The Design Process for a Medium Output Power Audio Amplifier

by Kenneth A. Kuhn
Sept. 20, 2009

Introduction

This note is an introduction to the process for designing a medium output power audio amplifier for electronics engineering students. The generic circuit for the amplifier is shown in Figure 1. The architecture of this circuit is useful for amplifier output powers up to several watts into an 8-ohm speaker. Higher powers require design extensions to the output stage to accommodate the higher currents and lower impedances. This circuit is capable of very high open-loop gain in the range of many hundreds of thousands. Such an amplifier is always used with a low value of closed loop gain so that the high open-loop gain can be used to minimize distortion. The output signal of the amplifier is inverted from the input. Because of the negative shunt feedback drives the impedance at the base of Q1 towards zero the input resistance of the amplifier is simply the value of R1. The output resistance of the amplifier is very small fraction of an ohm because of the shunt voltage feedback. The closed-loop gain of the amplifier is simply the ratio of R3 to R1.

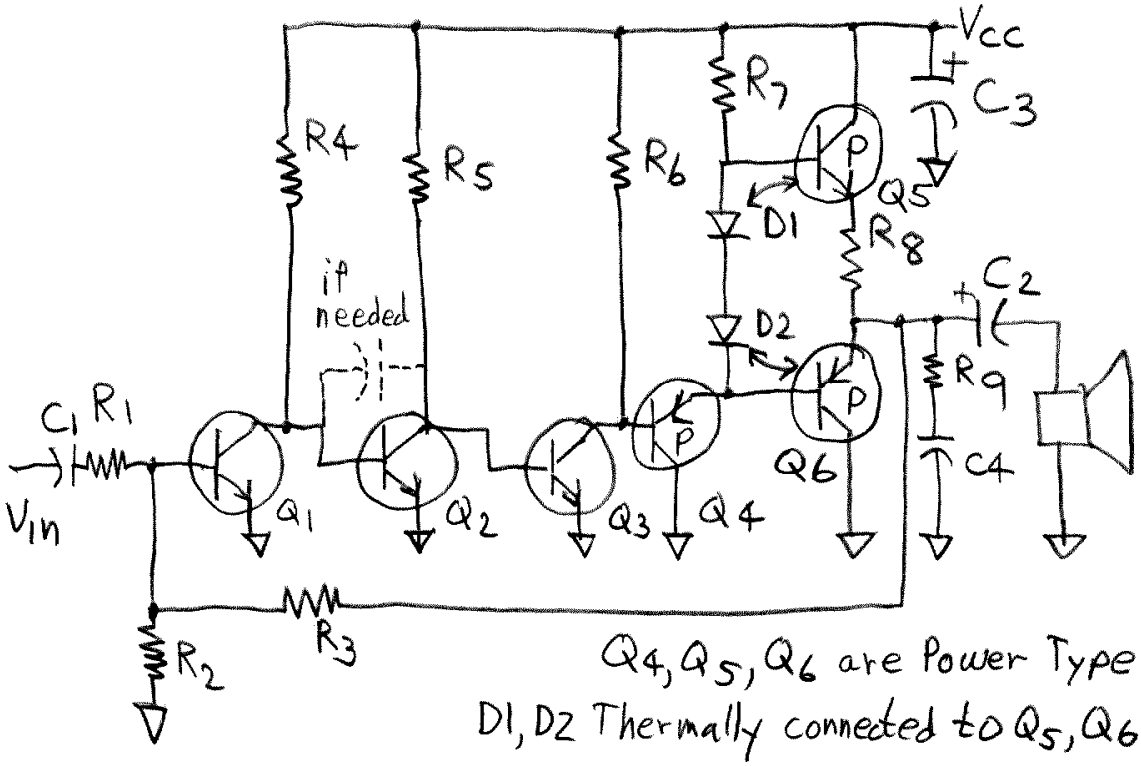

Figure 1: Simple Power Amplifier

The standard design process begins at the amplifier output and progresses to the amplifier input. A DC design is performed first to establish the resistor values and then an AC design is performed next to establish the capacitor values. It should be noted that some AC calculations are required at various steps to establish the DC operating points. The

AC signal amplitude only needs to be considered at the output stage as the signal levels at previous stages is very tiny because of the high gain. It should be noted that the design process lends itself to a spreadsheet which enables easy experimentation with various design choices to see the effect.

Notes about speakers

A speaker is an electrical transducer that converts AC electrical power to sound waves in the atmosphere. The typical design is a coil mounted inside an intense magnetic field and coupled to a large radiating area known as the speaker cone. Current in the coil reacts with the magnetic field to cause motion in the speaker cone thus producing sound. Common speakers have a nominal electrical impedance of 8 ohms although various impedances ranging from about 2 to over 100 ohms are available. The electrical impedance of the speaker is a complicated function of the air resistance experience by the speaker cone, the strength of the magnet, the number of turns of the coil, and other parameters of the speaker. It is not the electrical resistance of the coil although that is one component of the total. The electrical impedance of the speaker is mostly resistive over its useful frequency range but includes various inductive and capacitive reactances as a function of frequency and the presence of nearby acoustical structures. There is generally a mechanical resonance at some low frequency where the impedance can become fairly high compared to the nominal value. Speakers can only radiate effectively over a finite frequency band. The minimum useful frequency of the speaker is somewhat below the mechanical resonance and the upper useful frequency is typically several octaves higher. Speakers typically have low efficiencies in the single digit percent range. Thus most of the amplifier output power goes to heat in the speaker. There is no correlation between speaker efficiency and speaker quality (the ability to faithfully reproduce sound). Speakers are typically rated in dBa/watt which is not interpreted as dBa per watt but rather the sound pressure level in dBa for 1 watt at one meter distance. Common values are in the 80 to 90 dBa range. A higher rating means the speaker is more efficient. Note that a speaker rated at 80 dBa/watt would require 10 watts to produce a sound pressure level of 90 dBa and 100 watts to produce a sound pressure level of 100 dBa. All speakers have maximum power rating that must not be exceeded. Otherwise the speaker will be damaged from either heat or stress tears on the speaker cone.

Example Amplifier Specification

The example amplifier to be designed to illustrate the process will operate from a 12 volt DC power supply and drive an 8 ohm speaker. The amplifier will be designed for the maximum possible output power consistent with design practicalities. The design is performed assuming the amplifier will be operated with a low closed-loop gain so that the bias voltages are stabilized. The amplifier will not work open-loop. Some of the calculations are only approximate because all information is not very well known. Iteration can refine some of the calculations. High precision is not really necessary for

high gain systems that will be operated with a low closed-loop gain. Errors are divided by the large open-loop gain and thus become small. We try to be as accurate as practical but do not worry about extreme accuracy.

Output Stage Design

Design begins at the output. The output stage is dual class B meaning that only one transistor conducts at a time. Q5 conducts when driving the speaker with a positive signal and Q6 conducts when driving the speaker with a negative signal. Class B amplifiers are much more efficient than class A but have a non-linear issue known as cross-over distortion when the circuit switches between the amplifiers. A high open-loop gain can minimize this distortion. The efficiency of class B amplifiers can approach 78.5 percent at full output. Class B amplifiers also dissipate significantly less power at lower output levels than would a class A amplifier. Transistors Q5 and Q6 are power type transistors capable of supplying the required peak currents to the speaker. These transistors will typically be mounted in an appropriate heat sink so that the junction temperature rise from the power dissipation is not excessive. Otherwise the transistors will be destroyed from excessive heat.

Amplifier output power is normally rated in terms of a sine wave. The effective music power may be higher by an indeterminate amount depending on the wave shape. Thus, design is done assuming a sine wave as that is repeatable at any lab. There is a simple relationship for the maximum possible linear output power between the available supply voltage and the speaker impedance. This comes from the fact that the maximum possible peak-peak output voltage swing can not exceed the total power supply voltage and is typically around two volts or more less than the total power supply voltage. Since the rms value of a sine wave is the peak-peak value divided by the square root of 8 and power is the square of rms voltage divided by resistance we can write

$$Po_{max} = V_P^2 / (8 * R_L) \qquad \text{Eq. 1}$$

where
Po_{max} is the maximum possible output power in watts
V_P is the total power supply voltage minus amplifier limit voltages
R_L is the load resistance in ohms

If there is only a single power supply voltage the V_P is equal to that. Often there is a bipolar power supply that eliminates the need for an output coupling capacitor to the load. V_P in that case is the total difference in potential. Equation 1 is the analytical expression to determine the maximum output power given the existing supply voltage and the load resistance. Equation 1 can be inverted to become the design equation to determine the required supply voltage to produce a given amount of power to a given load.

Q4 is a common-collector amplifier that drives the common collector output stage, Q5 and Q6. The diodes, D1 and D2, roughly compensate for the base-emitter voltages of Q5 and Q6 and must be thermally coupled to the output transistors so that all junction temperatures are similar. R7 is the load resistor for Q4 and provides base current to Q5.

R8 is a critical resistor required to prevent thermal run-away of Q5 and Q6. Since the forward voltage across silicon semiconductor junctions has a negative temperature coefficient with temperature then the current through Q5 and Q6 tends to increase with temperature. The higher current causes both transistors to dissipate higher power thus leading to even more current, etc. Unless limited, this process will quickly lead to excessive current and likely destruction. As current increases the voltage drop across R8 increases which reduces the base-emitter voltages across Q5 and Q6 thus preventing the collector currents from becoming excessive. Assuming the diodes and transistors are properly thermally coupled the junction temperature of the transistors will always exceed that of the diodes by some amount. The resistance of R8 has to be determined to provide proper compensation for that temperature difference. At this point in the design it is impossible to calculate a good value for the resistance of R8 as there are too many unknowns concerning power dissipation, maximum operating temperature, and the ability of the heat sinks to remove heat from Q5 and Q6. An alternative to calculation is to choose a practical value. Later, that value can be checked to see if it works. The resistance of R8 must be large enough to do the job but not so large as to cause significant voltage division to the load resistance. This means that the maximum value of R8 for acceptable voltage division is roughly one-tenth of the load resistance. The smallest value for any useful effect is roughly one-thirtieth of the load resistance. It is better to start at the maximum and work down as practical after the amplifier is operational. This reduces the likelihood of destroying Q5 or Q6 in testing. For the example design, R8 is chosen to be 1 ohm – a conservative value.

The resistance of R7 has to be small enough to provide the peak base current to Q5. But the resistance of R7 can not be so small so that the loading on Q4 is excessive. What this means is that there must be some minimum specified voltage drop across R7 when the amplifier output voltage is at the positive peak. Choosing that voltage large reduces the maximum possible output power as the loaded output voltage swing is limited. Choosing that voltage small increases power dissipation in Q4 and causes the operating currents of Q1, Q2, and Q3 to be higher. There is no simple answer. Design experimentation is the only way to understand all of the ramifications. A practical value that works for a wide range of systems is to choose a voltage that is around one-tenth of the total power supply voltage. This is a good starting point and can be improved as more becomes known about the amplifier. This is another way of saying that design is an iterative process. For the example design the minimum voltage across R7 is chosen to be 1 volt.

Now it is possible to determine the maximum peak-peak output voltage of the amplifier. The positive peak voltage is the power supply voltage minus the specified minimum voltage across R7 minus the base-emitter voltage of Q5. The peak low voltage is roughly the sum of the base-emitter voltages of Q4 and Q5. For the design example the power supply is 12 volts and we chose 1 volt for the minimum voltage across R7 and we will

use a nominal 0.65 volts for the base-emitter voltages. This gives 12 – 1 – 0.65 – 0.65 – 0.65 = 9.05 volts peak-peak. But, we need to reduce this by the voltage division resulting from R8. Since R8 is only in the circuit for only the positive half of the waveform we will use only half of its resistance for the voltage division calculation to the load. The result is a maximum output voltage of 8.5 volts peak-peak, or 3 volts rms. This results in a maximum output power of 1.1 watts. Clipping will occur at higher output.

The peak output voltage across the speaker is half the peak-peak voltage just calculated or 4.25 volts for the example design. The peak emitter current of Q5 is thus 4.25 / 8 or 0.53 amperes. The beta of power transistors is usually lower than that of small signal transistors and is lower still when the collector base voltage is not very large as in this peak situation. For the example design the minimum beta of Q5 will be taken to be 40. Thus, the peak base current is 0.53 / 41 = 0.013 amperes. The resistance of R7 is then calculated to be 1 / 0.013 = 77 ohms. This is rounded to a standard value of 75 ohms.

The optimum DC bias voltage for the output is that which centers the signal swing between the maximum and minimum values. For the example design we calculate this as [(12 – 1 – 0.65) + (0.65 + 0.65)] / 2 = 5.83 volts. This will generally be close to but not exactly midway point of the total supply voltage.

Under quiescent conditions the voltage across R7 is the supply voltage minus the sum of the quiescent output voltage plus a base-emitter drop. For the example design this is 12 – (5.83 + 0.65) = 5.5 volts. Thus, the quiescent emitter current of Q4 is 5.5 / 75 = 0.073 amperes. For this level of current is it better to choose Q4 to be a power type transistor with a minimum beta of 40. The base current of Q4 will then be 0.073 / 41 = 0.00178 amperes.

Stage 4 design

The input resistance to the base terminal of either Q5 or Q6 (only one is active at any time) is the load resistance multiplied by beta plus one or 8 * 41 = 328 ohms. The dynamic resistance of the emitter is not known as it varies from negligibly small at high currents to large at low currents. For simplicity it was taken to be zero. The dynamic resistance at the emitter of Q4 is then the 328 ohms in parallel with R7 for a net resistance of 61 ohms. The dynamic emitter resistance of Q4 is 0.026 / 0.073 amperes = 0.35 ohms. The dynamic input resistance of the base of Q4 is then 41 * 61.35 = 2,515 ohms.

The resistance for R6 is a bit indeterminate. The resistance of R6 should be high enough so that it does not take too much collector current from Q3 but low enough so that the impedance to the base of Q4 is not too large. Thus R6 can be selected to be roughly three times the dynamic input resistance of Q4 to perhaps over ten times that resistance with little change in the operation of the amplifier. For convenience in the example design it will be chosen to be nominally ten times the base input resistance or 25,150 ohms and rounded down to a standard value of 24,000 ohms.

Stage 3 design

The quiescent collector voltage of Q3 (a small signal audio transistor such as the 2N3904) is the output bias voltage minus two base-emitter drops or 5.83 – 0.65 – 0.65 = 4.53 volts. The collector current of Q3 is the sum of the current through R6 and the base current of Q4. For the example design this is (12 – 4.53) / 24,000 + 0.00178 = 0.00211 amperes. The base current is 14.1 uA. The dynamic emitter resistance is 12.3 ohms and the dynamic input resistance of the base is 1,850 ohms.

Stage 2 design

There are a variety of approaches to the design of the direct coupled amplifier. The details will be the topic of another paper. For ease of design and also an approach that generally works well is to let the collector currents of Q1, Q2, and Q3 be nominally the same. Thus, the target collector current of Q2 is 0.00211 amperes. The resistance of R5 is then (12 – 0.65) / (0.00211 + 0.0000141) = 5,343 ohms which is rounded down to 5,100 ohms. The actual collector current of Q2 is then (12 - 0.65) / 5,100 – 0.0000141 = 0.0022 amperes. The base current is 14.7 uA. The dynamic emitter resistance is 11.7 ohms and the dynamic input resistance of the base is 1,764 ohms.

Stage 1 design

The target collector current for Q1 will be nominally 0.00211 amperes. The calculations for R4 are similar to that for R5 and result in R4 being chosen to be the standard value of 5,100 ohms. The base current is 14.7 uA. The dynamic emitter resistance is 11.7 ohms and the dynamic input resistance of the base is 1,764 ohms.

The resistance of R2 should be such that the current through it is substantially more than the nominal base current into Q1 in order to reduce the sensitivity of the bias to variations in beta. "Substantially more" is an indefinite term that can be interpreted to be nominally ten times. Thus, R2 = 0.65 / (10 * 14.7 uA) = 4,400 which is rounded to the standard value of 4,300 ohms.

The value of R3 is calculated to deliver exactly the sum of the current through R2 and the base current when the output voltage is at the calculated optimum quiescent level of 5.83 volts. R3 then calculates to be 31,200 ohms which is rounded to the standard value of 30,000 ohms.

Selection of R1 for gain

Because of the shunt feedback to the base of Q1 the impedance at that node is practically zero ohms. Thus the input resistance to the amplifier is the resistance of R1. For the

example design R1 was chosen to be 10,000 ohms. Because the signal current through R3 is the same as that through R1 the closed-loop voltage gain of the amplifier is -3. This results in a power gain of 11,250 or 40.5 dB.

Calculation of small-signal open-loop gain

For the small-signal common-emitter stages the term B/(B+1) occurs and is 0.9934 for a beta of 150.

Stage 1 voltage gain = 0.9934 * 5,100 / 11.7 = **433**

Voltage division, stage 1 to stage 2 = 1,764 / (5,100 + 1,764) = **0.257**

Stage 2 voltage gain = 0.9934 * 5,100 / 11.7 = **433**

Voltage division, stage 2 to stage 3 = 1,850 / (5,100 + 1,850) = **0.266**

Stage 3 voltage gain = 0.9934 * 24,000 / 12.3 = **1,926**

Voltage division, stage 3 to stage 4 = 2,515 / (24,000 + 2,515) = **0.095**

The voltage gain of stages 4 and 5 is slightly less than 1 and will be treated as 1. The product of the above terms is **2,345,000** – a very high gain!

Calculation of capacitor values

The coupling capacitors, C1 and C2, determine the low cutoff frequency of the amplifier. It is detrimental to the overall performance of the amplifier to design the low cutoff frequency to be significantly below the low cutoff frequency of the speaker. An excessively low amplifier cut-off frequency leads to increased distortion at high amplitudes because the amplifier must handle large voltage signals and higher drive currents that the speaker can not reproduce. It is better to design the amplifier low cutoff frequency to be somewhat below the low cutoff frequency of the speaker. That is easily achieved by determining C1 and C2 for a low cutoff frequency half that of the speaker. The cascade combination provides a net low cutoff frequency of nominally 0.78 the speaker low cutoff frequency. Given the speaker low cutoff frequency, FCL, then

C1 = 1 / (6.28 * FCL * R1) Eq. 2

C2 = 1 / (6.28 * RCL * RL) Eq. 3

For the example amplifier driving a small speaker with a low cutoff frequency of 100 Hz then C1 = 1 / (6.28 * 100 * 10000) = 159 nF which is rounded to the standard value of

150 nF. The value of C2 is 1 / (6.28 * 100 * 8) = 199 uF which is rounded to the standard value of 220 uF.

The purpose of C3 is to stabilize the supply voltage as the amplifier current varies with the signal waveform at high output amplitudes. The impedance of the 12 volt power supply is ideally zero but in reality may be an ohm or more. Power supply variations can couple back to the early stages via R4 and R5 and can degrade the performance of the amplifier with increased distortion or even oscillation. It is hard to calculate a value for C3 as there are numerous unknown variables. If the power supply impedance over the entire audio spectrum is very low (a small fraction of an ohm) then C3 may not be needed at all. If the power supply impedance is significant or increases with frequency over the audio spectrum then C3 may need to be hundreds or even thousands of microfarads to provide the required stability. Testing is the only way to reveal what is really needed. Unless it is known that the power supply impedance is low then plan on C3 being at least equal in capacitance to C2. Be prepared to make it significantly larger.

High Frequency Stabilization

There is a tendency for common-collector amplifiers to oscillate when the load is reactive. This problem typically occurs at some high frequency (100 kHz to over 1 MHz) for audio amplifiers depending on the load reactance. The solution is to provide a low resistance to be the dominate load at high frequencies. This is accomplished by the network, R9 and C4. R9 is typically in the range from around 5 ohms to 20 ohms and C4 is typically around 100 nF give or take. It is rare that enough information would be available to calculate these values. Fortunately, the correct range is broad and rough approximations generally work well. Note that the oscillation that might occur is from the output stage only. It is not a loop oscillation through the other stages. It is important that the resistance not be too small for its purpose is to be the damping mechanism. A capacitor only would increase the likelihood of oscillation.

Frequency Compensation

Multistage high gain amplifiers have accumulative phase shifts that always exceed 180 degrees at some high frequency. This is in addition to natural phase inversions that occur in common-emitter stages. Thus, the negative feedback path for low frequencies becomes positive feedback at some high frequency. If the loop gain is greater than one at that high frequency then the amplifier will oscillate. Interestingly and counter-intuitively at first, such an amplifier is more prone to loop oscillation when then closed-loop gain is low rather than high. The simple explanation is that closed-loop gain and loop gain are inversely related. A low closed-loop gain means that the feedback is little attenuated thus producing a high loop gain. The solution to this problem is to add a dominant pole to one of the inner stages such that the open-loop amplifier gain at any frequency where the phase shift is 180 degrees or more is less than one. It is very hard to predict in advance that this problem will occur. Testing is the only way to verify that the amplifier does or

does not have the problem. It is possible that some otherwise identical amplifiers will have the problem while others do not. An easy method to add a pole is to place a small feedback capacitance between the base and collector of Q2 as shown in Figure 1. The Miller effect for inverting amplifiers multiplies the effective capacitance by the stage gain. Thus, a tiny experimentally determined capacitance of perhaps 10 pF, give or take, can make the amplifier stable for low closed-loop gains. Ideally, this compensation may not be needed, but be prepared. Do not overcompensate the amplifier as that reduces the open loop gain in the upper portion of the audio band leading to increased distortion. The ideal compensation is as little as required.

Physical construction

The performance of a well designed amplifier can be seriously degraded by poor construction techniques. The physical layout of the circuit must be such that all internal loop areas are as small as possible. The power supply wires must be a tight transmission line and well bypassed to keep the impedance low throughout the audio spectrum and beyond. The speaker return as well as the collector of Q6 should go directly to the power supply source and not share any ground path to the rest of the amplifier. The reason for that is signal voltage drops in the ground can form parasitic feedback paths to earlier stages resulting in degraded amplifier performance including oscillation. The physical layout should minimize capacitive coupling between stages particularly across multiple stages. Ad hoc random haywire construction is bad. A carefully designed tight layout is key to proper operation.

Heat sink

The output transistors must dissipate heat as the speaker is driven. The power dissipated is purely from the laws of physics – voltage across the transistor multiplied by the current through it. The heat is not a result of "imperfections" in the transistors. The question is how much heat will be dissipated? For a sine wave output it can be shown (students should do this – it is not hard – only simple calculus) that the maximum power dissipation occurs when the peak-peak voltage output is 63.7 percent of the total power supply voltage. The amplifier is 50 percent efficient at this level so the dissipation is

$$P_{max} = 0.41 * V_P^2 / RL \qquad \text{Eq. 4}$$

This is the total power dissipation. The maximum power dissipated by Q5 or Q6 alone is half this value. It is interesting to observe that the amplifier runs cooler for output levels higher than this. Remember that this calculation is for a sine wave. An audio waveform is significantly different. However, Equation 4 is a good representation of the worst case power dissipation. If the thermal design is conservative then there is margin to accommodate the unknown error.

The junction temperature of Q5 or Q6 should never exceed 125 C or destruction will occur. Conservative thermal design practice is to target a maximum junction temperature of no more than 100 C. Lower is always better but may require an impractical heat sink. The data sheet for the power transistors will provide the thermal resistance from the junction to the case surface. There is also a small thermal resistance (typically in the 0.5 to 1 C/W range) in the mounting process of the transistor to a heat sink. The heat sink is typically rated in temperature rise per power dissipated in free air (i.e. unrestricted air flow, not inside an enclosure that limits heat exchange). It should be observed that the heat sink specification is the thermal resistance to ambient.

$$T_J = P_D *(\theta_{JC} + \theta_{CS} + \theta_{SA}) + T_A \qquad \text{Eq. 5}$$

where
T_J is the maximum desired temperature of the junction in C
P_D is the power dissipation in watts – in this case the maximum from Eq 5
θ_{JC} is the thermal resistance from the junction to the case in C/W
θ_{CS} is the thermal resistance from the case to the heat sink in C/W
θ_{SA} is the thermal resistance from the heat sink to ambient in C/W
T_A is the ambient temperature in C – this would be the maximum for the intended environment

The only parameter in Equation 6 that the engineer has control over is θ_{SA}. That term must be chosen to be small enough (i.e. the heat sink chosen large enough) so that the junction does not exceed the maximum desired temperature. It is not unusual for the calculation to produce a negative result – i.e. impossible – then some alternative approach needs to be determined. A thermal resistance of zero applies to a theoretical infinite heat sink. Real heat sinks will have positive values of thermal resistance. For design, Equation 5 is solved for the required thermal resistance from the heat sink to ambient and then a heat sink is selected that has a thermal resistance somewhat smaller.

Power supply requirements

The DC power supply must supply the power that is sent to the speaker and also the power dissipated in the amplifier electronics. The maximum DC power occurs at maximum output power and approaches 1.3 times the maximum output power. The actual value depends on how close the output stage can produce a peak-peak output voltage approaching the total power supply voltage.

Improvements

The amplifier circuit used for this example is intentionally simplistic for ease of analysis and design by students. There is room for improvements in a more professional design. Some sort of voltage stabilization system, either a hefty low-pass filter or voltage regulator could be added to the high side of R4 and R5 so that any supply fluctuations

resulting from dynamic output currents driving the speaker do not feedback to these early stages. Another improvement would be to design a specific upper cutoff frequency so that any high frequency signal beyond that which the speaker can reproduce is not amplified and fed to the speaker. That only results in increased heat for no useful purpose and can increase distortion. The output quiescent voltage is somewhat temperature dependent and more advanced circuitry could eliminate that. Instead of connecting directly to VCC, R7 might connect to the unregulated higher input voltage to the regulator that supplies VCC. This enables a higher output signal swing with better linearity and reduces the required operating current for Q4 as R7 can be a higher resistance.