Class D amplifier with AGC

GÖRAN NORDAHL

Master of Science Thesis in Radio Electronics Stockholm, Sweden 2005

IMIT/LECS-2005-13

Class D amplifier with AGC

Master of Science Thesis in Radio Electronics

Göran Nordahl

Stockholm 2005

Supervisor: Mathias Hemberg (ONE) Examiner: Prof. Håkan Olsson (KTH)

Abstract

The primary goal of this Master Thesis is to design an audio amplifier of class D as a pre-study for a future ASIC. The secondary goal is to equip the proposed design with an Automatic Gain Control (AGC), primary to suppress audio transients that can be harmful to the ear. Due to protection of the customer a limiter with variable clipping level is implemented instead of an AGC.

The ASIC is intended to replace an amplifier of class AB and some of its surrounding components in a product family currently in mass production. The main reason for the chosen amplifier topology is reduced power consumption. The product family is battery powered making efficiency a major concern. Class D offers in theory an efficiency of 100%. In reality the efficiency is lower, 85-90% is a more realistic value. Class AB only achieves 60-65%. Implementing the design in an ASIC will cut costs in the long run.

An amplifier of class D uses Pulse Width Modulation (PWM) to represent its output signal. The desired audio signal is restored by low pass filtering. There are two basic PWM modulation schemes; bipolar and unipolar PWM. Bipolar PWM offers the best audio performance but is less efficient than unipolar PWM. A combined modulation scheme that I call "biunipolar PWM" is implemented and evaluated in a prototype built using surface mounted components on a PCB. Measurements show that biunipolar PWM combines features of bipolar and unipolar PWM and is the way to go for the intended ASIC. Switching between unipolar and biunipolar PWM is easy. This opens up for the possibility with two selectable PWM modulation schemes for the ASIC. One efficient modulation scheme using unipolar PWM to save power. The second modulation scheme uses biunipolar PWM which achieves better audio performance at the cost of increased power consumption.

The efficiency for the prototype is low, approximately 50%. Implementing the amplifier in an ASIC will yield higher efficiency due to the ability to get custom made building blocks.

Contents

Ι	Ba	ckgrou	ınd														4
1	Intr	oducti	on														5
2	The	ory															7
	2.1	Ampli	fiers														7
		2.1.1	Class A														7
		2.1.2	Class B														9
		2.1.3	Class AB														9
		2.1.4	Class D														10
	2.2	Radio	Frequency Interference	e - F	RFI		 										11
	2.3	AGC	and Limiter														12
		2.3.1	Hard limiter														12
		2.3.2	Soft limiter														13
		2.3.3	Variable limiter				 •	•					•	•		•	13
3	Lite	rature	review														15
	3.1	Histor	у														15
	3.2	Transi	stors are not ideal swit	che	s.												15
	3.3	\mathbf{Pulse}	Width Modulation - P	WN	I.												16
		3.3.1	Bipolar PWM														17
		3.3.2	Unipolar PWM														18
		3.3.3	Biunipolar PWM			·	 •	•	 •	•		•	•	•	•	•	19
4	Mat	erials	and methods														21
	4.1	Buildi	ng blocks														21
		4.1.1	Analogue section														22
		4.1.2	Digital section														22
		4.1.3	H-bridge		• •	•	 •	•	 •	•	•	•	•	•	•	•	26
тт	Т.																07
11	In	plem	entation														27
5	Imp	lemen	tation														28
	5.1	Discre	te implementation														28

	5.2	Estimated power consumption	30
		5.2.1 Op-amps, comparators and H-bridge	30
		$5.2.2$ Logic gates \ldots	31
		5.2.3 RC-links	32
		5.2.4 Total supply power without load	32
	5.3	Estimated output power	32
	5.4	Estimated maximum efficiency	34
6	\mathbf{Res}	ults	35
	6.1	Efficiency	36
	6.2	Total Harmonic Distortion + Noise (THD+N)	37
	6.3	Variable limiter	40
	6.4	PWM generation	42
		6.4.1 Unipolar PWM	43
		6.4.2 Biunipolar PWM	46
7	Disc	cussion	48
	7.1	Efficiency	48
	7.2	THD+N	48
	7.3	Limiter	49
	7.4	PWM	50
8	Con	clusions	51
9	Rec	ommendations	53
\mathbf{A}	\mathbf{Sch}	ematic	57

Glossary of Terms AGC Automatic Gain Cont

AGC	Automatic Gain Control
ASIC	Application Specific Integrated Circuit
BJT	Bipolar Junction Transistor
CMOS	Complementary Metal Oxide Semiconductor
DMOS	Double Diffused Metal Oxide Semiconductor
DSP	Digital Signal Processor
MEMS	Micro Electro Mechanical System
MOSFET	Metal Oxide Semiconductor Field Effect Transistor
ONE	O-Network Engineering AB
PCB	Printed Circuit Board
PWM	Pulse Width Modulation
RFI	Radio Frequency Interference

Part I Background

Chapter 1

Introduction

Audio amplification comes in handy whenever you want to communicate over a greater distance than you can shout. There are many other areas where audio amplification plays a vital role. Movies with Charlie Chaplin works well without it, but Star Wars without sound only looks silly. Nowadays many electronic consumer products are supposed to be portable. Low weight and low power consumption has thus become very important. In many of these products the audio amplifier stands for a major contribution to the power consumption. Efficient audio amplification is therefore of great interest.

An ideal audio amplifier amplifies its input signal and deliver power to the load (normally a speaker) without any losses or distortion. In other words; it multiplies the input signal by a constant. Unfortunately the real world makes it hard to meet the requirement of both efficiency and linearity (i.e. low distortion) at the same time. Class A, B, AB and D are the most common audio amplifier topologies of today (class C exists, but is not intended for audio). Except for class D they all generate their output signal by wasting a varying amount of power as heat, i.e. they work in a way similar to linear voltage regulators. Class D works more like a switch mode power converter and as such it is much more efficient.

The primary goal of this Master Thesis is to design an audio amplifier of class D as a pre-study for a future ASIC. A secondary goal is to equip the proposed design with an Automatic Gain Control (AGC), primarily to suppress audio transients that can be harmful to the ear. The ASIC is intended to replace the current class AB amplifier and some of its surrounding components in one of ONE:s product families. The main reason for the chosen amplifier topology is reduced power consumption. Class D is more efficient than class AB. Implementing the class D amplifier in an ASIC will cut costs in the long run. Class D amplifiers has been around for a while and are available from several manufacturers but none fulfill all of ONE:s needs. The desired ASIC will do more than just providing efficient amplification. Except for suppressing harmful transients the AGC is also supposed to improve the audibility of spoken words in a noisy environment. The AGC is not part of this report due to protection of the customer. A limiter with variable clipping level is implemented instead.

This pre-study does not treat the implementation on silicon. Instead the proposed design is implemented using discrete components surface mounted on a PCB.

Chapter 2

Theory

2.1 Amplifiers

The desired function of an ideal audio amplifier is to multiply its input signal by a constant. In the real world the amplifier also introduces losses and adds distortion to the signal. Several amplifier topologies has emerged during the years but are only approximations to the ideal amplifier. Class A, B, AB and D are the most common topologies of today (class C exists, but is not intended for audio). Except for class D they all generate their output signal by wasting a varying amount of power as heat, i.e. they work in a way similar to linear voltage regulators. Class D works more like a switch mode power converter and as such it is much more efficient.

2.1.1 Class A

When trying to explain the operation of a very basic amplifier Ohms law is a good start:

$$U = R \cdot I \tag{2.1}$$

It clearly states that the voltage U over a resistor has linear dependence both of the resistance R and the current I flowing through it. If two resistors are connected in series they form a voltage divider (see figure 2.1). The available voltage is divided between the two resistors as their relative resistances:

$$V_{cc} = U_{1} + U_{2}$$

$$U_{1} = R_{1} \cdot I$$

$$U_{2} = R_{2} \cdot I$$

$$\frac{U_{1}}{U_{2}} = \frac{R_{1}}{R_{2}}$$

$$U_{2} = \frac{R_{2}}{R_{1}} \cdot U_{1}$$
(2.2)



Figure 2.1: Voltage divider

As seen in equation 2.2 any change of the relative resistances will affect the output voltage U_2 . When replacing one of the two resistors with a transistor the result is a very basic amplifier of class A:



Figure 2.2: Amplifier of class A

A small variation in the base-emitter voltage will produce an amplified, but inverted, replica at the output of the voltage divider made up by R_1 and Q_1 . As long as the transistor is working in its linear region this amplifier topology is very linear. The main drawback with class A is the required bias current for its operation, i.e. it conducts current at all times. Therefore class A is far from being an efficient amplifier topology. One might think that reducing the bias current would do the trick, but then there is no power available to be delivered to the load. An amplifier is most often intended to deliver lots of power to its load (low output impedance) and this requires a high bias current. Due to this amplifiers of class A are only used when linearity is the main concern and there are lots of power available to be wasted as heat. That is why class A most often is found in Hi-Fi audio equipment but seldom in battery powered equipment. A bias network, for simplicity not shown in figure 2.2, is also needed to set the transistors operating point, i.e. making sure it is in the middle of its linear region when the input signal is zero. The theoretical efficiency maximum for class A is only 25% [1], but this is only achieved when the output stage is almost driven to saturation. That is most often not the case. The crest factor quantifies the dynamic of a signal and is defined as the ratio between peak and average output power. The crest factor for a sine wave is 3dB [2]. For normal audio signals, such as music, the crest factor may vary between 6 dB and 24dB [2]. With normal audio signals class A does not achieve higher efficiency than typically 5-10% leaving 90-95% of the supplied power to be wasted as heat.

2.1.2 Class B

If both resistors of a voltage divider are replaced by transistors the result is called a push-pull stage. When the transistors are operated in strict time alternation this topology is called class B.



Figure 2.3: Amplifier of class B

Only one transistor conducts at the time amplifying one half of the waveform each. There is no bias current through the transistors thus greatly reducing the amount of wasted power. But there is a problem called cross-over distortion clearly seen and slightly exaggerated in figure 2.3. It is caused by non overlap between the two conducting states of the transistors. This topology is efficient, max 78.5% [1], but not very linear.

2.1.3 Class AB

The problem with cross-over distortion in class B can be eliminated if both of the transistors are biased to the region just above cutoff. A bias network with two diodes can do the trick, especially if they are of the same type as the transistors, i.e. similar or even equal forward voltage drops (see figure 2.4). Another possibility is using a so called rubber diode, i.e. a transistor configuration acting as a diode with a selectable forward voltage drop.

The bias network reduces or even eliminates the cross-over distortion. Close to zero crossings both transistors will conduct. This combined with the small current in the bias network makes this topology waste more power than class B. In most applications a slightly increased power dissipation, i.e. lower efficiency, is a small price to be payed for better linearity. As the name indicates this topology combines the benefits of class A and B and gives reasonable linearity and efficiency (60-65% [3]) at the same time. Class AB is very common in all kinds of electronic audio equipment.



Figure 2.4: Amplifier of class AB

2.1.4 Class D

Even the efficient class B topology suffers from losses in the transistors. A completely different approach is to operate the transistors of a push-pull stage as switches, i.e. quickly rushing through their linear region on the way back and forth from switched off to saturation (on). An ideal closed switch has no voltage drop but a current flowing through it. An ideal open switch conducts no current but has most often a voltage drop. Power is only wasted when there is a nonzero voltage and current present at the same time (P = UI). Theoretically a transistor operated as a switch wastes no power at all.

Class D is a more complex topology than amplifiers based on the basic voltage divider or push-pull principle, also known as (more or less) linear amplifiers. Class D requires more electronics to make it work. A basic approach to class D consists of a comparator comparing the input signal with a triangle wave whose pulse width modulated output signal is fed to a push-pull stage (see figure 2.5). PWM-signals contain considerable amounts of harmonics so low pass filtering is needed to restore the original signal. In other words; the audio signal is the instant mean value of the pulse train. By using two push-pull stages, a so-called H-bridge, and feeding them with both the PWM signal and its inverse twice the output voltage swing is achieved. As seen in equation 2.3 twice the voltage swing gives four times as much output power.

$$P = \frac{U^2}{R} \tag{2.3}$$

The main benefit of a class D amplifier is its high efficiency, theoretically as high as 100%. In reality the efficiency is lower due to switching losses, nonzero voltage drop over saturated transisors and power consumption in the electronics that generate control signals for the H-bridge. Distortion is



Figure 2.5: Amplifier of class D

mainly dependent on the frequency of the triangle wave, cross-over distortion in the push-pull stage and how quick the switches in the push-pull stage can be operated. But there is also a problem with switching artifacts from the digital parts interfering with the analogue input stage. In many applications, especially those who are battery powered, efficiency is much more important than linearity. In such applications class D is a good choice. An extra benefit is the reduced weight and space due to reduced need for cooling.

Class D is sometimes referred to as a digital amplifier, but that is not really the case. The input signal to an amplifier of class D is an analogue audio signal, not binary data that would be the case in a digital amplifier. The output stage works in a digital manner but that does not make the entire amplifier topology digital. Class D amplifiers are better described as switch mode power converters with a variable control voltage, i.e. audio, as input signal. Many switch mode power converters also use PWM to generate the desired output voltage. In contrast to class D the output signal is supposed to be a stable DC voltage (AC/DC or DC/DC converter) or a sinusoid (AC/AC or DC/AC converter).

2.2 Radio Frequency Interference - RFI

All kinds of electronic equipment can disturb or be disturbed by their surroundings. This is a problem known as Radio Frequency Interference - RFI. The class A, AB and B amplifier topologies generates almost no RFI by themselves. Class D transmits energy at the switching frequency and its harmonics regardless of what PWM modulation scheme is used (see figure 3.4). In many existing designs incorporating a class AB amplifier an amplifier of class D is put in as an upgrading replacement. Care need to be taken to reduce RFI. It is recommended to keep the connection wires to the speaker as short as possible, i.e. keeping the radiating elements to a minimum. Proper grounding is also required and bypass capacitors must be put close to the amplifier. A LC low pass filter is often used as an output filter to remove the harmonics, but filtering is not as easy as it might seem: Due to parallel capacitance of the inductor and series parasitic resistance and inductance of the capacitor a LC filter looses its desired function above a certain frequency (self-resonance). By reducing the high frequency content of the output signal before it is filtered the harmonics can be limited to frequencies below the self-resonance of the LC filter. The nonzero rise and fall times in the switching transistors are therefore not only to be considered a problem. The longer rise time the lower amount of energy at high frequencies, but long rise times also wastes power and adds distortion to the output signal. A nifty trick to reduce RFI, or at least spread its spectrum (perhaps primarily to make it look better to the certification institutes), is using a varying frequency for the triangle wave. This makes the harmonics look more like noise rather than steady peaks of the output spectra. The preferred solution for reducing RFI depends of the specific application. In some cases it is not really a big concern at all. Other components, such as a very fast CPU close by, may produce much more RFI making it unnecessary to reduce RFI from any other component within range.

2.3 AGC and Limiter

The main purpose of the AGC is to suppress audio transients that can be harmful to the ear. The AGC will also try to improve the audibility of speech in a noisy environment. This can for example be done by splitting the audio band in several frequency bands and treating them differently. That functionality can not be part of this report. Instead only a simple limiter is implemented (i.e. not a "real" AGC with feedback) to demonstrate the concept of suppressing harmful transients. The main difference between a limiter and an AGC is that an AGC adjusts the amplification to keep the output signal within bounds. A limiter does not work in such a nice and polite way. It simply makes it impossible for the output signal to go beyond a certain level. This nonlinear behavior is known as clipping and is a source of undesired distortion and intermodulation products.

2.3.1 Hard limiter

A simple limiter can be made out of just a few components (see figure 2.6). When the amplitude of the input signal reaches the forward voltage drop of either D1 or D2 (depends of the input signal polarity) it will conduct to ground thus preventing the output signal to grow any further. The resulting transfer function from in to out is referred to as hard clipping. A hard limiter does its job, at least in the sense that it makes it impossible for the output signal to go beyond a predefined amplitude. But it has a major drawback. Hard clipping sounds very bad due to all the generated distortion.



Figure 2.6: Hard limiter

2.3.2 Soft limiter

Soft clipping, i.e. reducing the amplification by a constant factor rather than shutting it off, sounds better. A soft limiter is preferably followed by a hard limiter as a last protection against harmful sound levels. There is no big difference between the hardware for a hard and soft limiter. A soft limiter is a hard limiter with the addition of an extra resistor:



Figure 2.7: Soft limiter

Instead of connecting the input signal to ground the diodes of a soft limiter connects the lower resistor (R2) to the upper resistor (R1) thus forming a voltage divider. The relative resistances in the voltage divider determines the reduction of the amplification (i.e. compression). Another way to achieve softer clipping is using germanium diodes instead of silicon diodes. Germanium diodes have a less sharp transition area between conducting and off compared to silicon diodes. The transfer function for a limiter with germanium diodes has more of a soft bend rather than a sharp knee in the area where the limiter kicks in.

2.3.3 Variable limiter

If the fixed forward voltage drop of a diode is a problem an inverting amplifier can be a solution (see figure 2.8) [4]. By adjusting the feedback network (R4/R3) of the inverting amplifier (i.e. the amplification) all values between zero and one diode forward voltage drop can be chosen as the clipping level. The inverting amplifier adds the inverted audio signal as a bias voltage to the diodes. If one forward voltage drop is not enough several diodes (small signal and/or zener) can be connected in series. Different values for R1 and R2 determines the steepness of the clipping function. A more advanced transfer function can be achieved by connecting several limiters with different clipping levels in series to form an approximation of the desired transfer function.



Figure 2.8: Variable limiter

Chapter 3

Literature review

3.1 History

In 1931 Loy Barton published "High Audio Output from Relatively Small Tubes" in the Institute of Radio Engineers proceedings. In this article he described a way of operating the two vacuum tubes of a push-pull stage in strict time alternation, i.e. class B. This basic theme has since been used in a great variety of designs. Class D amplifiers entered the scene in the late 1940s and early 1950s [5] and is based on a combination of PWM and Bartons idea. In class D the active elements are used as switches instead of as linear devices. Many ideas from AC/AC and DC/AC switch mode power converters are directly applicable to class D amplifier design. One of the main differences is that the output signal from an audio amplifier is not supposed to be a stable sinusoidal wave, a very desirable feature of DC/AC and AC/AC converters. It seems though as if class D amplifier designers have not studied power electronics very well until recently. For instance "filterless PWM" that has been introduced as something new by Texas Instruments and several others is very similar to what in the world of power electronics is known as unipolar PWM.

Early class D amplifiers used vacuum tubes as switches. Today the vacuum tubes have been replaced by semiconductors with much better performance. In fact, the first "class D like designs" was not intended for audio applications. They were used for motor control. Even switching frequencies as low as 50 kHz [5] lead to tremendous switching losses. In the beginning the idea of using such a device for audio amplification must have seemed awkward (most often a switching frequency of 250 kHz is used for audio).

3.2 Transistors are not ideal switches

Also semiconductors of today need improvements. A transistor has a non neglectable rise and fall time thus setting an upper limit for the switching frequency. During its transition time it operates in the linear region thus dissipating power. In other words; a transistor is not an ideal switch. Another problem is the nonzero voltage drop over a saturated transistor contributing to an additional power loss. In power electronics this voltage drop is usually neglectable due to the much higher supply voltages used. In low voltage audio amplifiers of class D this is more of a problem. Too long rise and fall times is not only a limiting factor for the switching frequency. It also distorts or even eliminates short PWM-pulses. It was not until recently that analogue circuitry could be integrated on the same chip as fast-switching DMOS power MOSFET:s [6]. This has eliminated the need for the discrete output power stage that can be found in older class D designs. In the future micro electromechanical (MEMS) switches may be quick enough to replace the transistors of the H-bridge. MEMS switches are tiny mechanical switches made out of silicon and are usually operated by electrostatic fields. In contrast to transistors MEMS switches do have a zero voltage drop when closed. This make them a better approximation of the perfect switch compared with transistors. MEMS switches are currently being developed for radio applications, e.g. for use in reconfigurable antennas or as TR-switches.



3.3 Pulse Width Modulation - PWM

Figure 3.1: Bipolar PWM

3.3.1 Bipolar PWM

The instantaneous output voltage from an amplifier of class D can only be one of two values. Either +Vcc/0V for a single push-pull stage or differential +Vcc/-Vcc if the output stage is a H-bridge (i.e. two push-pull stages simultaneously pushing and pulling in opposite directions). This might seem strange since a normal audio signal is made up by all possible levels in between. Instead of a voltage level the desired audio signal is represented by variations in the pulse widths of a pulse train. Demodulation, i.e. restoring the desired audio signal from the pulse train, is easily done by low pass filtering. The audio signal is the instant mean value of the pulse train.

As seen in figure 3.1 a Pulse Width Modulated signal is easily created by comparing a control voltage, in this case an audio signal (for this example only one period of a sinusoidal wave), with a triangle wave. The output signal is high when the amplitude of the input signal exceeds the amplitude of the triangle wave. A short pulse represents a low input voltage and a long pulse represents a high input voltage. The frequency of the triangle wave must be high enough to avoid the lowest harmonics of the PWM signal slipping into the audio band. The triangle wave frequency also determines the sampling rate of the input signal. According to Nyqvists sampling theorem a sampling frequency twice as high as the upper end of the audio band is good enough (i.e. 40 kHz). That would require a very steep output filter that almost only exists in theory to get rid of harmonics (i.e. expensive and quite large due to all needed components). Normally a triangle wave frequency at least ten to twelve times as high as the the upper end of the audio band is used. This reduces the requirements for the low pass filter needed for restoring the wanted output signal. The inductance of the speaker coil combined with the inertia of the speaker can be good enough, but most often an extra LC filter is used to smoothen the current, for instance a balanced 2-pole Butterworth. Another benefit is that a higher cut off frequency allows for components with smaller physical dimensions thus not wasting valuable space on the PCB. The PWM signal generated with this method is called bipolar PWM since the output voltages switches back and fourth between +Vcc and -Vcc.

The amplitude modulation ratio m_a of PWM is defined as:

$$m_a = \frac{\hat{V}_{audio}}{\hat{V}_{triangle}}[7] \tag{3.1}$$

In power converter applications it can sometimes be useful with overmodulated PWM, a phenomenon that occurs when the amplitude of the input signal exceeds the triangle wave (i.e. $m_a > 1$). Overmodulation leads to more energy at the fundamental frequency but the relationship between input signal and the low pass filtered output signal is not linear [8]. There are also more harmonics to get rid of with over modulated PWM. Overmodulation is therefore avoided in audio applications.

3.3.2 Unipolar PWM

Another approach to PWM is called unipolar PWM:



Figure 3.2: Unipolar PWM

In unipolar PWM the generated PWM signal switches between +Vcc/0 and -Vcc/0 depending of the polarity of the input signal. Unipolar PWM is composed out of two bipolar PWM channels (PWM A and B in figure 3.2) subtracted from each other. The difference between the two bipolar channels is that one of them is generated with the inverse of the input signal. The two channels are fed to one leg each of an H-bridge where the subtraction takes place. The total voltage swing is as high as for bipolar PWM (+/-Vcc, i.e. 2 Vcc), but since switching is done with half the voltage swing (Vcc) less current is pumped in and out of the output filter and speaker. This reduces resistive losses in the load. One of the main reasons for a LC filter as the output filter used with bipolar PWM is its ability to store the ripple current from one half cycle to the next. Otherwise the energy of the harmonics is lost as heat in the speaker. With unipolar PWM the ripple current is much lower making the LC filter superfluous. Unipolar PWM is therefore "filterless". Another good feature of unipolar PWM compared to bipolar PWM is frequency doubling of the harmonics (see figure 3.4) [9]. This makes it possible to halve the switching frequency compared with a corresponding bipolar solution and still avoiding harmonics getting into the

audio band. Halved switching frequency is the same as halved switching losses, so unipolar PWM is the preferred solution for high efficiency. This is what CROWN [10] has done in their "Balanced Current Amplifier" [11], but instead of simply calling it unipolar PWM they have come up with a name of their own.

3.3.3 Biunipolar PWM

Unipolar PWM suffers from cross-over distortion in the transition area where it goes from switching between +Vcc/0 and -Vcc/0, i.e. close to zero crossings of the input signal. A remedy to this problem is letting the amplifier output very narrow pulses of alternating polarity when the input signal is close to zero. One way of doing this is adding a small delay to one of the two PWM channels feeding the H-bridge (see figure 3.3). Close to zero this



Figure 3.3: Biunipolar PWM

configuration will act as bipolar PWM. When the input signal grows the small time shift between the channels becomes neglectable making the output signal unipolar. This is very similar, or perhaps identical, to the method in use by Texas Instruments in their most recent amplifiers of class D [12]. This modulation scheme will from here on be referred to as "biunipolar PWM". The similarities between biunipolar PWM and class AB are obvious. Bipolar PWM and class A gives the best audio performance. There exists more

efficient solutions such as unipolar PWM and class B respectively. Unipolar PWM with bipolar behavior at the zero crossings combines features from bipolar and unipolar PWM, just as class AB combines class A and B. Sadly the little time delay between the two PWM channels puts back the lowest harmonic at the same place as for bipolar PWM. Only unipolar PWM can take advantage of frequency doubling of the harmonics.



Figure 3.4: PWM in the frequency domain

The harmonics are not only important to get rid of due to audio quality matters. The energy of the harmonics can end up as heat in the speaker, or even worse; disturb other electronic equipment (see section 2.2 about RFI).

Chapter 4

Materials and methods

Unipolar PWM has benefits compared to bipolar PWM. By adding a small delay to one of the two bipolar channels that unipolar PWM is made up of, something what I call "biunipolar PWM" is achieved. This is very similar or identical to the modulation scheme in use by Texas Instruments in their latest family of class D amplifiers [12]. They call their modulation scheme "Filterless PWM". Biunipolar PWM provides smoother zero crossings than unipolar PWM, but is slightly less efficient. As mentioned in section 2.1.3 the analogy to class AB clear.

4.1 Building blocks



Figure 4.1: Building blocks

The entire design can be divided into a few building blocks as shown in figure 4.1. The input stage of this amplifier must be able to add signals from several signal sources and remove signal energy above the audio band (antialiasing). Unipolar and biunipolar PWM are generated from the input signal and its inverse (see figure 3.2 and 3.3). A differential amplifier is therefore needed before the transformation from an analogue input signal to a pulse width modulated pulse train. By using a H-bridge as the output stage twice the voltage swing is achieved. This is especially valuable when working with low voltage levels. One of the risks with the two push-pull stages of the H- bridge is shoot-through during switching, i.e. a state when both transistors of a push-pull stage conducts at the same time. Shoot-through can result in both heat and broken transistors. The control logic must be able to avoid shoot-through by not switching on a transistor before the other one within a push-pull stage has been switched off.

To avoid overmodulation the triangle wave must at all times have an amplitude exceeding the amplitude of the input signal. A simple limiter is one solution to that problem and also avoids audio transients that can be harmful to the ear. The desired ASIC will have an advanced AGC instead of a simple limiter. Except for suppressing harmful transients the AGC will also attempt to improve the audibility of spoken words in a noisy environment. The AGC is not part of this report due to protection of the customer.

4.1.1 Analogue section

This is the only pure analogue building block and as such it is sensitive to disturbances from its surroundings. Bypass capacitors and proper grounding are required but also board layout is important. By keeping the PCB-tracks and the connection wires as short as possible they will not be susceptible to electromagnetic interference from the surroundings. In a mixed analogue and digital design it is always a challenge to protect the analogue sections from digital switching artifacts. Especially in an integrated circuit signal coupling via the substrate is a problem. Separate analogue and digital ground planes helps to avoid current loops.

There will be several input signals to the amplifier. They will all be single ended. A simple Kirchoff-mixer (i.e. connecting all signals together to one point) is a start, but a better solution is using a summing amplifier. By adding a capacitor in parallel with the feedback network a first order low pass filter is achieved, i.e. a simple anti-alias filter. The amplified and low pass filtered mix of input signals is then fed via a limiter to a differential amplifier with just one purpose in life; converting single ended input to differential output. It is important that the two output signals have the same amplitude and are 180 degree out of phase with each other. Otherwise the subtraction of the two bipolar PWM channels will result in a distorted unipolar replica of the input signal. The limiter is of the variable kind with selectable clipping level described in section 2.3, but with the option to be of the simple kind with fixed clipping level. This is chosen by jumpers on the PCB.

4.1.2 Digital section

The input signals to the digital section are both analogue. Before any further processing they must be converted to digital signals, in this case a pulse width modulated signal train.

Analogue to PWM

The analogue to PWM conversion is carried out by two comparators comparing the input signals with a triangle wave. An audio signal represented by a PWM signal is much less sensitive to interference from its surroundings. The triangle wave can be generated by integrating a square wave. A simple ring oscillator is a quick and easy way to generate a square wave. An integrator is then the only thing needed to achieve the desired triangle wave. Figure 4.2 shows the complete schematic for a triangle wave generator.



Figure 4.2: Triangle wave generator

$$f = \frac{1}{2.2 \cdot R_2 \cdot C_1} Hz[13] \tag{4.1}$$

$$A = \frac{1}{R_3 \cdot C_2} \int V_{square} dt V[14]$$
(4.2)

The triangle wave is a very important part of the analogue to PWM conversion. Any imperfection will directly show up as a nonlinear relationship between the input and output of the amplifier. If the op-amp has too low slew rate the slopes of the triangle wave will differ from ideally straight lines. As a result the integrator will not have time enough to reach its desired peak voltage. In other words, high bandwidth is required to generate a perfect triangle wave. For best performance the op-amp must have a high slew rate, preferable 20 $V/\mu s$ or more. It is preferable with rail to rail output swing of the op-amps, i.e. an output swing equal to the supply voltage. Otherwise the op-amp will limit the maximum peak voltage of the triangle wave.

The logic gates for this prototype are of the well known HC-family. HC-MOS has a supply voltage range between 2V and 6V, except when used in oscillators. Then at least 3V must be used, otherwise there will not be enough output current to drive external components [15].

Control logic

Power is lost when both transistors of a push-pull stage conducts at the same time, a phenomenon known as shoot-through. This can be avoided by not switching on a transistor before the other one within a push-pull stage has been switched off. The control logic that handles this can be divided in two separate but equal channels, one for each of the two push-pull stages of the H-bridge. For biunipolar PWM the extra delay needed for one of the two channels is easily created by a RC-link proceeding the control logic of that channel. By adjusting the RC time constant $(T \approx RC)$ both unipolar (T = 0) and biunipolar $(T \neq 0)$ PWM can be achieved.

Each logic gate has a push-pull output stage very similar to the ones in the H-bridge. It might seem strange introducing several push-pull stages with shoot through of their own to avoid shoot-through in the two pushpull stages in the H-bridge. But there is one big difference; the H-bridge is supposed to deliver as much power as possible to the load. A logic gate is only supposed to supply enough current to maintain logic levels when loaded with one or a few input stages of other logic gates. In other words, the channel resistance of the H-bridge should be as low as possible while the output stage of the logic gate only require a channel resistance low enough to supply just one or a few other logic gates. This is useful in high power applications with large currents flowing through the H-bridge. In a low power application the control logic might consume more power than what is gained with eliminated shoot-through in the H-bridge. In the prototype the control logic is implemented for test purposes. For the final ASIC selectable control logic operation might prove useful. In a low power application the control logic can then be disabled to save power.

The control logic of the prototype are of the well known HC-family, i.e. standard CMOS (HCMOS). The push-pull stages of HC-logic has relatively low channel resistance. During switching more current than needed flow through their push-pull stages. But according to the manufacturer, in this case Philips Semiconductors, shoot-through is responsible of only 10% of the wasted power [15]. Instead the main power loss comes from charging and dischargning on-chip parasitic and load capacitances. A RC-link, e.g. the one in the triangle wave generator, introduces slower rise and fall times making shoot-through more of a problem. The final ASIC can use higher channel resistances for its push-pull stages and perhaps digital counters and shift registers instead of RC-links, thus wasting less power.

The control logic needs some way to detect "dangerous conditions" i.e. when switching takes place. When such a dangerous condition has been detected a transistor within a push-pull stage may not be switched on before the other one has been properly switched off. The time required for a transistor to be switched off is almost constant, so one solution to this problem is using a delay. As seen in figure 4.3 a switching detector can be made of a xor gate and a RC-link. As soon as the input signal changes the RC-link will delay this change to one of the two xor inputs. During the RC-delay the xor gate will output a logic one. This makes it easy to adapt to different types of switching elements (BJT, MOSFET etc).

The desired functionality can be summarized in a truth table where



Figure 4.3: Switching detector

T=Toggle, U=control signal for the upper switch in the push-pull stage and L=control signal for the lower switch:

Т	PWM	U	L
0	0	0	1
0	1	1	0
1	0	0	0
1	1	0	0

Table 4.1: Truth table for the control logic of one bipolar PWM channel

In this specific prototype PNP transistors are used as the upper switching elements in the push-pull stages of the H-bridge. Compared to controlling a NPN transistor PNP works in the opposite way. The inverted control signal must therefore be used, i.e. \overline{U} instead of U. This change is reflected in table 4.2. Equation 4.3 and 4.4 are the resulting boolean equations defining the control logic.

Т	PWM	\overline{U}	L
0	0	1	1
0	1	0	0
1	0	1	0
1	1	1	0

Table 4.2: Truth table that handles PNP/NPN in the H-bridge

$$\overline{U} = T + \overline{PWM} \tag{4.3}$$

$$L = T \cdot PWM \tag{4.4}$$

Due to gate delay in the xor gate a few inverters are inserted to ensure glitch free operation. The propagation delay of the inverters exceeds the propagation delay of the xor gate, so the toggle indicator is always ahead of the PWM signal. By preceding one of the two bipolar channels with a RC-link biunipolar PWM is achieved. The complete control logic for one bipolar PWM channel is shown in figure 4.4. The control logic for the other channel is identical except for the mentioned RC-link.



Figure 4.4: Control logic for one bipolar channel

4.1.3 H-bridge

The H-bridge of this prototype uses BJT:s instead of fast-switching DMOS power MOSFET:s as proposed in section 3.2. This does not matter much. Only a small change of the time constant in the control logic is needed to adjust for MOSFET:s.

Four resistors of 1 ohm are included in the H-bridge. They are used to detect shoot through. When the H-bridge and the control logic are verified to work the resistors can be replaced by 0 ohm.

Part II Implementation

Chapter 5

Implementation

With all building blocks known it is time to build a prototype to see if the proposed design will work. There are several ways to go for the implementation, but a design not involving solder feels unpleasant.

5.1 Discrete implementation

Here follows a brief survey of the implementation. A complete schematic can be found in appendix A. Photos of the assembled prototype are shown in figure 5.1 and 5.2.

The input stage is built around one of the two amplifiers within U103 (LMV358V). That amplifier is connected as a summing pre-amplifier with three input signals. An electret microphone connected to P101 can with P102 be chosen as one of three input signals. The other amplifier within U103 serves as the inverting amplifier of a variable limiter. The limiters functionality can be altered. R119 and P105 selects if the limiter shall be variable by letting the inverting amplifier apply bias to the diodes (D101/102)or not. R120 and R122 selects the variable clipping level. R114 and R123 determines the clipping "hardness". If P106 is left open the limiter will be disconnected from the signal path which can be useful during tests. U104 (LTC1992–2) is a differential amplifier used to transform single ended input to differential output. U104 is the last part of the analogue input stage. The differential output signal is then passed on to the digital section. A triangle wave generator is made up by U101(74HC04), U102 (LT6202) and a few passive components. It is exactly the same configuration as described in figure 4.2. The differential audio signal and the triangle wave are fed to one comparator each; U105 and U106 (LMV7219). The output signals from the comparators (U105 and U106) are two pulse width modulated signal trains. One of the two PWM channels can be delayed for a short time before being passed on to the control logic. The delay is chosen with an RC-link (R124/C110). Without any delay the resulting PWM modulation scheme will be unipolar, otherwise biunipolar. The control logic is identical for the two PWM channels feeding one leg each of the H-bridge. The functionality of the control logic has already been described in section 4.1.2. The only parameters to be altered are the delays generated by R113/C105 and R121/C109. Finally there is a H-bridge made up by Q101-104. The load (a speaker) is connected to P103.



Figure 5.1: Class D prototype. Notice the separate analogue and digital ground planes.



Figure 5.2: Class D prototype with comments

5.2 Estimated power consumption

A rough estimate of the power consumption is good enough for a comparison to the measured result. Except for a few passive components only the active components are part of the calculation. Most of the passive components do not contribute much to the total power consumption.

5.2.1 Op-amps, comparators and H-bridge

The power consumption for the op-amps, comparators and the H-bridge are listed in their data sheets.

- LT6202 (U102) Typ 2.3 mA Max 2.85 mA @ 3 V [16]
- LMV358 (U103) Typ 140 $\mu \mathrm{A}$ Max 340 $\mu \mathrm{A}$ @ 2.7 V [17]
- LTC1992-2 (U104) Typ 0.65 mA Max 1 mA @ 2.7 V [18]
- \bullet LMV7219 (U105, U106) Typ 0.9 mA Max 1.6 mA @ 2.7 V [19]
- MMBT2369A (Q103, Q104, $\approx 50\%$ duty cycle each) $I_B \approx 3mA$ [20]
- MMBT5771 (Q101, Q102, $\approx 50\%$ duty cycle each) $I_B \approx 3mA$ [21]

The prototype is intended for a supply voltage of 3 V. Most of the values in the data sheets are specified for a supply voltage of 2.7 V. The extra 0.3 V

is not likely to increase the supply current with more than a few mA. Sum: Typ 10.9 mA Max 13.4 mA + a few mA \approx 14-17 mA

5.2.2 Logic gates

Shoot-through is normally responsible for only 10% of the dissipated power in logic gates [15] (see also section 4.1.2). Instead charging and discharging on-chip parasitic and load capacitances stands for the major part. RClinks introduce slower rise and fall times making shoot-through more of a problem. The dynamic power dissipation for a logic device can be estimated with equation 5.1. For simplicity the RC-links are not part of the calculation.

$$P_D = C_{PD} \cdot V_{CC}^2 \cdot f_i + \sum \left(C_L \cdot V_{CC}^2 \cdot f_o \right) [15]$$
(5.1)

 C_{PD} : power dissipation capacitance per package C_L : total external load capacitance per output V_{CC} : supply voltage f_i : input frequency f_o : output frequency

The design consists of six logic devices; three 74HC04 and one each of 74HC08, 74HC32 and 74HC86. Their logic gates are listed in table 5.1 and sorted by input and output frequencies. The input capacitance C_I of the

Type	f_i 500 kHz	f_i 250 kHz	$f_o 500 \text{ kHz}$	$f_o 250 \text{ kHz}$	unused (0 Hz)
74HC04	2	11	2	11	5
74HC08	0	2	0	2	2
74HC32	0	2	0	2	2
74HC86	2	0	2	0	2

Table 5.1: Logic gates sorted by input and output frequencies

gates are used as C_L in the calculation for the preceding gate. Table 5.2 lists C_I and C_{PD} for the different types of gates used in the design. As seen in

Type	C_I	C_{PD}	ref
74HC04	$3.5~\mathrm{pF}$	21 pF	[22]
74HC08	$3.5 \ \mathrm{pF}$	10 pF	[23]
74HC32	3.5 pF	16 pF	[24]
74HC86	3.5 pF	30 pF	[25]

Table 5.2: C_I and C_{PD} for the gates in the prototype

table 5.1 almost all gates operate with a frequency of 250 kHz. This opens up for a simplified version of equation 5.1. A good approximation of P_D can

be obtained by summing all capacitances (C_{sum}) and assuming an average frequency (f_{avg}) . This reduces equation 5.1 to equation 5.2.

$$P_D = C_{sum} \cdot V_{CC}^2 \cdot f_{avg} \tag{5.2}$$

With all values known the power consumption of the logic gates can be calculated.

$$C_{sum} = 3 \cdot C_{PD \ 74HC04} + C_{PD \ 74HC08} + C_{PD \ 74HC32} + C_{PD \ 74HC86} + 19 \cdot C_I$$

= 3 \cdot 21 + 10 + 16 + 30 + 19 \cdot 3.5pF
= 185.5pF
$$P_D = C_{sum} \cdot V_{CC}^2 \cdot f_{avg}$$

= 185.5 \cdot 10^- 12 \cdot 3^2 \cdot 250 \cdot 10^3W
= 0.42mW
$$I_D = \frac{P_D}{V_{CC}}$$

= 0.14mA

5.2.3 RC-links

The four RC-links in the design are used to generate delays. R129/C112 determines the triangle wave frequency. R124/C110 adds the little time shift to one of the two bipolar PWM channels needed for generating biunipolar PWM. R113/C105 and R121/C109 are used to avoid shoot-through in one leg each of the H-bridge. The RC-links adds extra load and more shoot-through in the surrounding gates. They are judged to increase the supply current with a few mA.

5.2.4 Total supply power without load

A rough estimate of the total supply current without load is 15-18 mA at a supply voltage of 3 VDC. The power consumption $P_{unloaded}$ is then 45-54 mW.

5.3 Estimated output power

The fundamental-frequency voltage \hat{V}_{o1} seen to the left in figure 3.4 is the desired signal to be transferred to the load. The rms value, V_{o1} , is needed to calculate the output power. The amplitude modulation ratio m_a defined in section 3.3.1 determines the magnitude of the fundamental-frequency voltage:

$$\hat{V}_{o1} = m_a V_{cc} \quad (m_a \le 1) \quad [26]$$
(5.3)

Assuming a sinusoidal input signal the rms value is:

$$V_{o1} = m_a \frac{V_{cc}}{\sqrt{2}} \tag{5.4}$$

Combining equation 2.3 with 5.4 yields:

$$P = \frac{(m_a \frac{Vcc}{\sqrt{2}})^2}{R}$$
(5.5)

A speaker is not purely resistive but assuming so is a fairly good approximation. The output power is inversely proportional to the speaker impedance. A low impedance yields high output power but to the cost of increased power supply ripple. This amplifier is intended for a speaker impedance of 32Ω but will work well also with e.g. 8Ω . The intended supply voltage Vcc is 3V. Maximum output power is achieved with $m_a = 1$:

$$P_{max_{32\Omega}} = \frac{(1 \cdot \frac{3}{\sqrt{2}})^2}{32} \approx 140 mW \tag{5.6}$$

$$P_{max_{8\Omega}} = \frac{(1 \cdot \frac{5}{\sqrt{2}})^2}{8} \approx 560 mW$$
(5.7)

To compensate for the nonzero voltage drop over saturated transistors in the H-bridge equation 5.4 is adjusted to:

$$V_{o1} = m_a \frac{(V_{cc} - V_{CE(sat)_{PNP}} - V_{CE(sat)_{NPN}})}{\sqrt{2}}$$
(5.8)

Combining equation 2.3 with 5.8 yields:

$$P = \frac{\left(m_a \frac{(V_{cc} - V_{CE(sat)_{PNP}} - V_{CE(sat)_{NPN}})}{\sqrt{2}}\right)^2}{R}$$
(5.9)

According to the data sheets are $V_{CE(sat)_{PNP}} \approx 0.6 V$ [21] and $V_{CE(sat)_{NPN}} \approx 0.3 V$ [20]. This might seem much, but there flows quite large currents through the H-brigde explaining the high voltage drops.

$$P_{max_{CE(sat)32\Omega}} = \frac{\left(1 \cdot \frac{(3-0.6-0.3)}{\sqrt{2}}\right)^2}{32} \approx 69mW \tag{5.10}$$

$$P_{max_{CE(sat)8\Omega}} = \frac{\left(1 \cdot \frac{(3-0.6-0.3)}{\sqrt{2}}\right)^2}{8} \approx 276mW \tag{5.11}$$

With fast switching DMOS instead of BJT the voltage drop over saturated transistors is likely to decrease. The on-state resistance for DMOS can be as low as a few hundred m Ω [27]. This makes it possible to almost neglect the voltage drops in the H-bridge and thus achieve more output power.

5.4 Estimated maximum efficiency

The efficiency η is defined as:

$$\eta = \frac{P_{load}}{P_{supply}} \quad [1] \tag{5.12}$$

With values from section 5.2 and 5.3 an estimated efficiency can be calculated:

$$\eta_{max_{32\Omega}} = \frac{P_{max_{32\Omega}}}{P_{unloaded} + P_{max_{32\Omega}}} \approx 72\%$$
(5.13)

$$\eta_{max_{8\Omega}} = \frac{P_{max_{8\Omega}}}{P_{unloaded} + P_{max_{8\Omega}}} \approx 91\%$$
(5.14)

When the voltage drops due to the PNP and NPN transistors in the H-bridge are taken into account the efficiency is a bit lower:

$$\eta_{max_{CE(sat)32\Omega}} = \frac{P_{max_{CE(sat)32\Omega}}}{P_{unloaded} + P_{max_{CE(sat)32\Omega}}} \approx 56\%$$
(5.15)

$$\eta_{max_{CE(sat)8\Omega}} = \frac{P_{max_{CE(sat)8\Omega}}}{P_{unloaded} + P_{max_{CE(sat)8\Omega}}} \approx 84\%$$
(5.16)

It is important to note that the values above are calculated with $m_a = 1$, i.e. full output power. The efficiency is lower with $m_a < 1$.

Generally class D amplifiers intended for high output power gain higher efficiency than low power versions. This is due to the fact that there is always a constant power consumption in the electronics generating control signals for the H-bridge ($P_{unloaded}$). The more power delivered from the H-bridge the less $P_{unloaded}$ affects the total power consumption making the efficiency approach 100%.

Chapter 6

Results

Even if the output signal of an audio amplifier sounds good to the ear it does not imply that the design is perfect. Ernst Werner von Siemens once said "Messen ist Wissen", i.e. "To measure is to know". I am willing to agree. A measurement does not only express performance in numbers. It also helps the designer to track down and correct design flaws.

There are several interesting parameters to be measured. When it comes to class D efficiency is the most obvious one. The Total Harmonic Distortion + Noise (THD+N) is also of interest and can be measured both versus amplitude and frequency. For this prototype the difference between unipolar and biunipolar PWM is a lot more interesting than several measurements for each modulation form. THD+N is therefore only measured versus frequency with identical amplitudes for the two modulation schemes. The prototype has only one channel (mono) making crosstalk measurements impossible. The intended product will work as a set of two independent mono amplifiers separated a few decimeters away from each other. Crosstalk is therefore unlikely to be much of a problem.

One might think that the procedure when measuring linear (e.g. class A) and switched amplifiers (class D) are the same. That is true except for the simple fact that an audio analyser normally is not intended for frequencies above 20 kHz. The amplifiers of its input stage and signal processing stages are optimized for low noise and high linearity. A high slew rate is not really needed when dealing with the audio band. $5 - 10 V/\mu s$ is normally good enough [28]. The rich content of harmonics outside the audio band generated by class D may produce intermodulation products in these amplifiers. Slew rates of hundreds $V/\mu s$ or more is not unusual when it comes to class D. A quite sharp low pass filter with a flat passband is needed between the amplifier and the analyser to ensure valid measurements. Otherwise very high levels of distortion will be recorded.

Not only parameters related to performance are worth measuring. In this design both the generated PWM and the variable limiter are worth closer

looks. Imperfections to the PWM signal indicates problems with the control logic or the comparators. The limiter is not a very critical part of the design but it is interesting to verify its functionality.

6.1 Efficiency

The prototype set up for biunipolar PWM with a switching frequency of 240 kHz consumes 18.2 mA @ 3V without load. The current consumption for unipolar PWM with a switching frequency of 120 kHz is 17.8 mA @ 3V. The rough estimate in section 5.2 was 15-18 mA.

The switching nature of class D makes efficiency measurments a little more complicated than for linear amplifiers. Filterless PWM modulation schemes require the inductance and inertia of a speaker for its operation. When measuring the efficiency of linear amplifiers a resistor is good enough as load making it easy to calculate the output power. With class D the output power is calculated on basis of current and voltage:

$$\eta_{classD} = \frac{P_{load}}{P_{supply}} = \frac{V_{O(rms)} \cdot I_{O(rms)}}{V_{S(ave)} \cdot I_{S(ave)}} \quad [2]$$
(6.1)

 $I_{O(rms)}$ is measured by inserting a resistor with low resistance in series with the speaker. The rms voltage over the resistor $V_{R(rms)}$ is proportional to $I_{O(rms)}$:

$$\eta_{classD} = \frac{V_{O(rms)} \cdot \left(\frac{V_{R(rms)}}{R}\right)}{V_{S(ave)} \cdot I_{S(ave)}} \quad [2]$$
(6.2)

The test set up consists of:

- Mobile communications DC source (HP 66311A)
- $6\frac{1}{2}$ digit true rms multimeter (Agilent 34401A)
- Radio Communication Analyser (Rhode & Schwarz CMTA 54)
- 100 kHz Dual Channel Programmable Filter (Stanford Research Systems SR650)
- Resistor 0.111 Ω
- Speaker 32 Ω
- Class D amplifier

Class D efficiency is usually measured with a 1 kHz sinusoidal test signal [28] which in this case is generated by the Radio Communication Analyser. The signal level of the 1 kHz test signal is adjusted to achieve maximum output power, i.e. $m_a = 1$. $V_{S(ave)}$ and $I_{S(ave)}$ are measured with the averaging function of the power supply (HP 66311A). The low pass section of SR650 is used to filter out the fundamental frequency component V_O over the load. $V_{O(rms)}$ and $V_{R(rms)}$ are measured with the true rms multimeter.

$$\eta_{bi240kHz} = \frac{709 \cdot 10^{-3} \cdot \left(\frac{11.51 \cdot 10^{-3}}{0.111}\right)}{3 \cdot 48 \cdot 10^{-3}} \approx 51\%$$
(6.3)

$$\eta_{uni120kHz} = \frac{717 \cdot 10^{-3} \cdot \left(\frac{11.54 \cdot 10^{-3}}{0.111}\right)}{3 \cdot 46 \cdot 10^{-3}} \approx 54\%$$
(6.4)

6.2 Total Harmonic Distortion + Noise (THD+N)

The basic idea of THD+N measurements is feeding an amplifier with a low distortion sine wave (usually 1 kHz or a sweep over the entire audio band) and then removing the sine wave from the output signal with a sharp filter. The residue are simply the Total Harmonic Distortion and Noise. With an ideal amplifier there are no residues to be measured. Imperfections in real amplifiers can largely be quantified by the THD+N value. Both the input signal amplitude and frequency can affect the THD+N. The measurement bandwidth must be specified. Otherwise two different THD+N measurements can not be compared. For this prototype the difference between unipolar and biunipolar PWM is more interesting than THD+N for different amplitude levels. THD+N is therefore only measured versus frequency with identical amplitudes for the two modulation schemes. The amplitude is chosen to give maximum output power without clipping.

The test set up consists of:

- Audio Analyser (Rhode & Schwarz UPL, DC to 110 kHz)
- 100 kHz Dual Channel Programmable Filter (Stanford Research Systems SR650)
- Class D amplifier with a power supply of 3V

The audio analyser is set up to measure THD+N as a sweep over the entire audio band, i.e. 20 Hz to 20 kHz. The output signal from the audio analyser is fed to the input of the class D amplifier. As noted earlier the PWM output signal from class D consists of harmonics far outside the audio band. The harmonics must be removed with a low pass filter before the output signal from the class D amplifier is returned to the audio analyser. Filtering is done by the low pass section of SR650 (8-pole 6-zero elliptic) with the corner frequency set to 20 kHz. This ensures that there will be no intermodulation products in the amplifiers and signal processing stages of the audio analyser.



Figure 6.1: THD+N for biunipolar PWM. Switching frequency: 240 kHz



Figure 6.2: THD+N for unipolar PWM. Switching frequency: 240 kHz



Figure 6.3: THD+N for unipolar PWM. Switching frequency: 120 kHz

6.3 Variable limiter

A XY plot of the relationship between the input and output signal of the class D amplifier is a good way to visualise the functionality of the limiter. Just as for THD+N measurements low pass filtering is needed. Otherwise the XY plot will look distorted.

The test set up consists of:

- Oscilloscope (Tektronix TDS5104B DPO)
- Audio Analyser (Rhode & Schwarz UPL, DC to 110 kHz)
- 100 kHz Dual Channel Programmable Filter (Stanford Research Systems SR650)
- Class D amplifier with a power supply of 3V

The audio analyser is used as a signal generator feeding the input of the class D amplifier with a sine wave of constant frequency. It does not matter what frequency in use as long as it is within the audio band (20 Hz to 20 kHz), i.e. the intended frequency band for the amplifier. The output signal from the class D amplifier is low pass filtered with SR650 set up to a corner frequency of 20 kHz. The input signal to the class D amplifier is connected to the X-channel of the oscilloscope. The low pass filtered output signal is connected to the Y-channel. The oscilloscope is set up to XY mode. The clipping hardness of the limiter is constant. Only the clipping level is altered between the measurements. Normally a XY plot of a limiter looks like a single line. The output filter (SR650) adds a phase shift to the signal which explains the resulting elliptic form. The phase shift is frequency dependent so slightly different XY plots (more or less elliptic) are achieved with different frequencies. The clipping level is independent of the phase shift.



Figure 6.4: XY plot of the input and output signal. No clipping.



Figure 6.5: An input signal amplitude exceeding the clipping level $\mathbf{1}$



Figure 6.6: A lower clipping level compared to fig 6.5

6.4 PWM generation

As described in the literature review unipolar PWM is created by subtracting two bipolar PWM channels from each other (see section 3.3.2). Subtraction takes place in the H-bridge. Biunipolar PWM is generated in a similar way. The difference is that one of the two bipolar PWM channels has been slightly delayed.

The test set up consists of:

- Oscilloscope (Tektronix TDS5104B DPO)
- Audio Analyser (Rhode & Schwarz UPL, DC to 110 kHz)
- Class D amplifier with a power supply of 3V

The audio analyser is only used as a signal generator feeding the input of the class D amplifier with a sine wave of constant frequency. It does not matter what frequency in use as long as it is within the audio band (20 Hz to 20 kHz), i.e. the intended frequency band for the amplifier. Channel one and two of the oscilloscope are connected to one leg each of the H-bridge. The oscilloscope is set up to subtract the two channels and display the difference. Another way to visualise the output signal would be using a differential oscilloscope probe.





Figure 6.7: Unipolar PWM during positive input signal



Figure 6.8: Unipolar PWM during negative input signal



Figure 6.9: Input signal, unipolar PWM and low pass filtered output signal. Notice the phase shift introduced by the low pass filter (SR650).

6.4.2 Biunipolar PWM

The main difference between unipolar and biunipolar PWM is the behavior close to zero crossings. Biunipolar PWM outputs narrow pulses of alternating polarity to reduce crossover distortion.



Figure 6.10: Biunipolar PWM without any input signal



Figure 6.11: Biunipolar PWM during a slow zero crossing. Notice the bipolar behavior during the zero crossing.

Chapter 7

Discussion

7.1 Efficiency

Unipolar PWM offers higher efficiency than biunipolar PWM, but as seen in equation 6.3 and 6.4 is the difference small for this prototype (54 % and 51 % respectively). Switching losses obviously play a small role in low power class D amplifiers compared to the power consumption in the analog input stage. In high power class D amplifiers switching losses are more of a problem.

7.2 THD+N

Measurements of THD+N confirms that biunipolar PWM offers better performance than unipolar PWM (compare figure 6.1 with figure 6.2). Surprisingly halved triangle wave frequency yields even better performance, at least for audio signals below a few kHz (compare figure 6.3 with figure 6.2). This indicates a design problem with too long rise and fall times distorting the signal, i.e. too low bandwidth. In theory a high triangle wave frequency would yield better performance.

A closer look at the THD+N measurement of unipolar PWM at 120 kHz serves as good example when pinpointing out differences from the expected result (bold line in figure 7.1). With increasing input signal frequency the quotient between the triangle wave frequency and the input signal decreases. In other words; the number of samples per input signal cycle decreases. This explains why THD+N grows with frequency (clearly seen above 1 kHz). At D in figure 7.1 THD+N takes a quick jump from 0.8% down below 0.2%. This happens when the first dominant harmonic goes outside the measurement bandwidth.

High THD+N below 100 Hz (A in figure 7.1) is dependent of the capacitors that are used in the signal path to separate DC bias from different building blocks (e.g. C104 and C107 as seen in the attached schematic). They work as high pass filters and will not be needed when the design has



Figure 7.1: THD+N for unipolar PWM with a triangle wave frequency of 120 kHz compared to the expected result (bold line)

been verified to work. THD+N would then decrease as predicted by theory when going from 100 Hz toward 0 Hz.

Almost centered around 100 Hz (B in figure 7.1) there is a distinct area with increased THD+N. Its cause is not clear and requires further investigations. One possibility is some kind of interference with the power line frequency (50 Hz) that might have been introduced by the power supply (3 VDC from 230 VAC). This theory is supported by the fact that there is no distinct area with increased THD+N present in the measurement of biunipolar PWM. The measurements of unipolar and biunipolar PWM were conducted at different occasions and with different power supplies.

In this prototype signal ground is defined by a simple voltage divider (see R131 and R132 in the schematic found in appendix A). By adding a voltage follower signal ground will become more stable. This is expected to move the area around C in figure 7.1 closer to the desired (bold) curve.

7.3 Limiter

The variable limiter works as expected. The output signal does not grow beyond the limit set by the inverting amplifiers feedback network (i.e. amplification) and the forward voltage drop of the limiting diodes.

As noted the output filter (SR650) adds a phase shift which explains the

elliptic form of the XY-plots. If there would have been an AGC instead of a limiter the relationship between the input and output would have looked different. Instead of loosing its elliptic form at the outer ends the XY-plot of an AGC keeps its elliptic form no matter of the input signal amplitude. The AGC decreases the amplification when needed to keep the output signal within bounds. The shape of the audio signal is preserved which sounds better to the ear than clipping.

7.4 PWM

Both the generated unipolar and biunipolar PWM looks as expected. Close to zero crossings the biunipolar PWM outputs narrow pulses of alternating polarity. When the input signal amplitude grows both unipolar and biunipolar PWM switches with the same polarity as the input signal.

As noted in section 7.1 unipolar PWM offers a higher efficiency than biunipolar PWM. As already concluded the measurements of THD+N confirms that biunipolar PWM offers better audio performance than unipolar PWM. Compare figure 6.1 with figure 6.2 and notice that the overall THD+N is lower with biunipolar PWM.

Chapter 8

Conclusions

- There are several PWM schemes that can be used for amplifiers of class D. Bipolar PWM offers best linearity but is the least efficient modulation scheme. Unipolar PWM is efficient but suffers from cross over distortion. Biunipolar PWM combines features of both bipolar and unipolar PWM and offers a reasonable combination of linearity and efficiency. There is a clear analogy to the linear amplifier classes A, B and AB. Both unipolar and biunipolar PWM falls within a category called filterless PWM. As such they do not require an external low pass filter for optimal performance.
- Unipolar PWM is generated by subtracting two bipolar PWM channels from each other. One of the bipolar channels is generated with the inverted input signal. By adding a small delay to one of the channels biunipolar PWM is achieved. This opens up for the possibility of easily switching between the efficient unipolar PWM (no delay and halved switching frequency) and the more linear biunipolar PWM (small delay). This is useful in battery powered applications. As long as the batteries are ok biunipolar PWM can be used. Otherwise the amplifier can switch over to unipolar PWM and use halved switching frequency to save power.
- All building blocks work as expected but there are a few problems to take care of. Surprisingly a low switching frequency (120 kHz) led to lower THD+N than with 240 kHz. This indicates a problem with too long rise and fall times in the control logic or imperfections in the triangle wave generator. Another problem is the high power consumption in the analog part that generate control signals for the H-bridge. Implementing the amplifier in an ASIC will hopefully lead to a reduced power consumption due to the ability to get custom made building blocks.
- The efficiency of class D is theoretically 100%. In low power amplifiers

of class D the practical efficiency is degraded mainly by the power consumption for generating control signals for the H-bridge. In high power amplifiers of class D switching losses in the H-bridge are much more of a problem.

Chapter 9

Recommendations

The ASIC is intended to be incorporated in a product family with several members. Some of them are quite advanced and thus expensive. A DSP can among many other things easily handle the AGC functionality and generate control signals for the H-bridge. An ASIC is only really meaningful in applications where low cost and low power consumption are the main concerns, i.e. products at the lower end of the family tree.

One possibility is designing an ASIC that can work in two separate modes. One mode for stand alone operation in cheap products and one mode with reduced functionality for use with a DSP. It might seem strange using the ASIC at all when there already is a DSP in the application, but the DSP needs a H-bridge as its output stage. The cost for a H-bridge and the proposed ASIC is likely to be almost the same. The chip area needed to attach bonding wires defines a minimum area that is likely to be big enough for both a single H-bridge and an amplifier of class D with AGC.

The ASIC is supposed to have a simple serial CPU interface for monitoring and setting of audio parameters. The most important settings to be accessible are:

- Mixing and amplification of the three input channels, but also a general/global setting that affects them all.
- AGC parameters
- Equalizer to be able to compensate for different listening environments.
- Level for hard clipping, but with an AGC a limiter is not really needed except maybe as a final protection stage against harmful sound levels.
- Delay for one of the two bipolar PWM channels so that both unipolar and biunipolar PWM can be achieved.
- Delay for the transistors in the H-bridge to prevent shoot through.

• Adjustable triangle wave frequency to be able to take advantage of the frequency doubling of harmonics in unipolar PWM.

Bibliography

- Mark N. Horenstein. Microelectronic Circuits and Devices, pages 736– 737. Prentice Hall, 1996. ISBN 0137013353.
- [3] Steve Somers. Class D amplifiers Not 'just audio'. http://www.extron.com/technology/archive.asp?id=ts122001.
- [4] Chu Moy. Designing a limiter for headphone amplifiers. http://www.headwize.com/projects/, 1999.
- [5] Bill Schweber. CLASS D IC AMPS: ready for audio prime time. *EDN*, pages 79–90, January 1999.
- [6] Penton Media. Class-D amps raise low-power systems to the next level. *ED Online*, (7533), December 1998. http://www.elecdesign.com/Articles/ArticleID/7533/7533.html.
- [7] Ned Mohan, Tore M. Undeland, and William P. Robbins. *Power Electronics*, page 203. Wiley, second edition, 1995. ISBN 0471584088.
- [8] Ned Mohan, Tore M. Undeland, and William P. Robbins. Power Electronics, page 209. Wiley, second edition, 1995. ISBN 0471584088.
- [9] Ned Mohan, Tore M. Undeland, and William P. Robbins. *Power Electronics*, page 216. Wiley, second edition, 1995. ISBN 0471584088.
- [10] Crown. http://www.crownaudio.com/.
- [11] Crown. Reinventing the power amplifier BCA. http://www.crownaudio.com/pdf/amps/bcapaper.pdf, 1998.
- [12] Texas Instruments . TPA2005D1. http://wwws.ti.com/sc/ds/tpa2005d1.pdf, 2002.

- [13] Square wave oscillator (2). http://www.interq.or.jp/japan/seinoue/e_ckt29.htm.
- [14] P. Horowitz and W. Hill. The Art of Electronics, page 26. Cambridge University Press, 2nd edition, Sept 1989. ISBN 0521377099.
- [15] Philips Semiconductors. HC User Guide. http://www.philipslogic.com/support/techdocs/logic/pdf/hcguide.pdf, 1997.
- [16] LT6202. http://www.linear.com.cn/pdf/620234fa.pdf.
- [17] LMV358. http://www.national.com/ds/LM/LMV321.pdf.
- [18] LTC1992-2. http://www.linear.com.cn/pdf/1992-2i.pdf.
- [19] LMV7219. http://www.national.com/ds/LM/LMV7219.pdf.
- [20] MMBT2369A. http://www.fairchildsemi.com/pf/MM/MMBT2369A.html.
- [21] MMBT5771. http://www.fairchildsemi.com/pf/MM/MMBT5771.html.
- [22] 74HC04. http://www.semiconductors.philips.com/pip/74HC04.html.
- [23] 74HC08. http://www.semiconductors.philips.com/pip/74HC08.html.
- [24] 74HC32. http://www.semiconductors.philips.com/pip/74HC32.html.
- [25] 74HC86. http://www.semiconductors.philips.com/pip/74HC86.html.
- [26] Ned Mohan, Tore M. Undeland, and William P. Robbins. Power Electronics, page 217. Wiley, second edition, 1995. ISBN 0471584088.
- [27] FDP6035L. http://www.fairchildsemi.com/pf/FD/FDP6035L.html.
- [28] Bruce Hofer. Measuring switch-mode power amplifiers. Audio Precision, Oct 2003.

Appendix A Schematic

