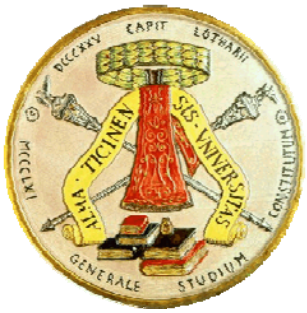
UNIVERSITÀ DEGLI STUDI DI PAVIA

FACOLTÀ DI INGEGNERIA

DIPARTIMENTO DI ELETTRONICA



DOTTORATO DI RICERCA IN MICROELETTRONICA
XXIII CICLO

CLASS-G HEADPHONES AMPLIFIER

Tutore:
Chiar.mo Prof. Rinaldo Castello

Coordinatore:
Chiar.mo Prof. Franco Maloberti

Tesi di dottorato di
Alex Lollio

Contents

Introduction

Today's portable devices present many challenges ranging from output power to maintaining high levels of efficiency. As devices continue to become more feature rich, consumers will continue to demand higher levels of performance along with minimal battery power consumption, putting further emphasis on proper design of key components such as the headphone amplifiers. Headphone amplifiers are migrating from Class AB to Class G technology. The differences and design advantages of each technology will be addressed in this work.

The first chapter shows an introduction on the human perception of a generic audio signal and a characterization of common audio signals in terms of amplitude and frequency distribution, and then it shows the most common system requirements for a headphone driver.

The first part of the second chapter illustrates the headphone moving coil transducer: its electrical model is the amplifier load, and it has a strong impact on the amplifier compensation network and the amplifier stability. At the end of this chapter class AB, class D and class G amplifiers are compared in terms of efficiency and linearity focusing on the advantages coming from the usage of the class G implementation.

The third chapter shows the implemented class G amplifier based on a novel switching approach realized by a very smooth handover between the stages. The switching distortion analysis presented in the first section gives guidelines for the component sizing. A prototype has been realized assuming the specifications of medium audio quality (the most part of the portable application market): the chosen architecture and the experimental results (comparing the key parameters

with the ones of similar amplifiers in literature and in commercial products) are shown in section four. The same class G core has been used to design another amplifier meeting the specifications for high audio quality players. The new specs have required an architectural change that is shown in the last section and this class-G implementation will be integrated, in Dec 2010, inside a novel Marvell audio codec.

1. Audio signals characterization

This chapter starts with an overview on the hearing mechanism, emphasizing the frequency response of the human ear. After that, in the second section, the most common system requirements for a headphone driver are defined (max output power, efficiency and linearity definitions). The last section of this chapter shows an analysis of the audio signals in terms of amplitude and frequency distribution. This is useful to define the whole system requirements in order to obtain the best performances for a realistic audio signal.

1.1 The hearing mechanism

There can be few branches of engineering science by such a basic division as the Subjectivist/rationalist dichotomy. Subjectivism is till a significant issue in the hi-fi section of the industry, but mercifully has made little headway in professional audio, where an intimate acquaintance with the original sound, and the need to earn a living with reliable and affordable equipment, provides an effective barrier against most of the irrational influences [1].

Most fields of technology have defined and accepted measures of excellence; car makers compete to improve MPH and MPG; computer manufacturers boast of MIPs (millions of instructions per second) and so on. Improvement in these real quantities is regarded as unequivocally a step forward. In the field of hi-fi, many people seem to have difficulty in deciding which direction forward is.

In evaluating the Subjectivist position, it is essential to consider the known abilities of the human ear. Contrary to the impression given by some

commentators, who call constantly for more psychoacoustical research, a vast amount of hard scientific information already exists on this subject, and some of it may be briefly summarized thus:

- The smallest step-change in amplitude that can be detected is about 0.3 dB for a pure tone. In more realistic situations it is 0.5 to 1.0 dB. This is about a 10% change [28].

- The smallest detectable change in frequency of a tone is about 0.2% in the band 500 Hz–2 kHz. In percentage terms, this is the parameter for which the ear is most sensitive [29].

- The least detectable amount of harmonic distortion is not an easy figure to determine, as there is a multitude of variables involved, and in particular the continuously varying level of program means that the level of THD introduced is also dynamically changing. With mostly low order harmonics present the just-detectable amount is about 1%, though crossover effects can be picked up at 0.3%, and probably lower. There is certainly no evidence that an amplifier producing 0.001% THD sounds any cleaner than one producing .005% [1].

- Interchannel crosstalk can obviously degrade stereo separation, but the effect is not detectable until it is worse than 20 dB, which would be a very bad amplifier indeed [1]

- Phase and group delay have been an area of dispute for a long time. As Stanley Lipshitz et al have pointed out, these effects are obviously

> perceptible if they are gross enough; if an amplifier was so heroically misconceived as to produce the top half of the audio spectrum three hours after the bottom, there would be no room for argument. In more practical terms, concern about phase problems has centered on loudspeakers and their crossovers, as this would seem to be the only place where a phase shift might exist without an accompanying frequency-response change to make it obvious. Lipshitz appears to have demonstrated [30] that a second order all-pass filter (an all-pass filter gives a frequency-dependant phase shift without level changes) is audible, whereas BBC findings, reported by Harwood [31] indicate the opposite, and the truth of the matter is still not clear. This controversy is of limited importance to amplifier designers, as it would take spectacular incompetence to produce a circuit that included an accidental all-pass filter. Without such, the phase response of an amplifier is completely defined by its frequency response, and vice-versa; in Control Theory this is Bode's Second Law [32], and it should be much more widely known in the hi-fi world than it is. A properly designed amplifier has its response roll-off points not too far outside the audio band, and these will have accompanying phase-shifts; there is no evidence that these are perceptible [1]

The picture of the ear that emerges from psychoacoustics and related fields is not that of a precision instrument. Its ultimate sensitivity, directional capabilities and dynamic range are far more impressive than its ability to measure small level changes or detect correlated low-level signals like distortion harmonics. This is unsurprising; from an evolutionary viewpoint the functions of the ear are to warn of approaching danger (sensitivity and direction-finding being paramount) and for speech. In speech perception the identification of formants (the bands of

harmonics from vocal-chord pulse excitation, selectively emphasized by vocal-tract resonances) and vowel/consonant discriminations, are infinitely more important than any hi-fi parameter. Presumably the whole existence of music as a source of pleasure is an accidental side-effect of our remarkable powers of speech perception: how it acts as a direct route to the emotions remains profoundly mysterious.

The human ear can nominally hear sounds in the range 20 Hz to 20,000 Hz (20 kHz). This upper limit tends to decrease with age, most adults being unable to hear above 16 kHz. The ear itself does not respond to frequencies below 20 Hz, but these can be perceived via the body's sense of touch [27].

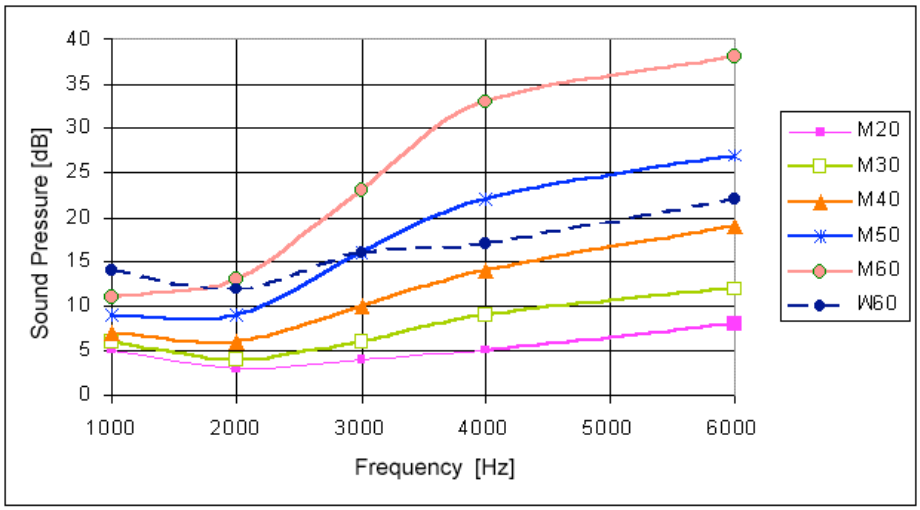


Figure 1: Thresholds of hearing for male (M) and female (W) subjects between the ages of 20 and 60

Frequency resolution of the ear is 3.6 Hz within the octave of 1,000–2,000 Hz. That is, changes in pitch larger than 3.6 Hz can be perceived in a clinical setting. However, even smaller pitch differences can be perceived through

other means. For example, the interference of two pitches can often be heard as a (low-) frequency difference pitch. This effect of phase variance upon the resultant sound is known as beating. The semitone scale used in Western musical notation is not a linear frequency scale but logarithmic. Other scales have been derived directly from experiments on human hearing perception, such as the mel scale and Bark scale (these are used in studying perception, but not usually in musical composition), and these are approximately logarithmic in frequency at the high-frequency end, but nearly linear at the low-frequency end.

The intensity range of audible sounds is enormous. Our ear drums are sensitive only to variations in the sound pressure, but can detect pressure changes as small as 2×10–10 atm and as great or greater than 1 atm. The ear can be exposed to short periods in excess of 120 dB without permanent harm — albeit with discomfort and possibly pain; but long term exposure to sound levels over 80 dB can cause permanent hearing loss.

A more rigorous exploration of the lower limits of audibility determines that the minimum threshold at which a sound can be heard is frequency dependent. By measuring this minimum intensity for testing tones of various frequencies, a frequency dependent absolute threshold of hearing (ATH) curve may be derived (see Figure 1). Typically, the ear shows a peak of sensitivity (i.e., its lowest ATH) between 1 kHz and 5 kHz, though the threshold changes with age, with older ears showing decreased sensitivity above 2 kHz.

The ATH is the lowest of the equal-loudness contours. Equal-loudness contours indicate the sound pressure level (dB), over the range of audible frequencies, which are perceived as being of equal loudness. Equal-loudness contours were first measured by Fletcher and Munson at Bell Labs in 1933 using pure tones reproduced via headphones, and the data they collected are called Fletcher-

Munson curves. Because subjective loudness was difficult to measure, the Fletcher-Munson curves were averaged over many subjects (see Figure 2).

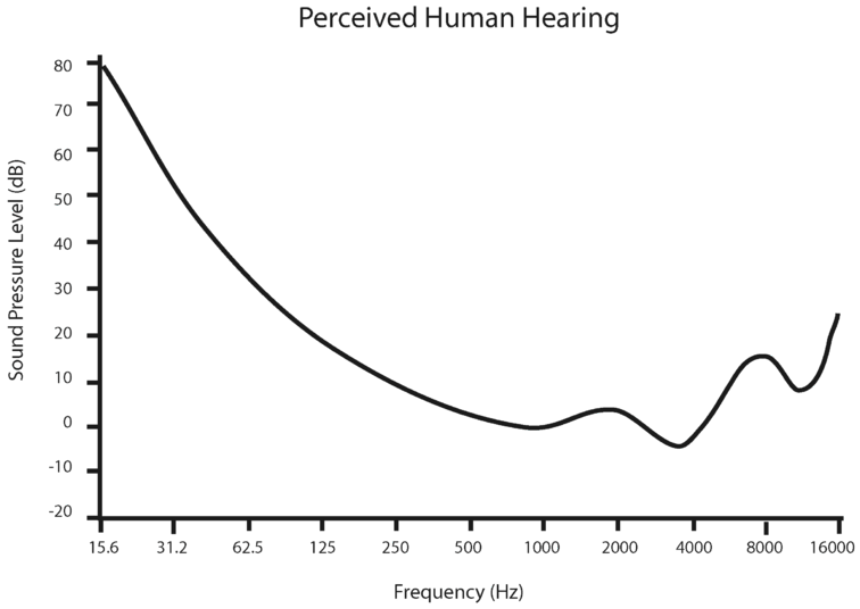


Figure 2: An equal-loudness contour. Note peak sensitivity between 2khz and 4khz, the frequency around which the human voice centers

Robinson and Dadson refined the process in 1956 to obtain a new set of equal-loudness curves for a frontal sound source measured in an anechoic chamber. The Robinson-Dadson curves were standardized as ISO 226 in 1986 (see Figure 3). In 2003, ISO 226 was revised as equal-loudness contour using data collected from 12 international studies.

1.1.1 Weighting filter

A-weighting is the most commonly used of a family of curves defined in the International standard IEC 61672:2003 and various national standards relating to the measurement of sound pressure level (as opposed to actual sound pressure). The weighting curves were originally defined for use at different average sound

levels, but A-weighting, though originally intended only for the measurement of low-level sounds (around 40 phon), is now commonly used for the measurement of environmental noise and industrial noise, as well as when assessing potential hearing damage and other noise health effects at all sound levels; indeed, the use of A-frequency-weighting is now mandated for all these measurements, although it is badly suited for these purposes, being only applicable to low levels so that it tends to devalue the effects of low frequency noise in particular [33].

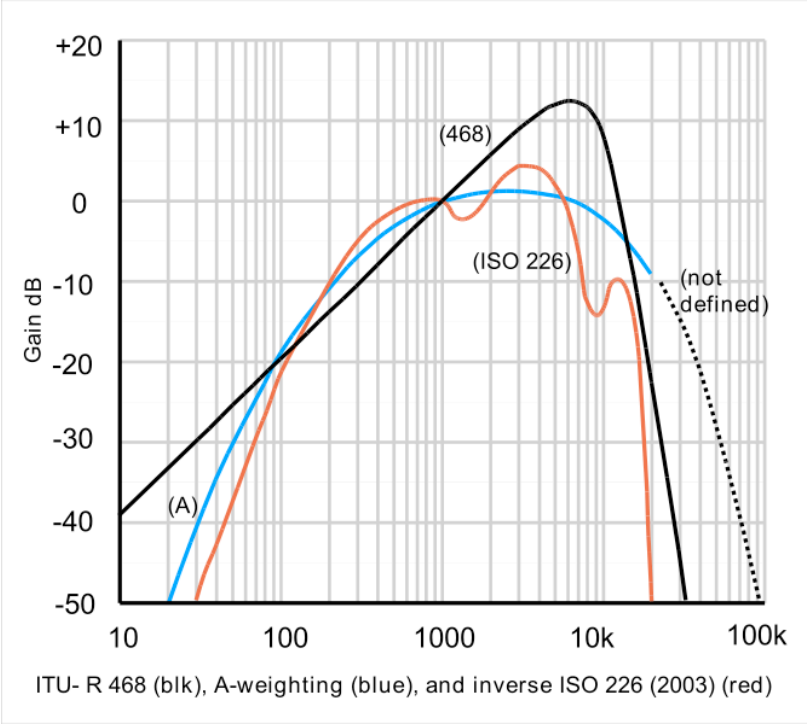


Figure 3: ITU-R 468 in black, A-weighting in blue and inverse ISO 226 in red.

An A-weighting filter is commonly used to emphasize frequencies around 3–6 kHz where the human ear is most sensitive, while attenuating very high and very low frequencies to which the ear is insensitive. The aim is to ensure that measured

loudness corresponds well with subjectively perceived loudness (see Figure 3). A-weighting is only really valid for relatively quiet sounds and for pure tones as it is based on the 40-phon Fletcher-Munson equal-loudness contour.

A-weighting filter is also used when measuring noise in audio equipment, especially in the U.S.A. In Britain, Europe and many other parts of the world, Broadcasters and Audio Engineers more often use the ITU-R 468 noise weighting (see Figure 3), which was developed in the 1960s based on research by the BBC and other organizations. This research showed that our ears respond differently to random noise, and the equal-loudness curves on which the A, B and C weightings were based are really only valid for pure single tones.

1.2 Audio signal characterization

It is important that the efficiency of audio amplifiers is measured correctly. Good test signals and adequate measurement procedures are crucial to make fair comparisons between amplifiers and reliably predict the dissipation in practical situations. This is a vital condition for judging the usefulness of new amplifier topologies. In fact, for the most part of the time, in headphone application, the output power is much less than the maximum possible. In this condition the quiescent consumption is more important than the consumption in middle output power level.

In literature, the efficiency of amplifiers is usually measured with sinusoidal signals. For headphone amplifiers based on a class D topology, this usually gives better results as for audio signals, as long as one bears in mind that the average output power of an audio amplifier while playing normal audio signals is much lower than its maximum sine output power. Also other high efficiency audio amplifiers need specific audio characteristics to obtain a high efficiency. Well

known topologies in this field are the class G and class H principles. The amplifiers described in [17][18][19][20][21][22] all use knowledge about either the amplitude or the frequency distribution of average audio signals. For this kind of amplifiers, measurements with sinusoids can give pessimistic results. The best signal would be a real audio signal, but this has several is advantages. The question is which audio signal should be taken. Speech? Music? What kind of music? This is not standardized. Furthermore, at least several seconds of audio are necessary to get a good impression, which is not very practical for simulations. Also, a music signal does not give stable readings on meters. Either the efficiency was measured indirectly by measuring heat sink temperatures [17], or an ad hoc measure is defined [21]. Another possibility is to use the IEC-268 'simulated program material' [22].In [23] and [24], the spectral distributions of program material were measured, and the latter also investigated whether the IEC test signal is useful for evaluating the power rating of loudspeakers. There is, however, no standard test signal intended for measuring or predicting amplifier efficiency.

1.2.1 Amplitude distribution

In order to compare test signals to realistic audio signals, it is necessary to define a test set of audio fragments. Due to the variation in volume in audio signals, the statistical parameters depend on the length of the time interval that is being analyzed. In [4] it has been chosen audio fragments with constant volume. Of course it should be noted that 'constant' is a relative measure, since the audio waveform itself is not constant. It has been chosen 80 fragments from various CD's, including classical music, pop music, jazz, hard rock, house, heavily compressed music, and speech signals. The length of each fragment is between 3 and 12 seconds. The volume during each fragment is constant. All fragments were converted to mono and normalized to full scale, with the highest sample just

clipping. The number of bits per sample was reduced to 8 to get smoother amplitude distributions. Because the fragments are normalized to full scale, this barely affects the sound impression.

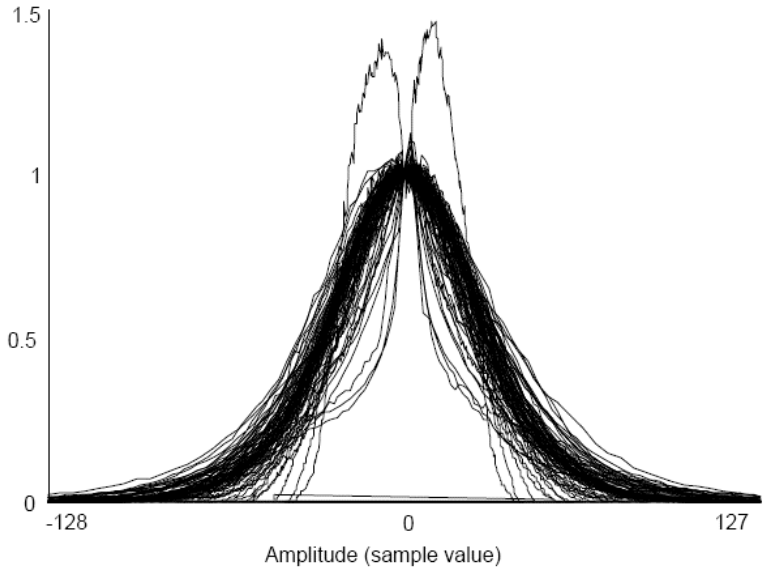


Figure 4: Amplitude distribution of 80 fragments normalized to 1@amplitude = 0 and then scaled to equal power

The amplitude distribution is determined by counting how many samples with a certain amplitude (28 = 256 levels) occur in one fragment. Figure 4 shows the amplitude distribution of all 80 fragments. It confirms that the shape of the amplitude distribution is Gaussian [23][19]. There are a few exceptions, though. Firstly, one curve has two peaks symmetrically around zero amplitude. This is the distribution of a fragment hard-core house music that contains purely synthesized sounds. Although this is an exceptional case, it shows the importance of realizing that certain audio characteristics can differ significantly from the average case. Secondly, we see some very narrow curves.

These are the distributions of speech signals. Due to the pauses inherent to spoken word, the distributions peak around zero amplitude. When discussing amplitude distributions, it is useful to critically examine the Peak-to-Average Ratio (PAR) [23][19]. It is widely acknowledged as a signal property, and identical to the traditional crest factor. Expressed in dB's, the PAR is defined as:

$$PAR = 20Log\left(\frac{U(t)_{max}}{U_{RMS}}\right) \qquad \mathbf{1}$$

Figure 5 shows the PARs of all fragments. Roughly, it is between 10dB and 20dB, with an average of 15dB. This means that -in order to be undistorted- the average audio fragment must have a power at least 12 dB below a full power sine wave.

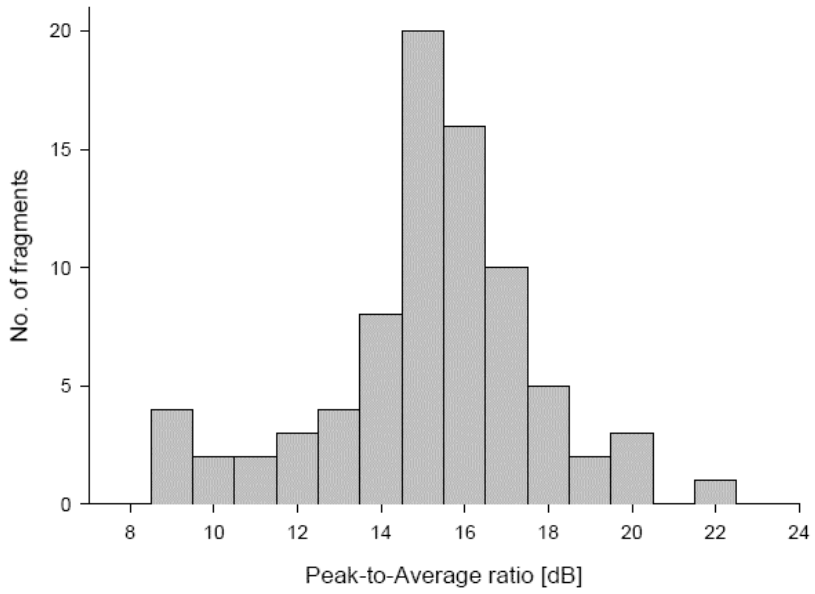


Figure 5: Peak to average ratio of all fragments of Figure 4

Often, the PAR is also used for calculating amplifier efficiencies, resulting in certain efficiency for a certain PAR of the signal. In that case it is assumed that every fragment is amplified to a level just below clipping. The result is that the amplifier dissipation strongly depends on the PAR. The reason for this is, that the average power (or U_{RMS}) also varies considerably, since $U(t)_{max}$ is the clipping point of the amplifier and therefore constant. In Figure 4, however, it can be seen that, when scaled to equal power, the amplitude distributions are almost the same. $U(t)_{max}$ varies, but since the high amplitudes near $U(t)_{max}$ are unlikely to occur, they hardly effect the total dissipation of the amplifier. When a fragment with a large PAR is amplified to equal power as a fragment with a low PAR, there will be some clipping, but this is barely perceptible in normal listening conditions. Only when we increase the volume a lot, the sound quality degrades. Subjective listening tests show that the PAR can be made as small as 6dB before most fragments sound really bad through clipping. A PAR of 6dB means that the output power is half the maximum sine power. From the above we conclude the following: Audio fragments of constant volume generally have a Gaussian amplitude distribution with an average PAR of 15dB. Concerning amplifier dissipation, average power is the most important variable, while the PAR does not play a significant role. Amplifier dissipation for Gaussian signals must be tested up to half the full sine power.

1.2.2 Frequency distribution

On the same audio fragments, a Fast Fourier Transform (FFT) was performed over the full length. A normal log-log Bode plot of the frequency content (Figure 6) does not provide very useful information.

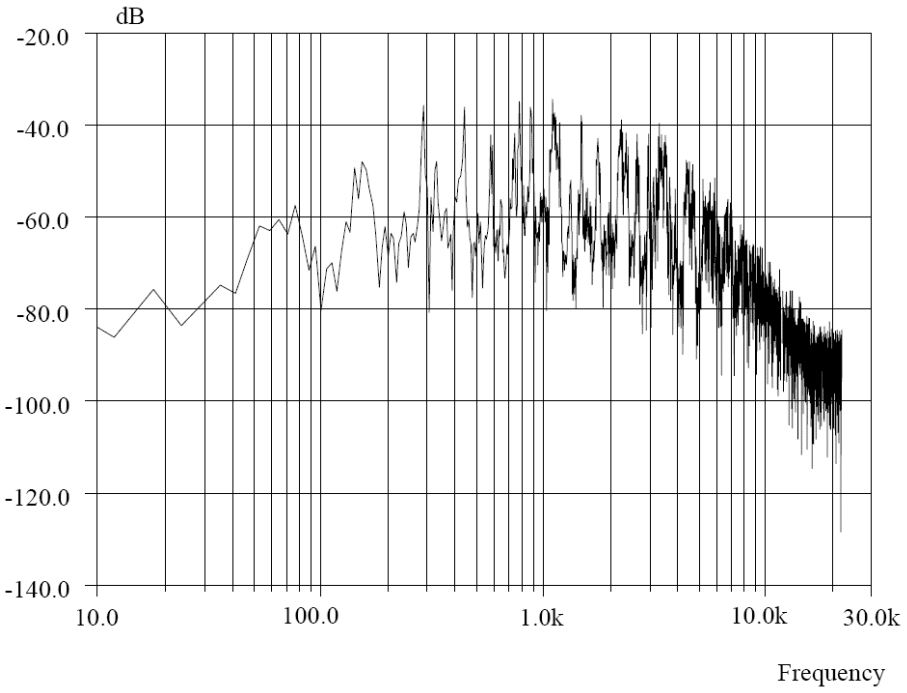


Figure 6: Traditional graph of a Fourier transform of a music fragment. Vertical scale dB's are relative to full scale for measurement bandwidth $2/T_{fragment}$.

Firstly, there is no need for a high accuracy, so it seems more logical to choose the vertical scale of the plot linear instead of logarithmic. Secondly, efficiency is a matter of power. When an amplifier has a better efficiency for certain frequencies, it is important to know how much power is present in those frequencies, not how much amplitude. So it's more useful to square the amplitudes. Finally, the squared FFT gives the power of the frequencies in the signal. The frequencies are linearly spaced. With a logarithmic frequency axis, a temptation exists to overemphasize the lower frequencies because they are relatively enlarged. A linear frequency axis might seem a logical choice, but since pitch perception is logarithmic in nature (every octave higher equals a factor two), it is preferable to use a logarithmic axis,

and plot the sum of the squared Fourier coefficients. An extra advantage is that the summation smoothes the curve.

Presented in this way, the frequency distribution is a line that starts at (almost) power = 0 at 20Hz, climbing to power = 1 at 20kHz. The frequency distributions of all fragments are shown in Figure 7. The average fragment is S-shaped, with a mid-frequency part corresponding to a straight line between (20Hz, 50Hz) and (3kHz, 20kHz). This does not come as a surprise when we realize that the notes in a musical scale are fixed factors in frequency apart, in which case a linear frequency distribution requires all notes to be equally loud.

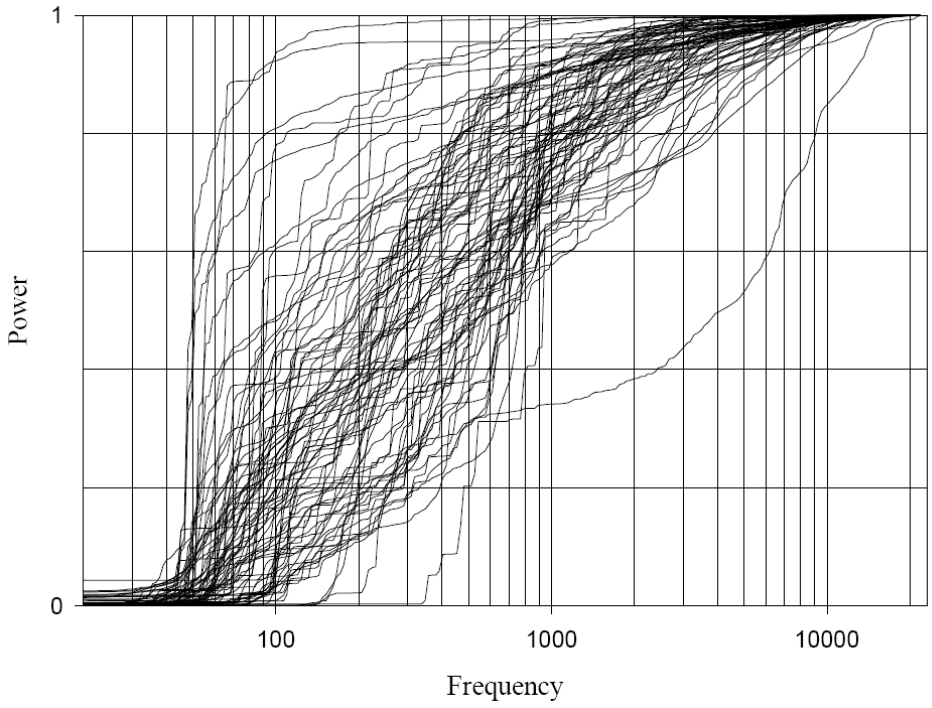


Figure 7: Frequency distribution of all audio fragments

In Figure 7, the fragments with much power in the lower frequencies have a house beat or a contrabass. The fragments with much power in the higher frequencies

mostly have electric guitars or synthesizers. One fragment in particular stands out because it contains much more high frequencies than the others. It is the intro of Melissa Etheridge's 'Like the way I do', containing a guitar and a tambourine.
A valid approximation is that, for the most part of the audio signals reported in Figure 7, the 80% of the power stays between 20Hz and 1kHz.

1.2.3 IEC-60268 test signal

The International Electrotechnical Commission (IEC) has defined a noise input signal representative for normal program material [25]. It is generated by a pink or white noise source followed by a filter. We will refer to this signal as the 'IEC signal', and investigate if it is useful for efficiency measurements. Figure 8 shows that the amplitude distribution of the IEC signal is Gaussian.

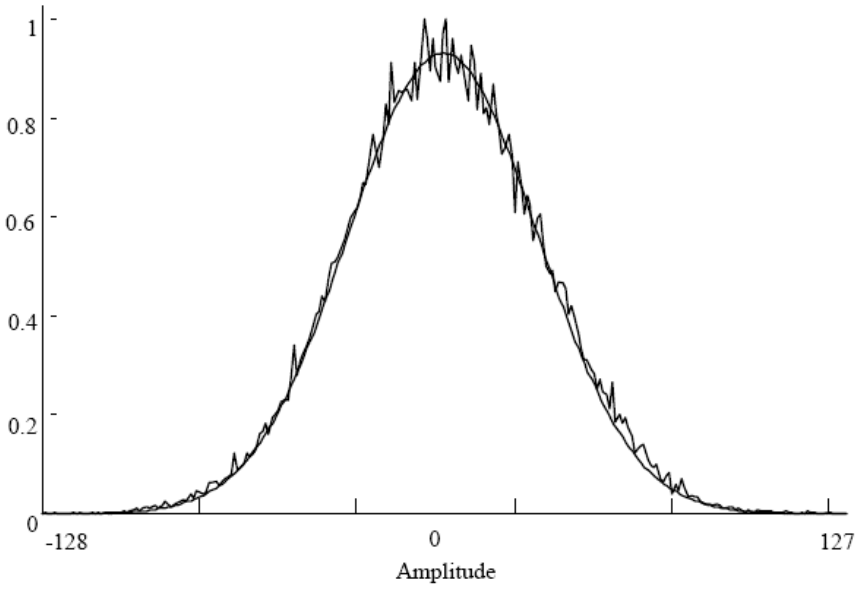


Figure 8: Amplitude distribution of the IEC-60268 test signal and a Gaussian curve as reference

This standard applies to sound systems of any kind, and to the parts of which they are composed or which are used as auxiliaries to such systems. This standard deals with the determination of the performance of sound system equipment, the comparison of these types of equipment and the determination of their proper practical application, by listing the characteristics which are useful for their specification and laying down uniform methods of measurements for these characteristics. The standard is confined to a description of the different characteristics and the relevant methods of measurement; it does not in general specify performance.

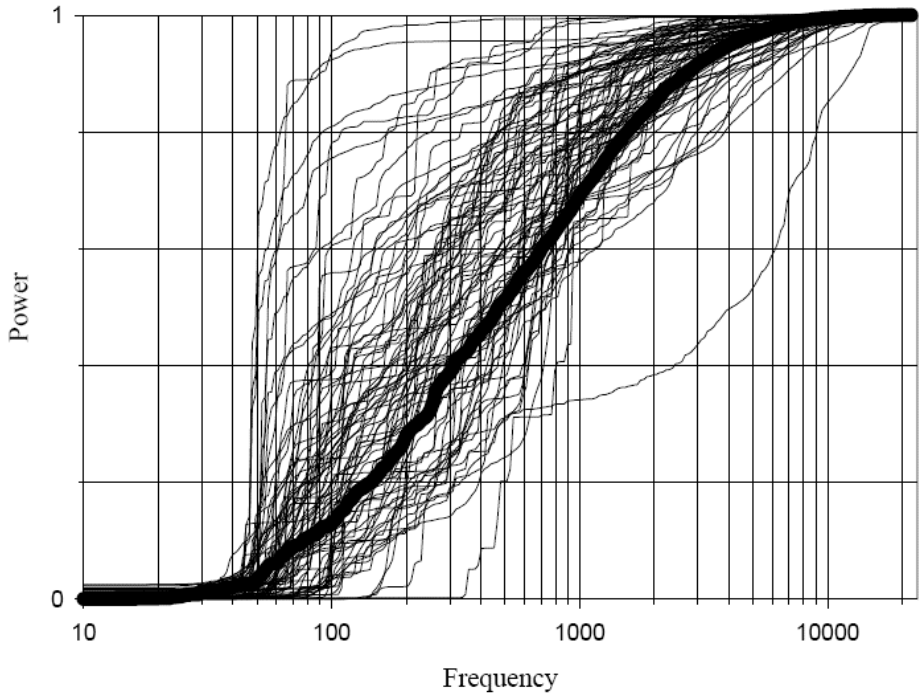


Figure 9: Frequency distribution of the fragments and of the IEC test signal (fat line)

Figure 9 shows the IEC signal frequency distribution, together with the distribution of the fragments. The IEC signal serves well as a typical audio fragment.

1.3 System requirements

Headphones are a pair of small loudspeakers situated close to the user's ears. They are also known as stereophones, headsets or, colloquially cans. The in-ear versions are known as earphones or earbuds. The user connects all of them to a signal source such as an audio amplifier, radio or CD player. The first difference between common loudspeakers and headphones is that in one case the ear is immersed in a propagating sound field and in the other it registers the SPL in a leak pressure chamber. In fact, the sound field of headphones is confined to a relatively small volume of up to about 30 cm^2.

1.3.1 Amplifier output power

Essentially, an audio amplifier is a normal voltage amplifier optimized for the amplification of audio signals. The limited frequency response of the ear sets the bandwidth limits: 20Hz - 20kHz, although most people are not able to hear 20kHz. Most power is concentrated in the mid frequencies, and occasionally in the low frequencies. Regarding the specification of the amplifier output power is important to remember that you have to increase it by an awful lot to make the amplifier significantly louder [1]. We do not perceive acoustic power as such – there is no way we could possibly integrate the energy liberated in a room. It is much nearer the truth to say that we perceive pressure. In fact, a generic sound source radiates power and this result in a sound pressure (SP). In particular, the Sound Pressure Level (SPL) is used to measure the sound intensity and it is defined as the local pressure deviation from the ambient (average, or equilibrium)

pressure caused by a sound wave [2]. Figure 10 shows the sound pressure diagram for a generic audible sound. It is well known that power in watts must be quadrupled to double sound pressure level (SPL) but this is not the same as doubling subjective loudness; this is measured in Sones rather than dB above threshold, and some psychoacousticians have reported that doubling subjective loudness requires a 10dB rather than 6dB rise in SPL, implying that amplifier power must be increased tenfold, rather than merely quadrupled [3].

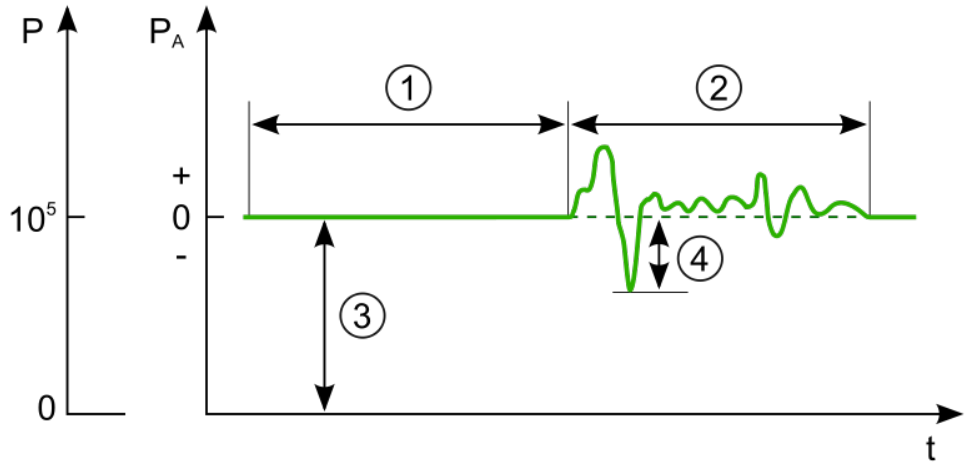


Figure 10: Sound pressure diagram: 1. silence, 2. audible sound, 3. atmospheric pressure, 4. instantaneous sound pressure

The ear has a very large dynamic range. To give an example: the ratio between the acoustic power of a rock concert and the sound of breathing can be as large as 10^{11}. This makes large demands on the dynamic range of the audio amplifier. To get an idea about the order of magnitude of amplifier output powers, refer to Table 1. The SPL's have been taken from [1]. Table 1 displays some situations in which audio power amplifiers can be used. The first column gives different sources of

sound, the second column shows the Sound Pressure in Pascal and the third column shows the Sound Pressure Level (SPL) in dB's. 0dB SPL is the hearing threshold and defined as 0.00002 N/m2.

Source of sound in air	Sound Pressure [Pa]	SPL [dB]	Power in Headphones (RMS)
calculated Krakatoa explosion at 100 miles (160 km) in air	20'000 Pa	180dB	-
Jet engine at 30 m	632 Pa	150dB	-
Hearing damage (ist.)	20 Pa	120dB	1W
Jack hammer at 1 m	2 Pa	100dB	10mW
Traffic on a busy roadway at 10 m	2×10^{-1} / 6.32×10^{-1} Pa	80/90dB	1 mW
Hearing damage (over long-term exposure, need not be continuous)	0.356 Pa	85dB	0.3mW
Passenger car at 10 m	2×10−2 / 2×10−1 Pa	60/80dB	30 uW
TV (set at home level) at 1 m	2×10−2 Pa	approx. 60 dB	1 uW
Normal conversation at 1 m	2×10−3 / 2×10−2 Pa	40/60dB	0.3 uW

Very calm room	2×10^{-4} / 6.32×10^{-4} Pa	20/30dB	1 nW
Auditory threshold at 1 kHz	2×10^{-5} Pa (RMS)	0dB	1 pW

Table 1: Example of Sound Pressure Level for different sources of sound and the equivalent power delivered to a 32Ohms Headphones load with 90dB/mW of sensitivity.

Now suppose we want to reproduce these SPL's with an audio amplifier and a couple of headphones. Assuming that the headphone has a load impedance of 32Ohms and an efficiency of 90dB/mW (possible values for commercial headphones), the needed speaker power can be calculated. Finally, audio signals have an average power that is considerably lower than their peak power, so, for undistorted sound, the maximum sine power rating of an amplifier should in average be 12dB higher than the average power delivered to the speaker. From Table 1 we conclude that audio amplifiers for headphones application must operate over a wide range of power levels. The maximum output power should be higher than 20mW in order to reproduce, with fidelity, sound showing high SPLs. The lower limit to the minimum delivered output power is limited by the noise floor. This is the reason because a headphone driver must have a very high Signal to Noise Ratio (SNR).

1.3.2 Efficiency

The efficiency of an audio amplifier is hardly important in systems that use audio amplifiers. In particular, in battery powered equipment, the dissipation should be minimal for the longest battery life time. The efficiency, η, of a generic system is defined as:

$$\eta = \frac{P_O}{P_{SUP}} = \frac{P_O}{P_O + P_{DISS}} \quad \textbf{2}$$

where PO is the output power and PDISS is the power dissipated by the amplifier. PSUP is the total power which is given by the supply to both the amplifier and the load. In literature the most common measurement graphs depict the efficiency of an audio amplifier as a function of output power as shown in Figure 11. A problem with these kinds of charts is that it is difficult to see how much the amplifier actually dissipates [4]. The dissipation of an amplifier in relation to the output power Po and the efficiency η is:

$$P_{DISS} = P_O(\frac{1}{\eta} - 1) \quad \textbf{3}$$

which makes it not very easy to see that the right amplifier in Figure 11 dissipates 50% more than the left one at full power. The fact that the left amplifier has a 50% higher quiescent power dissipation (which seriously affects the battery life of e.g. a portable radio) is not visible at all, since the efficiency is always zero at zero output power. Moreover, a headphone amplifier delivers a very small amount of the maximum output power for the most part of the time. In fact, in headphone application, it is more relevant the power consumption at a very low output power level than the one at middle power level. This means that the majority of the graph displays useless information.

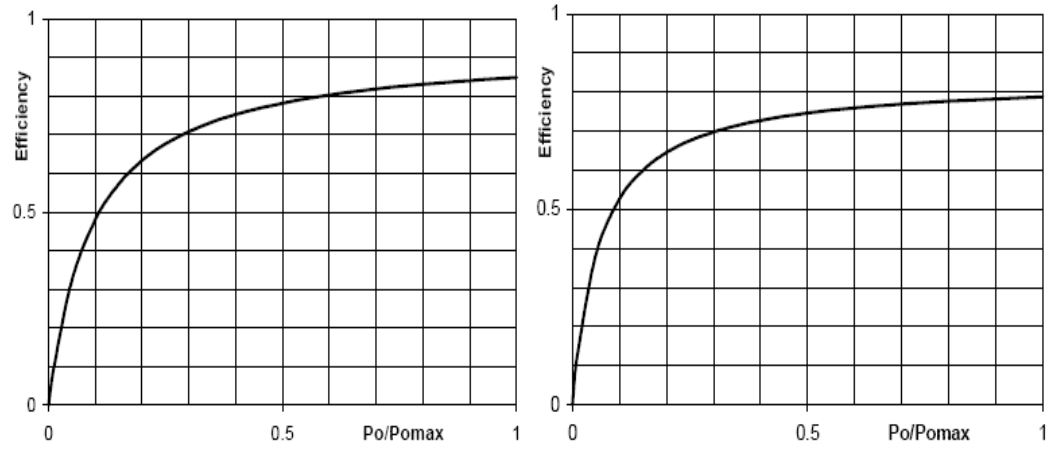


Figure 11: Simulated efficiency of two hypothetical audio amplifiers with different quiescent- and maximum dissipations.

From now on we will therefore use graphs as displayed in Figure 12. The dissipation for the whole power range is clearly visible thanks to the logarithmic x-axis, and also the maximum- and quiescent power dissipation can easily be observed.

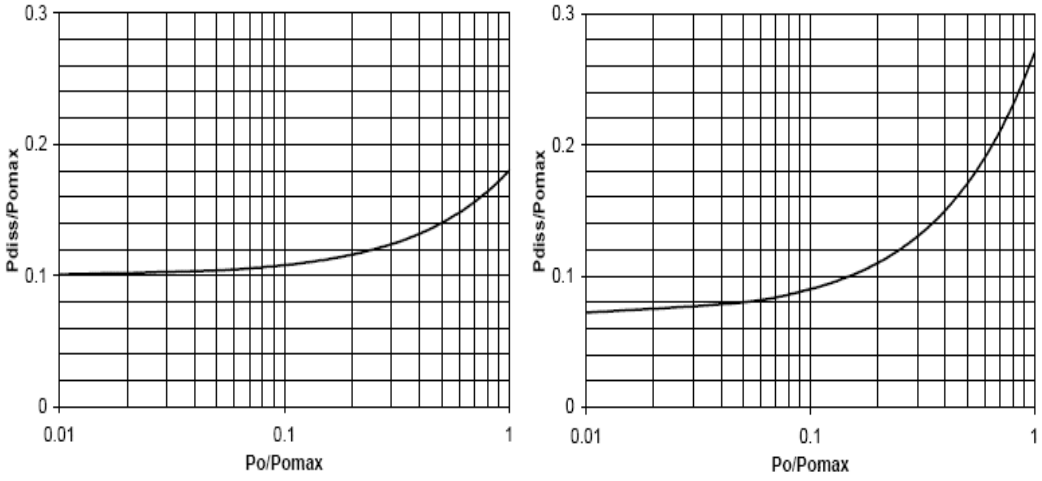


Figure 12: Dissipation of the two amplifiers in Figure 11

1.3.3 Distortion

Making a high efficiency audio amplifier would be a lot simpler if its distortion was not important. A class D amplifier on a low switching frequency can have an excellent efficiency, but its distortion will be too high. The design of class G amplifiers is complicated by switching distortion. Therefore, a low distortion is an important condition when judging efficiency. There are several types of distortion that can be measured:

Total Harmonic Distortion (THD)

When a sinusoidal signal is applied to a non-linear amplifier, the output contains the base frequency plus higher order components that are multiples of the base frequency. The Total Harmonic Distortion is the ratio between the power in the harmonics and the power in the base frequency. This can be measured on a spectrum analyzer. Most distortion analyzers, however, subtract the base signal from the amplifier's output and calculate the ratio between the total RMS value of the remainder and the base signal. This is called THD+N: Total Harmonic Distortion + Noise. Normally, the noise will be low compared to the distortion, but the noise of a noisy amplifier or the switching residues in a class D amplifier can give garbled THD figures. For a THD+N measurement, the bandwidth must be specified. For class D measurements, a sharp filter with a 20kHz corner frequency is necessary to prevent switching residues -that are inaudible- to show up in the distortion measurements.

InterModulation distortion (IM)

When two sinusoids are summed and applied to a non-linear amplifier, the output contains the base frequencies, multiples of the base frequencies and the difference

of (multiples of) the base frequencies. Suppose a 15kHz sinusoid is applied to an audio system that has a 20kHz bandwidth, and the THD+N needs to be measured. All the harmonics are outside the bandwidth and will be attenuated, resulting in too low a THD+N reading. The same situation occurs when the distortion analyzer has a 20kHz bandwidth. In these cases, an IM measurement can be a solution. The first standard was defined by the SMPTE (Society of Motion Picture and Television Engineers). A 60Hz tone and a 7kHz tone in a 4:1 amplitude ratio are applied to the non-linear amplifier. The 60Hz appears as sidebands of the 7kHz tone. The intermodulation distortion is the ratio between the power in the sidebands and the high frequency tone. Another common standard is defined by the CCITT (Comité Consultatif Internationale de Télégraphie et Téléphonie), and uses two tones of equal strength at 14kHz and 15kHz. This generates low frequency products and products around the two input frequencies, depending on the type (odd or even) of distortion.

Interface InterModulation distortion (IIM)

In this test, the second tone of an IM measurement set-up is not connected to the input, but to the output (in series with the load impedance) [5], [6].

Transient InterModulation distortion (TIM)

When a square wave is applied to an amplifier with feedback, its input stage has to handle a large difference signal, probably pushing it into a region that is less linear than its quiescent point. When a sinusoid is added to the square wave, the nonlinearity

induced by the edges of the square wave will distort the sinusoid, giving rise to TIM, also called transient distortion or slope distortion [7]. There are many ways of testing TIM and it remains unclear how much it adds to the existing

measurement methods. If the maximum input signal frequency during normal operation of an amplifier is limited to 20kHz, a 20kHz full power sinusoid is the worst case situation. When that generates little distortion, TIM will not occur [8].

Cross-over distortion

Cross-over distortion is generated at the moment the output current changes sign. At that moment, the output current gets supplied by another output transistor. The process of taking over generates distortion, visible as spikes in the residual signal of a THD measurement. This kind of distortion is notorious for its unpleasant sound (a small percentage error is quickly noticeable). Because it's usually present around zero amplitude, the impact on small signals can be relatively large.

Distortion summary

There is no consensus as to which distortion measurements are essential. In the ongoing search for the critical attributes that determine the 'sound' of an audio amplifier, many other mechanisms can play a role like reactive harmonic distortion [9], the spectrum of the distortion [10], non-linear crosstalk [IEC60268-1], memory effects [11], granularity distortion [12], and external influences like speaker cables [13], decoupling capacitors. It is unclear to what extent these concepts influence the 'sound' of an amplifier. Also, alternative measurement methods have been described, like measuring the difference between input and output of an amplifier for audio signals [14], or analyze the output signal in Volterra space [15]. The most used methods used to characterize the headphone audio amplifiers are the observation of the residual signal in a THD measurement in function of the output power and the signal frequency. However, it is acknowledged that THD measurements, taken with the usual notch type analyzer, are of limited use in predicting the subjective impairment produced by an

imperfect audio path [1]. With music, etc. intermodulation effects are demonstrably more important than harmonics. However, THD tests have the unique advantage that visual inspection of the distortion residual gives an experienced observer a great deal of information about the root cause of the non-linearity.

1.3.4 Power Supply Rejection Ratio

An important parameter in headphone application is the Power Supply Rejection Ratio (PSRR) which is defined as the ratio of the change in supply voltage (ΔVsupply) to the corresponding change in output voltage (ΔVout) of the device:

$$\mathrm{PSRR} = \frac{\Delta V_{supply}}{\Delta V_{out}} \qquad 4$$

In integrated headphone driver usually the negative supply voltage is generated by using a charge pump circuitry. This corresponds to a high noisy negative supply reference which could also affect the amp linearity. In fact, it is important that PSRR value is high enough to avoid distortion caused by variation of the supply (if you want higher linearity you need higher PSRR). During the design it is necessary to keep in mind this specification.

1.3.5 Crosstalk

In audio integrated amplifier another important parameter is the channels crosstalk. It is any phenomenon by which a signal transmitted on one channel creates an undesired effect in another channel [16]. In integrated circuit design, crosstalk normally refers to a signal affecting another nearby signal. Usually the coupling is capacitive, and to the nearest neighbor, but other forms of

coupling and effects on signal further away are sometimes important, especially in analog designs. For example a channel crosstalk could be conveyed through the integrated circuit substrate. There are a wide variety of possible fixes, with increased spacing, wire re-ordering, and shielding being the most common.

1.3.6 Noise

The noise there should be as little as possible without compromising other parameters. The noise performance of a power amplifier is very significant, especially in modern portable audio systems. The dynamic range of a system is the ratio, generally measured in dB, of its maximum undistorted output signal (THD < 1%) to its residual output noise or noise floor. Today, high-performance headphones audio amplifier needs a dynamic range up to 100 dB. The trend in modern audio codecs is toward increasing dynamic range by increasing resolution in available digital converters. However, analog signals accumulate noise and once noise is added to a signal, it's essentially impossible to remove it without altering or degrading the original signal. Therefore, noise and interference must be prevented along the entire signal path. This is a challenge in modern integrated audio codecs where the negative supply voltages are usually generated using internal noisy charge pumps. Moreover, a predictable amount of random or "white" noise is inherent in all electronic devices and must be expected. It is well known that the noise contributed of each stage of an amplifier decreases as the gain preceding the stage increases, implying that the first few stages in a cascade are the most critical [26]. In addition the feedback resistors used to close the opamp loop are a source of thermal noise. For this reason, their value must be chosen as low as possible to maintain their thermal noise value under the amplifier noise floor.

1.4 Conclusions

In this chapter we have seen that to characterize an audio amplifier we have to use the aweighted filter in order to simulate the hearing response. After that, we have seen the audio characterization of common real audio signals and finally, we have reported the key parameters for an audio amplifier and, finally.

1.5 References

[1] D. Self, Audio Power Amplifier Design Handbook, 4th ed. Oxford, U.K.: Newnes-Elsevier, 2006

[2] Wikipedia contributors, "Sound pressure," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Sound_pressure&oldid=359766171 (accessed May 5, 2010).

[3] Moore, B. J. An Introduction to the Psychology of Hearing, Academic Pressm 1982, pp. 48-50.

[4] Ronan van der Zee, High Efficiency Audio Power Amplifiers; design and practical use, 1999.

[5] E.M. Cherry and G.K.Cambrell, "Output resistance and intermodulation distortion of feedback amplifiers", J. Audio Eng. Soc. Vol. 30, No.4, pp. 178-190, April 1982

[6] Matti Otala, Jorma Lammasniemi "Intermodulation at the amplifierloudspeaker interface", Wireless World, pp. 45-7 Nov. 1980 and pp. 42-4 Dec. 1980

[7] M. Otala "Transient distortion in transistorized audio power amplifiers" IEEE Trans. Audio Electroacoust., Vol. AU-18, pp. 234-9, Sept. 1970

[8] Peter Garde "Amplifier First-Stage Criteria for Avoiding Slew-Rate Limiting", J. Audio Eng. Soc. Vol. 34, No. 5, pp. 349-53, May 1986.

[9] Klaas Bult "Analog CMOS square-law circuits", PhD Thesis University of Twente, ISBN 90-9002025-X, Jan. 1988.

[10] Ben Duncan "Spectrally challenged: the top 10 audio power chips", Electronics world + wireless world, pp. 804-10, Oct. 1993.

[11] Perrot, Gérard "Measurement of a Neglected Circuit Characteristic", 100th convention of the Audio Engineering Society, Copenhagen, preprint #4282, 11-14 May 1996.

[12] Douglas Self, “Ultra-Low-Noise Amplifiers & Granularity Distortion” Journ. Audio Eng.Soc, pp. 907- 915, Nov. 1987.
[13] Duncan, B., “Measuring speaker cable differences”, Electronics world, pp.570-1, July/Aug. 1996.
[14] Collins, Andrew R. “Testing Amplifiers With A Bridge”, Audio, pp 28-32, March 1972.
[15] Martin J. Reed and Malcolm O.J. Hawksford “Comparison of Audio System Performance in Volterra Space”, 103rd convention of the Audio Engineering Society, New York, preprint #4606, September 1997.
[16] Wikipedia contributors, "Crosstalk (electronics)," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Crosstalk_(electronics)&oldid=346789948 (accessed May 5, 2010)
[17] E. Botti, T. Mandrini, F. Stefani "A High-Efficiency 4x20W Monolithic Audio Amplifier for Automobile Radios Using a Complementary DMOS BCD Technology", IEEE J. of solid state circuits, Vol. 31, No. 12, pp. 1895-901, Dec. 1996.
[18] T. Sampei, S. Ohashi, Y. Ohta, S. Inoue, "Highest efficiency and super quality audio amplifier using MOS power FETs in class G operation", IEEE Transactions on Consumer Electronics, Vol. CE-24, No. 3, pp.300-7, Aug. 1978.
[19] Frederick H. Raab, "Average efficiency of class-G power amplifiers", IEEE Transactions on Consumer Electronics, Vol. CE-32, No. 2, pp.145-50, May 1986.
[20] Jørgen Arendt Jensen “A New Principle for a High-Efficiency Power Audio Amplifier for Use with a Digital Preamplifier”, J. Audio Eng. Soc., Vol. 35, No. 12, pp. 984-93, Dec. 1987.
[21] Karsten Nielsen “High-Fidelity PWM-Based Amplifier Concept for Active Loudspeaker Systems with Very Low Energy Consumption”, Journal of the Audio Eng. Soc., Vol. 45, No. 7/8, pp. 554-70, July/Aug. 1997.
[22] R.A.R. van der Zee and A.J.M. van Tuijl "A High Efficiency Low Distortion Audio Power Amplifier", 103rd convention of the Audio Engineering Society, New York, preprint #4601, Sept. 1997.
[23] Peter John Chapman “Programme Material Analysis”, 100th convention of the Audio Engineering Society, Copenhagen, preprint #4277, May 1996.
[24] R.A. Greiner, Jeff Eggers " The Spectral Amplitude Distribution of Selected Compact Discs" Journal of the Audio Eng. Soc., Vol. 37, pp. 246-75, April 1989.
[25] IEC (International Elektrotechnical Committee), Publication IEC60268-16, third edition, 2003.
[26] B Razavi, RF Microelectronics, Prentice Hall, Upper Saddle River, NJ, 1998

[27] Wikipedia contributors, "Psychoacoustics," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Psychoacoustics&oldid=355561791 (accessed May 6, 2010).
[28] Harris, J D Loudness discrimination J. Speech Hear. Dis. Monogr. Suppl. 11, pp. 1–63.
[29] Moore, B C J Relation between the critical bandwidth k the frequency-difference limen Journ. Acoust. Soc. Am. 55, p. 359.
[30] Lipshitz et al, On the audibility of midrange phase distortion in audio systems JAES, September 1982, pp. 580–595.
[31] Harwood, H Audibility of phase effects in loudspeakers Wireless World, January 1976, pp. 30–32.
[32] Shinners, S Modern control system theory and application publ. Addison-Wesley, p. 310.
[33] Wikipedia contributors, "A-weighting," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=A-weighting&oldid=355697561 (accessed May 6, 2010).

2. Headphone amplifier

The first section of this chapter starts with a brief introduction on the different technologies used to realize headphones transducers. After that, it shows detailed information on the most used one, the “dynamic” driver. The sensitivity parameter of the headphones, reported in this section, is used to derive the noise floor requirement for the headphones amplifier. Moreover, the first section shows the electrical model of the moving coil transducer which is useful for dimensioning the amplifier compensation network and for studying the amplifier stability. The last section shows a comparison between class AB, class D and class G headphone amplifier in terms of linearity, efficiency and number of external components. This section emphasizes the advantages of using class G amplifier as headphone driver.

2.1 Headphones transducers

Headphones are a pair of small loudspeakers situated close to the user's ears. They are also known as stereophones, headsets or, colloquially cans. The in-ear versions are known as earphones or earbuds. The user connects all of them to a signal source such as an audio amplifier, radio or CD player. The first difference between common loudspeakers and headphones is that in one case the ear is immersed in a propagating sound field and in the other it registers the SPL in a leak pressure chamber [1]. In fact, the sound field of headphones is confined to a relatively small volume of up to about 30 cm^2.

There are several technologies used to realize a headphone transducer and they employ one or more of several methods of sound reproduction:

- The **moving coil driver**, more commonly referred to as a "dynamic" driver [2] is the most common type used in headphones. Figure 13 shows the cross section of the moving coil transducer. The operating principle consists of a stationary magnetic element affixed to the frame of the headphone which sets up a static magnetic field (the main components of this type of magnet are either neodymium or a ferrite composite). The diaphragm of the headphone is attached to a coil of wire (voice coil) which is immersed in the static magnetic field. The diaphragm is actuated by the attached voice coil, when an audio current is passed through the coil. The alternating magnetic field produced by the current through the coil reacts against the static magnetic field in turn, causing the coil and attached diaphragm to move the air, thus producing sound. This transducer has a relatively large cavity and for this reason cannot be expected to be accurate above 1 kHz.

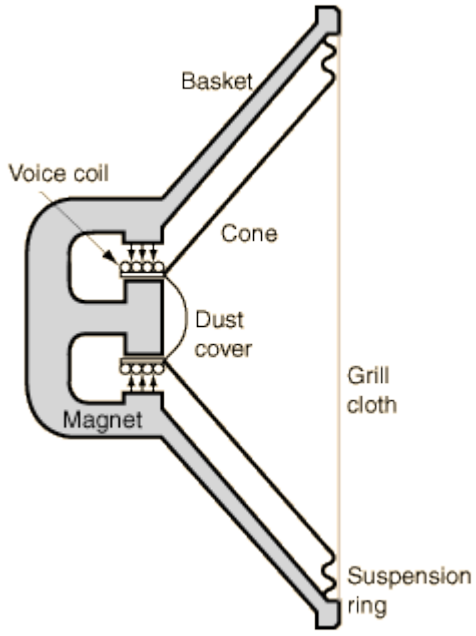


Figure 13: Cross section of the moving coil transducer

- **Piezoelectric speakers** are transducers that rely on the piezoelectric effect instead of a moving wire coil to generate sound [3]. The small form factor associated with piezoelectric speakers is quite attractive to the portable electronics market where size and weight are significant factors. These thin ceramic loudspeakers require a significant amount of voltage drive to achieve respectable sound pressure levels. This higher voltage requirement poses a problem for the speaker amplifiers typically used in handheld applications, because those amplifiers cannot generate such high voltages.
- **Electrostatic** drivers consist of a thin, electrically charged diaphragm, typically a coated PET film membrane, suspended between two perforated metal plates (electrodes) [3]. The electrical sound signal is applied to the electrodes creating an electrical field; depending on the polarity of this field, the diaphragm is drawn towards one of the plates. Air is forced through the perforations; combined with a continuously changing electrical signal driving the membrane, a sound wave is generated. Electrostatic headphones are relatively uncommon.
- An **electret** driver functions along the same electromechanical means as an electrostatic driver [2]. However the electret driver has a permanent charge built into it, whereas electrostatics have the charge applied to the driver by an external generator. Electret headphones, like electrostatics are relatively uncommon.

As we have seen, there are several types of transducers which can be used in headphone application, however, up to the 99% of the commercial products use the moving coil transducer.

There are several specifications related to headphones, for example: the frequency response, the diffuse field equalization, the distortion, the sensitivity (loudness)

and the impedance. In particular, we are focusing on the sensitivity and on the impedance because they have an important role during the amplifier design.

The sensitivity is a measure of headphone efficiency in dBs SPL per milliwatt of input. A low number means that the headphones need more power to sound as loud as those which have a higher sensitivity. Headphones for portables need to be fairly sensitive because of the lower power output of portable stereos. Modern dynamic headphones have sensitivity ratings of 90 dB or more. This parameter is useful to define the maximum output power of the amplifier and its noise floor (headphones with very high sensitivity need an amplifier which has a minimal output noise).

The impedance is a measure of the nominal headphone load and stated in ohms. Both consumer and professional headphones generally have impedances of less than 100 ohms (32 Ohms). There are professional models rated at 200 ohms or more to minimize loading effects on distribution amplifiers which are often drive a whole bank of headphones at one time. As the sensitivity the value of the nominal impedance is an important parameter, which must be taken into account, during the amplifier design. In fact, less impedance means less amplifier gain and this could affect the amplifier stability. With this scope, Table 2, shows a comparison between different headphone products highlighting their sensitivity, impedance and price.

Model name	Sensitivity	impedance	Headphone	Price
Koss UR 20	97dB/mW	32	Full-size	20$[1]
Sennheiser HD 201	92dB/mW	24	Full-size	30$[1]
Koss UR 29	103dB/mW	100	Full-size	30$[1]

[1] Best price from www.headphone.com in date March 1 2010

Sennheiser HD 205	112dB/mW	32	Full-size	60$[1]
Sennheiser HD 438	112dB/mW	32	Full-size	90$[1]
Sony MDR-V500DJ	102dB/mW	40	Full-size	100$[1]
iHarmonix iHX Platinum ev-Series	102dB/mW	16	In-ear	34$[1]
Sennheiser CX 300 II-Black	113dB/mW	16	In-ear	40$[1]
Maximo iMetal iM590 -Black	100dB/mW	16	In-ear	50$[1]
Sennheiser CX 380	113dB/mW	16	In-ear	60$[1]
Grado iGi Earphones	105dB/mW	24	In-ear	90$[1]
Etymotic ER-6i Black	107dB/mW	16	In-ear	87$[1]
Sennheiser HD 202	99dB/mW	32	Earpad	30$[1]
DENON AH-D301K	100dB/mW	28	Earpad	35$[1]
Sennheiser HD 218	92dB/mW	24	Earpad	45$[1]
DENON AH-D501K	103dB/mW	28	Earpad	45$[1]
Koss Porta Pro	101dB/mW	60	Earpad	50$[1]
Sony MDR-XB300	100dB/mW	24	Earpad	50$[1]
Koss UR 20	97dB/mW	32	Closed	20$[1]
Sennheiser HD 201	92dB/mW	24	Closed	30$[1]
DENON AH-D301K	100dB/mW	28	Closed	35$[1]
Sony MDR-G75LW	106dB/mW	24	Closed	40$[1]
Sennheiser HD 218	92dB/mW	24	Closed	45$[1]
Beyerdynamic DT 235 Black	95dB/mW	32	Closed	60$[1]
Lenntek Sonix 3 w/	105dB/mW	16	Headset	70$[1]

Microphone				
Sennheiser PXC 350	92dB/mW	150	Noise canc.	200$[1]
DENON AH-NC732	105dB/mW	40	Noise canc.	250$[1]
Sennheiser PMX 80	121dB/mW	64	Behind the Neck	50$[1]
Koss KSC 75	101dB/mW	60	Clip-on	20$[1]
Sony MDR-Q68LW	108dB/mW	24	Clip-on	30$[1]
Klipsch Image S2	106dB/mW	18	In-ear	49$[1]
Jivo Jellies In-Ear Headphones	108dB/mW	32	In-ear	8$[2]
Soyntec Netsound 250 Headphones	107dB/mW	16	In-ear	9$[2]
Soyntec Netsound 220 Headphones	108dB/mW	16	In-ear	4$[2]
Sony MDR-J10L	104dB/mW	16	Clip-on	11$[2]
JVC HA-KX100	103dB/mW	16	In-ear	51$[3]
Shure SE530 Sound Isolating	119dB/mW	36	In-ear	220$[3]
Koss KSC9 Sportclip	102dB/mW	32	Clip-on	10$[3]
Sennheiser HD 25-1	120dB/mW	70	Full-size	200$[4]
MGD01	80dB/mW	8	Clip-on	-

Table 2: comparison between some headphone products highlighting the sensitivity, the impedance and the price.

[2] Price from http://www.headphoneworld.com in date March 1 2010
[3] Price from http://www.headphoneworld.com in date March 1 2010
[4] Price from http://pro-audio.musiciansfriend.com in date March 1 2010

The most part of the commercial headphones have 32 Ohms of load impedance. However, in the table, we can found headphones with 8 Ohms and with 150 Ohms of load impedance. Moreover, we can see that the headphones in commerce show a very wide range of sensitivity: starting from 80dB/mW to 121dB/mW. Both the sensitivity and the load impedance vary significantly and this forces the amplifier to have a very high dynamic range (higher than 105dB). Let's make two examples: in the first one, we suppose to have headphones with 100dB of sensitivity. We have to evaluate the maximum output noise which allows no audible Sound Pressure Level at the output (SPL = 0dBs). The output noise power is defined as

$$\mathrm{Pout,noise} = \frac{\text{Noise Spectral Density}\left[\mathrm{V^2/Hz}\right]\mathrm{BW}\left[\mathrm{Hz}\right]}{\text{impedance}\left[\Omega\right]} \qquad \textbf{5}$$

and the Sound Pressure Level ca be written as

$$\mathrm{SPL} = \text{Sensitivity}\left[\mathrm{dBs@Pout} = 1\mathrm{mW}\right] + \mathrm{Pout}\left[\mathrm{dBm}\right] \qquad \textbf{6}$$

We want zero sound pressure level and this means

$$\mathrm{Pout}\left[\mathrm{dBm}\right] = -\text{Sensitivity}\left[\mathrm{dBs@Pout} = 1\mathrm{mW}\right] \qquad \textbf{7}$$

From previous expression we obtain the output power expressed in Watt:

$$\mathrm{Pout}\left[\mathrm{W}\right] = 10^{\text{sensitivity}/10} = 1\mathrm{pW} \qquad \textbf{8}$$

The noise spectral density results equal to $32pV^2$, and the output noise, expressed in Volt RMS, results $5{,}65uV_{RMS}$.

In the second example we suppose to have headphones with 92dB/mW of sensitivity and with a load impedance of 150Ohms (in the table we can find the Sennheiser PXC 350 which shows the same specifications). We want to listen a song with an average SPL of 85dB. It can be a valid assumption that the music has usually a Peak to Average Ratio of 15dB (see chapter 1). This means that the maximum Sound Pressure Level is 100dB. From equation [[2 we can derive the maximum output power which has to be delivered to the load in order to reproduce the music without distortion, which results:

$$\mathrm{Pout}\left[\mathrm{dBm}\right]_{MAX} = \mathrm{SPL}_{MAX} - \mathrm{Sensitivity}\left[\mathrm{dBs@Pout} = 1\mathrm{mW}\right] = 8\mathrm{dBm} \qquad \textbf{9}$$

The output power of 8dBm is delivered on 150Ohms of load and the resulting RMS output voltage is

$$V_{RMS} = \sqrt{10^{\mathrm{dBm}/10}\,\mathrm{mW}\cdot R_L} = 972\mathrm{mV} \qquad \textbf{10}$$

Summarizing we have a noise floor of $5{,}65uV_{RMS}$ and a maximum output voltage swing (without distortion) of almost $1V_{RMS}$. We can evaluate the Signal To Noise Ratio (SNR) of the amplifier, which results

$$\mathrm{SNR} = 20\mathrm{Log}\left(\frac{V_{n,out}}{V_{max,out}}\right) = 20\mathrm{Log}\left(\frac{5{,}65uV}{1V}\right) = 105\mathrm{dB} \qquad \textbf{11}$$

In this section we have seen that in the commerce we can find headphones showing a wide range of different performance (in term of load impedance and sensitivity). For this reason, a fundamental requirement for a headphones amplifier, used in hi-fi audio systems, is a very high dynamic range (higher than 105dB). In particular, the maximum delivered output power on 32 Ohms should be higher than 30mW and the output noise floor should be lower than $5uV_{RMS}$.

2.1.1 Load electrical model

As we have seen in the previous section, the most common driver type is an electromechanical transducer using a voice coil rigidly connected to a diaphragm. The moving system of the heapdhone (including the cone, cone suspension, spider and the voice coil) has a certain mass and compliance. This is most commonly likened to a simple mass suspended by a spring that has a certain resonant frequency at which the system will vibrate most freely.

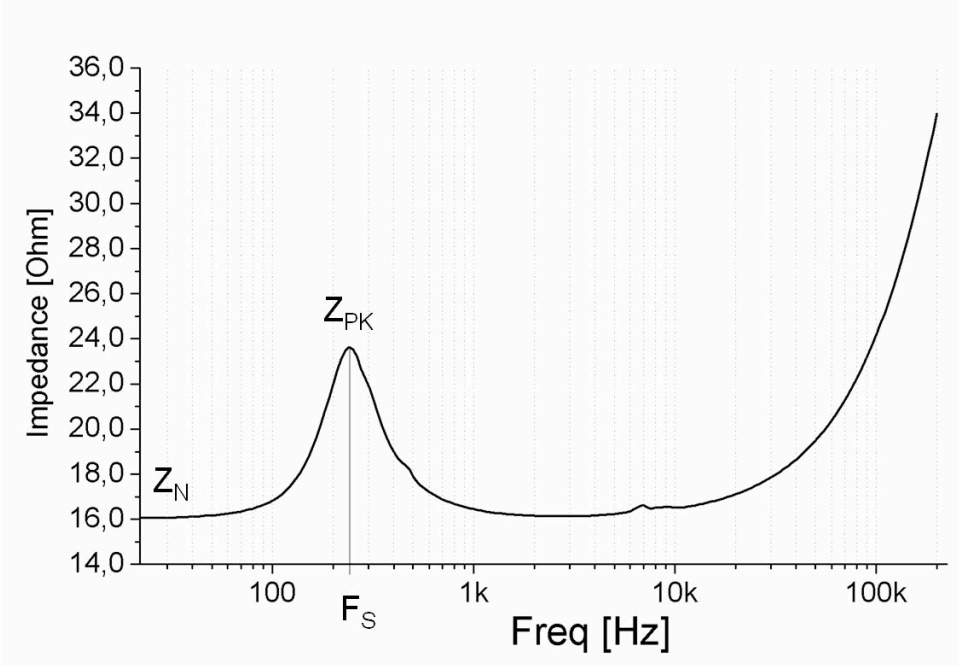


Figure 14: Module of the electrical impedance of headphone load versus frequency.

This frequency is known as the "free-space resonance" of the speaker and is designated by Fs. At this frequency, since the voice coil is vibrating with the maximum peak-to-peak amplitude and velocity, the back-emf generated by coil motion in a magnetic field is also at its maximum [4]. This causes the effective electrical impedance of the speaker to reach a local maximum, shown as Z_{PK} in the Figure 14. For frequencies just below resonance, the impedance rises rapidly as the frequency approaches Fs and is inductive in nature.

Amplifiers present relatively low impedance to the loudspeaker which acts as a generator when a coil is moving in a magnetic field. This is the so-called back EMF. The ratio between the loudspeaker impedance and the amplifier's impedance at a particular frequency provides damping (ie, energy absorption) for the back EMF generated by a driver. In practice, this is important to

prevent ringing or overhang which is, essentially, a free vibration of the moving structures in a driver when it is excited (ie, driven with a signal) at that frequency. Many people claim to be able to hear the effects in music and speech reproduction, however, this phenomenon is not usually taken into account when designing headphone amplifier because the free-space resonance varies significantly from a product to another.

We have seen the electrical behavior of the headphone load into the audio bandwidth (20Hz-20kHz), however, during the design of a headphones driver, it is much important to study it at higher frequency. In fact, if this is not taken into account during the design, the amplifier may oscillate when connected to the load. Figure 14 shows that when the signal frequency is higher than 20kHz the impedance increases significantly. This is due to the fact that the moving coil transducer has an inductive nature. For this reason, outside the audio band (f>20kHz), the moving coil transducer can be simply modeled as an inductance in series with a resistance. The headphone speaker is always connected to the amplifier through a cable and this cable introduces a distributed capacitance connected to the output node of the amplifier and to ground. Figure 15 the electrical model of the headphone load at frequencies higher than 20kHz and Figure 16 shows the module of the impedance of Figure 15 versus frequency.

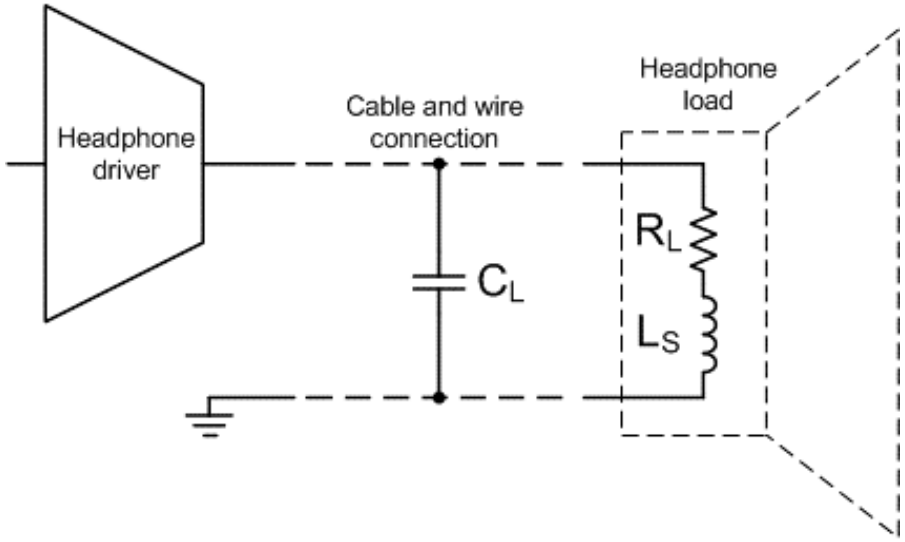


Figure 15: model of the headphones load when the signal frequency is higher than 20kHz.

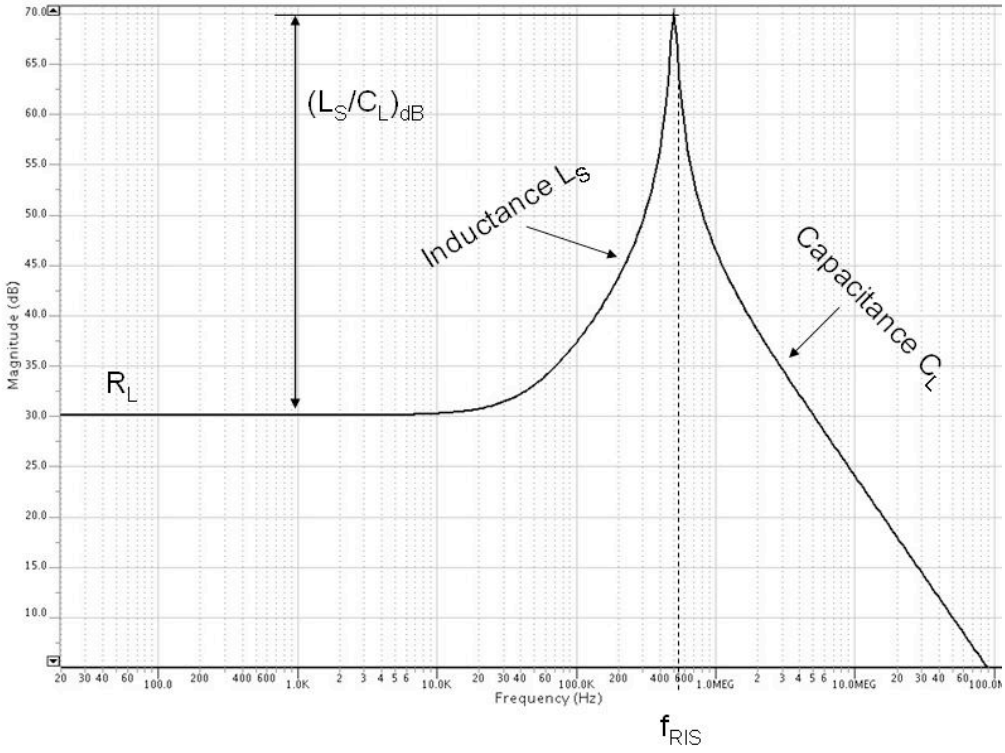


Figure 16: Bode diagram of the module of the load impedance of Figure 15

The graph shows a resonant frequency F_{RIS} which values $1/\sqrt{L_S \cdot C_L}$.

At $f=F_{RIS}$ the load impedance, Z, becomes only resistive and it values

$$Z(f_{RIS}) = \frac{L_S/C_L}{R_L} \qquad \textbf{12}$$

When the signal frequency is higher than F_{RIS} the load impedance becomes capacitive and it values $1/sC_L$.

When studying the amplifier stability we can distinguish three different cases:

1. The amplifier unity gain frequency, f_T, is below F_{IRS} and the load impedance, Z, can be approximated with the resistance R_L.
2. f_T is about F_{RIS} and the load impedance can be approximated with the inductance L_S in series to the resistance R_L.
3. The worst case for the amplifier stability happens when the amplifier unity gain frequency, f_T, is higher than F_{RIS}. In this condition, when the amplifier gain is close to 0dB, the load impedance is only capacitive.

Different products have different load impedance and resonant frequency F_{RIS}. The best choice, when designing the amplifier, is always to consider the worst case condition. This means that, from the stability point of view, we can simplify the load impedance only to the capacitance C_L. In fact, an important parameter for a headphones driver is the maximum allowed load capacitance which must be high enough to guarantee the amplifier stability in all the possible cases. Usually a class AB amplifier used in headphone application can drive a load capacitance up to 200pF.

2.2 Class AB amplifier

Most headphones amplifiers nowadays are in class AB, which is not really a separate class, but a combination of A and B. If an amplifier is biased into Class-B, and then the bias further increased, it will enter AB [4]. For outputs below a certain level both output devices conduct, and operation is Class-A. At higher levels, one device will be turned completely off as the other provides more current, and the distortion jumps upward at this point as AB action begins. Each device will conduct between 50% and 100% of the time, depending on the degree of excess bias and the output level. The class AB maximum theoretical efficiency is limited by the efficiency of the class B. Assume a class B amplifier with power supplies $+V_S$ and $-V_S$. Load resistance is R_L. Figure 17 shows how the amplifier dissipates.

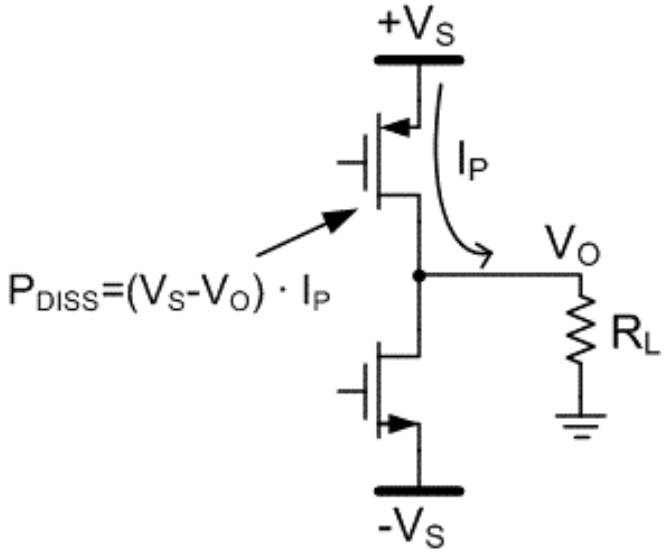


Figure 17: Dissipation in a Class AB amplifier when $V_O>0$

The classic way of calculating the efficiency of a class B amplifier assumes a rail to- rail sine wave at the output V_O. See Figure 18.

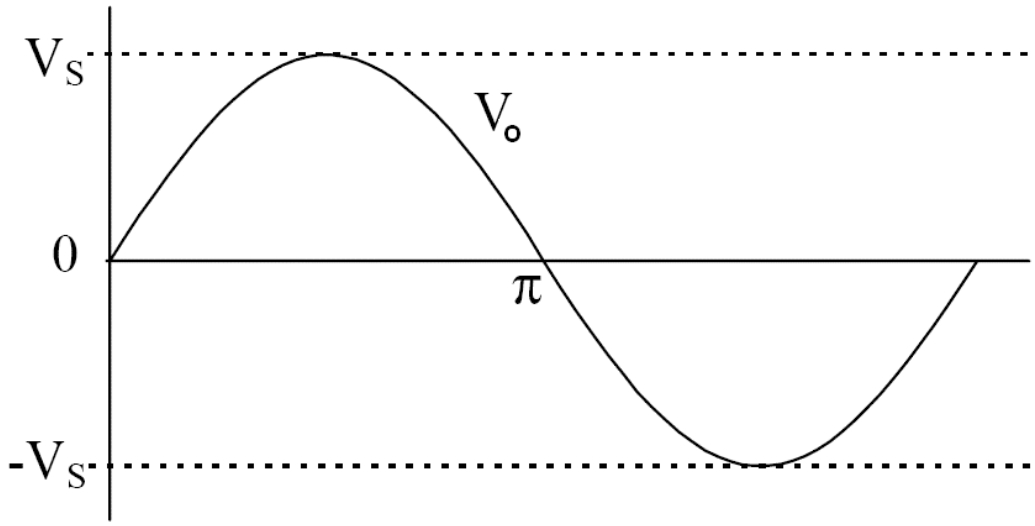


Figure 18: Rail-to-rail sine wave

The efficiency over any number of periods is equal to the efficiency over a quarter of a period. The efficiency is defined as η = Po/Pi. Over a quarter of a period:

$$P_O = \int_0^{\pi/2} V_O^2/R_L \cdot \sin^2(\vartheta)\, d\vartheta \qquad \textbf{13}$$

and

$$P_i = \int_0^{\pi/2} V_S \cdot V_O/R_L \cdot \sin(\vartheta)\, d\vartheta \qquad \textbf{14}$$

the expression of the theoretical efficiency results

$$\eta = \frac{P_O}{P_i} = \frac{V_O}{V_S} \cdot \frac{\pi}{4} \qquad 15$$

which yields a maximum efficiency of ¼ π = 78.5%.

A low distortion class AB amplifier has a relatively low complexity and requires almost no external components if integrated. The efficiency for audio signals however, is quite low, in particular when the output voltage is low (see equation 11). In the first chapter we have seen that the most part of the time a headphone amplifier delivers to the load a small fraction of the maximum output power. For example, let's assume that the voltage supply V_S is equal to 1.8V. The output power of a headphone amplifier should be less than 0.1mW on 32Ohms. This corresponds to a sinusoid with an amplitude V_O equal to

$$V_O = \sqrt{P_O \cdot 2 \cdot R_L} = 80\text{mV} \qquad 16$$

The resulting theoretical efficiency is 3.5% which is a very poor value. Other classes of amplifier shows better efficiency, however the class AB is the most linear and it is also the most simple to realize. Moreover it requires almost no external components. Those are the reasons because, up today, it results as the most used in headphone applications. The problem is that modern cellular phones incorporate hands-free operation, MP3 music playback and DMB reception and the users may wish to activate these features for many hours. A low efficiency amplifier obviously depletes the battery in a short time and, for this reason, class AB architecture starts to be not so suitable in modern systems and actually the designers are looking for more efficiently architectures.

2.3 Class D amplifier

In the Class D amplifier the input signal is converted to a sequence of higher voltage output pulses [6]. The averaged-over-time power values of these pulses are directly proportional to the instantaneous amplitude of the input signal. The frequency of the output pulses is typically ten or more times the highest frequency in the input signal to be amplified. The output pulses contain inaccurate spectral components (that is, the pulse frequency and its harmonics) which must be removed by a lowpass passive filter. The resulting filtered signal is then an amplified replica of the input.

These amplifiers use pulse width modulation, pulse density modulation (sometimes referred to as pulse frequency modulation) or more advanced form of modulation such as Delta-sigma modulation (for example, in the Analog Devices AD1990 Class-D audio power amplifier). Output stages such as those used in pulse generators are examples of class D amplifiers. The term Class D is usually applied to devices intended to reproduce signals with a bandwidth well below the switching frequency.

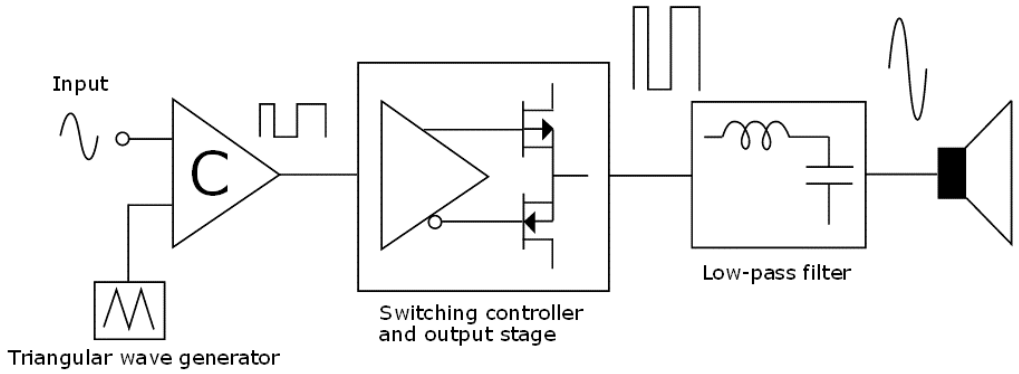


Figure 19: Block diagram of a basic switching or PWM (Class-D) amplifier.

Class D amplifiers can be controlled by either analog or digital circuits. The digital control introduces additional distortion called quantization error caused by its conversion of the input signal to a digital value.

The main advantage of a class D amplifier is power efficiency. Because the output pulses have fixed amplitude, the switching elements (usually MOSFETs) are switched either completely on or completely off, rather than operated in linear mode. A MOSFET operates with the lowest resistance when fully-on and thus has the lowest power dissipation when in that condition, except when fully off. In fact, when operated in a linear mode the MOSFET has variable amounts of resistance that vary linearly with the input voltage and the resistance is something other than the minimum possible, therefore more electrical energy is dissipated as heat.

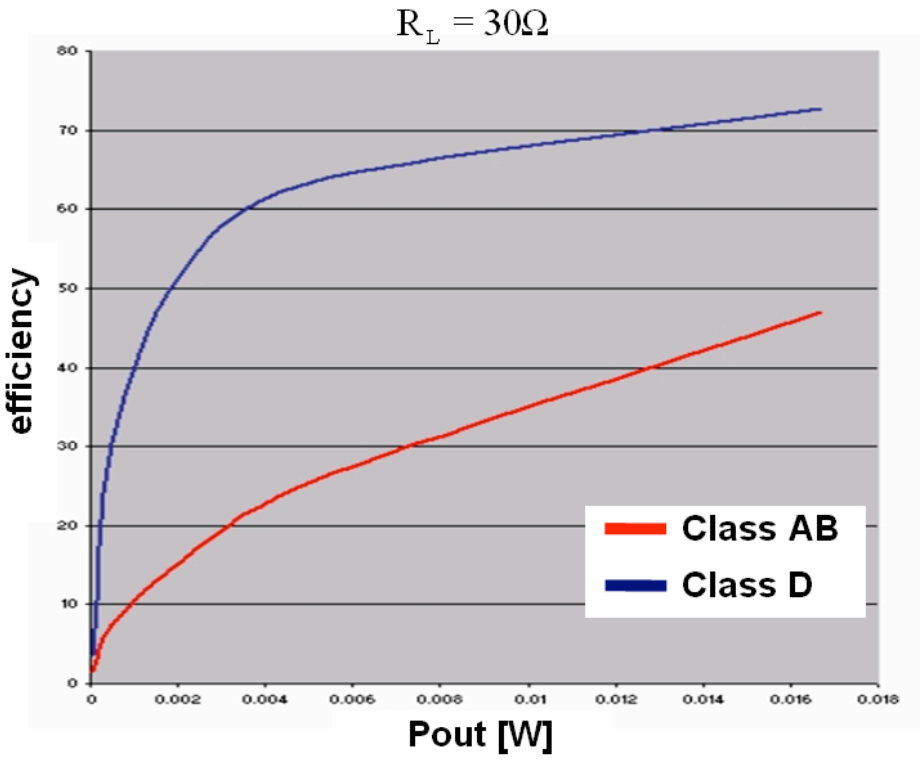


Figure 20: Class D and Class AB headphone amplifiers efficiency versus the output power.

Figure 20 shows the efficiency of a class D amplifier used in headphone application [7]. We can see that the efficiency of the class D is much higher than the class AB one, in particular when the output power stays between 1mW and 14mW. It's worth noticing that the RMS output power in headphone application stays usually below 0.5mW per channel (1mW stereo). In this condition we can see that also the class D efficiency is below 30%.

The drawbacks of a class D amplifier are:

- **Noise**: in a class D amplifier the high frequency power switching introduces high frequency noise which must be taken into account during the design. In fact this noise can be coupled with the substrate affecting the overall chip performance. Moreover, the switching noise must also be prevented from corrupting the audio signal path in the analogue portion of the chip. Power supply decoupling, which is important in any class D design if switching noise is not to degrade circuit operation, takes on a greater importance in the context of mixed signal chip design.
- **Linearity**: Class D amplifier is usually less linear than class AB one. However, the latest class D designs are now able to achieve very low total harmonic distortion comparable to high quality class AB performance [wolfson reference].
- **Electromagnetic Interferences (EMI)**: is a disturbance that affects an electrical circuit due to either electromagnetic conduction or electromagnetic radiation emitted from an external source. The high frequency switching output signal needs to be attenuated using a low pass filter. If the speaker is close to the amplifier it is possible to implement such filter using the coil inductor (speaker amplifier). However, in headphone application, the transducer is far away from the amplifier

output node and this means that it is necessary the usage of an external low pass filter (see Figure 21).

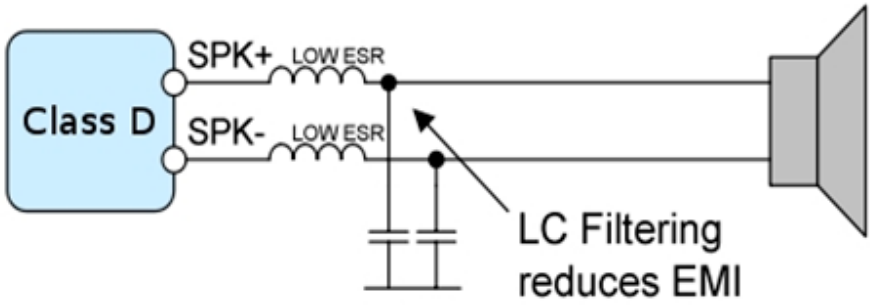


Figure 21: Class D external low pass filter to reduce the EMI.

The EMI drawback is the most important: usually into the datasheet of integrated class D amplifier, the producer gives a large number of board layout suggestions in order to limit the EMI emissions. However, those suggestions seem to be not enough. In fact, some system designers have seen problems with switching noise during certain modes of operation; for example when the handset is being operated as an FM radio receiver [wolfson classD]. An accumulation of system design parameters may conspire to promote excessive interference in audio output. In a Wolfson product, WM9001-WM9081, there is a dynamically selectable class D or class AB amplification modes in order to prevent this noise coupling [8].

2.4 Class G amplifier

A class G is a high-efficiency analog amplifier (without EMI problems) that tries to bring together the best of class AB and class D. Figure 22 shows a schematic block representation of a class G amplifier.

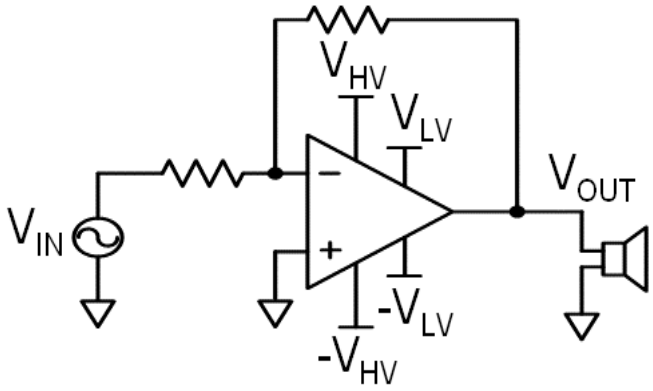


Figure 22: Simplified block diagram of a class G amplifier

It uses two voltage supply rails (VHV is the high voltage rail and VLV is the low voltage rail) and switches to the appropriate supply as required by the instantaneous output voltage level. The class G topology is a modification of another Class of amplifier (normally Class B or Class AB) to increase efficiency and reduce power dissipation. Class G takes advantage of the fact that musical and voice signals have a high crest factor with most of the signal content at lower amplitudes. A class G device uses a minimum of two different supply rails. The device operates from the lower supply until output headroom becomes an issue [4]. At this point the device switches the output stage to the higher supply rail. Once the output signal drops below a predetermined level, the device switches back to the lower rail. Power dissipation is greatly reduced for typical musical or voice sources. The core of a class G amplifier is the switching circuitry, which should enable a smooth handover of the load driving between the lower voltage supply and the higher one. The switching point is defined as the output voltage level where the amplifier switches from one supply to the other. We define two switching point levels called VLV-VTH and –VLV+VTH, where VLV is the

value of the amplifier low voltage supplies and VTH is a threshold voltage (see Figure 23).

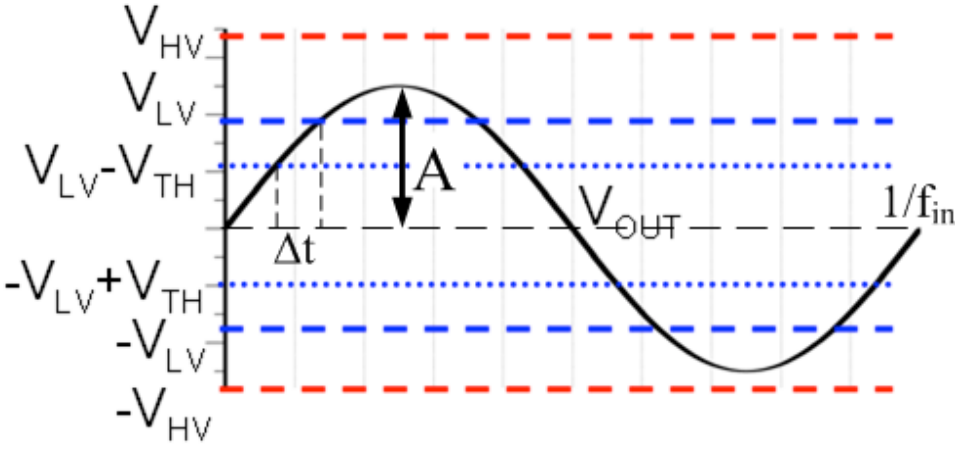


Figure 23: Class G output voltage versus time. The y-axis shows the supply voltage rails and the switching point levels

There are two basic ways in which class G amplifiers are realized [2]: the "series" implementation (shown in Figure 24(a)) and the "parallel" implementation (shown in Figure 24(b)). Thanks to its simplest switching circuit, the most common implementation is the series one. It uses a single output stage connected to both the low voltage and to the high voltage supply rails respectively through diodes and switches. When the output voltage is lower than the switching point (VLV-VTH) the switches are open and the amplifier is connected only to the low voltage supply through the diodes.

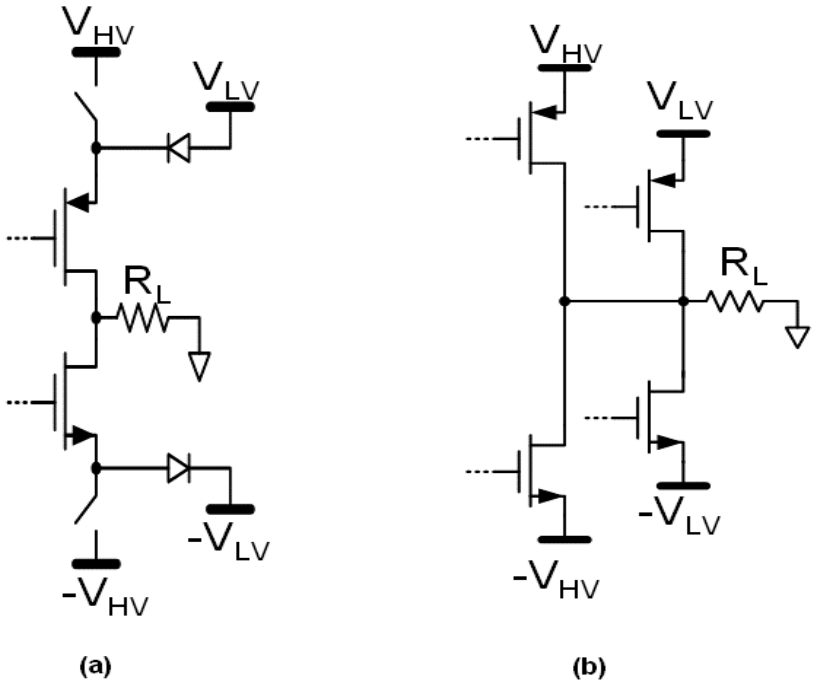


Figure 24: Class G topologies, (a) serial and (b) parallel.

When the output voltage exceeds the switching point the switches are closed and the push pull output stage is supplied by the high voltage rails. The diodes are used to prevent any current flowing between the high and low voltage supply rails. The main limitation of the series topology is due the diodes in series with the low voltage supply that make this implementation unsuitable for low voltage application. In fact, the turn-on voltage of the diode puts a strong limitation both on the minimum value of VLV and on the switching point distance from VLV.

On the other hand, in the parallel topology, shown in Figure 24(b), there are two output stages working in parallel and there is nothing between the power transistors and the supplies. This fact poses no constraints on the minimum value

of VLV and allows placing the switching point very close to VLV, giving higher efficiency.

2.4.1 Efficiency and distortion

Class G efficiency depends largely on the source material (music or voice) and the characteristics of the signal. If the output amplitude remains at a level where the class G device operates from its lower supply rail, then power dissipation does decrease compared to the other architectures (class B and class AB) that can only operated from a fixed, higher voltage supply. For real-world Class G amplifiers, the maximum efficiency occurs when operating under the lowest supply rails as opposed to operating at peak output power on the higher rails due to biasing conditions, current x voltage loss, and IR losses in FETs. Figure 25 shows how a musical output may look [9]. In this example we can intuitively see the efficiency improvements in a class G implementation in respect with a class AB one.

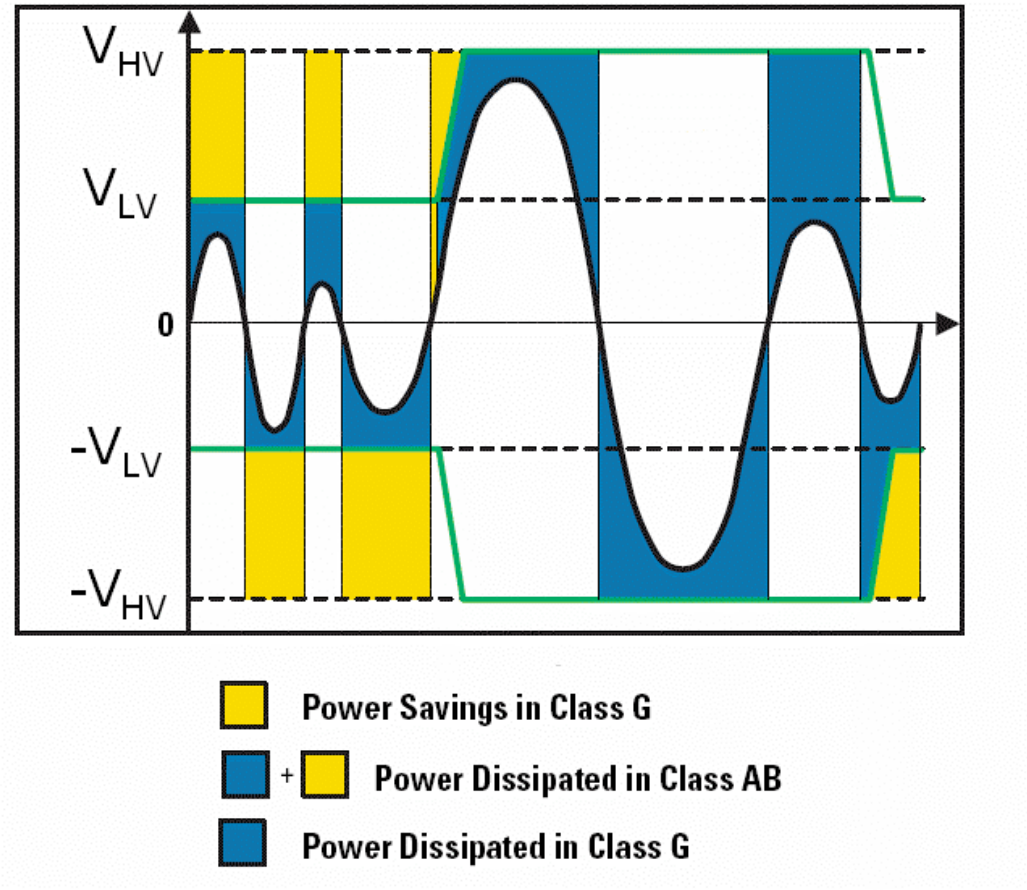


Figure 25: Music output example. The figure shows intuitively the Class G efficiency improvement.

The classic way of calculating the efficiency of a class G amplifier assumes a rail to- rail sine wave at the output V_O. See Figure 26. Let's assume to have a sinusoidal output signal with peak amplitude equal to A. We define α as

$$\alpha = \sin^{-1}\left(\frac{V_{LV} - V_{TH}}{A}\right) \qquad 17$$

where V_{LV}-V_{TH} is the switching point level. The low voltage stage drives the load when $\theta<\alpha$, $\pi-\alpha<\theta<\pi+\alpha$ and $2\pi-\alpha<\theta<2\pi$. In the remaining cases the high voltage stage drives the load.

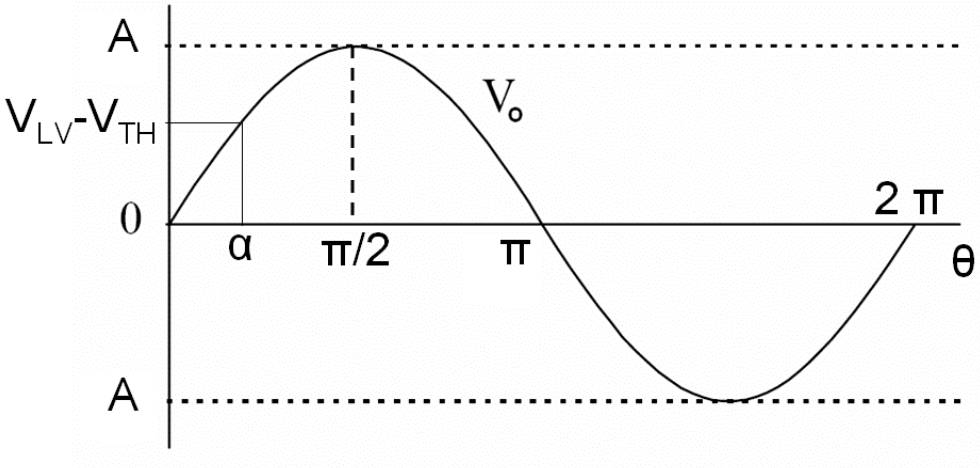


Figure 26: Class G output voltage with amplitude equal to A versus θ = ωt.

The efficiency over any number of periods is equal to the efficiency over a quarter of a period. The efficiency is defined as η = Po/Pi. Over a quarter of a period:

$$\eta = \frac{P_O}{Pi} = \frac{A^2/(2R_L)}{\frac{1}{2\pi}\left[4 \cdot V_{LV} \cdot \int_0^{\alpha} A\sin(\theta)d\theta + 4 \cdot V_{HV} \cdot \int_{\alpha}^{\pi/2} A\sin(\theta)d\theta\right]} \qquad 18$$

which results

$$\eta = \frac{A \cdot \pi}{4V_{LV}(1-\cos(\alpha)) + 4V_{HV}\cos(\alpha)} \quad \textbf{19}$$

Using eq.(13-15) we obtain the efficiency of a class G amplifier in relation of the output signal amplitude. From eq. 15, assuming a load resistance equal to 32 Ohms, it is possible to derive the expression of the class G efficiency as a function of the output power, which results:

$$\eta = \frac{\pi \cdot \sqrt{2P_O R_L}}{4V_{LV}(1-\cos(\alpha)) + 4V_{HV}\cos(\alpha)} \quad \textbf{20}$$

where P_O is the output power.

Figure 27 shows the efficiency of a class G amplifier versus the output power for different switching point levels and compared with that of a class AB able to handle the same maximum power. The core of the parallel topology is the switching circuit which must simultaneously enable low distortion and high efficiency. If the output voltage of the class G amplifier is below the switching point level, the ratio between the efficiency of the class G and the class AB is approx equal to the ratio between the value of the high and low supplies:

$$\frac{EFF(ClassG)}{EFF(ClassAB)} \approx \frac{V_{HV}}{V_{LV}} \quad \textbf{21}$$

Music has a large peak-to-average power ratio which can vary from 10 dB for compressed rock to 30 dB for classical music[3]. This implies that the output power is below the peak level for most of the time. As a consequence power efficiency can improve significantly taking advantage of class-G operation.

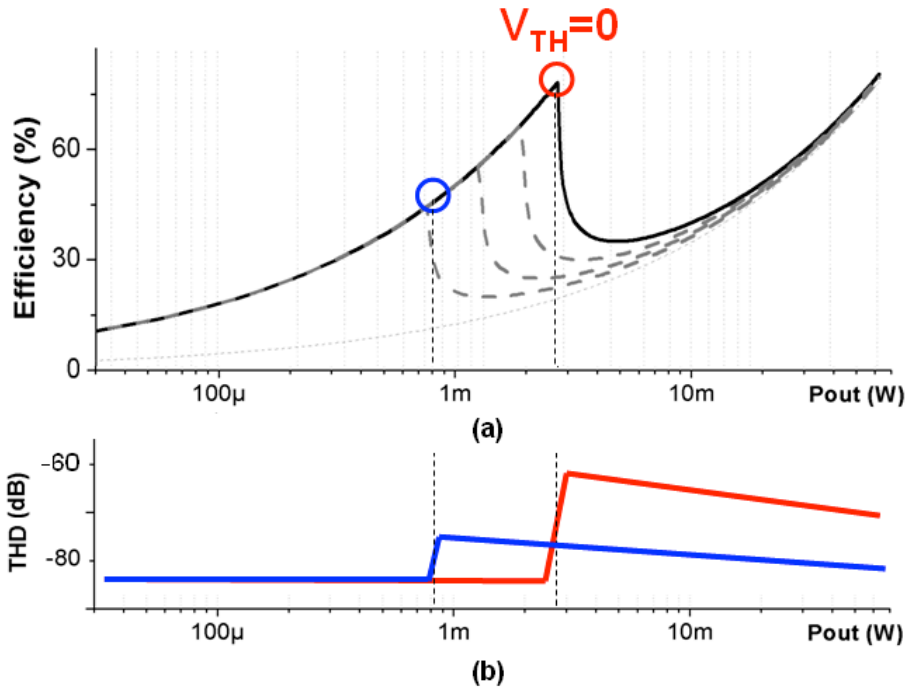


Figure 27: Class G efficiency (a) and the Total Harmonic Distortion versus the output power at different switching point level (b)

Figure 27 shows the strong impact of the switching point level on the efficiency [10]. The switching point should be set as close as possible to the low voltage supply VLV (ideally VTH should be equal to zero). However, if VTH is too small, the power transistors supplied by the low voltage rail have to work in deep thriode region for a significant portion of time affecting the Total Harmonic Distortion (THD) of the amplifier as shown in Figure 24(b). The value of VTH needs to be chosen as a trade off between efficiency and linearity.

The drawbacks of using a class G amplifier in respect with a class AB one are mainly three:

- It shows less linearity than a class AB. A part of distortion is related to the switching between the supply rails which is not present in a class AB amplifier.
- It needs more area than a class AB: class G uses two output stages and in this implementation the power transistors require double area than that of class AB.
- The main drawback of the class G is that it needs two voltage supply rails while class AB requires only a voltage supply rail. In most cases it is necessary to generate those additional supply voltages and, for this purpose, it can be used a buck converter to generate the positive low voltage supply starting from the high voltage one and it can be used an inverting charge pump to generate the negative low voltage supply starting from the positive one (see Figure 28). Those additional blocks are necessary to implement the low voltage supply rails and they require additional pins on the package and a certain number of external components.

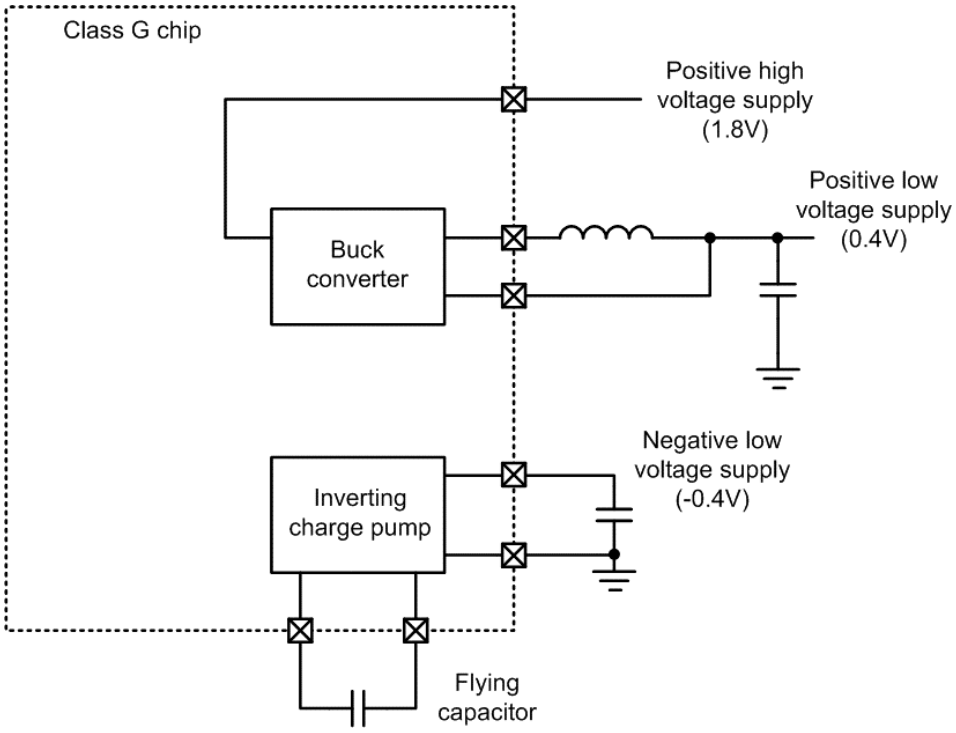


Figure 28: Class G amplifier needs an additional supply voltage rails in respect with a class AB. The figure shows an example of implementation where the low voltage supply rail is internally generated using a buck converter and a charge pump.

Despite the high number of the pin count necessary in this implementation (6 additional pins and 4 additional external components compared to a class AB implementation), the class G solution seems to be the most suitable for headphone applications. In fact, the last released products of the most important producers of headphone drivers include class G amplifiers.

2.5 Conclusions

The first section of this chapter has shown the procedure to derive the requirement of noise floor (<6uV_{RMS}) for headphone driver starting from the headphones

sensitivity (100dB/mW). We have seen that the SNR, referred on the maximum output power of 30mW, has to be higher than 105dB in order to fulfill the market requirements. In the last section of this chapter we have compared different architectures of headphones driver and we have seen that the class G approach is able to bring together the efficiency of the class D and the linearity of the class AB. We have chosen the parallel topology of class G because it obtains better performance in terms of efficiency than the serial one.

2.6 References

[1] D. Self, Audio Power Amplifier Design Handbook, 4th ed. Oxford, U.K.: Newnes-Elsevier, 2006

[2] Wikipedia contributors, "Headphones," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Headphones&oldid=380553860 (accessed August 27, 2010).

[3] Wikipedia contributors, "Loudspeaker," Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Loudspeaker&oldid=379711622 (accessed August 27, 2010).

[4] Elliot sound products website, Effects of source impedance on loudspeakers, http://sound.westhost.com/index2.html (accessed August 25, 2010)

[5] Ronan van der Zee, High Efficiency Audio Power Amplifiers; design and practical use, 1999.

[6] Sánchez Moreno, Sergio Class D Audio Amplifiers - Theory and Design [2] - Contains material on the theory and design of Class-D amplifiers.

[7] White paper, Class D amplifiers empower mobile multimedia. http://www.wolfsonmicro.com/documents/uploads/misc/en/Class_D_Amplifiers_Empower_Mobile_Multimedia.pdf, (accessed August 27, 2010)

[8] Wolfson, "1-W Mono Class D/AB speaker driver", http://www.wolfsonmicro.com/products/audio_amplifiers/WM9001/, (accessed August 26, 2010)

[9] National Semiconductor, "PowerWise Class G versus Class AB Headphone Amplifiers", http://www.national.com/vcm/national3/en_US/resources/signal_path_designer/national_sp_designer118.pdf, (accessed September 1, 2010)

[10] Alex Lollio, Giacomino Bollati, Rinaldo Castello, “A Class-G Headphone Amplifier in 65nm CMOS Technology”, Solid-State Circuits, IEEE Journal of, vol. 45, no.12, Dec 2010

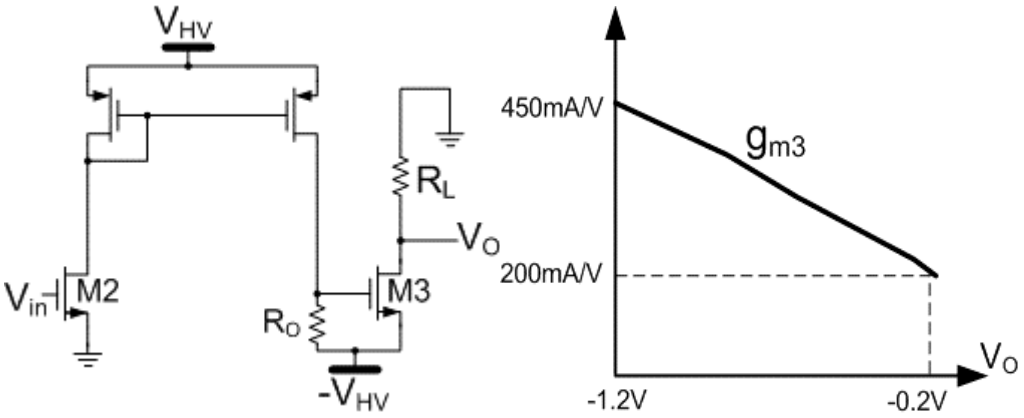
3. Class G amplifier

This chapter shows the implementation details of the "parallel" class G headphone driver. The first section shows the class G switching principle which represents the real innovation of this work. The very smooth handover between the stages, obtained with this implementation, enables both high linearity and high efficiency. Moreover, the first section shows also a mathematical modeling of class G amplifier useful for the evaluating of the switching distortion and the ability of the loop to reject it. The results obtained in this section are used to optimize the overall distortion after compression by the feedback loop. The second section shows the switching speed limitation at high frequency and a practical solution to overcome the problem. Section three shows the amplifier architecture and a detail on the compensation technique. Section four shows the experimental results including a comparison with the state of the art. In the last section there is an improved version of the class G amplifier (which uses one more stage in respect to that reported in section 3) showing better THD and SNR performance. The implementation reported in the last section has been integrated into a novel Marvell audio codec where the class G amplifier plays an important role on the overall performance.

3.1 Class G switching principle

At the end of previous chapter we have compared the class G serial topology, of Figure 24(a), with the parallel one, of Figure 24(b), and we have chosen the parallel one because it is more suitable for the implementation in low voltage

systems. In this class G topology, the switching approach is the most important point and it represents the real innovation of this work [12]. The goal in a class G amplifier design is to find a switching principle which maximizes the efficiency and limits the distortion introduced by the switching operation. In this Section the switching circuit proposed and its effects on the amplifier linearity are presented. Using a simple 2 stages model we will study the distortion introduced by the switching circuit and the capability of the loop to reject it. In order to simplify the discussion, the class G amplifier (based on a parallel topology) is derived from a class A amplifier. The obtained conclusions can, however, be easily extended to a push-pull class AB output stage.



(a) (b)

Figure 29: Simple class A amplifier (a) and the graph of the transconductance of the output stage gm3 in relation with the output voltage using R_L=32Ω (b).

The following equation describes the low frequency input-output transfer function of the two stages class A amplifier shown in Figure 29(a):

$$V_O = -g_{m2} R_O g_{m3} R_L \cdot V_{in} \qquad 22$$

Starting from this class A amplifier we want to build a simple class G one which switches from one supply rail to the other one. To do this, let's start to split transistor M3 into two transistors M3L and M3H as shown in the top of Figure 30 (at this point we don't consider the differential pair shown in the figure).

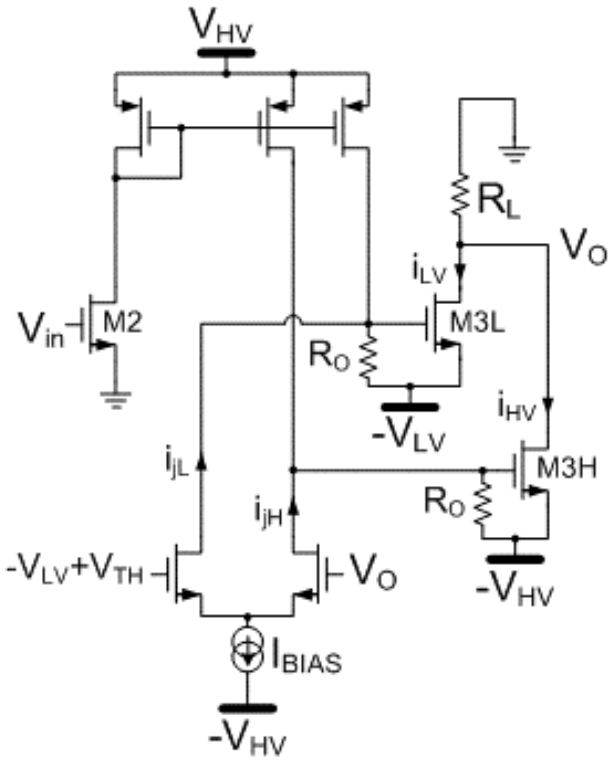


Figure 30: Class G output stage including the switching circuit

The input-output transfer function is given by the following equation:

$$V_O = -g_{m2} R_O \left[g_{m3L}(V_O) + g_{m3H}(V_O) \right] R_L \cdot V_{in} \qquad 23$$

where gm3L(Vo) and gm3H(Vo) represent the transconductance of the low and high voltage output stages as a function of the output voltage (as long as the load

in the band of interest is purely resistive the dependence of gm3 on output current translate in a dependence on the output voltage).

Let's first suppose to implement an ideal switching between the output stages in such a way that the switching operation does not introduce any extra distortion. This is obtained if the total equivalent transconductance of the output stages of the circuit shown at the top of Figure 30 given by gm3L(Vo)+gm3H(Vo) is equal to the tranconductance gm3(Vo) shown in Figure 29(b). In a real implementation this condition is difficult to satisfy, however, it will be used only to derive the linear model shown at the end of this section and the analysis carried out on this simplified model, will shows that the sensitivity of the switching distortion to variations of the gm3 value is very low. Furthermore, the two output stages have the same topology making not difficult to keep the variation of gm3 during switching smaller than that due to the output current variation (typically the main source of non linearity in a class AB amplifier). It will be shown that the difference between the simulated behavior of the real amplifier connected in unity gain configuration as shown in Figure 34(a) and the model is acceptably small.

The following equations describe a possible behavior of gm3L and gm3H as a function of the output voltage that satisfy the ideal condition gm3L(Vo)+gm3H(Vo)=gm3(Vo):

$$g_{m3L}(V_O) = \begin{cases} g_{m3}(V_O) & V_O < -V_{LV} + V_{TH} - \delta \\ g_{m3}(V_O) \cdot f_1(V_O) & (-V_{LV} + V_{TH}) - \delta < V_O < (-V_{LV} + V_{TH}) + \delta \\ 0 & V_O > -V_{LV} + V_{TH} + \delta \end{cases} \quad (24)$$

$$g_{m3H}(V_O) = \begin{cases} 0 & V_O < -V_{LV} + V_{TH} \\ g_{m3}(V_O) \cdot (1 - f_1(Vo)) & (-V_{LV} + V_{TH}) - \delta < V_O < (-V_{LV} + V_{TH}) + \delta \\ g_{m3}(V_O) & V_O > -V_{LV} + V_{TH} \end{cases} \quad (25)$$

Where (-VLV+VTH)-δ and (-VLV+VTH)+δ represent two voltage levels such that the switching point (-VLV+VTH) is in the middle between them. The output voltage range 2δ is where the switching operation takes place. The expression f1(VO) defines how the value of gm3L evolves during switching and, for our purposes, doesn't need to be explicitly known. Figure 31 shows a graphical representation of these equations where it can be immediately seen that gm3L(Vo)+gm3H(Vo)=gm3(Vo). Unfortunately this ideal situation cannot be realized in practice because the handover between one stage and the other requires an additional circuit able to force such a transition.

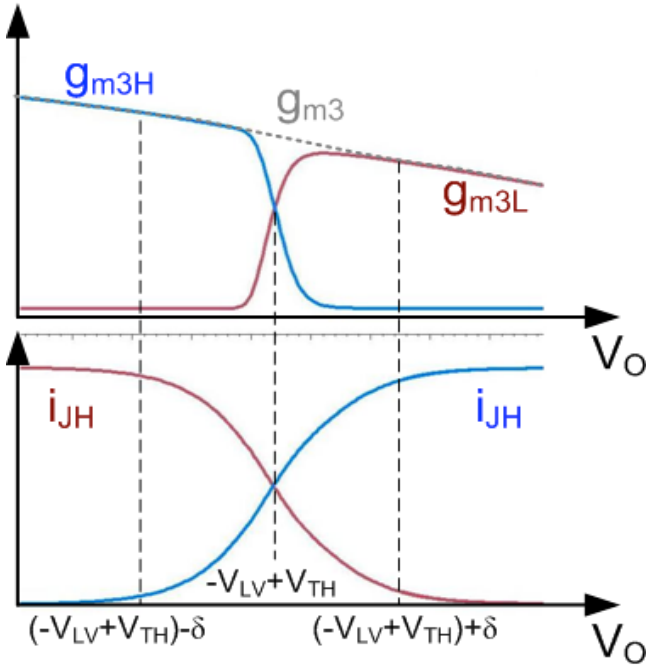


Figure 31: graphical relation between g_{m3H}, g_{m3L}, i_{JH} and i_{JL} of Figure 30 versus the output voltage.

Figure 30 shows this circuit which is implemented as a differential pair that compares the output voltage to the switching point voltage (–VLV+VTH). The differential pair sends its bias current, IBIAS, either to the gate of M3L or to the gate of M3H according to the result of the comparison. The value of IBIAS and the transconductance of this differential pair determine the speed of the transition. Let's assume that gm3L and gm3H still have the expression written in equation (24) and (25). In similar way we can write the equations representing the current IJL injected into the gate of M3L and the current IJH injected into the gate of M3H:

$$i_{jL}(V_O) = \begin{cases} 0 & V_O < -V_{LV} + V_{TH} \\ I_{BIAS} \cdot f_2(Vo) & (-V_{LV} + V_{TH}) - \delta < V_O < (-V_{LV} + V_{TH}) + \delta \\ I_{BIAS} & V_O > -V_{LV} + V_{TH} \end{cases} \quad \textbf{26}$$

$$i_{jH}(V_O) = \begin{cases} I_{BIAS} & V_O < -V_{LV} + V_{TH} \\ I_{BIAS} \cdot (1 - f_2(Vo)) & (-V_{LV} + V_{TH}) - \delta < V_O < (-V_{LV} + V_{TH}) + \delta \\ 0 & V_O > -V_{LV} + V_{TH} \end{cases} \quad \textbf{27}$$

The expression f2(VO) defines how the value of gm3H evolves during switching and, as f1(V0), doesn't need to be explicitly known.

Figure 31 shows a graphical representation of the switching parameters (transconductances and currents) versus the output voltage. From the simple class G implementation of Figure 30 we obtain the following input-output characteristic:

$$V_O = -\left[\left(V_{in}\, g_{m2} + i_{jL}\right) R_O\, g_{m3L} + \left(V_{in}\, g_{m2} + i_{jH}\right) R_O\, g_{m3H}\right] R_L$$

$$V_O = -V_{in} \cdot g_{m2} R_O R_L (g_{m3L} + g_{m3H}) - R_O R_L \left[g_{m3L} \cdot i_{jL} + g_{m3H} \cdot i_{jH}\right]$$

$$V_O = -V_{in} \cdot g_{m2} R_O g_{m3} R_L - R_O R_L \left[g_{m3L} \cdot i_{jL} + g_{m3H} \cdot i_{jH}\right] \quad \textbf{28}$$

If we combine the equation (24-28) we obtain the following relation

$$V_O = -V_{in} \cdot g_{m2} R_O g_{m3} R_L - R_O R_L g_{m3}(Vo) \cdot i_j(V_O) \quad \textbf{29}$$

where

$$i_j(V_O) = \begin{cases} I_{BIAS}\left(2f_1f_2 + 1 - f_2 - f_1\right) & (-V_{LV} + V_{TH}) - \delta < V_O < (-V_{LV} + V_{TH}) + \delta \\ 0 & \text{elsewhere} \end{cases} \quad \textbf{30}$$

It is interesting to note that the input-output transfer function for the model shown in Figure 32 is the same as that expressed by equation (29).

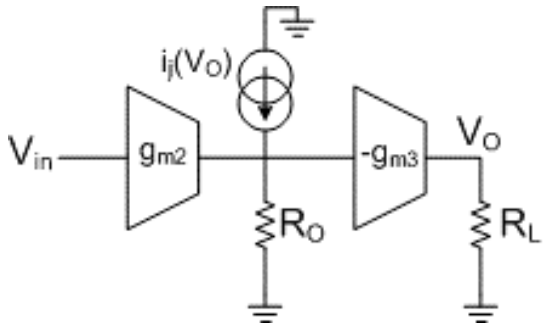


Figure 32: Amplifier model of Figure 30 used to analyze the switching distortion

This means that it is possible to model the distortion introduced by the switching circuit with a current source (connected to the input of the last stage of the linear circuit of Figure 32) that injects an unwanted signal at the harmonics of the input. This is consistent with what is done in literature [1] to model the non linearity due to the last stage of a class AB amplifier. Such a model will be used to understand the dependence of the switching distortion on the circuit parameters in order to optimize the overall class G amplifier performance. Figure 33(a) shows the real class G amplifier connected in a unity gain feedback configuration while Figure 33(b) shows the time domain representation of the output voltage and some key currents all obtained through simulation.

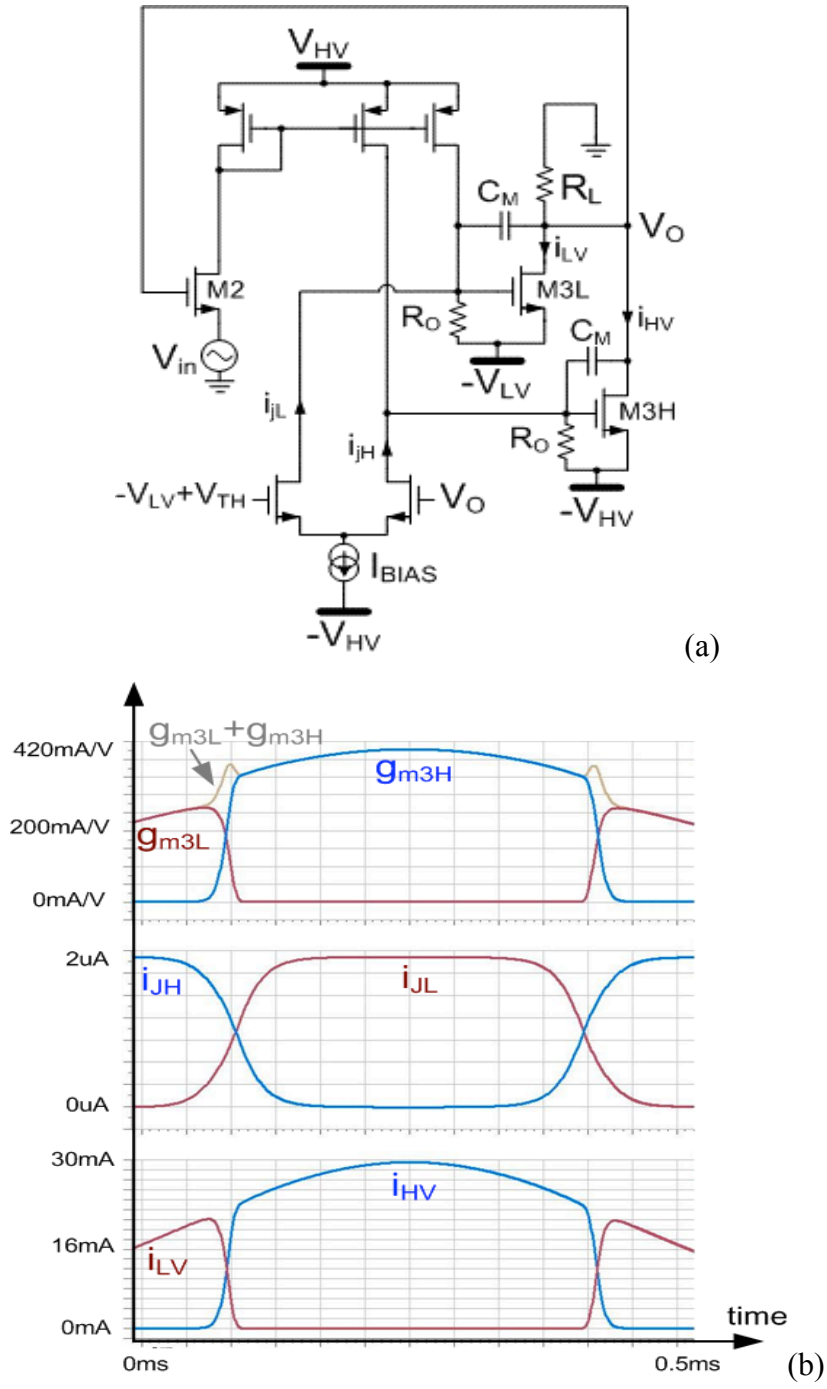


Figure 33: Class G amplifier closed in a unity gain feedback loop (a) and the simulated parameter versus time (b)

A design target is to size the two output stages to minimize the gm3L(Vo)+gm3H(Vo) variation due to switching as shown in Figure 31, nonetheless the circuit of Figure 33(a) has been intentionally sized to have a significant variation of gm3L(Vo)+gm3H(Vo) (as shown in Figure 33(b)). This was done to show that this is not a strict requirement for the model to remain valid as confirmed by the simulated results shown later in this Section.

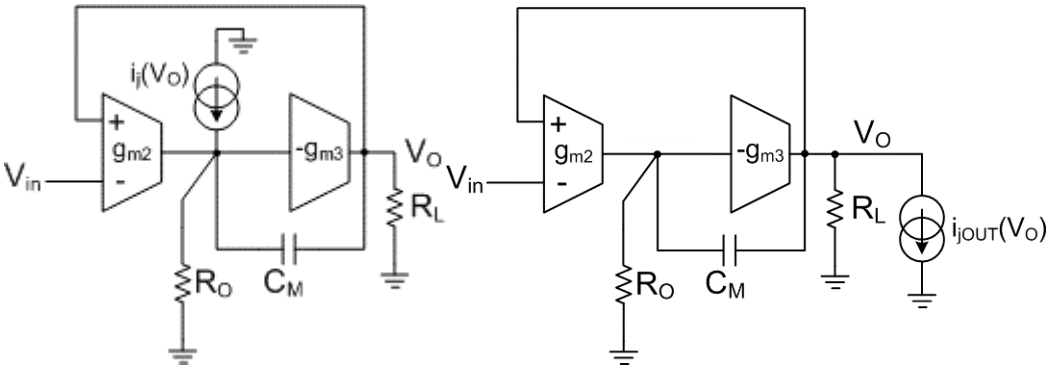


Figure 34: Amplifier model of Figure 33(a) used to analyze the amplifier rejection to the switching distortion (a) and, amplifier model used to analyze the amplifier rejection to the last stage non linearity (b).

Using the model of Figure 34(a), the relation between the injected currents ij and the output voltage can be computed obtaining the following expression:

$$\frac{V_O}{i_j(V_O)} \cong \frac{\dfrac{g_{m3}R_L R_O}{1+g_{m2}g_{m3}R_L R_O}}{1+s\dfrac{C_M(1+g_{m3}R_L)R_O}{1+g_{m2}g_{m3}R_L R_O}} \qquad 31$$

Assuming $g_{m2}\,R_O\,g_{m3}\,R_L >> 1$ and $g_{m3}\,R_L >> 1$ equation (31) becomes:

$$\frac{V_O}{i_j(V_O)} \cong \frac{1}{g_{m2}} \cdot \frac{1}{1 + s\frac{C_M}{g_{m2}}} \qquad 32$$

Equation (32) expresses the ability of the amplifier to reject the switching distortion. The closed loop residual distortion is inversely proportional to gm2 and directly proportional to IBIAS (since ij is proportional to IBIAS). In headphone application, the amplifier bandwidth of a two stage amplifier is approximately given by gm2/CM and needs to be much higher than the signal bandwidth (20kHz). This means that the distortion introduced by the switching circuit is, to first order, insensitive to the value of CM, gm3 and RL. The result is quite different from the rejection of the distortion due to the output stage (that is the main source of distortion in a class AB stage). In fact, in this case, the distortion can be modeled as a current source, $i_{JOUT}(Vo)$, connected to the output node of the amplifier [1] (see Figure 34(b)). The transfer function from i_{JOUT} to Vo can be calculated and simplified assuming:

$g_{m2}R_O g_{m3}R_L >> 1$, $g_{m3}R_L >> 1$ and $f_{in} >> 1/(R_O C_M)$

The result is given in eq. (33):

$$\frac{V_O}{i_{jOUT}(V_O)} \cong \frac{1}{g_{m3}} \cdot \frac{s\frac{C_M}{g_{m2}}}{1 + s\frac{C_M}{g_{m2}}} \qquad 33$$

To minimize distortion in a class AB, gm3 needs to be the highest allowed by power consumption constraints and the closed loop bandwidth needs to be the

highest allowed by stability requirements. Eq. (33) doesn't give constraints on the absolute values of gm2 and CM but only on their ratio. Eq. (32), instead, shows that the ability of the loop to reject switching distortion is directly proportional to the absolute value of gm2. The requirement of a high gm2 value has negligible effect on the power consumption but forces to use a compensation capacitor CM much bigger than the one that could be used in a class AB amplifier with a significant increase in silicon area.

We have simulated the THD of the simplified class G amplifier for a 1kHz input signal with 400mV peak to peak amplitude centered around -500mV. The simulations were carried out under the following operating conditions –VHV=-1.8V, -VLV=-0.9V –VLV+VTH=-700mV and assuming the following parameter values gm2=300uA/V, cm=5pF, RL=32Ohm, IBIAS=1uA, gm3=300mA/V (at DC point).

We have chosen a switching point relatively far from the low voltage supply (VTH=200mV). This was done to insure that the overall distortion is primarily caused by the switching operation. In this way we are able to validate the model of Fig. Figure 34(a). On the other hand, if the switching point is chosen very close to the low voltage supply (i.e. VTH close to zero), the distortion caused by the non linearity of the last stage becomes dominant. Figure 35 shows the dependence of THD on various design parameter values (IBIAS, gm2, RL, CM).

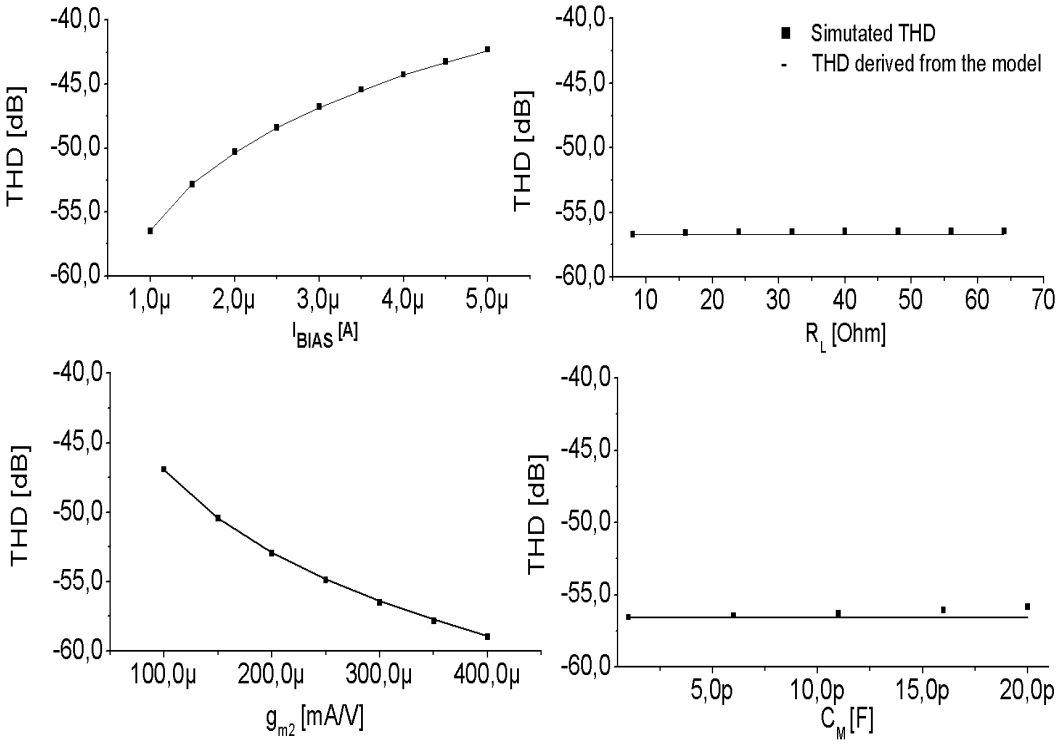


Figure 35: Comparison between the simulated THD and the model of Figure 34(a)

The plots display both simulated results and analytical calculation obtained using equation (32). The expression for the current Ij (necessary to calculate the absolute value of the distortion for the model of Figure 34(a)) has not been analytically derived. Instead, the first calculated value of each plot has been taken equal to the simulation result obtained from the circuit of Figure 33(a) and the others have been calculated starting from these points. It can be seen how the simulation results validated the equations derived from the model of Figure 34(a). In particular the plot of THD as a function of CM shows that, as predicted by eq. (32), the switching distortion is independent from the CM value. This is different from what happens to the output stage distortion (predicted by eq. (33)).

3.2 Switching speed limitation

In the previous Section it has been shown that the distortion introduced by the switching circuit is proportional to the switching current IBIAS. Therefore to achieve high linearity IBIAS should be taken as small as possible while still insuring proper operation. The first limit to the minimum usable value for IBIAS is determined by matching requirements. For the amplifier of Figure 33(a) a mismatch associated with the PMOS transistors in the top mirrors causes an error current which appear at the same points as the injected switching current. If the switching current value (IBIAS) is smaller than the maximum equivalent mismatch current of the PMOS transistors, the switching circuit may not be able to switch on and off the output stages as required to implement class G operation.

The second potential limit to the minimum value of IBIAS is defined by the switching speed requirements. This limit can significantly affect the amplifier linearity, and, for this reason, a circuit which overcomes the problem has been introduced.

The speed limitation is due to the fact that the output MOS transistor M3L is switched off by pulling down its gate with the small current IBIAS. Since the capacitive load at this gate is significant (being the sum of the gate capacitance of M3L and of the Miller capacitance CM) the switching time can be very long.

Also, M3L needs to be switched off when the output voltage is close to the switching point i.e. when it is delivering its maximum current and its gate voltage is close to VHV. Furthermore, in order to keep M3L off even when the output voltage approaches –VHV its gate should be pulled all the way down to –VHV. In fact, if this is not the case, the terminal of M3L, connected to the output, starts to behave as a source and a cross conduction current can flow from –VLV to –VHV.

Due to all the above constraints the required voltage variation on the gate of M3L and the corresponding switching time Δt are related by the following equation:

$$\Delta V_{GateL} = 2 \cdot V_{HV} = \frac{I_{BIAS}}{Cgate + C_M} \cdot \Delta t \qquad 34$$

The maximum allowable value for Δt that still insure proper switching (no cross conduction) is given by the time during which the output voltage at the maximum audio frequency (20kHz) goes from –VLV+VTH to –VLV.

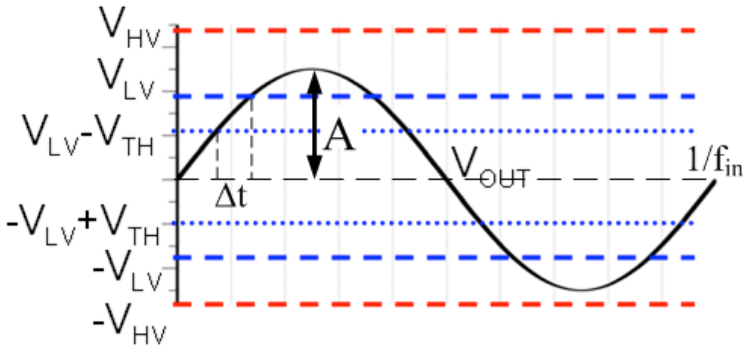


Figure 36: Output voltage versus time including the supply voltages and the switching point levels

Figure 36 shows a graphical representation of the maximum allowable value for the time interval Δt while eq.(35) gives its value:

$$\Delta t = \frac{V_{TH}}{A \cdot 2\pi f_{in}} \qquad 35$$

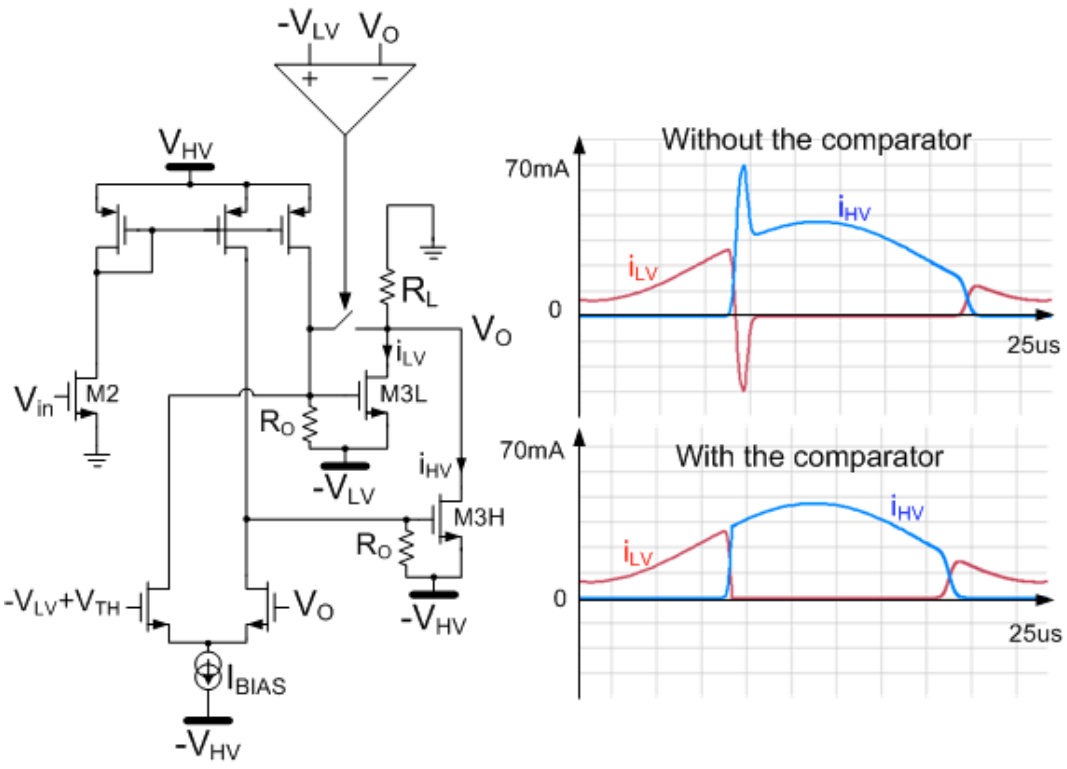
Combining equations (34) and (35) it is possible to obtain an expression for the minimum allowed switching current (IBIAS):

$$I_{BIAS} = \max\left(\frac{\Delta V_{GateL}}{V_{TH}} 2\pi f_{in} \;\; A \cdot C_M\right) = \frac{2V_{HV}}{V_{TH}} 2\pi f_{in,max} \;\; V_{HV} \cdot C_M \quad \textbf{36}$$

This represents the minimum value for IBIAS which still ensure a correct switching operation in the worst case condition that correspond to the max gate voltage swing (ΔVGateL=2V_{HV}) and max output voltage slope (i.e. max signal frequency and max output signal amplitude V_{HV}).

For this design, equation (36) suggests a value of IBIAS of 40uA that is much higher than the value necessary to satisfy the matching requirements (1uA).

In order to use a value of IBIAS smaller than the one required using equation (36), which means less distortion, we have implemented a hard switching circuit which immediately switches off M3L when the output voltage falls below the negative low voltage supply -VLV. Figure 37(a) shows the implementation of the class G amplifier including such hard switching circuit which is made up by a comparator and a switch (realized using an NMOS transistor). If the output voltage Vo is higher than –VLV the switch is open and the amplifier works as described in the previous section. When Vo falls below –VLV the switch is turned ON hard by the comparator and the gate of M3L is immediately connected to the output node Vo forcing the gate-source voltage of M3L to be equal to zero. The hard switching circuit doesn't introduce any additional distortion because it acts when the VDS of M3L is equal to zero i.e. when M3L is not contributing any current to the output. The plot at the top of Figure 37(b) shows the output currents without the hard switching circuit assuming a small IBIAS value (used to reduce switching distortion).



(a) (b)

Figure 37: Class G amplifier schematic including the hard switching circuit (a) and the output currents versus time using or not the comparator (b)

A current glitch flows between the supplies (-VLV and –VHV) when the output voltage crosses the switching point. This current means wasted power and has a negative impact on the overall distortion. The plot at the bottom of Figure 37(b) shows that the glitches have been removed. However, the limited speed with which the gate of M3L can be turned ON (that still depends only on IBIAS/(Cgate+CM) since the hard switching circuit is of no help) affects the shape of the output currents that appear to be non symmetric. These non symmetrical output currents only slightly affect the linearity (even at high frequency) and the efficiency and have been accepted in this design. As a

consequence the value of IBIAS has been defined only by matching requirements. To minimize the amplitude of current glitches on the output node the comparators have been designed to have a rise/fall time smaller than 10us and an offset smaller than 5mV. The residual offset doesn't affect significantly the linearity because the harmonics fall outside the audio bandwidth.

3.3 Amplifier architecture

The implemented class G headphone driver uses a differential input and single ended output topology. The number of stages to be used depends on the linearity requirements: three stages are enough to satisfy the target THD of -80dB at 1kHz. Figure 38 shows the chosen architecture. The amplifier main path is composed of three stages compensated using the nested Miller techniques [2] (look at the appendix for more information about the compensation technique).

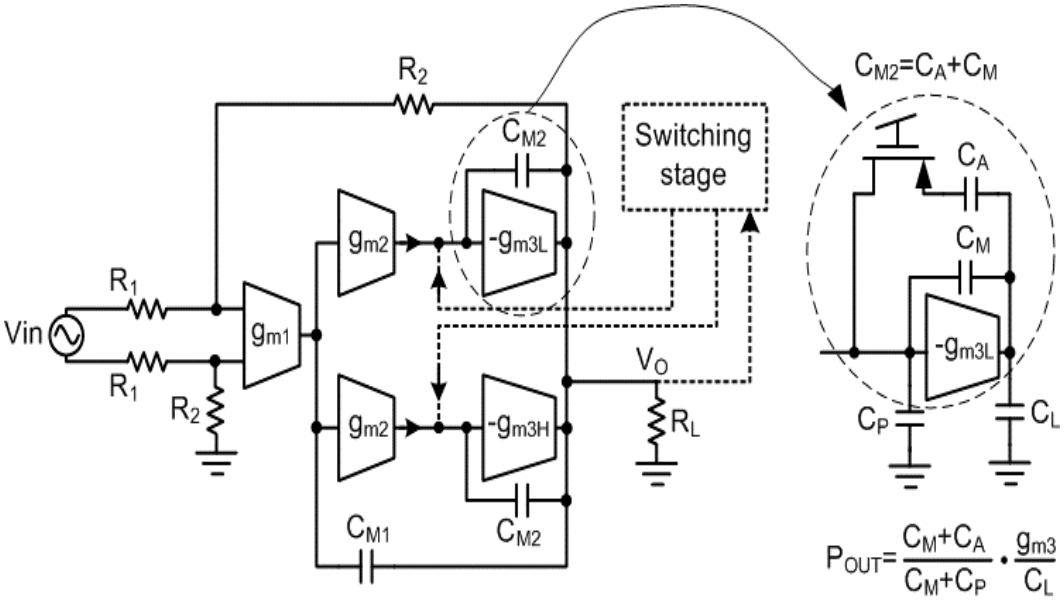


Figure 38: Amplifier architecture (three stage opamp)

In order to save quiescent power consumption, the last stage gm3 is compensated using an active-feedback compensation technique [10] as shown in Figure 38. The first stage, gm1, drives two parallel paths. One is made up by the cascade of gm2 and gm3L and the other one by the cascade of gm2 and gm3H. The switching stage alternatively enables one of these two parallel paths according to the output voltage level relative to the threshold level. Only the gm3L stage is supplied by the low voltage rail VLV while the rest of the circuit is supplied by the high voltage rail VHV.

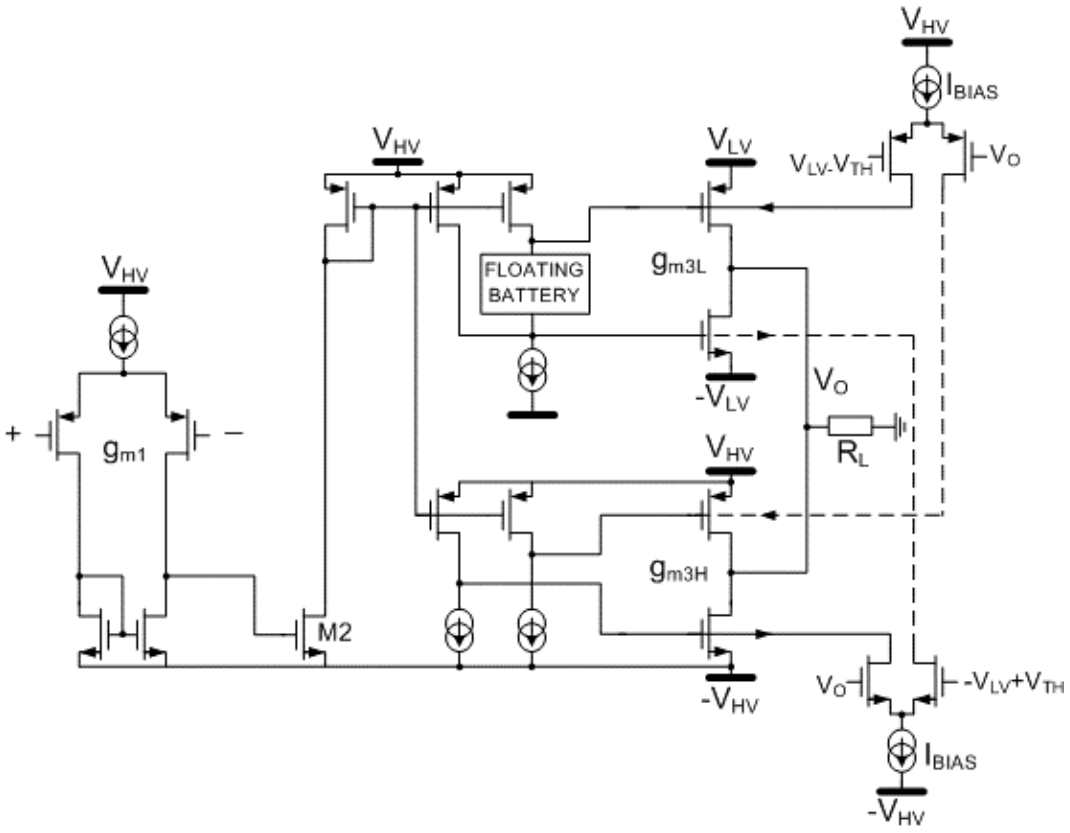


Figure 39: Transistor level implementation of the class G opamp

Figure 39 shows a somehow simplified transistor level implementation of the amplifier. The first stage, gm1, is a conventional PMOS differential pair that also

implements the differential to single ended conversion. The second stage is made up by the NMOS transistor M2 that drives a PMOS current mirror with one input and two pairs of output. The floating battery, shown in Figure 39, is a circuit used to control the quiescent bias current of the push-pull low voltage output stage, gm3L [3]. The complete output stage is composed by two push-pull class AB paths working in parallel (for simplicity, in the previous sections only the pull-down part was shown). The switching principle is the same as the one reported in the previous section. The switching circuit is composed by two differential pairs: one of them acts on the NMOS power transistors and the other one acts on the PMOS power transistors. For simplicity the hard switching circuit has not been shown: it is built with two comparators and two switches which rapidly switch off the NMOS and PMOS low voltage power transistors when the output voltage falls below –VLV or rise above VLV.

3.3.1 Switching distortion analysis

The switching distortion analysis that was carried out in Section 3.1 for the case of a two stages amplifier can be easily extended to a three stage topology. Fig. 13 shows the simplified linear model corresponding to this case.

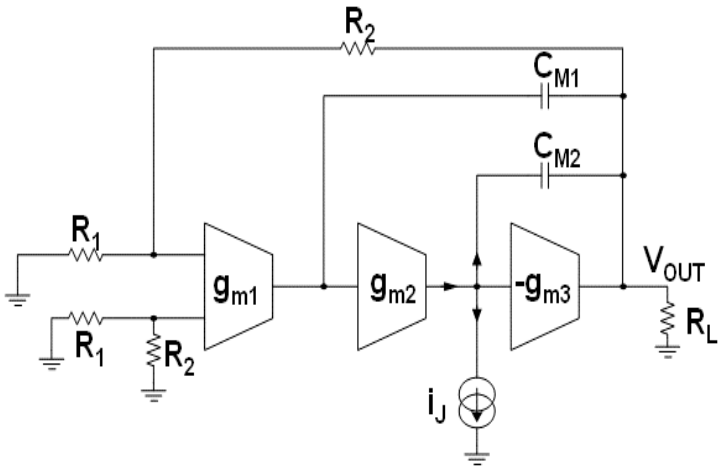


Figure 40: Simplified model of the class G amplifier shown in Figure 39

The transfer function from the switching current to the output is given by

$$\frac{V_O}{i_j} \cong -\frac{1/(\omega_t \cdot g_{m2})}{1 + s/\omega_t + \frac{s^2}{\omega_t}\frac{C_{M2}}{g_{m2}}\frac{1 + g_{m3}R_L}{g_{m3}R_L}} \qquad 37$$

Where $\omega_t = g_{m1}/C_{m1} \cdot R_1/(R_1 + R_2)$ is the closed loop -3dB angular frequency of the amplifier that, in this design, is much higher than the signal bandwidth (ωt=2π230krad/s versus an audio bandwidth of BW=20kHz). Moreover, for stability reasons, $\frac{g_{m2}}{C_{M2}}\frac{g_{m3}R_L}{1 + g_{m3}R_L}$ has been chosen twice as high as ωt. It follow that equation (37) can be approximated by equation (38) with very good accuracy.

$$\frac{V_O}{i_j} \cong -\frac{1}{\omega_t \cdot g_{m2}} \qquad 38$$

The simple relation reported in equation (38) expresses with good approximation the ability of the used three stage amplifier to reject the unwanted current injection Ij associated with the switching circuit. From equation (38) we can see that the amount of switching distortion compression operated by the amplifier is proportional to the value of gm2 as in the case of a two stages amplifier.

As for the simpler two stage amplifier of Section 3.1 the THD of the complete amplifier has been obtained via simulation applying a 1kHz 600mVpp tone centered at 0V at its input under the following operating conditions VHV=1.8V, VLV=0.4V and assuming the following parameters gm1=80uA/V, gm2=120uA/V, gm3=100mA/V (simulated in DC operating point), cm1=10pF, cm2=7pF. Again the switching points have been chosen far from the low voltage supply VLV (VTH=200mV) to avoid saturation effects of the low voltage stage.

Following the same procedure described in Section 3.1 the model of Figure 40 has been validated evaluating the distortion sensitivity to the key circuit parameters and comparing the results with the simulations on the real circuit. Figure 41 shows that the comparison results validate the expression given by equation 38.

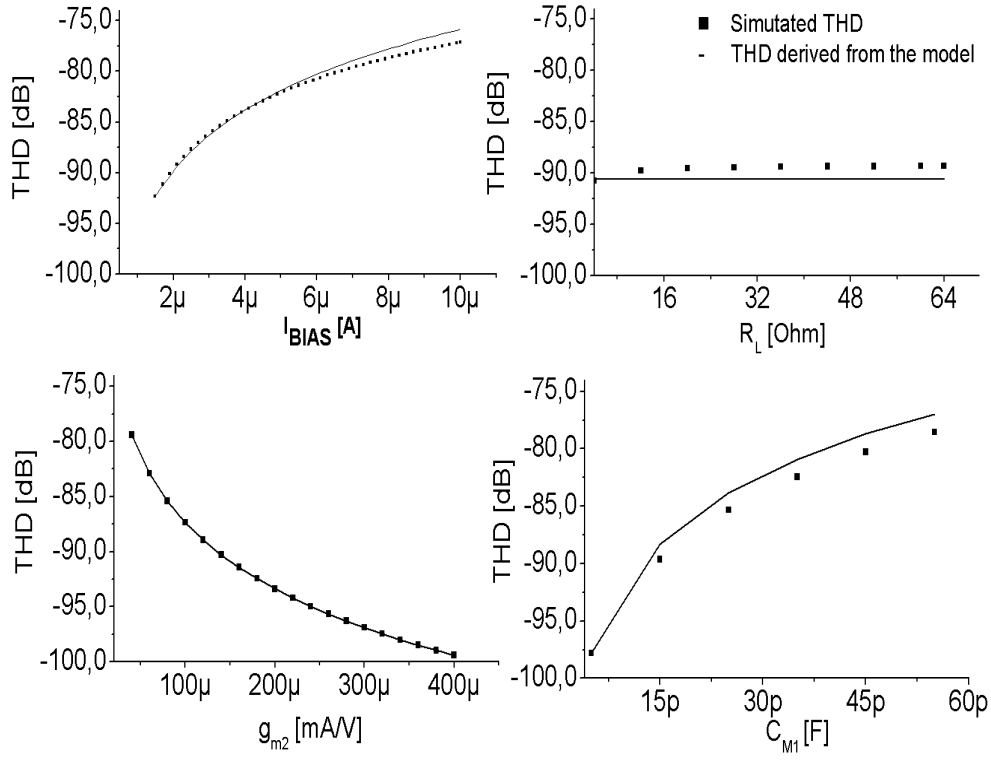


Figure 41: Comparison between the simulated THD of the class G opamp of Figure 39 and the model of Figure 40

3.4 Experimental results

To achieve a high integration level, placing the headphone amplifier on the same die with the baseband processor would be advantageous. In this case the digitally intensive nature of the processor forces the usage of a deeply scaled CMOS technology, such as 65nm or below.

The prototype class G amplifier reported here (designed to work using a high voltage supply VHV=1.4V and a low voltage supply VLV=0.35V) has been

realized using the 1.8V devices available in 65nm CMOS. For reliability reasons no transistor should see (even during transients) a gate to source and gate to drain voltage that exceed the maximum value allowed by the process. Since the amplifier output voltage swings from -1.4V to 1.4V, the 1.8V output transistors need to be cascaded and the biasing circuit should be carefully designed to ensure that no device is stressed beyond its limit even during power on/off transients.

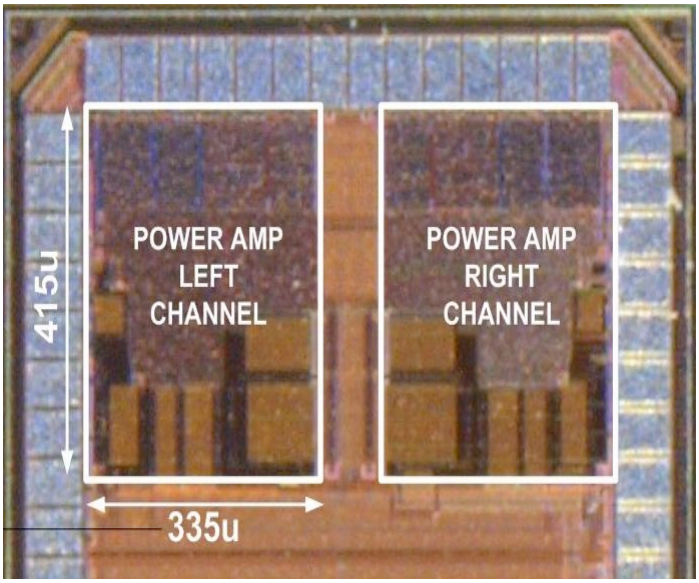


Figure 42: Die micrograph

The die micrograph is shown in Figure 42 and corresponds to a silicon area of 0.14mm^2 per channel. We have integrated both audio channels. Table 3 shows the values of gm's in the three stages and that of the compensation capacitors. CM2 represents the value of each compensation capacitors used across the four output transistors: this means that the total capacitance is four times the CM2 reported. As explained in Section 3.1 the area occupied is much bigger than the area necessary to compensate a class AB amplifier.

Parameter	Value
Unity Gain Bandwidth	230kHz
DC gain	About 100dB
Gm1	35uA/V
Gm2	200uA/V
Gm3	7mA/V
C_{M1}	12pF
C_{M2} (C_A+C_M) four C_{M2} cap are used	18pF (15pF+3pF)
R_1=R_2	25k

Table 3: gm's and currents in three stages and compensation capacitors

The switching point has been set only 50mV away from the low voltage supplies making it possible to scale VLV down to 0.35V while still having good efficiency up to an output power of 1.5mW on a 32Ω load. Figure 43 shows the efficiency and the Signal to Noise + Distortion Ratio (THD+N) versus the output power that displays the classical non monotonic behavior of class G amplifiers.

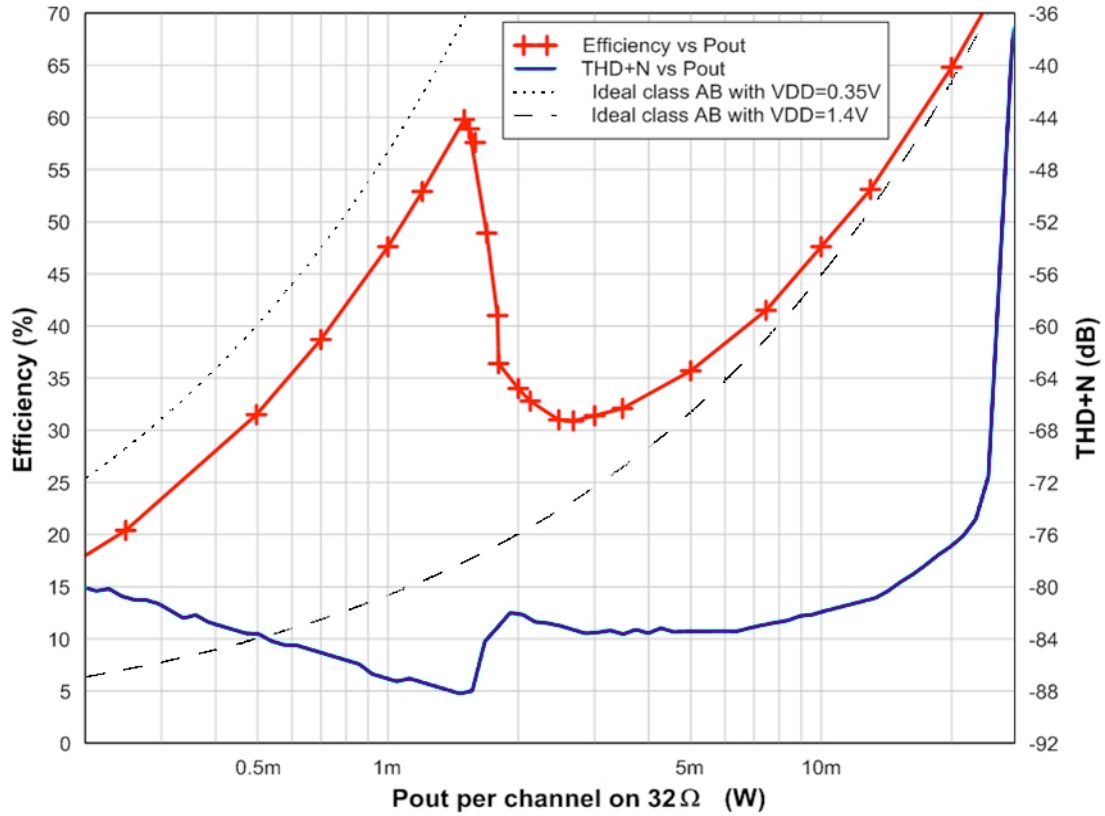


Figure 43: Total Harmonic Distortion and efficiency versus output power delivered on 32Ω

Better than 30% efficiency is achieved from Pout=0.5mW to the maximum deliverable output power of 30mW, THD+N for a 1kHz tone is better than 80dB in the entire operating range and reaches -88dB just below the power level for which class G operation begins to occur at the signal peaks. As expected, beyond this point the THD+N plot shows a slight degradation (a step of about 6dB occurs at the output amplitude corresponding to the switching point). Figure 44 shows the power dissipation versus the delivered output power (per channel). A key target for a headphone driver is to minimize the power dissipation when the delivered output power is 0.1mW and 0.5mW into a 32Ω load.

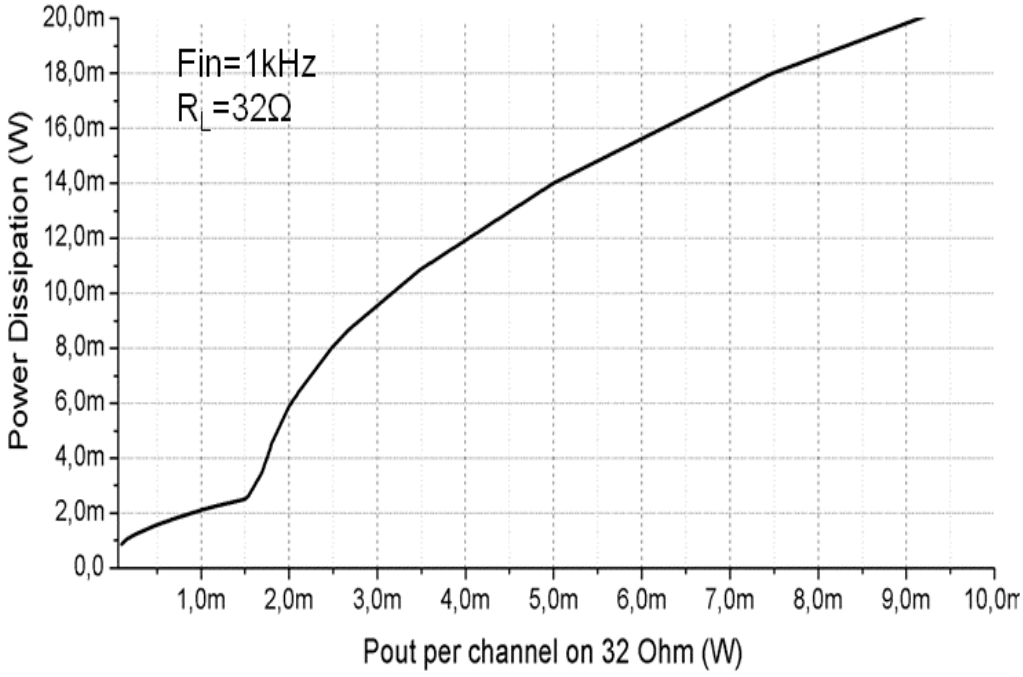


Figure 44: Amplifier power consumption versus the delivered output power on 32Ω

This is because such power levels define the range of operation of a typical headset speaker used to reproduce MP3 audio signals. In those conditions, the proposed driver dissipates much less than today state of the art circuits [4][5][6][7][8][9]. When the output power is below 1.5mW the amplifier uses only the low voltage stage thanks to the low VTH used and the power dissipation is less than 3mW as shown in Figure 44. To deliver an output power exceeding 1.5mW requires also the use of the high voltage stage. As shown by the plot from this point on the power dissipation increases at a much higher rate with the delivered power.

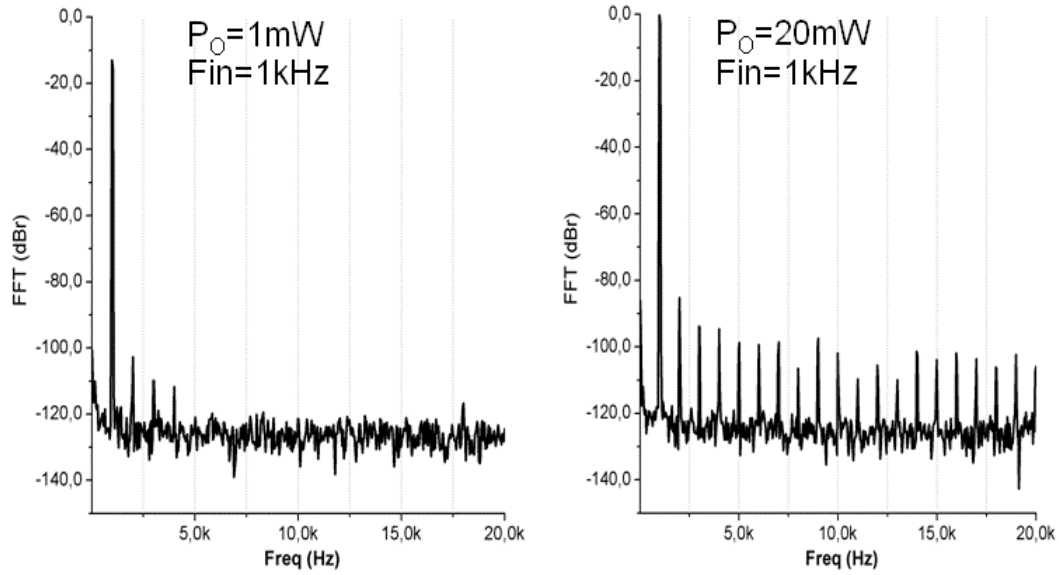


Figure 45: Output Spectrum at different output power level

Figure 45 shows the output spectra at two different output power levels. In Figure 45(a) the output power is 1mW (this level is below the switching threshold and the amplifier works always in class AB) producing a very clean spectrum. On the other hand, in Figure 45(b) the output power is 20mW. In this condition the amplifier uses both the output stages (class G operation) and the spectrum includes also the harmonic distortion introduced by the switching operation. In both cases of Figure 45, the distortion is dominated by the second harmonic.

Figure 46 shows a plot of THD versus frequency for two different output power levels (2mW and 5mW). In both cases distortion increases with frequency up to fin=10kHz (above this frequency distortion appears to decrease because THD has been evaluated only in the signal band integrating both noise and harmonics up to 20kHz). In the frequency range of interest the maximum THD degradation with respect to low frequency is only 10dB and this is a very good result for a class G amplifier [6][7][8].

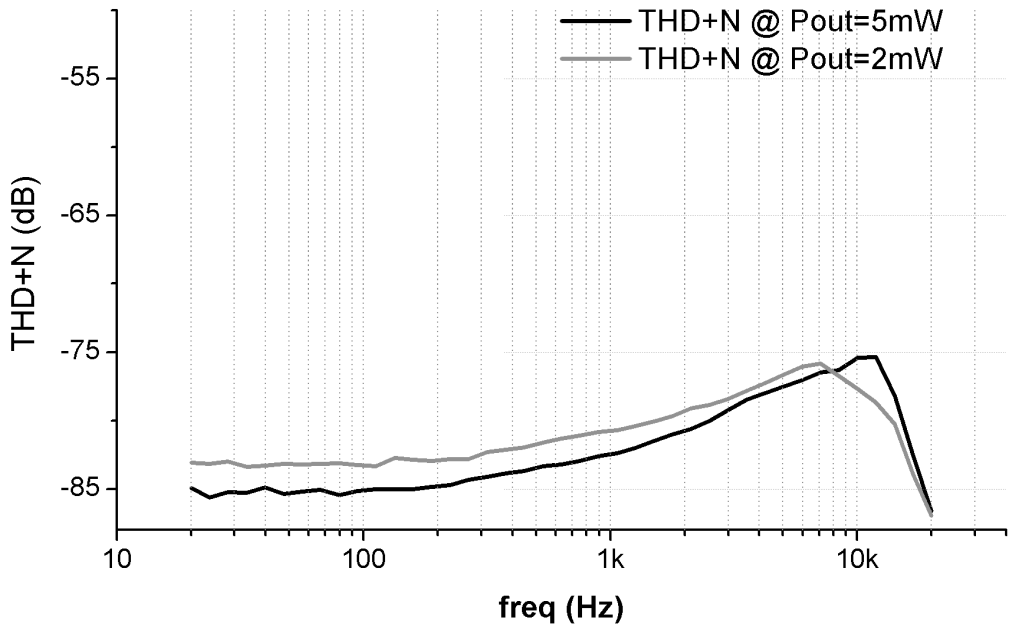


Figure 46: Total Harmonic Distortion versus frequency

This work has been compared with class AB, D, H and G amplifiers reported in technical literature or in commercial datasheets. In Table 4 the performance summary of the amplifier and a comparison with class AB and D amplifiers taken from technical literature is shown. In this table the problem of the supply voltage generation is not considered. For a fair reading of this table, it must be taken into account that the class G amplifier, compared with class AB and D amplifiers, needs an additional supply rail (±VLV). The negative high voltage supply, instead, is widely used to avoid external decoupling capacitor (whose linearity is a limit for the audio quality).

Parameter	This work (Class G)	JSSC 09 (Class AB) [4]	ESSCIRC 06 (Class AB) [5]	ISCAS 09 (Class D) [9]
Technology	65nm	130nm	65nm	0.13um
Supply voltage	±1.4V ±0.35V	±1V ±0.6V	2.5V	3.6V
Quiescent power (per channel)	**0.41mW**	1.2mW	12.5mW	1.8mW
Peak load power (16Ω)	**90mW**	40mW	53.5mW	50mW
THD+N @ PRMS (32Ω)	-80dB @ 16mW	-84dB @ 10mW	-68dB @ 27mW (16Ω)	-80mW @ 10mW
SNR A-weighted	**101dB**	92dB (un-weighted)	-	96dB

Table 4: Class G performance summary compared with two recent papers

The quiescent power consumption is 0.41mW which is 3 times less than the lowest reported. The peak deliverable power onto a 16 Ohms load is 90mW. The total harmonic distortion is 80dB when the delivered output power is 16mW. The signal to noise ratio (A-weighted) is 101dB.

Table 5 shows the performance summary compared with recent class H and G commercial products. Unfortunately the datasheets do not give the performance of the stand alone amplifier. For example, reference [6] includes the power consumption of two charge pumps while references [7][8] include power consumption of one charge pump and one buck converter. The buck converter is used to generate VLV starting from VHV and the charge pumps are used to generate the negative supply voltages –VLV and –VHV. In order to make a fair

comparison the power consumption of the charge pumps and of the buck converter need to be estimated.

Parameter	This work (Class G)	MAX97200 (Class H) [6]	TPA6141 (Class G) [7]	LM48824 (Class G) [8]
Supply voltage	1.4V with two charge pumps + 1 buck	1.5V with one charge pump	3.6V with 1 charge pump + 1 buck	3.6V with 1 charge pump + 1 buck
Quiescent power (per channel)	**0.41mW + 0.3mW (2 CPs + 1 buck)**	1.05mW	2.16mW	1.62mW
PSUP @ PL=0.1mW	**0.87mW + 0.4mW**	-	4.5mW	3.24mW
PSUP @ PL=0.5mW	**1.63mW + 0.6mW**	-	7.2mW	5.58mW
Peak load power (16Ω)	**90mW -> 70mW (CP RON=2.5Ω)**	90mW	50mW	74mW
THD+N @ PRMS (32Ω)	-80dB @ 16mW	-87dB @20mW	-80dB @20mW	-69dB@20mW
SNR A-weighted	100dB	105dB	105dB	102dB

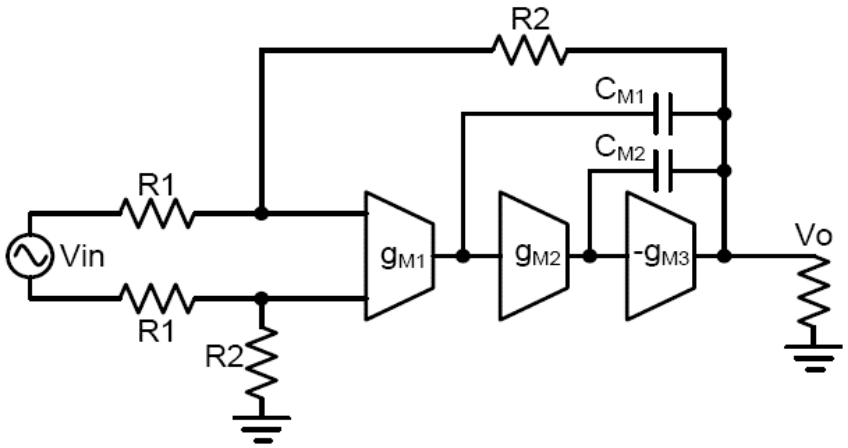
Table 5: Class G performance summary compared with recent commercial products

Conservative power estimation for the two inverting charge pumps, with 2.5Ohm of series resistance, is 0.6mW (0.3mW per channel). This power is fairly independent from the delivered current. On the other hand Buck converter efficiency increases with the delivered current. We assumed quite a low efficiency that should be easily achievable: 50% when delivering the quiescent current (0.4mA) resulting in 0.1mW power dissipation and 70% when the output power is between 0.1mW and 0.5mW. Table 5 shows the overall power consumption of the reported class G amplifier taking into account also the contribution of the charge pumps and the buck.

3.5 Improved design for high audio quality

The amplifier shown in the previous Sections has been designed to meet the requirements for standard audio quality that fit the needs of the majority of the portable applications. A new version has been developed to cover the high audio quality specifications needed for smart phone and high quality MP3 players. While the improvement in terms of linearity is quite small (THD@1kHz = 80 dB to 85 dB), the SNR needs to increase significantly (SNR = 100dB to 110dB).

The SNR required improvement could be achieved simply resizing the circuit previously shown.



$$f_T = \frac{g_{M1}}{2\pi C_{M1}} \cdot \beta \quad \text{where} \quad \beta = \frac{R_1}{R_1 + R_2}$$

Figure 47: Block diagram of a three stages amplifier showing its cut off frequency f_T

Figure 47 shows the block diagram of a generic three-stage amplifier. As reported in the figure, the amplifier cut off frequency, f_T, is proportional to the ratio g_{M1}/C_{M1}. Amplifier noise performance needs to be improved increasing the value of g_{M1} with consequent power consumption increase. Furthermore, the maximum acceptable value of the cut off frequency is set by stability requirements and cannot be increased. For this reason the noise performance increase requires an increase C_{M1} to hold the g_{M1}/C_{M1} ratio.

	3-stages	3-stages improved
SNR@ $1V_{RMS}$	**100dB**	**110dB**
C_{M1}	**15pF**	**260pF**
C_{M2}	**4x18pF**	**4x18pF**

P_Q	**0.41mW**	**0.55mW**

Table 6 shows the increasing in area and power consumption due to the increasing of the SNR of 10dB in a three-stage amplifier

Table 6 shows the increasing in area and power consumption due to a 10dB increasing of the SNR. We can notice that C_{M1} is more than 17 times bigger paying a lot in terms of occupied area. For this reason we have implemented a different solution as shown in Figure 48. Two additional gain stages, g_{M11} and g_{M12}, working in parallel to the first gain stage g_{M1} are used in this architecture.

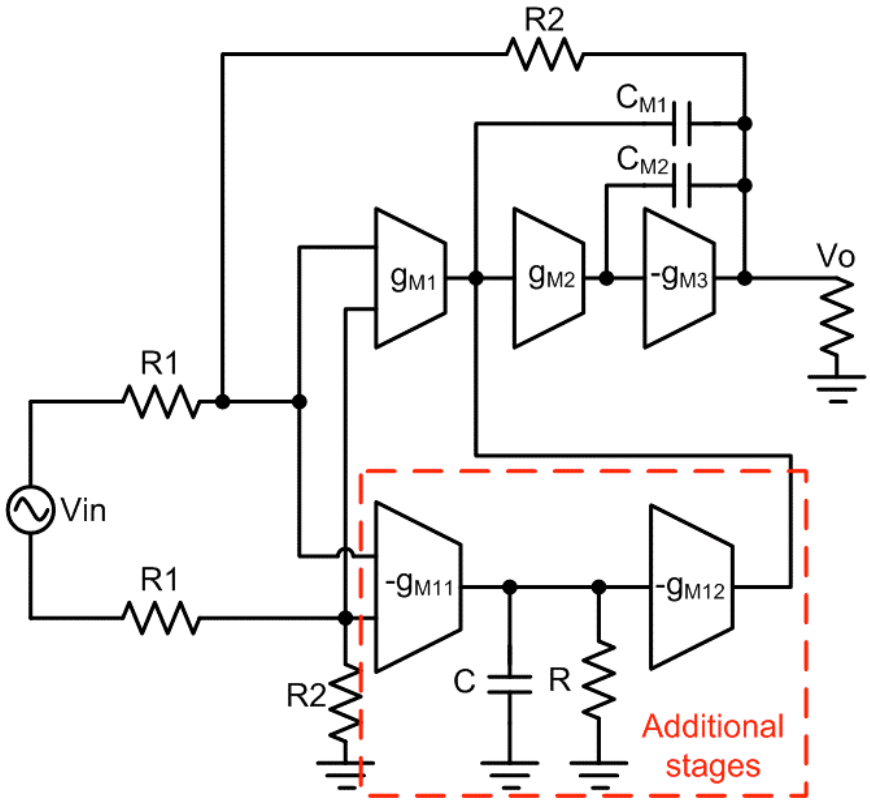


Figure 48: Four-stage class-G with feed forward.

If we study the open loop gain of amplifier shown in Figure 48 we obtain the following result:

$$G_{LOOP} = \beta \cdot \frac{1 + \frac{g_{M1}}{g_{M12}} \cdot \frac{C}{g_{M11}} s}{s \frac{C_{M1}}{g_{M11} R g_{M12}} (1 + sCR)\left(1 + s C_{M2} \frac{1 + g_{M3} R_L}{g_{M3} R_L}\right)} \quad \textbf{39}$$

The impedance at the output of g_{M1} and g_{M2} is very high and can be considered infinite for the study of the GLOOP in the band of interest. We can notice that the GLOOP comprises a zero which depends on the g_{M12}/g_{M1} ratio and on the g_{M11}/C ratio. Figure 49 shows, in continuous line, the GLOOP plot of the four-stage FF solution (see Figure 48) while, in dotted line, the GLOOP plot of the three-stage solution (see Figure 47). The Bode plot shows the graphical representation of the zero written in equation (39).

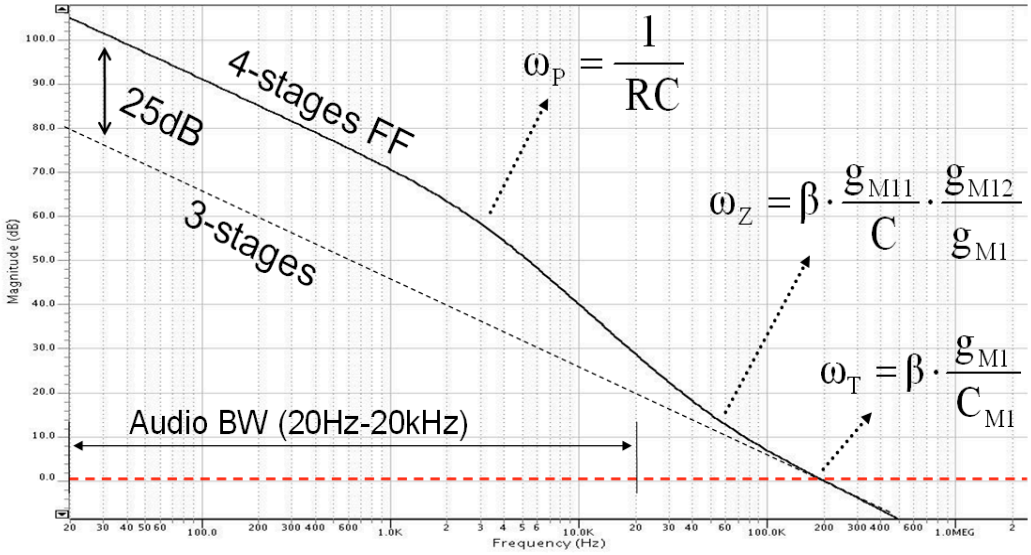


Figure 49: the continuous line represents the GLOOP plot of the amplifier shown in Figure 48 while the dotted line represents the GLOOP plot of the amplifier shown in Figure 47

The benefit given by the 4-stage implementation is a higher open loop gain inside the audio bandwidth (the cascade of g_{M11} and g_{M12} is higher that g_{M1}) without increasing the fT value (fT=200kHz) as shown in Figure 49. As in the three-stage architecture the noise is dominated by the first stage that in this case is gM11.

In the three-stage approach, g_{M1} needs to be equal to 600uA/V and C_{M1} needs to be equal to 260pF. In the four-stage solution, g_{M11} dominates the noise performance and it has to be equal to 600uA/V. We chose g_{M12}/g_{M1}=3u/60u=0.05, and the capacitor C equal to 55pF. In this way, the frequency of the zero f_Z is 43kHz. In the four-stage solution, g_{M1} is not anymore the major noise contributor and can be scaled down to 60uA/V and, consequently, C_{M1} can be equal to 26pF (10 times smaller than in the three-stage approach).

We took another advantage using the four-stage solution, which is the switching distortion compression. As we have seen in the previous section, in the three-stage

scheme the switching distortion is proportional to $\omega/(g_{M2}\ \omega_T)$ while, in the four-stage implementation, the switching distortion is proportional to $\omega^2/(g_{M2}\ \omega_T\ \omega_Z)$. Taking advantage of this fact, in the four-stage solution, it is possible to reduce the g_{M2} value maintaining the same linearity performance. The first advantage of reducing g_{M2} is the quiescent power saving, while the second (and more important) advantage is that the same g_{M2}/C_{M2} ratio can be achieved with a smaller C2 value.

	3-stage improved	**4-stages FF**
g_{M2}	**200uA/V**	**55uA/V**
C_{M2}	**4x18pF**	**4x5pF**
THD@1kHz	**-82dB**	**-85dB**

Table 7: Comparison between the three-stage solution and the four-stage FF one in terms of THD, gm2 and Cm2.

Table 7 shows a comparison of the two solutions in terms of g_{M2}, C_{M2} and THD values. We can notice that, the four-stage FF solution shows better linearity (-85dB instead of -82dB) using almost four times smaller C_{M2} capacitor area (20pF instead of 72pF).

	3-stages	**3-stages improved**	**4-stages FF**
SNR@1V_{RMS}	100dB	**110dB**	**110dB**
C_{TOT}	87pF	332pF	**101pF**
P_Q	**0.41mW**	0.55mW	0.6mW

THD@1kHz	-82dB	-82dB	**-85dB**

Table 8: Comparison between the first class G implementation (called 3-stages in the table), the 3-stage improved and the 4-stage FF.

To summarize, the four-stage FF architecture, as we can see in Table 8, achieves the same noise performance (and a small distortion improvement) using three times less compensation capacitors area and paying only 10% power increase.

3.6 Conclusions

In the first section of this chapter we have presented the working principle of the class G amplifier focusing on the novel switching approach, and, in second section, we have presented some details about the circuital implementation. In third section we have seen the implemented class G amplifier in 65nm CMOS technology, and, as we have seen in the fourth section, the presented amplifier results the best in class in terms of power consumption. In last section we have presented some improvements of the class G amplifier in order to fulfill the most aggressive performance requirements for headphone amplifiers. The basic idea shown in this chapter could be used to realize other electronic blocks, such as speakerphone, earphone, and ADSL line driver.

The Class G switching principle, described in this chapter, is a Marvell patent pending (A. Lollio, G. Bollati, R. Castello, Ref No. MP3391).

3.7 References

[1] G. Palumbo, S. Pennisi, “High-Frequency Harmonic Distortion in Feedback Amplifiers: Analysis and Applications”, IEEE Transactions on Circuits and Systems, vol. 50, no. 3, March 2003

[2] R. Eschauzier, J. Huijsing, Frequency Compensation Techniques for Low-power Operational Amplifiers, Boston, MA: Kluwer, 1995

[3] WCM Benirie, KJ de Langen, JH Huijsing, “Parallel Feedforward Class-AB Control Circuits for Low-Voltage Bipolar Rail-to-Rail Output Stages of Operational Amplifier”, Analog Integrated Circuits and Signal Processing, Vol. 8, 1995, pp. 37-48

[4] Vijay Dhanasekaran; Jose Silva-Martinez; Edgar Sanchez-Sinencio, "Design of Three-Stage Class-AB 16Ohm Headphone Driver Capable of Handling Wide Range of Load Capacitance," Solid-State Circuits, IEEE Journal of , vol.44, no.6, pp.1734-1744, Jun 2009

[5] P. Bogner, H. Habibovic and T. Hartig, ‘‘A High Signal Swing Class AB Earpiece Amplifier in 65nm CMOS Technology,’’ Proc. ESSCIRC, pp.372-375, 2006

[6] Maxim, ‘‘Low-Power, Low-Offset, Dual Mode, Class H DirectDrive Headphone Amplifier’’ Rev. 1; 3/10, accessed on Jun. 25, 2010 http://datasheets.maxim-ic.com/en/ds/MAX97200.pdf

[7] Texas Instrument, ‘‘Class-G Directpath Stereo Headphone Amplifier,’’ 3/09, accessed on Jul. 7, 2009 < http://focus.ti.com/lit/ds/symlink/tpa6141a2.pdf>

[8] National Semiconductor ”Class G Headphone Amplifier with I2C Volume Control,” August 31,2009, accessed on Jan. 25, 2010 < http://www.national.com/ds/LM/LM48824.pdf >

[9] Pillonet, G., et al,”A 0.01% THD, 70dB PSRR Single Ended Class D using variable hysteresis control for Headphone Amplifiers”, ISCAS 2009 pp.1181-1184.

[10] B. Ahuja, “An improved frequency compensation technique for COMS operational amplifiers”, IEEE J. Solid-State Circuits, vol. 37, pp. 1077-1084, Sept. 1990

[11] A. Bosi, G. Cesura, F. Rezzi, R. Castello, J. Chan, S. Wong, O. Carnu, T. Cho, “A VDSL2 Analog Front End in 0.15um CMOS with Integrated Line Driver”, ISSCC, proofread 2009

[12] Alex Lollio, Giacomino Bollati, Rinaldo Castello, “A Class-G Headphone Amplifier in 65nm CMOS Technology”, Solid-State Circuits, IEEE Journal of, vol. 45, no.12, Dec 2010

4. Conclusions

In the First chapter we have seen the definitions of the most important system requirements for headphone amplifiers (output power, efficiency and linearity) and a characterization of the real audio traces in terms of amplitude and frequency distribution (these information are useful during the amplifier design).
Second chapter shows the electrical modeling of the headphone load which is very important for dimensioning the amplifier compensation network and for studying the amplifier stability. We have seen the procedure to calculate the noise floor requirement for headphone amplifiers starting from the sensitivity of the speaker. The last section of the second chapter shows that the class G solution is the most suitable in headphone application. In fact it has efficiency comparable to that of class D (without showing EMI problems) and linearity comparable to that of class AB. Moreover, last section of this chapter shows that the "parallel" class G topology is the most suitable implementation in low voltage systems.
Third chapter shows the working principle of the class G amplifier focusing on the novel switching approach and it shows the amplifier mathematical model useful to evaluate the feedback compression of the switching distortion. The very smooth handover between the stages, obtained with this solution, enables both high linearity and high efficiency. This chapter shows in details the implemented class G amplifier in 65nm CMOS tech. The presented class G dissipates 0.41mW of quiescent power consumption which is three times less than the lowest reported in literature. In last part of this chapter we have seen some improvements of the class G amplifier in order to fulfill the most aggressive performance requirements for headphone amplifiers. This improved class G version will be implemented in

Dec 2010 into a novel Marvell audio codec. The basic idea shown in this chapter could be used to realize other electronic blocks, such as speakerphone, earphone, and ADSL line driver.

The Class G switching principle, described in this work, is a Marvell patent pending (A. Lollio, G. Bollati, R. Castello, Ref No. MP3391).

5. Appendix: compensation techniques for headphones amplifiers

In a modern headphone driver it is important to minimize the quiescent power consumption and, for this reason, the transconductance of the power transistors change, with the output signal level, more than 1 order of magnitude. This fact has a strong impact on the whole amplifier compensation. This appendix shows some compensation techniques able to stabilize the amplifier in the different operating conditions.

5.1 Nested Miller compensation technique

Let's suppose to implement an headphone amplifier composed by three stages and compensated using the nested Miller technique [1].

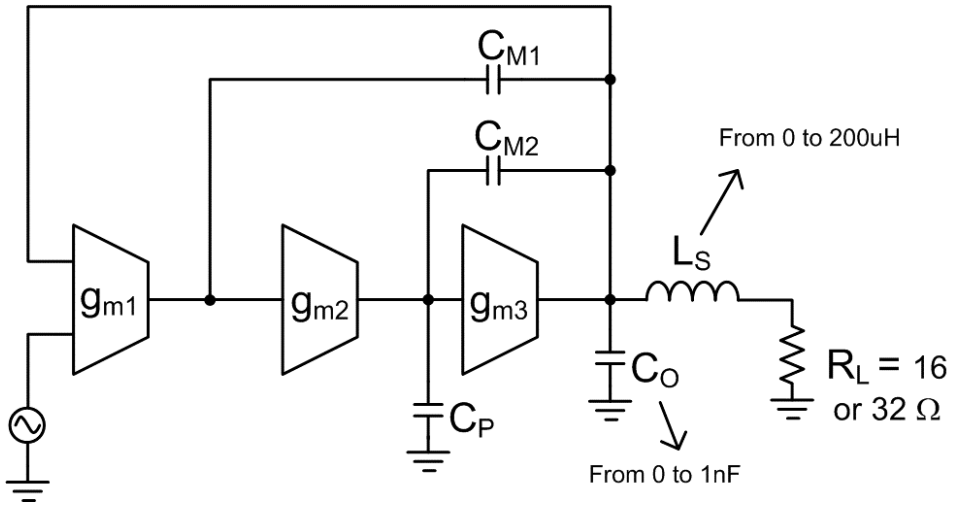


Figure 50: Three stages headphone amplifier compensated using nested Miller technique

Figure 50 shows the whole amplifier including the compensation capacitors, CM1 and CM2, the parasitic capacitance, CP, at the input of the last stage (gm3) and the electrical model of the typical headphone load (given by CO, LS and RL).
The value of gm3 is a function of the output voltage and can vary from 10mA/V @ Vout=0 to 500mA/V @ Vout=1V. This variability of gm3 determines a movement of the position of the amplifier poles.
In the chapter 2, we have seen that there are two worst cases when studying the amplifier poles position:

- The first is considering only a capacitive load Co
- The second is considering only a resistance load RL

In fact there is a wide variability of the position of the zero given by Ls/RL and it is necessary to guarantee the stability of the amplifier in all the possible scenarios.
The worst case associated to the output pole Pout happens when we have only the output capacitive load. In this scenario, we can derive, looking at Figure 50, the following relation:

$$\mathrm{Pout} = \frac{C_{M2}}{C_{M2} + C_P} \frac{g_{m3}}{C_O} \qquad \textbf{40}$$

The position of the second pole, P2, varies with the load condition. If we consider only a capacitive load we obtain

$$\mathrm{P2,max} = \frac{g_{m2}}{C_{M2}} \qquad \textbf{41}$$

Otherwise, considering only the resistance RL as the load, we obtain

$$\mathrm{P2} = \frac{g_{m2} g_{m3} R_L}{(1 + g_{m3} R_L) C_{M2} + C_P} \qquad \textbf{42}$$

If we have $g_{m3}R_L \ll 1$, equation (42) becomes

$$P2,\min = \frac{g_{m2}g_{m3}R_L}{C_{M2} + C_P} \quad \textbf{43}$$

Thanks to the high loop gain, the first pole, P1, is quite insensitive to the value of gm3 and it results approx equal to gm1/CM1 in all the operating conditions.
In order to be stable, with 60 degree of phase margin, we have to satisfy the following relation:

$$P1 < \frac{1}{2}P2 < \frac{1}{4}Pout \quad \textbf{44}$$

that means

$$P1 < \frac{1}{2}P2,\min \quad \rightarrow \quad \frac{g_{m1}}{C_{M1}} < \frac{1}{2}\frac{g_{m2}g_{m3}R_L}{C_{M2} + C_P} \quad \textbf{45}$$

and

$$P2,\max < \frac{1}{2}Pout \quad \rightarrow \quad \frac{g_{m2}}{C_{M2}} < \frac{1}{2}\frac{C_{M2}}{C_{M2} + C_P}\frac{g_{m3}}{C_O} \quad \textbf{46}$$

Both the expressions (45) and (46) pone a constraint on the minimum allowable value of gm3 and, consequently, they pone a constraint on the minimum value of quiescent power consumption.

5.2 Active cascode compensation technique

The active cascode compensation technique [2] is effective to reduce the power consumption and to overcome the limitations of the nested Miller approach (equations (45)(46)).

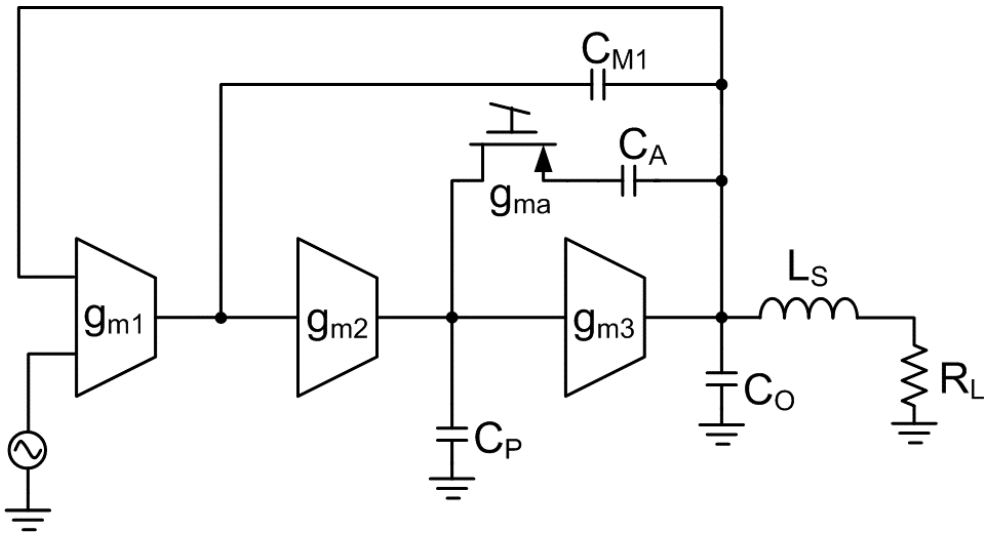


Figure 51: Three stages amplifier compensated using the active cascode technique

Figure 51 shows the block diagram of the three stages amplifier compensated using the active cascode technique. In this case the expression of Pout, assuming only a capacitive load C_O, becomes:

$$\text{Pout} = \frac{C_A}{C_P}\frac{g_{m3}}{C_O} \qquad \textbf{47}$$

If we choose $C_A/C_P > C_M/(C_M+C_P)$, the output pole, Pout, is pushed at higher frequency than the nested miller one without increasing the value of gm3, thus, saving quiescent power consumption.

There is another interesting effect of the active compensation approach which regards the position of the second pole P2: if we consider only a capacitive load on the output node we obtain:

$$P2,\max = \frac{g_{m2}}{C_A} \quad \textbf{48}$$

Otherwise, if we consider only a resistance as load, we obtain

$$P2 = \frac{g_{m2}g_{m3}R_L}{g_{m3}R_L C_A + C_P} \quad \textbf{49}$$

And if we assume $C_A g_{m3} R_L << C_P$ equation (49) becomes

$$P2,\min = \frac{g_{m2}g_{m3}R_L}{C_P} \quad \textbf{50}$$

Let's suppose C_A of equations (48) and (50) to be equal to C_{M2} of equations (41) and (42). Under this assumption we can see that the maximum value of P2 is g_{m2}/C_{M2} for both approaches (the nested miller technique and the active cascode technique). However, the minimum value of P2 for the nested miller approach is $g_{m2}g_{m3}R_L/(C_{M2}+C_P)$ while for the active cascode one is $g_{m2}g_{m3}R_L/C_P$. This means that the movement of P2 in relation with the load conditions is less in the active cascode approach instead of the nested Miller one.

5.2.1 Stability of the active cascode compensation

We have seen that the active cascode compensation technique has significant advantages in respect to the classical nested Miller one. The principal drawback is related to stability. In order to understand the problem we have to study the loop gain of the inner loop of the active cascode approach and, using the schematic of Figure 52, we obtain

$$GLOOP = \frac{C_A}{C_P}\frac{g_{m3}}{sC_O}\frac{1}{1+sC_A/g_{ma}} \quad \textbf{51}$$

Equation (51) shows a unity gain frequency equal to g_{m3}/C_O and a second pole g_{ma}/C_A. As we have seen in the first section of this appendix, the value of gm3

varies significantly with the output signal level and the value of C_O varies with different load conditions. In fact, there are DC operating conditions where the phase margin is very poor (few degrees). Having poor phase margin means to have a significant ringing of the output voltage in the transient response.

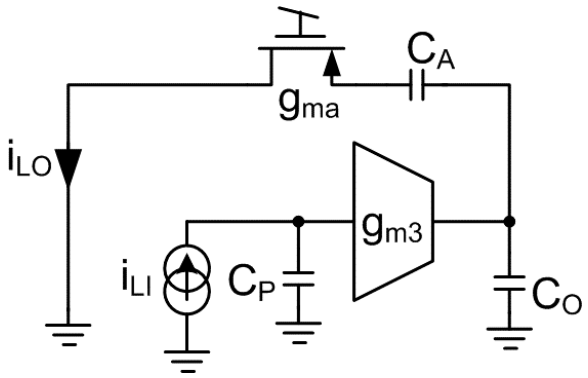


Figure 52: Study of the inner loop gain of the active cascode compensation technique

To avoid this problem it is necessary to guarantee a minimum output capacitance $C_{O,MIN}$. In this way, the maximum unity gain frequency of the inner loop becomes $g_{m3,MAX}/C_{O,MIN}$ (where $_{gm3,MAX}$ is the maximum value of gm3 varying the output signal level). To have a good phase margin we have to satisfy the following relation

$$\frac{g_{m3,MAX}}{C_{O,MIN}} < \frac{g_{ma}}{C_A} \qquad 52$$

This solution has two problems:

- It is necessary to use an external fixed output capacitor $C_{O,MIN}$ in order to guarantee the stability.
- It is necessary to spend a significant amount of quiescent current to have a high value of gma ($g_{m3,MAX}$ can be very high).

The way to solve those problems is to adopt the improved cascode compensation technique described in the following section.

5.3 Improved cascode compensation technique

The improved cascode compensation [3] enables a good phase margin in all the operating conditions without pone constraint on the minimum output capacitance CO. The solution shakes together the nested miller approach and the active cascode one (seen in previous sections).

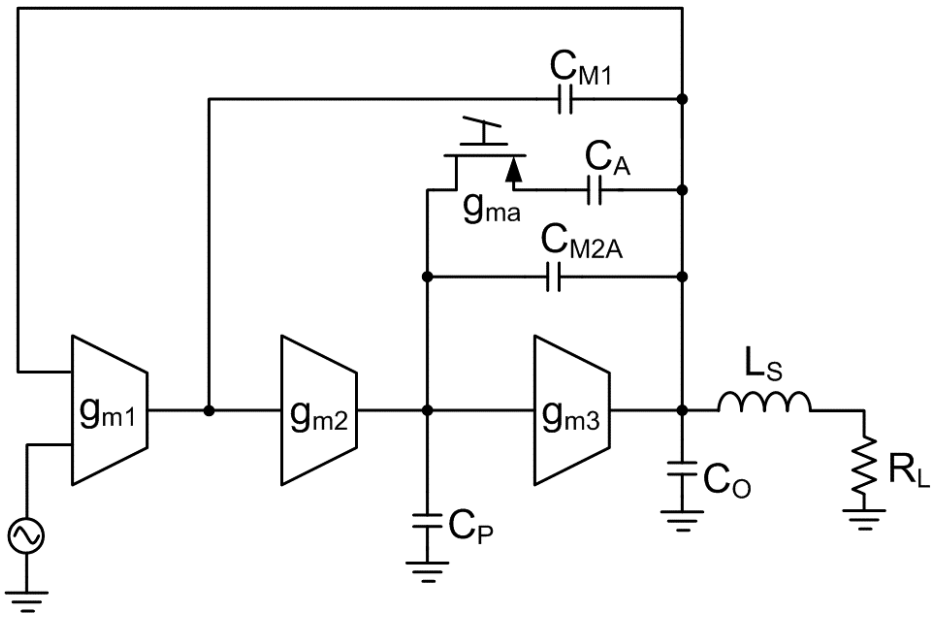


Figure 53: Three stage headphone driver compensated using the improved cascode technique.

Figure 53 shows a three stage headphone driver compensated using the improved cascode technique. The frequency of the output pole Pout, considering as load only C_O, is

$$Pout = \frac{C_{MA} + C_A}{C_{MA} + C_P} \frac{g_{m3}}{C_O} \quad \textbf{53}$$

If we choose $(C_A+C_{MA})/(C_P+C_{MA}) > C_M/(C_M+C_P)$, the output pole, Pout, is pushed at higher frequency than the nested miller one without increasing the value of gm3, thus, saving quiescent power consumption. However, the output pole Pout is pushed at a lower frequency in respect to the active cascode approach (see equation (47)). Only if C_A is much higher than C_{MA} and C_{MA} is much higher than C_P the two solutions give the same result.

If we consider only a capacitive load on the output node, the second pole of the amplifier, P2, is:

$$P2, \max = \frac{g_{m2}}{C_A} \quad \textbf{54}$$

Otherwise, if we consider only a resistance as load, we obtain

$$P2 = \frac{g_{m2} g_{m3} R_L}{g_{m3} R_L C_A + C_P + C_{MA}(1 + g_{m3} R_L)} \quad \textbf{55}$$

And if we assume $C_A g_{m3} R_L << C_P$ equation (16) becomes

$$P2, \min = \frac{g_{m2} g_{m3} R_L}{C_P + C_{MA}} \quad \textbf{56}$$

Let's suppose C_A+C_{MA} of equations (54) and (56) to be equal to C_A of equations (48) and (50) to be equal to C_{M2} of equations (41) and (43). Under those assumptions we can see that the maximum value of P2 is g_{m2}/C_{M2} for all approaches (the nested miller technique, the active cascode technique and the improved cascode technique). The minimum value of P2 for the nested miller approach is $g_{m2}g_{m3}R_L/(C_{M2}+C_P)$, for the active cascode one is $g_{m2}g_{m3}R_L/C_P$ and for

the improved cascode approach is $g_{m2}g_{m3}R_L/(C_P+C_{MA})$. This means that the movement of P2 in relation with the load conditions is less in the improved cascode approach instead of the nested Miller one but it is higher in respect to the active cascode technique. In fact the improved cascode technique is a solution which stays between the nested miller (more conservative) and the active cascode one (more aggressive).

5.3.1 Stability of the improved cascode compensation

Figure 54 shows the circuit used to calculate the loop gain of the improved cascode compensation described in the previous section. Let's suppose that the zero in the right plane given by $gm3/C_{MA}$ falls in a frequency much higher than the bandwidth of the compensation network (it is a valid assumption because the parasitic output capacitance C_O is more than one order of magnitude than C_{M2}). Under this assumption, the loop gain results:

$$GLOOP = \frac{C_A}{C_{MA}} \frac{1}{(1+s\frac{C_A}{g_{ma}})(1+s\frac{C_O}{g_{m3}}\frac{C_{MA}+C_P}{C_{MA}})} \qquad \textbf{57}$$

Equation (18) shows a loop gain with two poles, one at g_{ma}/C_A and the other at g_{m3}/C_O $(C_{MA}+C_P)/C_{MA}$, and a DC gain equal to C_A/C_{MA}. Figure 55(a) shows an example of Bode diagram of the loop gain reported in equation (57). The plot shown in Figure 55(a) corresponds to the situation with output voltage equal to zero. In this condition, gm3 assumes its lowest value. However, when the output voltage increases the gm3 value increases to, and the two poles, shown in Figure 55(a), could swap their position.

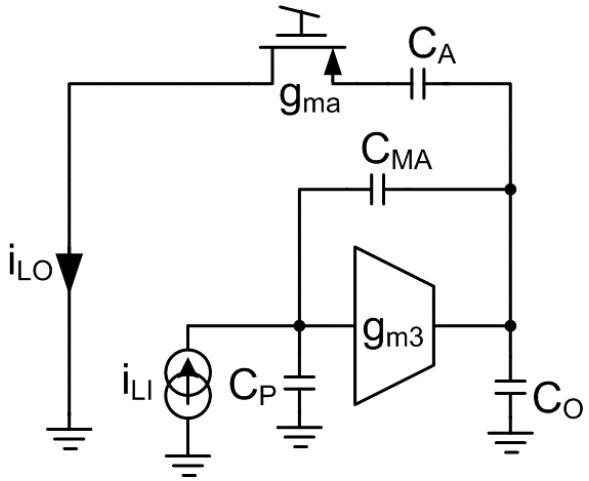
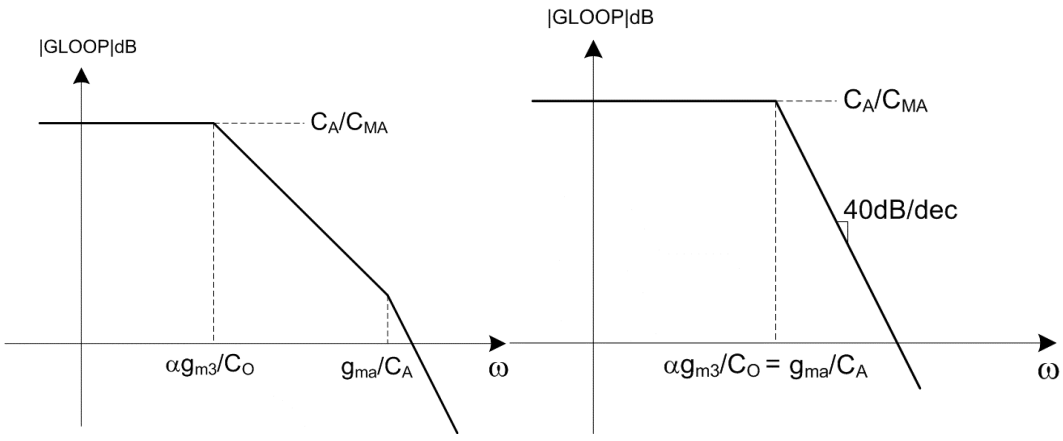


Figure 54: Study of the inner loop gain of the improved cascode compensation technique

From the phase margin point of view, the worst case happens when the two poles of Figure 55(a) are at the same frequency. Figure 55(b) shows this situation.



(a) (b)

Figure 55: Bode plot of the loop gain of circuit shown in Figure 54

In this condition, we have mathematically derived the expression of the phase margin (PM), which results:

$$PM = 180° - 2\tan^{-1}(1+\sqrt{C_A/C_{MA}}) \qquad \textbf{58}$$

It is interesting to note, in equation (58), that the phase margin is a function of the DC loop gain C_A/C_{MA}. Figure 56 shows the phase margin as a function of C_A/C_{MA}.

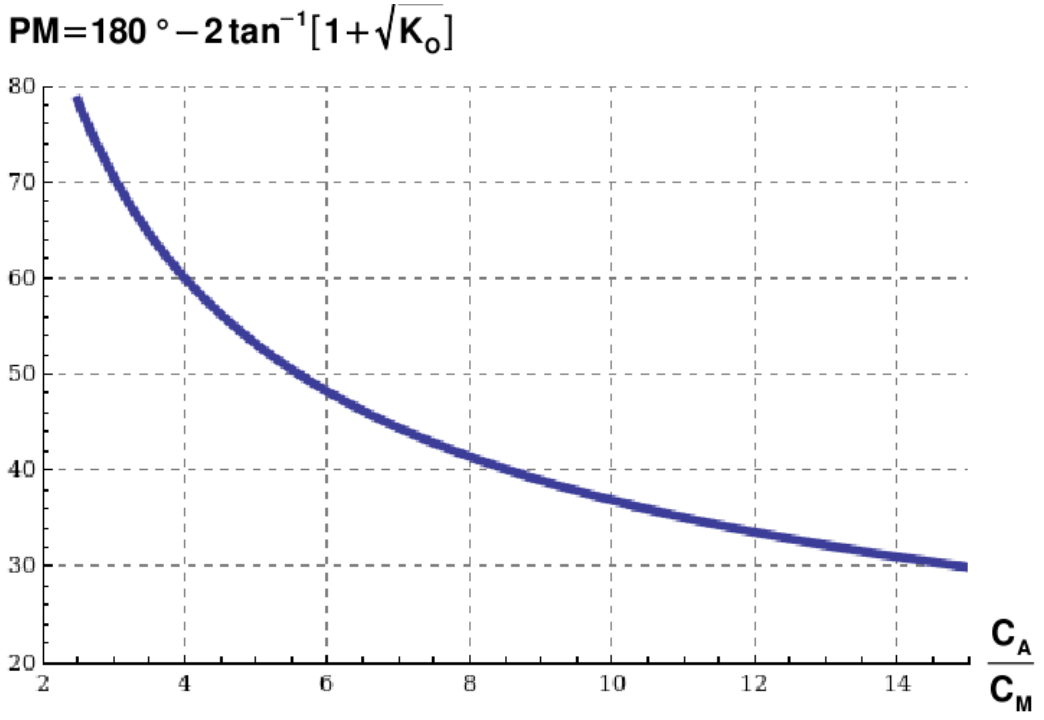


Figure 56: Phase margin of the circuit shown in Figure 54 in the worst case condition versus the open loop DC gain, C_A/C_{MA}

From this graph we can see that if we choose an open loop DC gain, C_A/C_{MA}, equal to 4 we have a phase margin which, only in the worst case condition, reaches 60 degrees and in all the other conditions is better than 60 degrees. The

improved cascode approach is less aggressive than the active cascode one, described in the previous section, but it ensures a good phase margin (and consequently a good transient response) independently from the load conditions and the operating point.

References

[1] R. Eschauzier, J. Huijsing, Frequency Compensation Techniques for Low-power Operational Amplifiers, Boston, MA: Kluwer, 1995

[2] B. Ahuja, "An improved frequency compensation technique for COMS operational amplifiers", IEEE J. Solid-State Circuits, vol. 37, pp. 1077-1084, Sept. 1990

[3] H. Lee and P. K. T. Mok, "Active-feedback frequency compensation technique for low power multistage amplifiers," IEEE J. Solid-State Circuits, vol. 38, no. 3, pp. 511–520, Mar. 2003