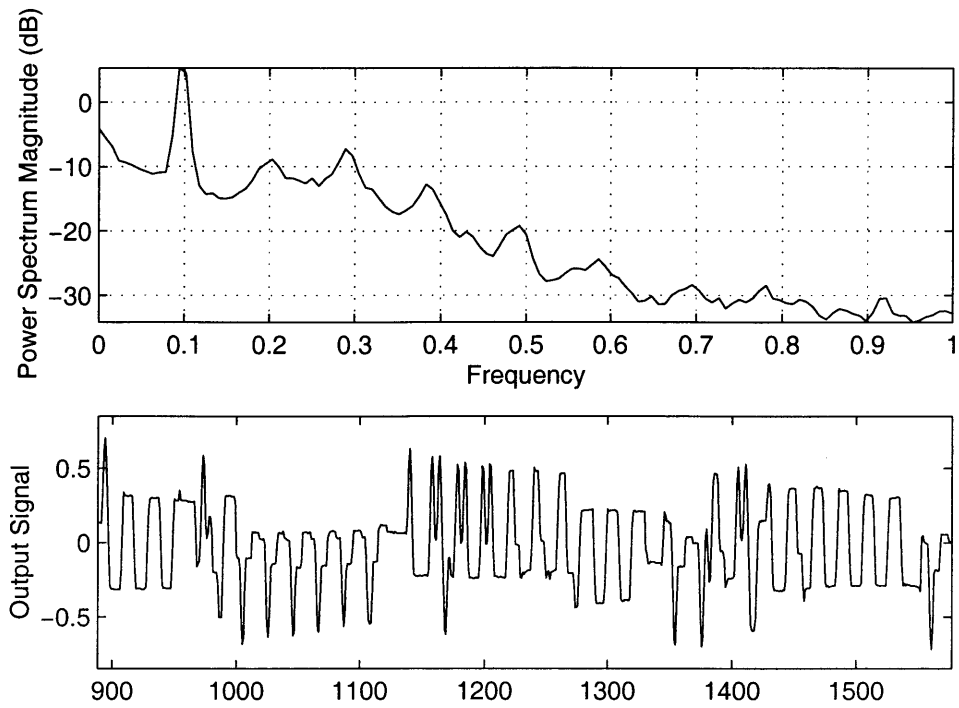
Figure 8-1: Spectrum and Output of DT Converter

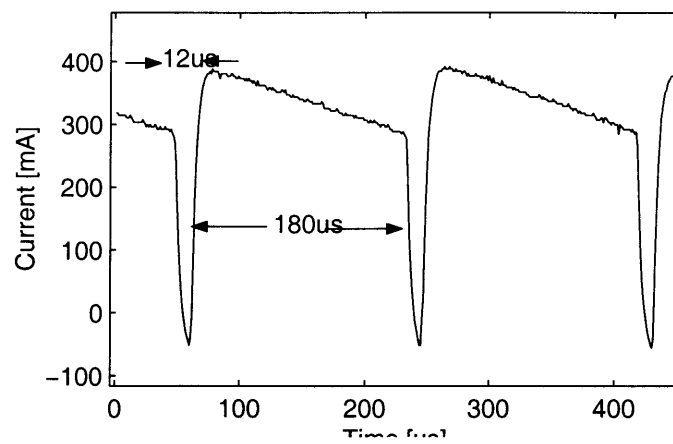Figure 8-2: Transformer Secondary Current Waveform

100

8.0.2 CT System Results

Figure 8-3 shows the measured THD at the transducer for a 50kHz an input amplitude range of $[3V..15V]$, which corresponds to the output peak-to-peak range of $[37V..200V]$. For inputs below $6V$ the minimum output step size, as determined by error threshold $\alpha = 1V$ and loop delay, is still too large for the signal, thus distortion increases dramatically.
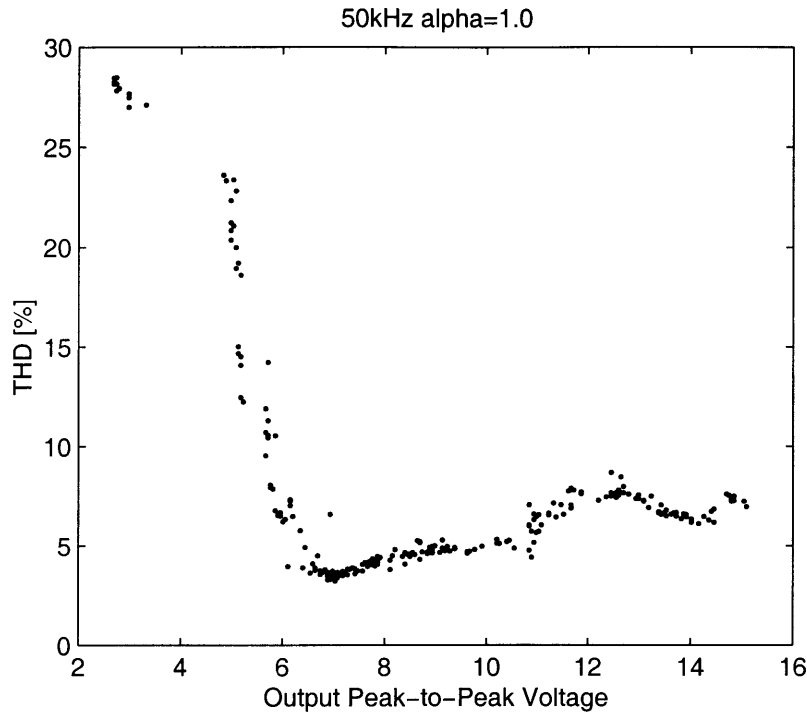

Figure 8-3: THD vs.Amplitude, f=50kHz, $\alpha = 1.0$

By decreasing the threshold α from the 1.0V just shown to 0.75V, shown in Figure 8-4, distortion is reduced by 2-3% across most of the amplitude range. In addition, as the amplitude of the input signal is reduced the distortion doesn't spike up until a 6V, whereas for $\alpha = 1.0$ in the prior figure it occurs at 7V.

Naturally the distortion is inversely related to the error threshold α. Figures 8-5 - 8-7 show total harmonic distortion of the output signal as related to frequency for $\alpha = [0.75, 1.0, 1.25]$.
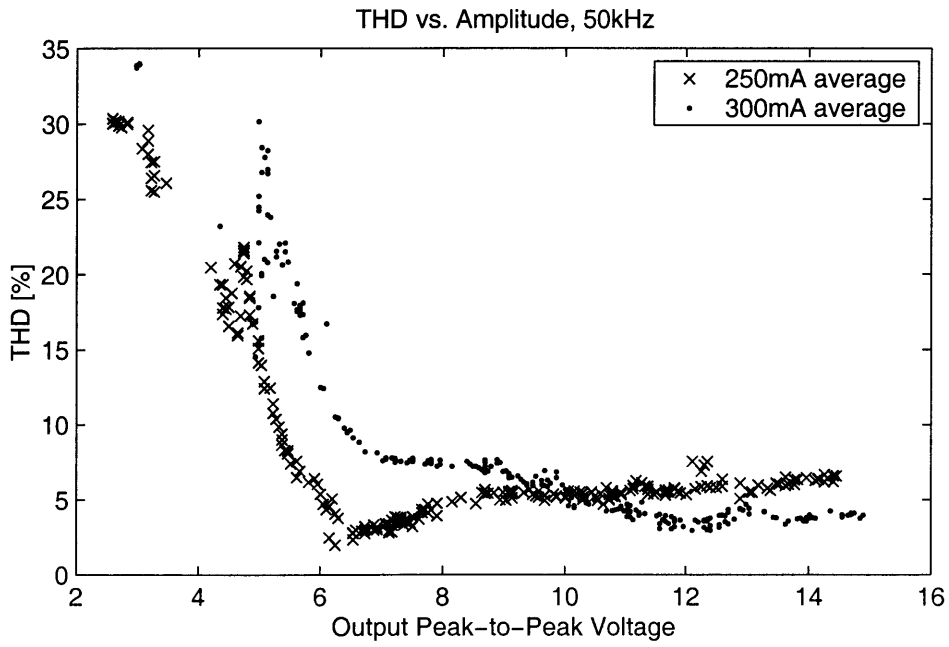
Figure 8-4: THD vs. Amplitude, f=50kHz, $\alpha = 0.75$

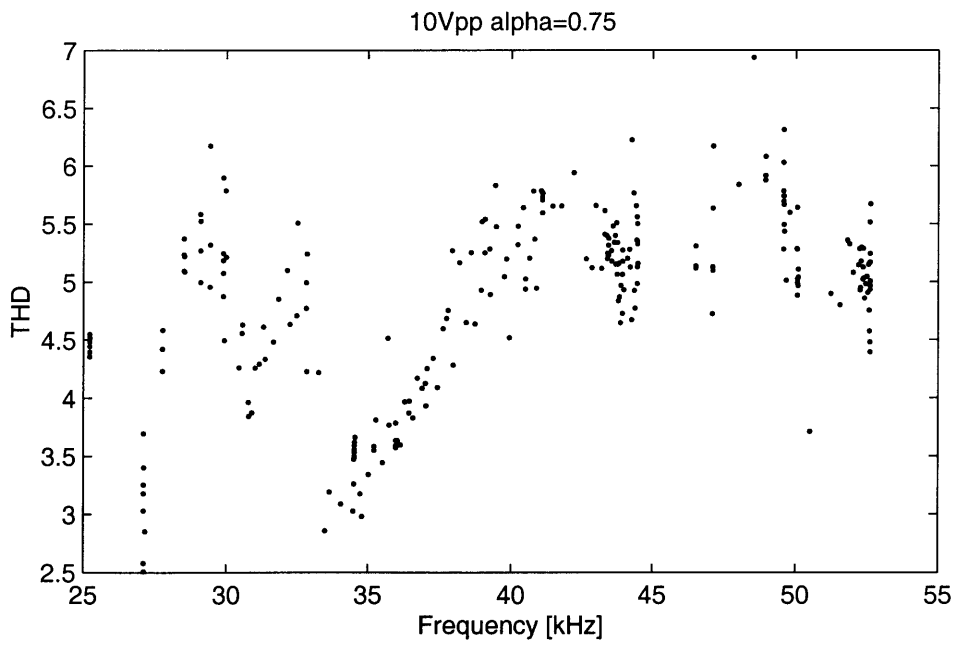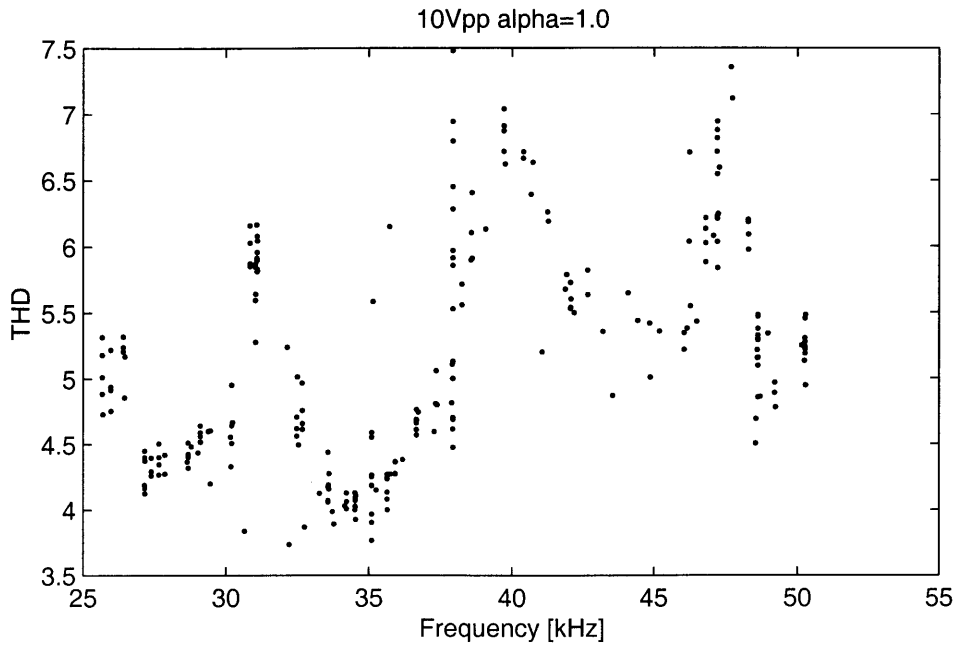Figure 8-5: THD vs. Frequency for $\alpha = 0.75$

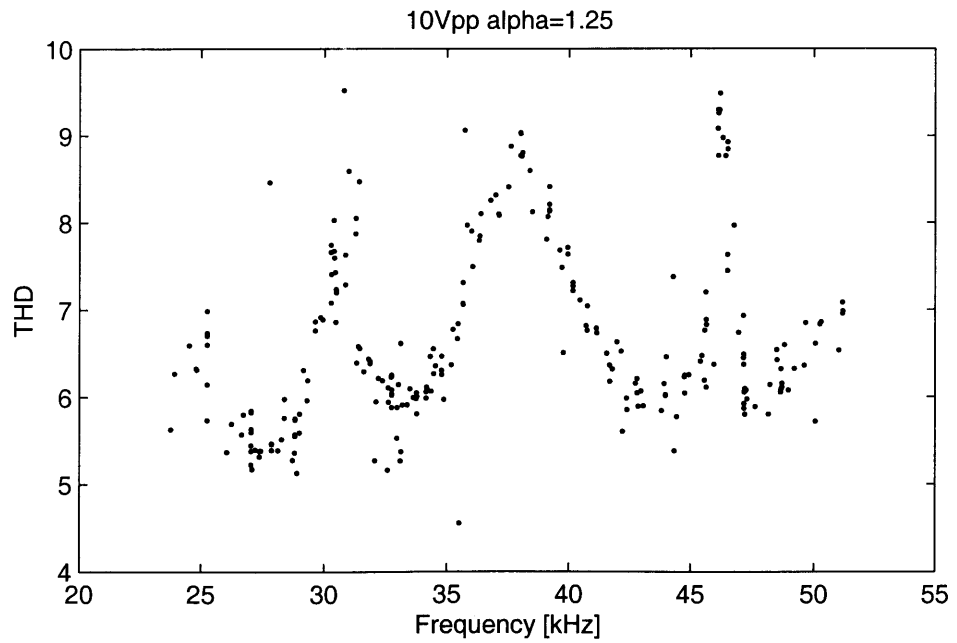Figure 8-6: THD vs. Frequency for $\alpha = 1.0$

Figure 8-7: THD vs. Frequency for $\alpha = 1.25$

103

For a switching curve based on the error function $s = (e + \lambda\dot{e})$ we use the transfer function of the error implemented as $H_e = (1 + R_{err}C_{err}s)$ (see Figure 7-18 and Section 6). The values of $R_{err} = 5.1k\Omega$ and $C_{err} \in \{22pF, 100pF, 220pF\}$ were used and distortion versus frequency for each is shown in Figure 8-8.
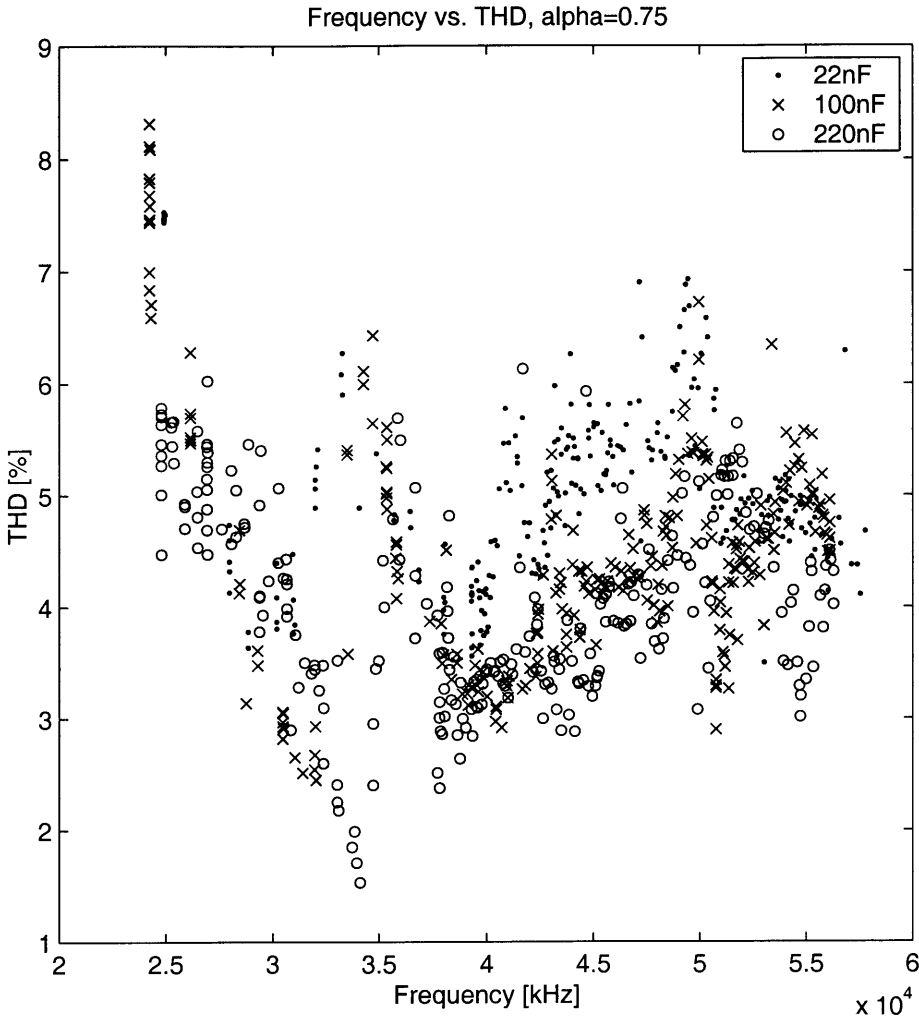

Figure 8-8: THD vs. Frequency for Different Switching Curves

The current waveforms of the 1.0mH inductor are shown in Figure 8-9 as measured across the inductor sense resistor. The current ramps down from 400mA to 150mA every $200\mu s$. From this we can estimate the average DC voltage presented to the current source by the switch R_{DS} voltage drops

104

$$V_{DC} = \frac{L\,\Delta I}{\Delta T} = \frac{1mH\,250mA}{200\mu s} = 1.25V$$

The current decay period of $200\mu s$ is 20 times longer than the period of the input signal, and 100 times longer than transistors operating at 500kHz. The DC source is then connected positively across the inductor for about $21\mu s$ (difficult to see on the figure).
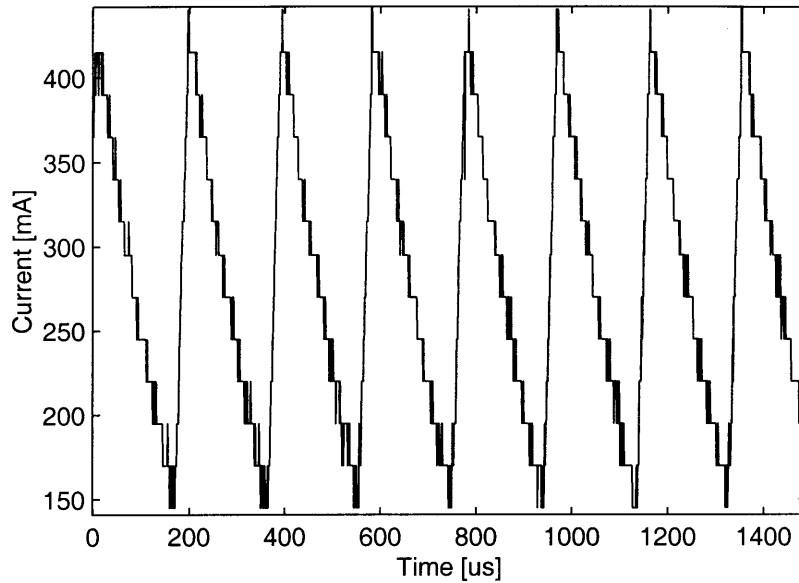
Figure 8-9: Current Waveform of Inductor

The transducer load waveforms as measured after the output amplifier are shown in Figure 8-10 for a 46kHz signal with the *alpha*-threshold set to a modest 1.0V. Remember that the output amplifier has a first-order low-pass filter set to 100kHz as shown in Figure 7-14.

An electrostatic ultrasound transducer was connected as a load and a wide-band B&K ultrasound microphone was used to measure the sound pressure at approximately 1m away. These curves are shown in Figure 8-11. The smooth filtering of the
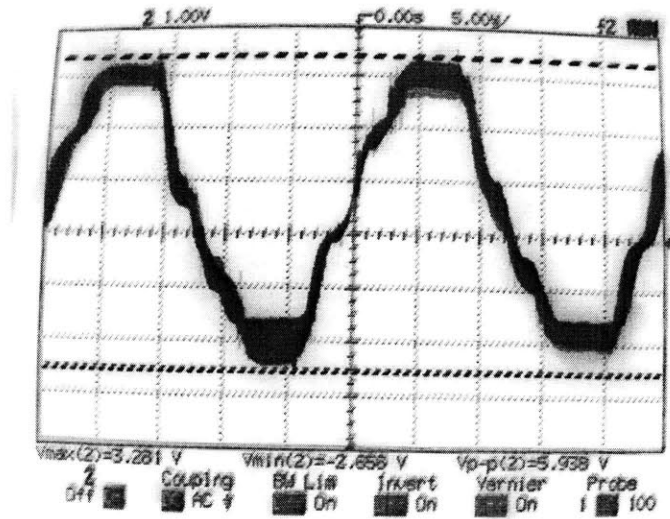
Figure 8-10: Ultrasound Transducer Input Waveform

60mA	XESS Evaluation board with Xilinx FPGA
20mA	quiescent current of analog control
36mA	current source shorted across H-bridge
100mA	100Vpp 50kHz (approx 1.2Watts)
150mA	200Vpp 50kHz (approx 1.8Watts)

Table 8.1: Power Consumption by Amplifier

waveforms is due to physical transfer functions of the load (see Section 7.3).

The power consumed by this setup off of 12V supplies is accounted for in Table 8.0.2.

Figure 8-11: Ultrasound Measurement at Microphone

8.1 Conclusion

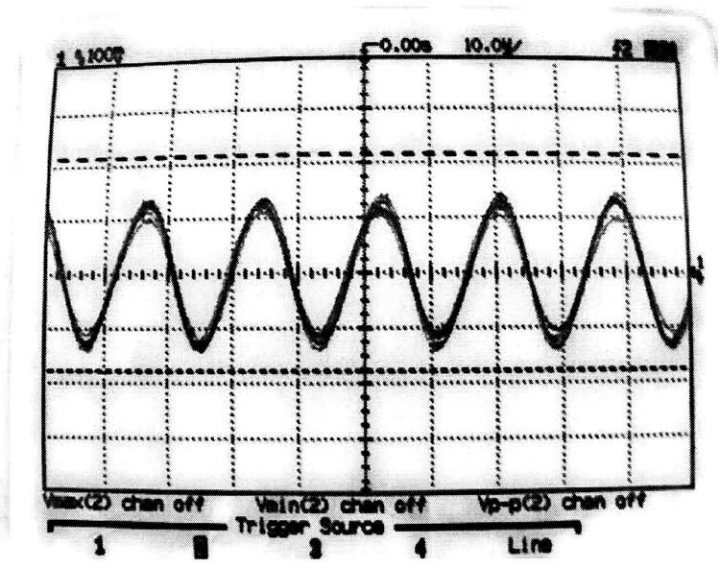The constant-slope integrating converters described in this thesis have proven to be extremely viable options for capacitive transducer drives. Most notable was the $200V_{pp}$, 50kHz operation of a 13.5nF capacitive transducer at 1.8W! Prior methods consume over 30Watts for the same results, showing over an order of magnitude of power savings. However, unlike the conventional methods which increase dissipation linearly with frequency and with amplitude squared, dissipation of the converter discussed in this thesis is to a first order completely independent of both frequency and amplitude!

Several control schemes and current-source implementations were considered and built; the measurements were given in the previous section. A few conclusions can be made about how to build and operate this efficient amplifier.

8.1.1 Control Schemes

The discrete-time implementation was a bit of a disappointment. The output and spectrum are shown in Figure 8-1. The OSR was simply not high enough to move noise energy out of the signal band. For example, if we have a 50kHz signal on a 500kHz converter – there are ten clock periods per signal period, uniformly distributed over one sinusoid interval. Furthermore, between each interval we can only move a fixed amount up or down (Figure 4-5 shows this picture pretty well).

The PWM algorithm explored purely in simulations shows much promise (see Figure 5-11 for a 50kHz output). Additionally, the control logic for the implementation is extremely simple. However, PWM methods are inherently non-optimal, especially as the the output frequencies near the switching frequencies.

The asynchronous controller is clearly the hero. We design a switching curve designed to track the input signal; the error between the input and output ends up as a nonlinear limit cycle due to loop delay and minimum switching time requirements of the transistors. Theoretically, the error can be reduced to zero as the implementing devices approach divine operation.

108

8.1.2 Current Source

Both a flyback-style transformer and an inductor were built and tested as viable current sources. The transformer current source had an inductance about an order above the inductor-version, so the waveforms of current decay can't be fairly compared.

Figure 8-9 exhibits quite a large range of operating currents, and performance could be improved by using a larger inductor. However, the distortion curves measured using this value speaks for the robustness of the control loop. The discrete-time control scheme is much more sensitive to variations in loop gain, of which I_{DC} is a factor. The danger of allowing such low current levels is that it imposes an upper limit on output amplitude. A larger inductor can decay proportionally less over the same current cycle interval, and from Equation 7.3 we can see that the duty cycle is independent of inductor value. In fact, the only quantity that varies unfavorably with increasing inductance is cost.

The transformer currernt source allows for expansion towards multiple-slope converters (i.e. Figure 7-6). However, for single-slope converters such as the one built for this thesis, the inductor source is preferrable because there are no "re-energizing" breaks during which output is either frozen or supplied by a secondary source. The inductor current source is maintained by a simultaneous buck converter with a control loop slow enough to ignore any of the signal frequencies.

8.1.3 Overall

The conclusion of this thesis is that integrating switching converters for capacitive loads are the solution for high-frequency, high-voltage, broadband operation. The switching amplifier takes up very little space – a magnetic storage device, a few opamps and a logic chip – so that it can be built onto the trandsucer assembly itself. It also uses very little power, allowing for battery-powered operation.

Bibliography

[AF66]   Michael Athans and Peter Falb. *Optimal Control*. McGraw-Hill, 1966.

[Ape]   Apex Microtechnology. *App Note 25: Driving Capacitive Loads*.

[Ber95]   Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, 1995.

[Coo94]   Peter A. Cook. *Nonlinear Dynamical Systems*. Prentice Hall, 1994.

[FL68]   Irmgard Flugge-Lotz. *Discontinuous and Optimal Control*. McGraw-Hill, 1968.

[Inta]   International Rectifier. *App Note 950: Transformer-Isolated Gate Driver Provides very large duty cycle ratios*.

[Intb]   International Rectifier. *App Note: Power MOSFET Basics*.

[Intc]   International Rectifier. *Use Gate Charge to Design the Gate Drive Circuit for Power MOSFET and IGBTs*.

[Jay84]   Nuggehally S. Jayant. *Digital Coding of Waveforms*. Prentice-Hall, 1984.

[KSV91]   John G. Kassakian, Martin F. Schlect, and George C. Verghese. *Principles of Power Electronics*. Addison-Wesley, 1991.

[Nat]   National Semiconductor. *LM2622 600kHz/1.3MHz Step-Up PWM DC/DC Converter Datasheet*.

[OS99]   Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, Inc., 2 edition, 1999.

[PBSA87] John J. Paulos, Gregory T. Brauns, Michael B. Steer, and Sasan H. Ardalan. Improved signal-to-noise ratio using tri-level delta-sigma modulation. In *IEEE Proc. ISCAS*, pages 463–466, 1987.

[Rab99]   Shahriar Rabii. *The design of low-voltage, low-power, Sigma-Delta modulators*. Kluwer Academic Publishers, 1999.

[Rya82]   E. P. Ryan. *Optimal Relay and Saturating Control System Synthesis*. Institution of Electrical Engineers, 1982.

[SL91]   Jean-Jacques E. Slotine and Weiping Li. *Applied Nonlinear Control*. Prentice-Hall, Inc., 1991.

[Tel]   TelCom Semiconductor, Inc. *TC4467/8/9 Datasheet*.

[Tre92]   Michael Treseler. *Designing State Machine Controllers Using Programmable Logic*. Prentice-Hall, Inc., 1992.