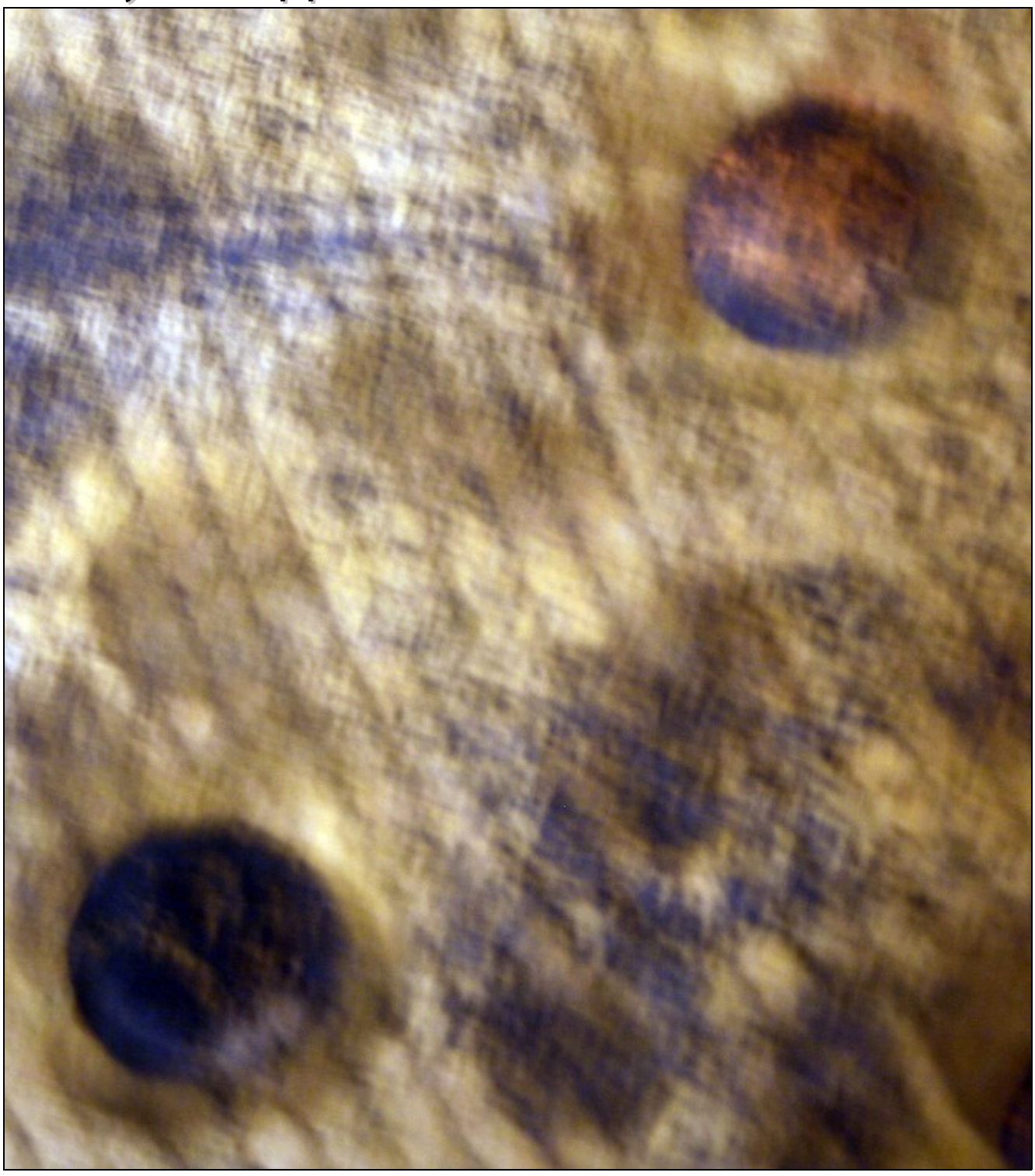
Semiconductor Devices:

Theory and Application. 2E



James M. Fiore

Semiconductor Devices:
Theory and Application. 2E

by

James M. Fiore

Version 2.0.8, 24 October 2024

Published by James M. Fiore via dissidents

ISBN13: 978-1796543537

For more information or feedback, contact:
James Fiore, Professor Emeritus

oer@jimfiore.org
jfiore@mvcc.edu

For the latest revisions, related titles, and links to low cost print versions, go to:
www.mvcc.edu/jfiore or my mirror sites www.dissidents.com and www.jimfiore.org

YouTube Channel: Electronics with Professor Fiore

Cover art, *Ride*, by the author

Preface

Welcome to the second edition of *Semiconductor Devices*, an open educational resource (OER). The goal of this text, as its name implies, is to allow the reader to become proficient in the analysis and design of circuits utilizing discrete semiconductor devices. It progresses from basic diodes through bipolar and field effect transistors. The text is intended for use in a first or second year course on semiconductors at the Associate or Baccalaureate level. In order to make effective use of this text, students should have already taken coursework in basic DC and AC circuits, and have a solid background in algebra and trigonometry along with exposure to phasors. Calculus is used in certain sections of the text but for the most part it is used for equation derivations and proofs, and is kept to a minimum. For students without a calculus background these sections may be skipped without a loss of continuity.

An OER companion laboratory manual is also available. It features nearly 30 exercises that parallel the topics presented in this text. For continued study, a follow-on OER text and lab manual, *Operational Amplifiers and Linear Integrated Circuits*, *Third Edition* is available. Several other electrical OER titles are available as well.

I cannot say enough about the emerging Open Educational Resource movement and I encourage budding authors to consider this route. While there are (generally) no royalties to be found, having complete control over your own work (versus the "work for hire" classification of typical contracts) is not to be undervalued. Neither should contributions to the profession nor the opportunity to work with colleagues be dismissed. Given the practical aspects of the society in which we live, I am not suggesting that people "work for free". Rather, because part of the mission of institutes of higher learning is to promote and disseminate formalized instruction and information, it is incumbent on those institutions to support their faculty in said quest, whether that be in the form of sabbaticals, release time, stipends or the like. It is also my opinion that no one should be deprived of a higher education due to lack of funds. A society that truly cares for its citizens would institute free college tuition.

If you have any questions regarding the text or lab manual, or are interested in contributing to the project, do not hesitate to contact me. Finally, please be aware that the most recent versions of all of my OER texts and manuals may be found at my MVCC web site as well as my mirror site: www.dissidents.com

Acknowledgements

For their continued support, my family and friends. For their input on draft versions, my colleagues (esp. Ben Glallard, Hill Bunt, Deb Bocker, Hal Paulko, Thom Timas) and students. For *on-going sanity reassurance*, Bernie Sanders and Mike Lofgren. For *unintentional diversionary humor elicitation*, Wally and pals.

Finally, a serious *thank you, thank you, thank you* to all of the people who have created useful software and other OER tools without which this project would have been impossible. When we each do a little, we all gain a lot. This text was created using several free and open software applications including Open Office and Dia. Some graphs and device curves were created using SciDAVis. Screen imagery was often manipulated though the use of XnView. The free TINA-TI circuit simulator was used in Chapter 17, new for the second edition.

- Jim Fiore, 2017/2022

For Karen

Does humor belong in academia?

Table of Contents

1 Semiconductor Fundamentals

1.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Define the term *semiconductor*.
- Describe the differences between conductors, semiconductors and insulators in terms of atomic energy levels.
- Describe the atomic structure of mono-crystalline silicon.
- Detail the effect of *doping* on a silicon crystal.
- Describe the differences between *P material* and *N material*.
- Draw the energy level diagrams for P- and N-type materials.

1.1 Introduction

First, a Little History

Just as the late eighteenth through nineteenth centuries are known as the industrial age due to the rise of mechanization, the twentieth century can be referred to as the beginning of the electronic age. The first half of the century was dominated by electronic vacuum tubes that made possible devices such as radio, television, radar and long distance telephone. The technology of the vacuum tube was displaced mid-century by the introduction of *solid-state semiconductors*. The first working prototype transistor was invented at Bell Labs in 1947 by John Bardeen, Walter Brattain and William Shockley. This device, properly referred to as a *point contact transistor*, was quickly superseded by the *bipolar junction transistor*, a major topic of this text.

Commercial production of the transistor and related devices improved the performance of existing applications and made possible a range of new ones. Semiconductors proved to be smaller, lighter, more reliable and less expensive to build than their vacuum tube counterparts. The last 30 or so years of the century saw the rapid expansion of the *integrated circuit* where numerous transistors are combined in a single device. Initially such a device may have contained the equivalent of a dozen or so individual semiconductor devices, but today that number has grown to the billions[1]. This extreme density has given rise to now common-place applications such as cell phones, GPS devices, laptop computers, tablets and our global communications infrastructure.

1 It is worth noting that the construction of an integrated circuit does *not* involve the creation and interconnection of millions or billions of single discrete transistors. Instead, the manufacturing process builds all of the transistors simultaneously, rather like a layer cake.

The science writer, Arthur C. Clarke, once observed that "Any sufficiently advanced technology is indistinguishable from magic". Indeed, although today the typical citizen living in an industrialized country makes use of numerous electronic devices each day (sometimes without even being aware of it), they typically have scant knowledge of how these devices "work their magic". Obviously there is no magic, only the application of scientific principles mixed with human ingenuity. Further, just as it is true that many more people can use a cell phone than design one, it is also true that there is a greater need for people who can design, manufacture and maintain devices based on semiconductors than for people who design the semiconductors themselves. The scope of this text, then, focuses on the operation and application of semiconductor devices rather than the design of the semiconductors themselves.

Variable Naming Convention

One item that often confuses beginning students of almost any subject is nomenclature. Before we begin our discussion of semiconductor devices it is important that we decide upon a consistent naming convention. Throughout this text we will be examining numerous circuits containing several passive and active components. We will be interested in a variety of parameters and signals. In order to keep confusion to a minimum we will use the following conventions in our equations for naming devices and signals.

R	Resistor (DC, or actual circuit component)
r	Resistor (AC equivalent, where phase is 0 or ignored)
C	Capacitor
L	Inductor
Q	Transistor (Bipolar or FET)
D	Diode
V	Voltage (DC)
v	Voltage (AC)
I	Current (DC)
i	Current (AC)

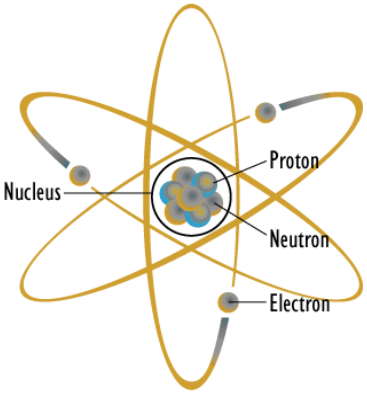
Resistors, capacitors and inductors are differentiated via a subscript that usually refers to the active device to which it is connected. For example, R_E is a DC bias resistor connected to the emitter of a transistor while r_C refers to the AC equivalent resistance seen at a transistor's collector. C_E refers to a capacitor connected to a transistor's emitter lead. Note that the device related subscripts are always shown in upper case, with one exception: If the resistance or capacitance is part of the device model, the subscript will be shown in lower case to distinguish it from the external circuit components. For example, the AC dynamic resistance of a diode would be called r_d. If no active devices are present or if several items exist in the circuit, a simple numbering scheme is used, such as R_1. In very complex circuits a specific name will be given to particularly important components, as in R_{source}.

Voltages are normally given a two-letter subscript indicating the nodes at which it is measured. V_{XY} is the DC potential from node X to node Y while v_{XY} indicates the AC signal appearing across node X to node Y. A single-letter subscript, as in V_X, indicates a potential relative to ground (in this case from node X to ground). The exceptions to this rule are power supplies, that are given a double letter subscript indicating the connection point (V_{CC} is the collector power supply), and particularly important potentials that are directly named, as in v_{in} (AC input voltage) and V_{R2} (DC voltage appearing across R_2). If an equation for a specific potential is valid for both the AC and DC equivalent circuits, the uppercase form is preferred (this makes things more consistent with circuits that are directly coupled, and thus can amplify both AC and DC signals). Currents are named in a similar way but generally use a single subscript referring to the measurement node (I_X is the DC current flowing through a conductor into or out of node X). All other items are directly named. By using this scheme, you will always be able to determine whether the item expressed in an equation is a DC or AC equivalent, its approximate circuit location, and other factors about it.

1.2 Atomic Structure

In our effort to understand the operation of semiconductors, a fundamental question we might ask is "What is the internal structure of an atom?" Please understand that it is nonsensical to ask what an atom might "look like" because its components are all smaller than the shortest wavelengths of light that humans can see. Instead, we simply need a model to explain its observed behavior.

Figure 1.1
Planetary atomic model: Pretty, well-known and wrong.
Image source (modified)



Perhaps the most prolific model in the popular imagination is the *planetary model* shown in Figure 1.1. In this model, the core, or *nucleus*, is drawn at the center and contains positively charged protons and non-charged neutrons. Revolving around this core are negatively charged electrons, each following a nice, regular, planar path much like a planet around the sun. Unfortunately for us, this model is starkly incorrect, although it has found use as a symbol for nuclear regulatory agencies and a DEVO album cover from the 1970s.

Before we come up with a more accurate and useful model, let's take a closer look at the sub-components; namely the proton, neutron and electron. First off, most of the mass of any given atom is from the protons and neutrons. Protons and neutrons have similar masses, about 1.67E−24 grams each. The mass of an electron is roughly 2000 times smaller. The radius of a proton is approximately 0.87E−15 meters and the mean distance to the nearest electron is about 5.3E−11 meters. This means that this electron is about 60,000 times farther away from the proton than the size of said proton. To put this into perspective, that's roughly the same as the ratio between a golf ball and a sphere with a radius of 3/4ths of a mile or 1200 meters. This would be the case for a hydrogen atom as it consists of a single proton and electron. The magnitude of this ratio is not much different for other substances, including things

like crystalline carbon (diamond) and quartz (a molecule of silicon and oxygen) that are very hard and solid. If you think about that for a moment, you realize that the idea of "solidity" is in some ways an illusion because the vast majority of what we call "something" is really just empty space. For example, chances are that you are sitting down while reading this. You probably feel your buttocks pressed against the chair. Both of these things are considered solid yet at the atomic level the vast majority of both items is nothingness. In reality, the feeling of solidity is just the result of the interaction of atomic forces between the two. So if someone suggests that you might have a bit too much to spare in the *department of the posterior*, you can inform them that it's really nothing.

One of the major issues with the planetary model is the idea that electrons whirl around the nucleus in stable, planet-like orbits. That's simply not true. First, the electron inhabits a region of 3D space, it does not simply move through a plane. Second, due to the Heisenberg Uncertainty Principle, we can't precisely plot the position and trajectory of a given electron. The best we can do is make a plot of where the electron is *likely to be*. This is called a *probability contour*. Imagine that you could record the position of an electron relative to the nucleus. A moment later you record its new position, a moment after that you record the next position, and on and on for thousands of measurements. If you attempted to plot them all, you would wind up with a cloud of dots around the nucleus. This cloud is referred to as an *orbital*. You wouldn't know how the electron got from one position to the next but you would get a general idea of where it was likely to be. Do not confuse *orbital* with *orbit* (like a planetary orbit). They are two different beasties.

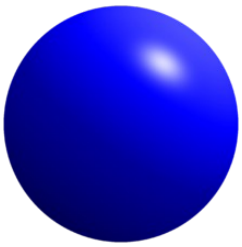
There are several potential orbitals. Due to quantum physics, only certain orbitals are allowed. The permissible electron energy levels are first grouped into *shells*, then *subshells* and finally *orbitals*. It is important to remember that **orbitals indicate the electron energy level**. That is, a higher orbital implies a higher energy level. Further, orbitals fill in first from lowest energy level to highest energy level. These are important ideas that we will leverage in future discussions.

Shells are denoted by their *principal quantum number, n*; 1, 2, 3, etc. The higher the number, the more subshells it can contain. *Subshells* are organized by their orbital shape and are designated by letters, the first four being *s*, *p*, *d*, and *f*. Shell *1* contains only subshell *s* while shell *2* contains subshell types *s* and *p*. Shell *3* contains subshell types *s*, *p* and *d*, and so on.

Thus, we see designations such as *1s*, *2s* and *2p*. These subshells may also have variations within them. There is one variation on *s*, three variations on *p*, five variations on *d*, etc. **These variations are the orbitals and each orbital can hold a maximum of two electrons.**

Putting this all together, we find that the first shell can contain a maximum of two electrons: two in the single *s* subshell orbital (*1s*). The second shell can contain a maximum of eight electrons: two in the *s* subshell (*2s*) plus two in each of the three

Figure 1.2
Electron probability contour for innermost orbital, 1s.
Image source

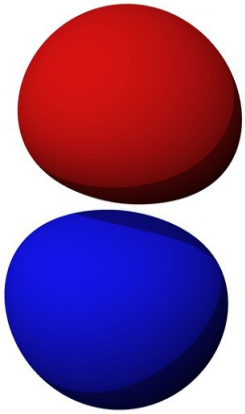


p subshell orbitals (*2p*). In like manner the third shell can contain a maximum of 18 electrons: two in *3s*, six in *3p* and two in each of the the five *d* subshell orbitals (*3d*). You can condense this into a simple formula, $2n^2$, where *n* is the shell number.

Figure 1.2 shows the electron probability contour of the innermost orbital, namely *1s* (i.e., principle quantum number *1*, subshell *s*). As you can see, it is spherical in shape. The nucleus is located at the center, obscured here. All *s* orbitals are similarly spherically shaped although the internals change. *1s* is the lowest energy orbital.

Orbitals are not limited to simple spherical shapes. Higher order orbitals can take on a variety of forms. Figure 1.3 shows the electron probability contour for the *2p* orbitals (recall there are three *p* variations, one each oriented along the X, Y and Z axes). The nucleus is situated in the small void between the two lobes. Obviously, this is nothing like the well-behaved elliptical orbits of planets around the sun. Probability contours can be very complex. For the highest orbitals, especially when combined with the lower orbitals, the contour combinations can become reminiscent of the sculptures of a deranged clown forming herds of imaginary balloon animals.

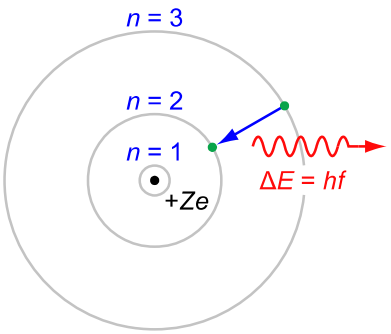
Figure 1.3
Electron probability contour for orbital 2p.
Image source



As interesting as these graphics are, they are cumbersome to work with. Consequently, a more functional graphic is called for. Such a device is the *Bohr model*, named after Danish physicist Niels Bohr. An example is shown in Figure 1.4.

It is important to understand that the Bohr model is an energy description of the atom, not an attempt to mimic its physical appearance or structure. The nucleus is placed at the center. It is surrounded by concentric rings that represent the electron shells. The higher the number, the larger the ring and the greater the energy level. If an electron were to move from a higher level to a lower level, the energy difference is radiated out. This could be in the form of heat or light. This is a point worth remembering. For example, this transition is what makes light emitting diodes (LEDs) function. The inverse is also possible, namely that by absorbing energy, an electron can move into a higher orbital. This is an equally powerful concept, as we shall soon see.

Figure 1.4
Generic Bohr model.
Image source



Using the Bohr model we can create diagrams to represent individual elements. For example, copper has an atomic number of 29 meaning that it has 29 protons and 29 electrons. The electron shell configuration is 2-8-18-1. That is, the first three shells are completely filled and there is a single electron in the fourth shell. This single outer electron is only loosely bound and thus makes copper a very good conductor. The Bohr model for copper would simply show four rings, the first three being filled and with a single electron in the fourth ring.

Figure 1.5 shows the Bohr model of an atom of Silicon, atomic number 14, with an electron shell configuration of 2-8-4. In this version, the individual electrons are drawn in each shell and the atomic number is indicated at the nucleus. Again, please do **not** imagine this representing individual electrons orbiting the nucleus in lanes. This is an energy level depiction.

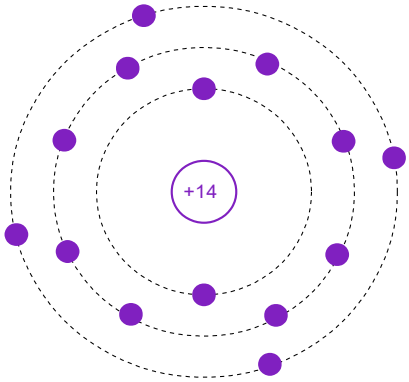


Figure 1.5
Bohr model of Silicon.

Often, it is useful to simplify this model further by omitting the filled inner shells. Also, the atomic number is replaced by the number of electrons in the outermost, or *valence*, shell. This is shown in Figure 1.6. The valence shell is particularly important as it gives insight into the general behavior of the material.

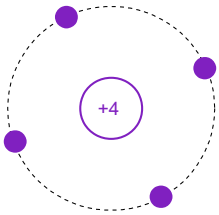


Figure 1.6
Simplified Bohr model of Silicon.

As an alternative, sometimes we will "straighten out" the Bohr model so that it simply shows the energy levels graphically as lines or bands, and without counting specific electrons. This is depicted in Figure 1.7.

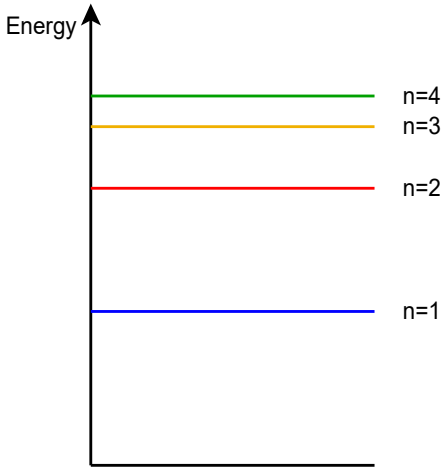


Figure 1.7
Energy level diagram.

1.3 Crystals

We used silicon in the preceding example on purpose. The fact that it has a half-filled valence shell with four electrons puts it in a special place. As is, it's neither a great conductor nor a superior insulator. With some attention to detail, it will become a *semiconductor*. Silicon is not the only material that can be used for semiconductors. In fact, many of the earliest semiconductors were made from germanium and currently we make semiconductors from other materials. Silicon, however, remains the source of most semiconductors today.

It is possible for pure silicon to be arranged in a mono crystaline structure. That is, all of the silicon atoms align in a very specific, well-ordered manner, without any voids or breaks in the pattern. As silicon has only four electrons in its valence shell, four more electrons would be needed to obtain stability (i.e., eight electrons in the outer shell). In the crystal, any given atom of silicon effectively "shares" an electron from its four closest neighbors through a *covalent* bond (meaning "with or among the valence"). Each atom does this, therefore each atom is tightly bound to its neighbors. This is illustrated in Figure 1.8 using simplified Bohr models. Note the color coding that indicates the sharing.

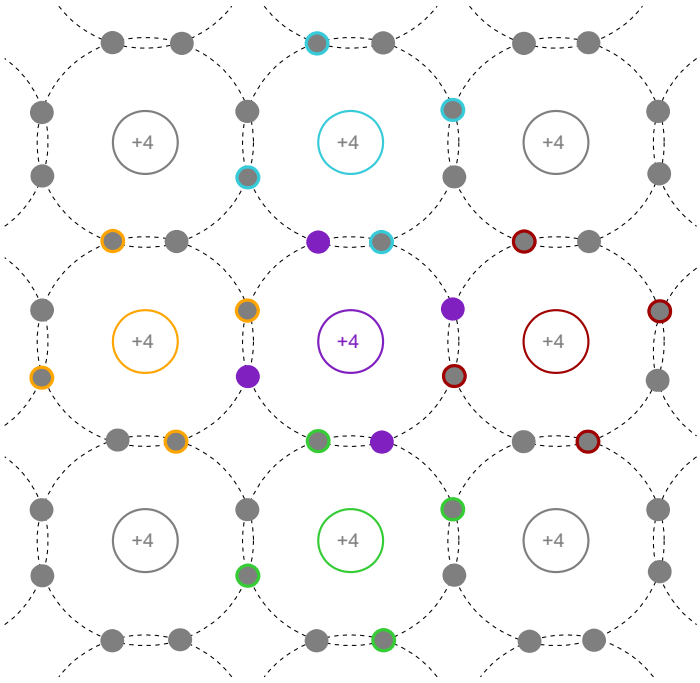


Figure 1.8
Silicon crystal covalent bonding.

Remember, this is an energy diagram. We are not trying to indicate that individual valence electrons are zipping between atoms in figure eight patterns. Indeed, this diagram is drawn flat whereas a real crystal is not a simple sheet, but is three dimensional with varying thickness.

While it would be a practical impossibility to draw a highly realistic representation of atoms in the crystal, given a few liberties we can draw something that at least comes a little closer to reality. We start by representing each silicon atom as a ball and the covalent bond as a connecting tube. Recalling that each atom must be bound to four others in a regular, equal pattern, we come up with the drawing of Figure 1.9. Notice that the overall structure is essentially that of a cube. Further, at the center of each face of the cube there exists an atom of silicon. Therefore, we say that the crystal structure is *face centered cubic*.

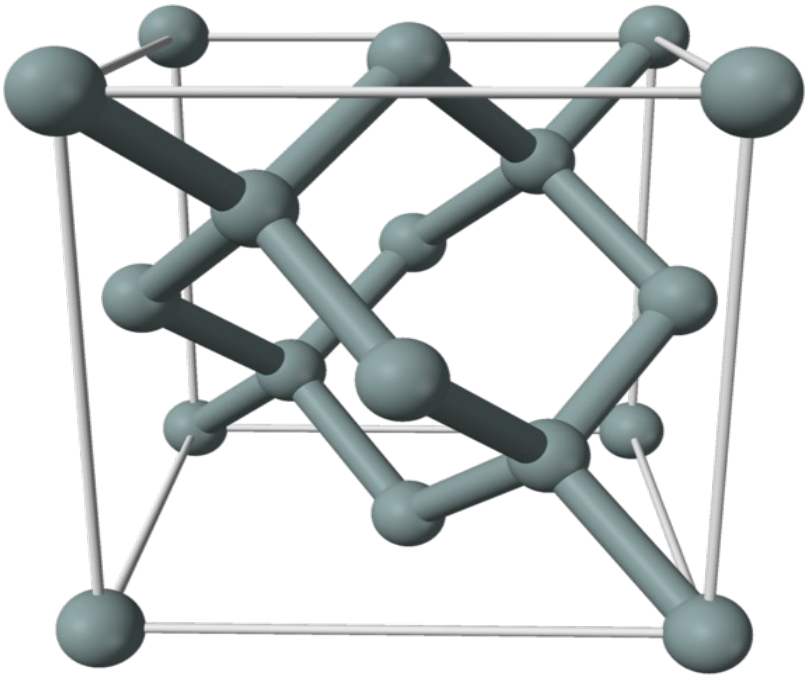


Figure 1.9
Representation of silicon crystal structure.
Image source

An interesting thing happens in the crystal when we examine the energy levels. With a single atom we would expect to see an energy diagram like that of Figure 1.7. That is, discrete, permissible steps. Within a crystal, though, each atom is affected by those around it. This causes slight changes in the energy levels. Taken as a whole, all of these individual variations cause the discrete levels to blur into broader bands. If we were to examine the valence and conduction energy levels, instead of discrete, thin lines we'd see the thicker bands as illustrated in Figure 1.10. These bands still represent permissible electron energy levels, it's just that now there is a continuum rather than a discrete level. There will still be non-permissible or forbidden zones between these regions. A forbidden zone is referred to as a *band gap*.

Associated with this idea is the concept of the *Fermi level*, named after physicist Enrico Fermi. Basically, the Fermi level is the energy level in a given material at which there is a 50% probability that it is filled with electrons. In other words, levels below this value tend to be filled with electrons and levels above tend to be empty. If the Fermi level lies within a band, the material will be good a conductor. On the other hand, if the Fermi level lies between two widely separated bands, the material will be a good insulator. If the Fermi level is between bands that are relatively close, the material is a semiconductor.

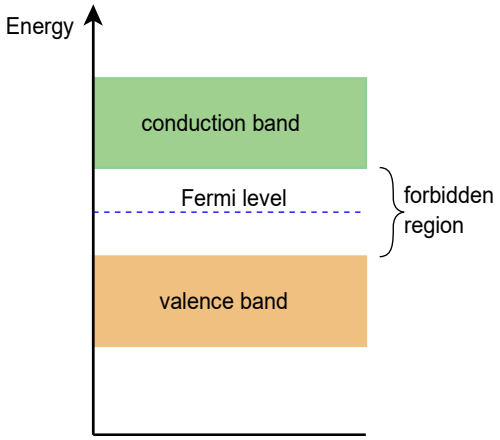


Figure 1.10
Energy diagram for an intrinsic semiconductor.

Figure 1.10 shows the energy bands for an *intrinsic semiconductor*, such as an ideal silicon crystal. The term *intrinsic* simply means that there are no impurities in the crystal. Between the valence band and conduction band is an impermissible or forbidden region. This is a band gap. In practical terms you can think of the band gap as the amount of energy that needs to be applied to an electron in order to move it from the valence band to the conduction band. The value of the band gap will depend on a variety of factors, the precise material being used for the semiconductor is of particular importance.

Without any external energy applied (i.e., isolated and at absolute zero), the crystal lattice is stable and there is no electron movement through the crystal. As we add thermal energy, it is possible for valence electrons to jump up to the *conduction band*. At this point, the electron can "wander" through the crystal in the manner depicted in Figure 1.11.

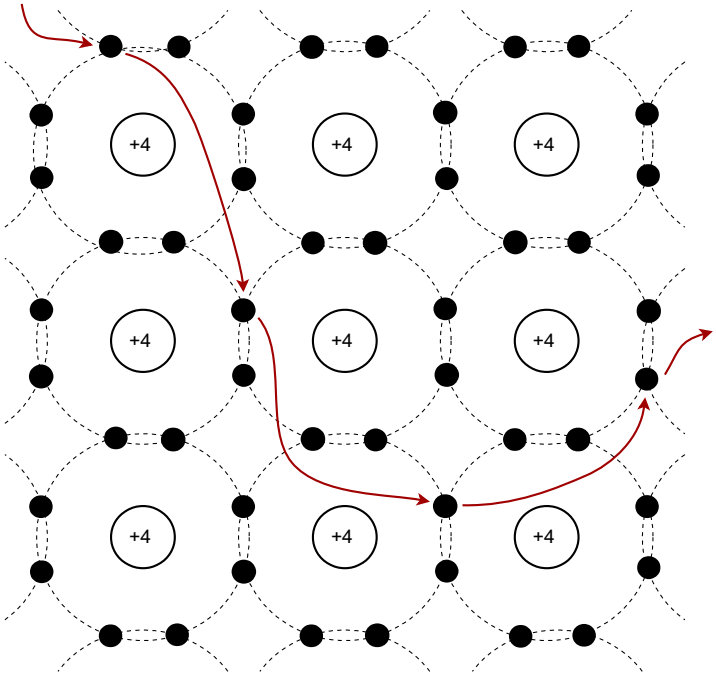
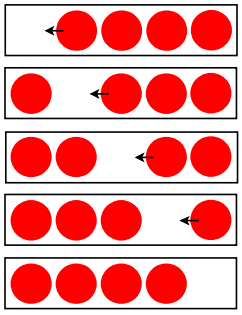


Figure 1.11
Electron movement in a crystal.

Here is how this happens: Because the thermal energy causes the electron to jump to the higher energy level of the conduction band, it leaves behind a "hole", that is, a place devoid of an electron. Now that the hole exists, it provides a place for another electron to "fall into". The higher the temperature, the greater the number of freed electrons and the greater the number of corresponding holes. We now have thermally-induced electron movement. We can also look at this from the opposing perspective, namely that we have an equal magnitude but opposite direction "hole flow". If you find this idea hard to grasp, simply look at Figure 1.12. Each horizontal bar contains four dots representing electrons. In the topmost bar there is an empty space (a hole) to the extreme left. When the leftmost electron moves into this hole it fills it in a process called *electron-hole recombination*, which of course, sounds much more impressive than it really is. The result is the second bar. We repeat this process of moving an electron right to left as we traverse down the diagram. Eventually we end up with the four electrons packed together toward the left. Finally, instead of focusing on the dots, focus instead on the negative space (the empty white bit). Moving from top to bottom, the hole moves left to right, in the opposing direction.

Figure 1.12
Electron versus hole flow.



Just as we think of the movement of electrons as a movement of negative charge, then the movement of holes can be thought of as a movement of positive charge. We can say that the electron is the carrier of negative charge while the hole is the carrier of positive charge.

Before moving on to the next section, it is important to remember that in an intrinsic (pure) semiconductor, the number of thermally produced electrons and holes will be equal. Also, even at room temperature the total number will also be quite small compared to the number of electrons in the crystal.

1.4 Doped Materials

By themselves, intrinsic semiconductors are not of particular use. They are neither good conductors nor insulators, and their conduction is largely dependent on temperature. We can alter the properties of the material by introducing foreign substances or impurities into the crystal. These impurities are also known as *dopants*. A crystal with an added dopant is referred to as an *extrinsic* semiconductor or *doped* material. The amount of impurity added is generally small, perhaps in the neighborhood of one part per million. The dopant may be added through a gaseous diffusion process where the crystal is heated in an oven and the dopant added in gaseous form. Over a period of time the impurities will diffuse or "seep into" the target crystal. An alternate approach is ion implantation. In this method the impurities are accelerated and quite literally smash into the target, dislodging and replacing some of the original atoms in the crystal.

There are two different types of semiconductors possible. One is called *N-type material*, and the other, *P-type material*. Unsurprisingly, the N stands for Negative and the P stands for (you guessed it) Positive. N-type material is created by adding pentavalent impurities, that is, a dopant with five electrons in its outer shell. Examples include phosphorus, arsenic and antimony. In contrast, P-type material is created by adding a trivalent impurity, one with three electrons in its outer shell. Possible trivalent impurities include boron, gallium and indium.

N-Type Material

Figure 1.13 shows a model of a silicon crystal with a pentavalent impurity at its center. Compared to an ordinary silicon atom that would have four electrons in its outer shell, the pentavalent impurity creates an extra, or *donor*, electron. Thus, the crystal has a net negative charge and is referred to as N-type material. The energy level of the donor electrons is just below the bottom of the conduction band. In other words, the difference between the donor level and the bottom of the conduction band is much, much smaller than the band gap itself. Therefore it is relatively easy for these donor electrons to jump into the conduction band, becoming free ionized electrons and leaving behind ionized holes[2].

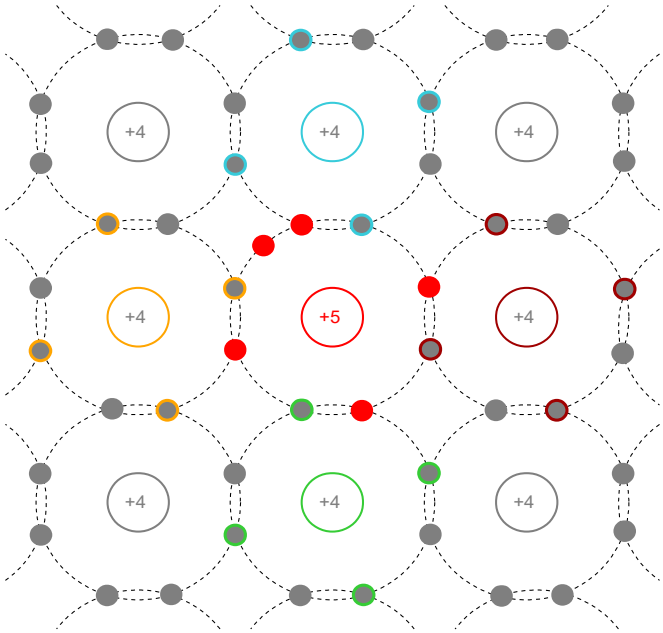


Figure 1.13
Crystal with added pentavalent impurity (N-type).

2 An *ion* is an atom or molecule that does not have a neutral net charge, i.e., the numbers of protons and electrons are not equal. If it loses electrons, resulting in a net positive charge, it is called a *cation.* If it gains electrons resulting in a net negative charge it is called an *anion*.

Compared to the undoped intrinsic crystal, the doped extrinsic crystal exhibits a relatively high number of free electrons. As you might surmise, this enhances the conductivity of the material, and the greater the doping level, the greater the enhancement. Earlier it was mentioned that both electrons and holes can serve as charge carriers. Because the number of free electrons is significantly larger than the number of holes in N-type material, electrons in N-type material are referred to as the *majority charge carrier* (or more simply, the majority carrier) while holes are referred to as the *minority charge carrier* (or minority carrier).

The extra electrons add to the number of filled energy states and, being of higher energy than the valence electrons, push the Fermi level to a higher value. Remember, the Fermi level represents the point where 50% of states would be filled, so if we add states above this, then the new 50% point must be higher than the former level. This is illustrated in Figure 1.14. Note how close the donor level is to the conduction band and that the Fermi level has been pushed up, away from the valence band and closer to the conduction band. This will be of great significance in up-coming discussions on semiconductor devices.

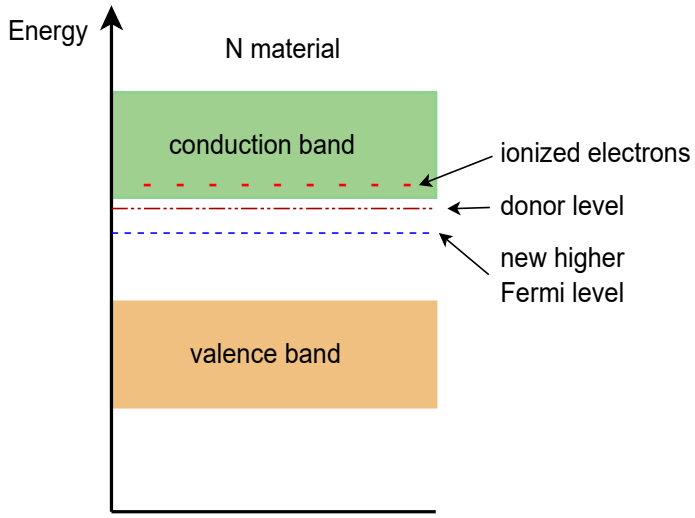


Figure 1.14
Energy band diagram of N-type semiconductor.

P-Type Material

In a similar manner, if we introduce a trivalent impurity, our crystal model now features a hole; a location where an electron is lacking. For this reason, trivalent impurities are sometimes called *acceptors*. The resulting crystal model is illustrated in Figure 1.15.

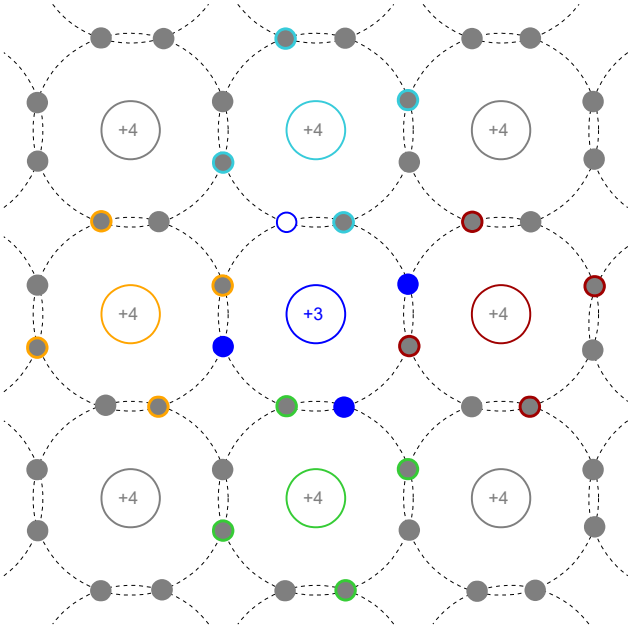


Figure 1.15
Crystal with added trivalent impurity (P-type).

The resulting situation is essentially the reverse of that of the N-type material. Figure 1.16 shows the energy band diagram for our new P-type material. In this case, the Fermi level has been pushed down, closer to the valence band.

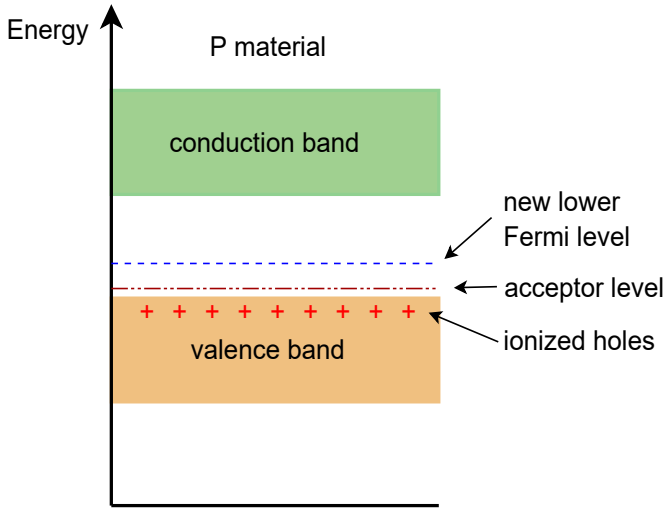


Figure 1.16
Energy band diagram of P-type semiconductor.

In P-type material, holes out number free electrons. Consequently, *holes* are referred to as the *majority carrier in P material* while electrons take on the role of minority charge carrier.

As with N-type material, the greater the amount of trivalent impurity added, the greater the overall effect. By itself, a doped crystal can be used to create a resistor. The resistivity of the material is a function of the doping level. By setting the cross-sectional area, length and doping level, we can create well-defined resistor values. If this was all we could do with semiconductors then we could say two things: first, the solid state semiconductor revolution would not exist; and second, this text would be very short. The interesting bits arrive when we combine both N- and P-type materials into a single device, as we shall begin to see in the next chapter.

Summary

In this chapter we have examined the basic structure of atoms. This includes the concept of electron shells and permissible energy states. We have used both the Bohr model of the atom and the corresponding energy band diagrams.

Crystals such as silicon show a very ordered three dimensional structure that relies on strong covalent bonds. The crystal tends to "fuzz" or broaden the permissible energy levels into thicker energy bands. Further, the crystal exhibits a modest energy gap, or band gap, between the valence band and conduction band. This gap is much smaller than the gap seen in insulators, and therefore the material is referred to as a semiconductor, being somewhere between a true conductor and a true insulator.

The electrical characteristics of a pure, or intrinsic, semiconductor crystal can be altered by adding impurities or dopants. A doped crystal is referred to as an extrinsic crystal. If a pentavalent dopant is added, there will be a surplus of electrons and a raising of the Fermi level. The new crystal is called N-type material. In contrast, if a trivalent dopant is added, there will be a surplus of holes and a lowering of the Fermi level. The new crystal is called P-type material. In N-type material, electrons are the majority charge carrier and holes are the minority charge carrier. In P-type material, holes are the majority carrier while electrons serve as the minority carrier.

Review Questions

1. Describe the differences between a conductor, an insulator and a semiconductor.
2. Define the terms *Fermi level*, *valence band*, *conduction band* and *band gap*.
3. What is the fundamental difference between an *intrinsic crystal* and an *extrinsic crystal*?
4. What is meant by the term *doping*?
5. What is the effect of donor and acceptor impurities on the Fermi level?

The author Kurt Vonnegut once said, "Never index your own book".

One beauty of an electronic text is that you don't have to.

2 PN Junctions and Diodes

2.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Describe and diagram the *energy hill* for a PN junction.
- Discuss the different kinds of diodes available and their uses: rectifier, Zener, LED, photodiode and varactor.
- Detail the device characteristics exhibited by different diode types.
- Graph the forward- and reverse-bias operation regions of diodes.
- Determine the effective resistance of a diode under specific conditions.
- Solve basic DC resistor-diode circuits for various system voltages and currents.

2.1 Introduction

Having investigated the characteristics of extrinsic N-type and P-type materials in the prior chapter, we shall continue by examining what happens when these two materials are combined into a single device. It is critical to understand that when we combine P- and N-type materials, we do not do so through simple mechanical means. That is, we do not in some way solder, weld, bolt, friction-fit, glue or duct tape[3] one type of material to another. Rather, we must maintain a single piece of mono-crystalline silicon, not a poly-crystalline amalgam of individual pieces. This can be achieved via a diffusion or ion implantation technique that is applied repeatedly to a single piece of silicon crystal. This will leave regions or zones in the crystal that are N-type or P-type. In fact, it is quite possible to have a region of one type completely embedded within a region of the opposite type as we shall see in later chapters.

By creating a single zone of N material adjacent to a zone of P material, we wind up with the *PN junction*. The PN junction is arguably the fundamental building block of solid state semiconductor devices. PN junctions can be found in a variety of devices including bipolar junction transistors (BJTs) and junction field effect transistors (JFETs). The most basic device built from the PN junction is the *diode*. Diodes are designed for a wide variety of uses including rectifying, lighting (LEDs) and photodetection (photodiodes). We shall begin by examining the basic structure and operation of the PN junction. This will include a look at the many different kinds of diodes available. To assist with circuit analysis, a series of simplified models will be created and investigated. We shall use these models to solve a number of example circuits that feature the many diode variations available.

3 They say it has 1001 uses but this ain't one of them.

2.2 The PN Junction

If we were to create a region of N material abutting a region of P material in a single crystal, an interesting situation occurs. Assuming the crystal is not at absolute zero, the thermal energy in the system will cause some of the free electrons in the N material to "fall" into the excess holes of the adjoining P material. This will create a region that is devoid of charge carriers (remember, electrons are the majority charge carrier in N material while holes are the majority charge carrier in P material). In other words, the area where the N and P materials abut is depleted of available electrons and holes, and thus we refer to it as a *depletion region*. This is depicted in Figure 2.1. The excess electrons of the N material are denoted by minus signs while the excess holes of the P material are denoted with plus signs. At the interface, the free electrons have recombined with holes. When an electron recombines, it leaves behind a positive ion in the N material (shown here as a circled plus sign) and produces a negative ion in the P material (shown as a circled minus sign).

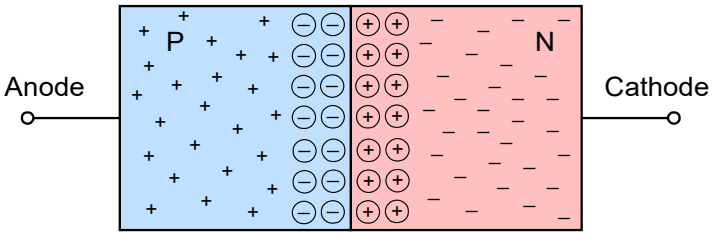


Figure 2.1
PN junction.

We now have a region depleted of charge carriers and this will have an effect on the ability to establish a flow of current through the device. We have, in essence, created an *energy hill* that will need to be overcome.

To understand the concept of the energy hill, recall that in the prior chapter it was discovered that doping an intrinsic crystal would shift the Fermi level. For N material, the Fermi level is shifted up, toward the conduction band. In contrast, for P material the Fermi level is shifted down, nearer to the valence band. When two dissimilar regions adjoin, as in the case here, the energy bands will adjust so that the Fermi levels are consistent. Effectively, this causes the bands of the P material to rise relative to the bands of the N material. The interface between the two appears as a hill, and this is the aforementioned depletion region. This situation is depicted graphically in Figure 2.2. Compare this energy diagram to the energy diagrams for N material and P material presented in the prior chapter. By simply aligning the Fermi levels, it should be clear how we arrive at the new energy diagram.

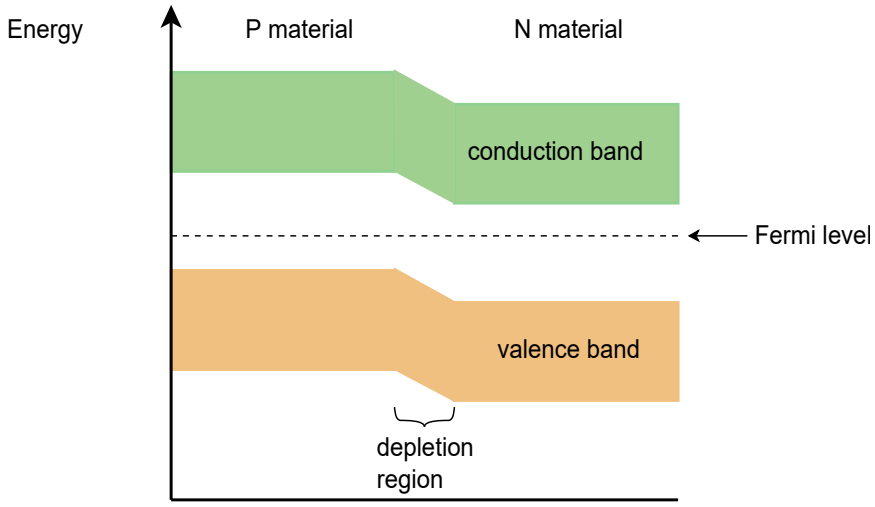


Figure 2.2
Energy bands in PN junction.

Now let's consider what happens if we were to connect this device to an external voltage source as shown in Figure 2.3. Obviously, there are two ways to orient the PN junction with respect to the voltage source. This version is termed *forward-bias*.

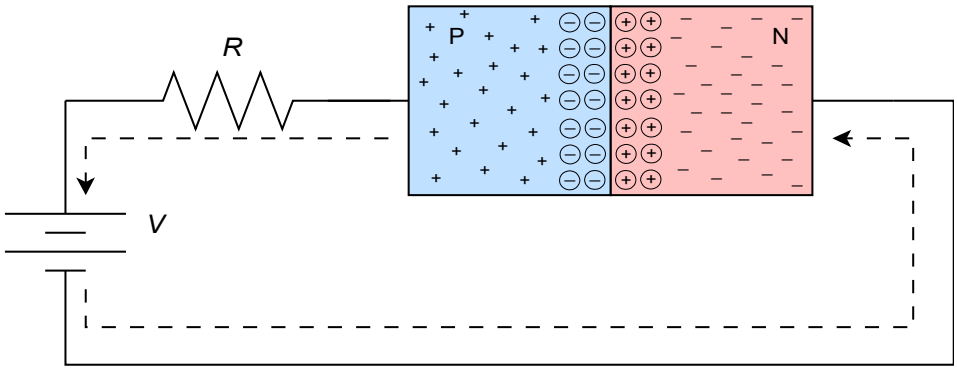


Figure 2.3
PN junction connected to external voltage source.

Forward-Bias

The dotted line of Figure 2.3 shows the direction of electron flow (opposite the direction of conventional flow). First, electrons flow from the negative terminal of the battery toward the N material. In N material, the majority carriers are electrons and it is easy for these electrons to move through the N material. Upon entering the depletion region, if the supplied potential is high enough, the electrons can diffuse into the P material where there are a large number of lower energy holes. From here, the electrons can migrate through to the positive terminal of the source, completing the circuit (the resistor has been added to limit maximum current flow). The “trick” here is to assure that the supplied potential is large enough to overcome the effect of

the depletion region. That is, a certain voltage will be dropped across the depletion region in order to achieve current flow. This required potential is called the *barrier potential* or *forward voltage drop*. The precise value depends on the material used. For silicon devices the barrier potential is usually estimated at around 0.7 volts. For germanium devices it is closer to 0.3 volts while LEDs may exhibit barrier potentials in the vicinity of 1.5 to 3 volts, partly depending on the color.

Another way of thinking about this is that the addition of the voltage source "flattens" the inherent energy hill of the junction. Once the applied forward-bias voltage is at least as big as the hill, current can flow easily.

Reverse-Bias

If the voltage source polarity is reversed in Figure 2.3, the behavior of the PN junction is altered radically. In this case, the electrons in the N material will be drawn toward the positive terminal of the source while the P material holes will be drawn toward the negative terminal, creating a small, short-lived current. This has the effect of widening the depletion region and once it reaches the supplied potential, the flow of current ceases. In essence, we have increased the size of the energy hill. Further increases in the source voltage only serve to make the situation worse. The depletion region simply expands to fill the void, so to speak. Ideally, the PN junction acts like an open circuit with an applied reverse-bias voltage.

This asymmetry in response to a supplied potential turns out to be extraordinarily useful. Perhaps the simplest of all semiconductor devices is the *diode*. In its basic form a diode is just a PN junction. It is a device that will allow current to pass easily in one direction but prevent current flow in the opposite direction.

Shockley Equation

We can quantify the behavior of the PN junction through the use of an equation derived by William Shockley.

$$I = I_S\left(\mathrm{e}^{\frac{V_D q}{n k T}} - 1\right) \qquad (2.1)$$

Where
I is the diode current,
I_S is the reverse saturation current,
V_D is the voltage across the diode,
q is the charge on an electron, 1.6E−19 coulombs,
n is the quality factor (typically between 1 and 2),
k is the Boltzmann constant, 1.38E−23 joules/kelvin,
T is the temperature in kelvin.

At 300 kelvin, q/kT is approximately 38.6. Consequently, for even very small forward (positive) voltages, the "−1" term can be ignored. Also, I_S is not a constant. It increases with temperature, approximately doubling for each 10 C° rise (more on this in a moment).

If we plot the Shockley equation using typical values for a silicon device, we arrive at the curve shown in Figure 2.4. This plots the junction current as a function of the forward (positive) device voltage. It is a representative curve only. While all silicon diodes will exhibit this same general shape, the precise value of current for a specific voltage will vary depending on the device design.

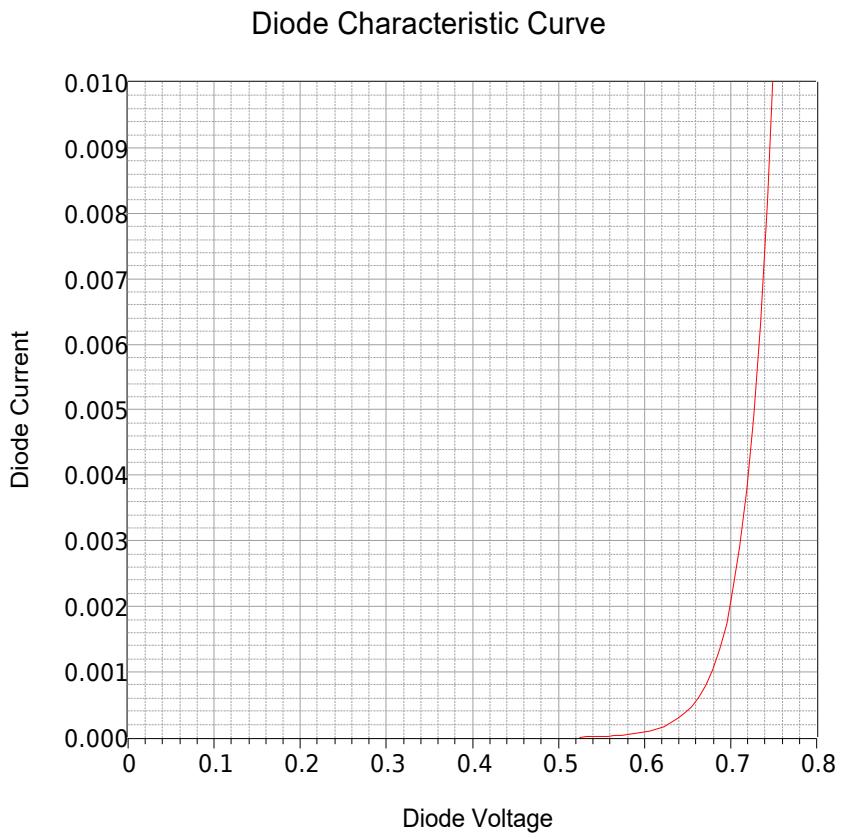


Figure 2.4
Characteristic curve of forward-biased silicon PN junction.

For potentials below about 0.5 volts, the current is virtually non-existent. Above this value, the current rises rapidly, becoming nearly vertical after approximately 0.7 volts. If the plot was recreated using a higher temperature, the effect would be to shift the curve to the left (i.e., a higher current for a given voltage).

If we were to alter the graph to use a logarithmic current scale rather than a linear scale, the graph of Figure 2.5 results. The resulting straight line plot shows clearly the logarithmic relationship between the diode's voltage and current.

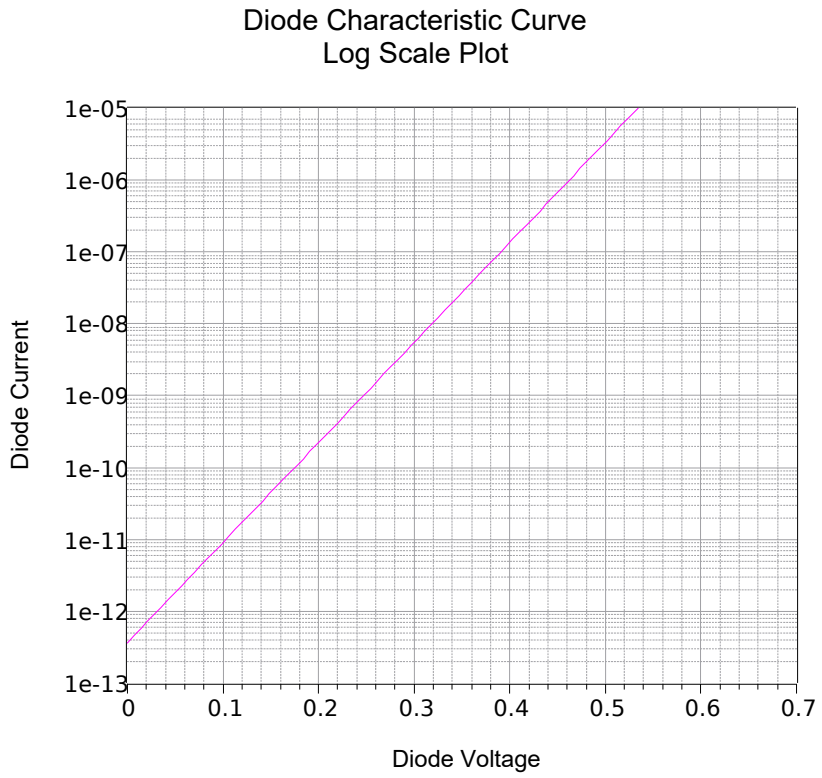


Figure 2.5
Characteristic curve of forward-biased silicon PN junction using log scale.

For negative voltages (reverse-bias) the Shockley equation predicts negligible diode current. This is true up to a point. The equation does not model the effects of *breakdown*. When the reverse voltage is large enough, the diode will start to conduct. This is shown in Figure 2.6. In the first quadrant we see the same general shape we found in Figure 2.4. V_F is the forward "knee" voltage (roughly 0.7 volts for silicon). I_R is the reverse saturation current (ideally zero but in reality a very small amount of current will flow). V_R is the reverse breakdown voltage. Note that the current increases rapidly once this reverse voltage is reached.

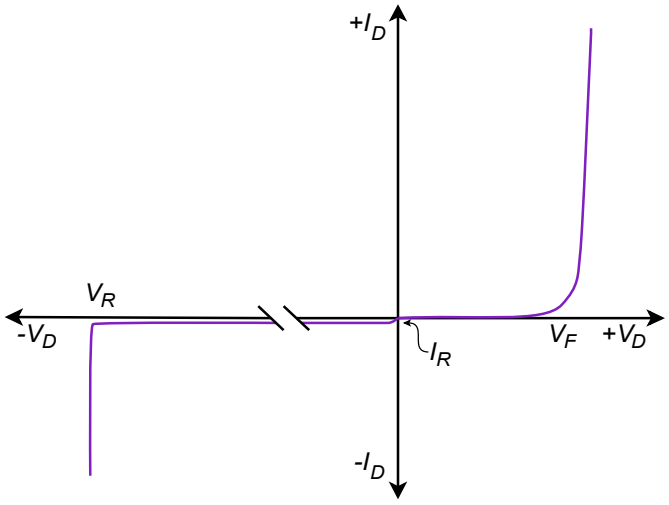
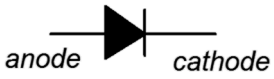


Figure 2.6
Simplified forward and reverse I-V curve for diode.

In general, diodes should not be operated in the breakdown region (the exception being Zener diodes). There are two mechanisms behind this phenomenon. The *Zener effect*, named after Clarence Zener, predominates when the doping levels are high and produces breakdown voltages below roughly five or six volts. It is due to the production of a very high electric field across the depletion region which then results in the production of a high current through electron tunneling. In devices using lower levels of doping, *avalanche* dominates. In this instance, a high electric field accelerates the free electrons to the point where they can impact surrounding atoms and create new electron-hole pairs, thus creating new free electrons that can repeat the process, resulting in a rapid increase of current.

Figure 2.7
Diode schematic symbol (ANSI).

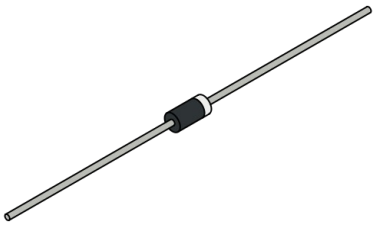


The schematic symbol for a basic switching or rectifying diode is shown in Figure 2.7. This is the ANSI standard which predominates in North America. The P material is the *anode* while the N material is the *cathode*[4]. As a general rule for semiconductor schematic symbols, arrows point toward N material. In this case, the arrow also points in the direction of easy conventional current flow. Figure 2.8 shows an alternate schematic symbol, the IEC international standard, the difference being that it is in outline form without the body being filled in.

Figure 2.8
Alternate diode schematic symbol (IEC).

When it comes to physical device packaging, small and medium current and power devices for through-hole mounting include the DO-35 and DO-204, with the model number stamped on the body. The typical size is comparable to a 1/4 to 1/8 watt resistor. As seen in Figure 2.9 the cathode end is denoted by a band, reminiscent of the bar on the schematic symbol. Surface mount packages are also available. Devices handling higher currents and powers often come in stud or bolt styles such as the DO-4 shown in Figure 2.10. These packages facilitate mounting to a metal plate or heat sink to help dissipate the excess heat.

Figure 2.9
DO-204 case.



2.3 Diode Data Sheet Interpretation

A data sheet for the popular 1N4148 switching diode is shown in Figure 2.11. The 1N4148 is designed for high speed operation required in high frequency signal applications but also finds use in a variety of general purpose applications that do not require very high current or power handling.

Figure 2.10
DO-4 case.
Courtesy of Vishay Intertechnology, Inc.

Some of the key features include a four nanosecond switching speed, a maximum reverse voltage of 100 volts and a 450 milliamp maximum forward current (with short single pulses as high as four amps being possible). Power dissipation is 500 milliwatts.

Referring to Figure 2.11b, the variation in reverse current with regard to temperature is obvious. This also verifies the “doubles every 10 C°” rule-of-thumb.

4 *Cathode* is often denoted by a *k*. This is likely due to the word's Greek root, *kathodos*.

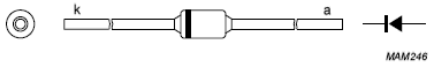
High-speed diodes 1N4148; 1N4448

FEATURES

- Hermetically sealed leaded glass SOD27 (DO-35) package
- High switching speed: max. 4 ns
- General application
- Continuous reverse voltage: max. 100 V
- Repetitive peak reverse voltage: max. 100 V
- Repetitive peak forward current: max. 450 mA.

APPLICATIONS

- High-speed switching.



The diodes are type branded.

Fig.1 Simplified outline (SOD27; DO-35) and symbol.

DESCRIPTION

The 1N4148 and 1N4448 are high-speed switching diodes fabricated in planar technology, and encapsulated in hermetically sealed leaded glass SOD27 (DO-35) packages.

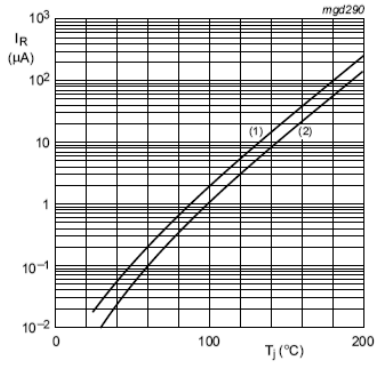
MARKING

TYPE NUMBER	MARKING CODE
1N4148	1N4148PH or 4148PH
1N4448	1N4448

ORDERING INFORMATION

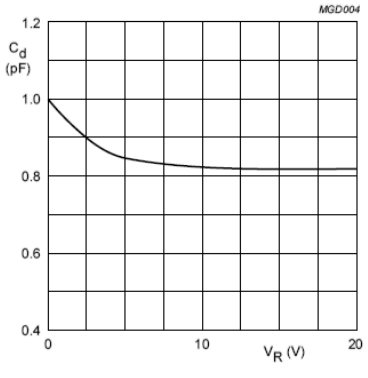
TYPE NUMBER	PACKAGE		
	NAME	DESCRIPTION	VERSION
1N4148	–	hermetically sealed glass package; axial leaded; 2 leads	SOD27
1N4448			

Figure 2.11a
1N4148 data sheet.
Courtesy of NXP Semiconductors.



(1) V_R = 75 V; typical values.
(2) V_R = 20 V; typical values.

Fig.5 Reverse current as a function of junction temperature.



f = 1 MHz; T_j = 25 °C.

Fig.6 Diode capacitance as a function of reverse voltage; typical values.

Figure 2.11b
1N4148 data sheet (continued).

Figure 2.11c
1N4148 data sheet (continued).

LIMITING VALUES

In accordance with the Absolute Maximum Rating System (IEC 60134).

SYMBOL	PARAMETER	CONDITIONS	MIN.	MAX.	UNIT
V_{RRM}	repetitive peak reverse voltage		–	100	V
V_R	continuous reverse voltage		–	100	V
I_F	continuous forward current	see Fig.2; note 1	–	200	mA
I_{FRM}	repetitive peak forward current		–	450	mA
I_{FSM}	non-repetitive peak forward current	square wave; T_j = 25 °C prior to surge; see Fig.4			
		t = 1 μs	–	4	A
		t = 1 ms	–	1	A
		t = 1 s	–	0.5	A
P_{tot}	total power dissipation	T_{amb} = 25 °C; note 1	–	500	mW
T_{stg}	storage temperature		−65	+200	°C
T_j	junction temperature		–	200	°C

Note

1. Device mounted on an FR4 printed-circuit board; lead length 10 mm.

ELECTRICAL CHARACTERISTICS

T_j = 25 °C unless otherwise specified.

SYMBOL	PARAMETER	CONDITIONS	MIN.	MAX.	UNIT
V_F	forward voltage	see Fig.3			
	1N4148	I_F = 10 mA	–	1	V
	1N4448	I_F = 5 mA	0.62	0.72	V
		I_F = 100 mA	–	1	V
I_R	reverse current	V_R = 20 V; see Fig.5		25	nA
		V_R = 20 V; T_j = 150 °C; see Fig.5	–	50	μA
I_R	reverse current; 1N4448	V_R = 20 V; T_j = 100 °C; see Fig.5	–	3	μA
C_d	diode capacitance	f = 1 MHz; V_R = 0 V; see Fig.6	–	4	pF
t_{rr}	reverse recovery time	when switched from I_F = 10 mA to I_R = 60 mA; R_L = 100 Ω; measured at I_R = 1 mA; see Fig.7	–	4	ns
V_{fr}	forward recovery voltage	when switched from I_F = 50 mA; t_r = 20 ns; see Fig.8	–	2.5	V

THERMAL CHARACTERISTICS

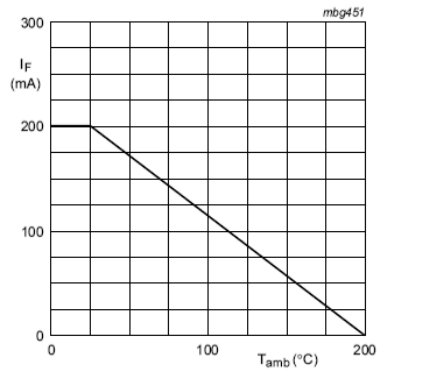
SYMBOL	PARAMETER	CONDITIONS	VALUE	UNIT
$R_{th(j-tp)}$	thermal resistance from junction to tie-point	lead length 10 mm	240	K/W
$R_{th(j-a)}$	thermal resistance from junction to ambient	lead length 10 mm; note 1	350	K/W

Note

1. Device mounted on a printed-circuit board without metallization pad.

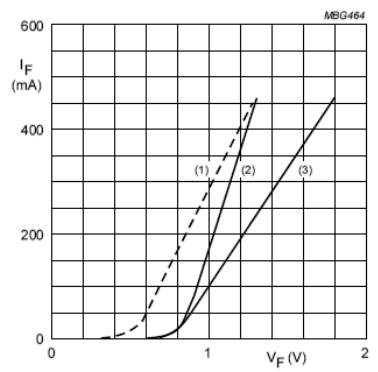
Figure 2.11d
1N4148 data sheet (continued).

GRAPHICAL DATA



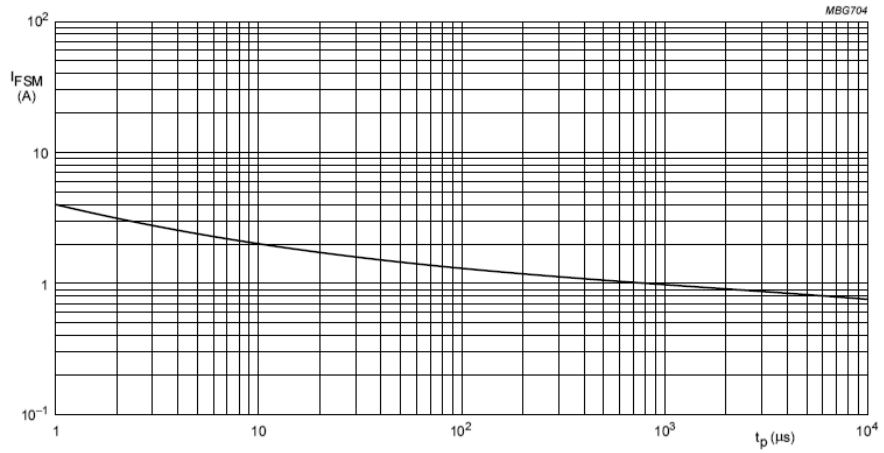
Device mounted on an FR4 printed-circuit board; lead length 10 mm.

Fig.2 Maximum permissible continuous forward current as a function of ambient temperature.



(1) T_j = 175 °C; typical values.
(2) T_j = 25 °C; typical values.
(3) T_j = 25 °C; maximum values.

Fig.3 Forward current as a function of forward voltage.



Based on square wave currents.
T_j = 25 °C prior to surge.

Fig.4 Maximum permissible non-repetitive peak forward current as a function of pulse duration.

Power derating and permissible pulse amplitudes can be seen in Figure 2.11d. Finally, note the variation in the forward voltage curves due to temperature. As stated previously, for a given current, an increase in temperature results in a lower forward voltage. Also, at room temperature, we see a knee voltage of approximately 0.7 volts.

2.4 Diode Circuit Models

One thing is very clear from the characteristic curve of the diode: It is not a linear bilateral[5] device, quite unlike a resistor. Consequently, we cannot use the superposition technique to solve diode circuits unless we have *a priori* knowledge about it, that is, whether or not it is forward- or reverse-biased. For example, we can imagine a circuit comprised of two voltage sources, resistors and a diode. By itself, one of the voltage sources might forward-bias the diode while the other would reverse-bias it. Obviously, a diode cannot be both forward and reverse-biased at the same time.

A second problem we face with circuit analysis is the added complexity of the Shockley equation. For speed and ease of computation we find it useful to model the diode with simpler circuit elements. Three diode models are shown in Figure 2.12.

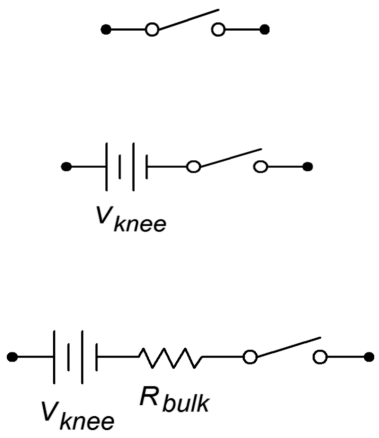


Figure 2.12
Simplified diode models. Top to bottom: first, second and third approximations, increasing in accuracy.

The first approximation is the simplest of the three. It treats the diode as a simple dependent switch: the switch is closed if the diode is forward-biased and open if it is reverse-biased. The second approximation adds the effect of the forward voltage. V_{knee} is the "turn-on" potential required to overcome the energy hill. It would be 0.7 volts for a silicon device. The third approximation is the most accurate of the three. A close look at the characteristic curve of Figure 2.4 shows that once the knee voltage is reached, the curve does not transition to a perfect vertical line. Instead, there remains some positive, non-infinite slope. That is, the voltage continues to increase, although modestly, with further increases in current. We can approximate this effect as a small resistive value, R_{bulk}. The three corresponding I-V plots are shown in Figure 2.13. Compare these to Figure 2.6 and note the increasing accuracy.

5 The I-V plot is not a straight line (linear) and the forward and reverse quadrants are not identical (bilateral).

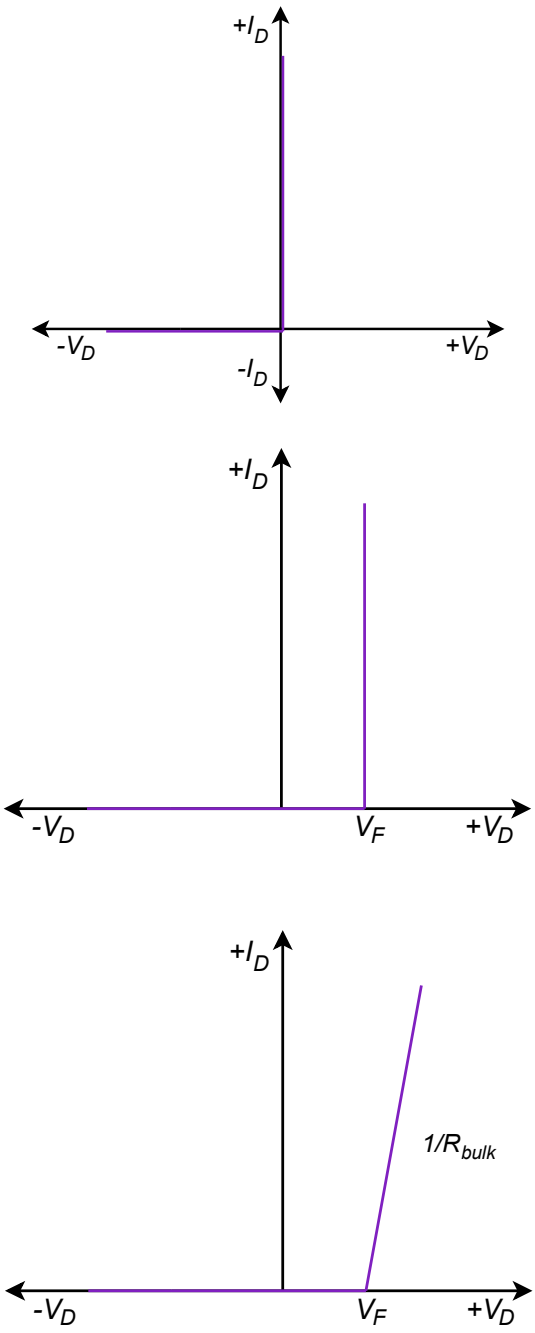


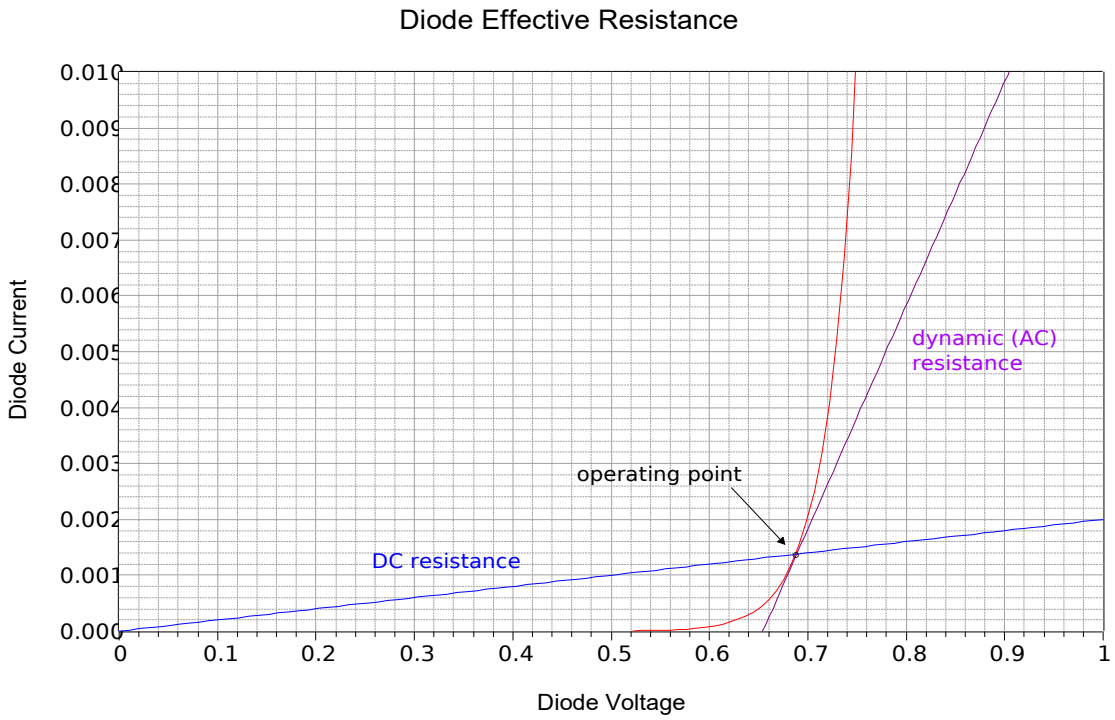
Figure 2.13
I-V curves for simplified diode models.
Top to bottom: first, second and third approximations.

In many applications the second approximation will yield sufficiently accurate results and we will tend to make greatest use of it. Just remember that these are behavioral models; don't think that there are literally 0.7 volt sources or little resistors in the diodes.

It should be noted that R_{bulk} does not represent the "diode resistance" per se, rather, it models a minimum value. There really is no such thing as a singular diode resistance. We can, however, talk about the *effective resistance* of a diode in a particular circuit in both DC and AC terms.

The key to understanding this concept is to remember that resistance is a linear function, a straight line on a I-V graph. Therefore, we need to find a straight line "fit" for the diode curve. Two possibilities are shown in Figure 2.14.

Figure 2.14
Diode effective resistance for DC and AC.



The red curve is the diode's characteristic curve (arbitrary current values are shown). For some particular DC circuit, a specific current will flow through the diode which will produce a particular voltage, denoted on the graph as the *operating point*. If we simply compute the ratio of that voltage to the driving current, we wind up with a resistance. This is the effective DC resistance of the diode under these circuit conditions and is represented by the blue line. That is, the reciprocal of the slope of the blue line is the effective DC resistance. Obviously, if we shift the operating point along the red diode curve, the slope of the intersecting blue line changes and therefore we arrive at a new DC resistance. The higher the current, the lower the effective DC resistance.

Instead of just DC, the diode might see a combination of DC and AC signals. Visualize this as adding a small AC variation on top of the DC. We can imagine the operating point moving along the red diode curve, back and forth about the operating point. If we divide the small AC voltage variation by its associated AC

current variation, we wind up with the AC equivalent resistance, also known as the *dynamic resistance*. Graphically, we can think of this as finding the slope of a line that is tangent to the operating point (the purple line). This will in fact, be an average value across the AC variation. It should also be apparent that the effective AC resistance must be smaller than its DC counterpart because the AC approximation (purple line) must be steeper than the DC approximation (blue line)[6].

It is time for a few illustrative examples.

Example 2.1

Consider the resistor-diode circuit of Figure 2.15. Assume the voltage source is 12 volts and the resistor is 2 kΩ. Further, assume the diode is silicon and its bulk resistance is 10 Ω. Using the three diode approximations, compute the circulating current.

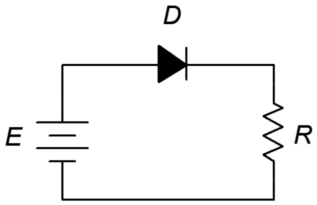


Figure 2.15
Schematic for Example 2.1.

First, note that the diode is forward-biased. This must be the case because there is a single voltage that is larger than the knee voltage and its positive terminal is attached to the diode's anode. No matter which approximation we use, Kirchhoff's voltage law (KVL) must be true so it will be a matter of summing the available voltage drops versus resistance(s).

Using the first approximation:

Here we assume the diode is a closed switch. Consequently all of the source voltage must drop across the single resistor.

$$I = \frac{E}{R}$$
$$I = \frac{12\,\text{V}}{2\,\text{k}\Omega}$$
$$I = 6\,\text{mA}$$

Using the second approximation:

In this instance we include the knee voltage.

$$I = \frac{E - V_{knee}}{R}$$
$$I = \frac{12\,\text{V} - 0.7\,\text{V}}{2\,\text{k}\Omega}$$
$$I = 5.65\,\text{mA}$$

6 The dynamic resistance of a PN junction may be approximated as 26 mV/$I_{junction}$. This will be shown in an upcoming chapter.

Using the third approximation:

The most accurate of the three, we include both the knee voltage and bulk resistance.

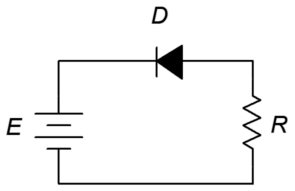
$$I = \frac{E - V_{knee}}{R + R_{bulk}}$$
$$I = \frac{12\,\text{V} - 0.7\,\text{V}}{2\,\text{k}\Omega + 10\,\Omega}$$
$$I = 5.622\,\text{mA}$$

In this particular case the difference between the second and third approximations is less than 1%. It is also worth noting that the third approximation predicts a diode voltage of slightly more than 0.7 volts (approximately 0.756 volts) due to the additional potential across the bulk resistance.

Example 2.2

Determine the circulating current for the circuit in Figure 2.16. Also find the diode and resistor voltages. Assume the power supply is 20 volts, the diode is silicon and the resistor is 2 kΩ.

Figure 2.16
Schematic for Example 2.2.



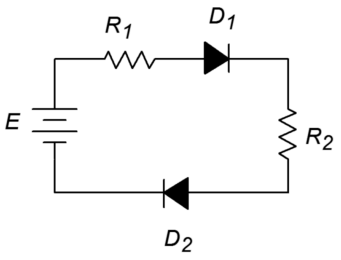
This problem is deceptively easy. Note that the positive terminal of the source is connected to the cathode. As there are no other sources in the circuit, the diode must be reverse-biased. The model for a reverse-biased diode is an open switch and the circulating current in an open circuit is zero. Therefore, the resistor voltage must also be zero and values for the knee voltage and bulk resistance are not needed. In order to satisfy KVL, the diode voltage will equal the source of 20 volts (+ to − from cathode to anode).

The only time this would not be the case is if the reverse breakdown voltage of the diode is less than the 20 volt source. In that case the diode voltage would equal the breakdown voltage with the remainder of the source voltage dropping across the resistor.

Example 2.3

Determine the circulating current for the circuit in Figure 2.17. Also find the diode and resistor voltages. Assume the power supply is 9 volts, the diodes are silicon and R_1 = 1 kΩ, R_2 = 2 kΩ.

Figure 2.17
Schematic for Example 2.3.



According to KVL, the applied source must equal the sum of the voltage drops across the resistors and diodes as this is a single loop. Both diodes are forward-biased (conventional current entering the anodes).

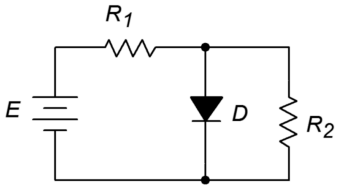
$$I = \frac{E - V_{knee1} - V_{knee2}}{R_1 + R_2}$$
$$I = \frac{9\,\text{V} - 0.7\,\text{V} - 0.7\,\text{V}}{1\,\text{k}\,\Omega + 2\,\text{k}\,\Omega}$$
$$I = 2.533\,\text{mA}$$

Note that if either diode was reversed, there would be no current flow and all of the source potential would drop across the reversed diode.

Example 2.4

Determine the source current and resistor voltages for the circuit in Figure 2.18. Also find the resistor voltages if the diode polarity is reversed. Assume the power supply is 10 volts, the diode is silicon and the resistors are 1 kΩ each.

Figure 2.18
Schematic for Example 2.4.



As D and R_2 are in parallel they must have the same voltage drop. Also, the diode is forward-biased. Therefore, the voltage across R_2 must be approximately 0.7 volts, leaving 9.3 volts to drop across R_1. The current through R_1 is the source current.

$$I = \frac{E - V_D}{R_1}$$
$$I = \frac{10\,\text{V} - 0.7\,\text{V}}{1\,\text{k}\,\Omega}$$
$$I = 9.3\,\text{mA}$$

If the diode is reversed it behaves as an open switch. The circuit reduces to a simple 1:1 voltage divider, each resistor dropping half of the supply, or 5 volts each.

Computer Simulation

To verify our results, Example 2.4 is simulated. The circuit is captured as shown in Figure 2.19a. This particular example is shown in Multisim although any decent quality simulator will do. The very common 1N4148 switching diode is used here. Another popular choice would be the 1N914 switching diode or a 1N400X series rectifier.

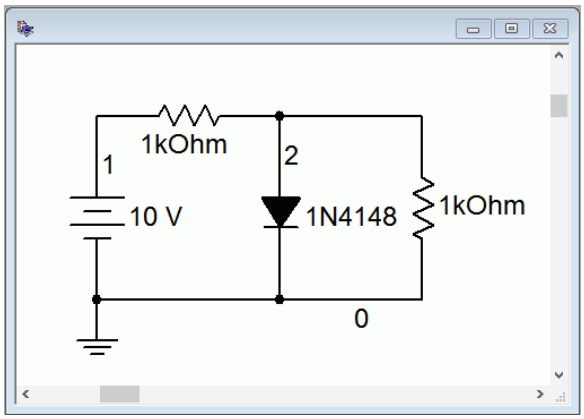


Figure 2.19a
The circuit of Example 2.4 in Multisim.

Next, a DC Operating Point analysis is performed. The results are shown in Figure 2.19b. Note that the diode potential is just under the 0.7 volt approximation. From this we can deduce that the voltage drop across the first resistor must be slightly more than 9.3 volts, producing a current slightly more than 9.3 mA.

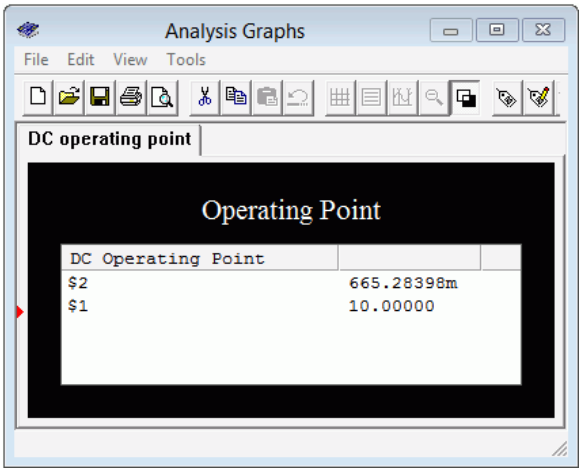
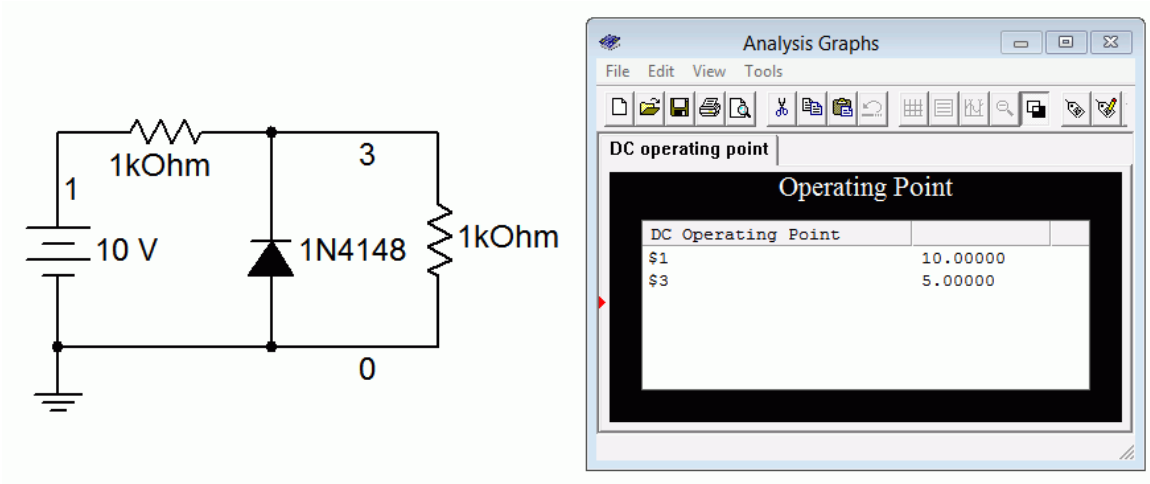


Figure 2.19b
DC Operating Point simulation results for the circuit of Example 2.4.

Finally, Figure 2.19c shows the results when the diode is reversed in the circuit. The second resistor (node 3 to ground) shows 5 volts as expected. Therefore, the first resistor must also be dropping 5 volts.

Figure 2.19c
Simulation of Example 2.4 using reversed diode orientation.



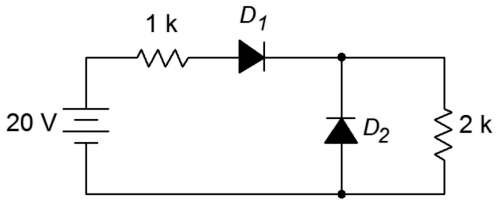
Before moving on to another topic, let's take a look at a somewhat more involved example using multiple diodes.

Example 2.5

Determine the diode and resistor voltages for the circuit in Figure 2.20. Assume the diodes are silicon.

The first thing to notice is that D_1 is forward-biased while D_2 is reverse-biased. Therefore, the 20 volt source must equal the drop across D_1 and the two resistors. D_2 will take on whatever the drop across the 2 kΩ works out to as they are in parallel.

Figure 2.20
Schematic for Example 2.5.



$$I = \frac{E - V_{D1}}{R_1 + R_2}$$

$$I = \frac{20\,\text{V} - 0.7\,\text{V}}{1\,\text{k}\,\Omega + 2\,\text{k}\,\Omega}$$

$$I = 6.433\,\text{mA}$$

Note that virtually no current flows down through D_2 as it is reverse-biased. Using Ohm's law, the drop across the first resistor is 6.433 volts and for the second resistor, 12.867 volts.

2.5 Other Types of Diodes

Diodes have been designed to exploit different aspects of PN junctions. Besides the basic use as a switching or rectifying device, diodes are available for voltage regulation, variable capacitance, illumination and light sensing. The schematic symbols for a number of popular diode types are shown in Figure 2.21. Note the similarities of the symbols. The "bar" portion represents the cathode for all of them.

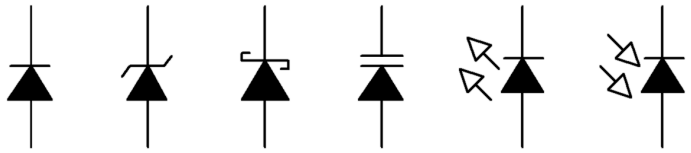


Figure 2.21
Diode schematic symbols:
a) switching or rectifying
b) Zener
c) Schottky
d) varactor
e) LED
f) photodiode

Zener Diode

The Zener diode behaves like an ordinary signal diode when forward-biased. Normally, though, Zeners are used in a reverse-bias condition. From our previous discussion, recall that if the reverse potential is high enough, a diode can go into breakdown, causing a rapid increase current. This was caused by either of two effects, Zener conduction or avalanche. The Zener diode takes advantage of this in order to produce a stable voltage[7]. Zeners are specified by their reverse potential (generically referred to as the "Zener voltage") and are designed to handle larger currents and powers than the average signal diode. Zener voltages are standardized in much the same manner as resistors so values such as 3.9 volts, 5.1 volts and 6.8 volts are to be expected. The Zener voltage is measured at I_{ZT}, the Zener test current. A lower current may not fully push the diode into conduction resulting in a lower than expected diode potential.

Figure 2.22
Zener diode schematic symbol.

Instead of modeling the Zener as an open switch when reverse-biased, instead we model it as an open when its voltage is less than the rated voltage, and as a voltage source equal to the rated value if its voltage tries to exceed that value. When analyzing Zener-based circuits, the first thing to do is determine if the diode is forward-biased. If it is then treat it like an ordinary switching diode. If, on the other hand, it is reverse-biased then treat it like an open switch. If the resulting diode voltage is greater than the Zener voltage then recompute the circuit but this time mentally replace the Zener with a voltage source equal to the Zener voltage. Our next example will illustrate this method.

7 Although they are called *Zener* diodes, they rely on either the Zener or avalanche effects, depending on the magnitude of the voltage.

Example 2.6

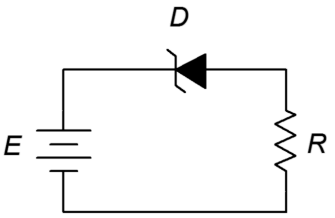
Determine the circulating current for the circuit in Figure 2.23. Also find the diode and resistor voltages. Assume the power supply is 9 volts, the Zener voltage is 5.1 volts and the resistor is 3.3 kΩ.

The diode is reverse-biased. If we treat it as an open then it would drop the entire source voltage, or 9 volts. This is greater than the Zener potential so the device must be in Zener conduction. This means that conventional current will flow relatively easily in a clockwise direction. The voltage across the diode will equal the rated value of 5.1 volts, + to − from cathode to anode. By KVL the resistor drop must be 9 V − 5.1 V, or 3.9 volts.

$$I = \frac{E - V_{Zener}}{R}$$
$$I = \frac{9\,\text{V} - 5.1\,\text{V}}{3.3\,\text{k}\,\Omega}$$
$$I = 1.182\,\text{mA}$$

If the diode was flipped in orientation then it would be forward-biased and show the expected 0.7 volts with 8.3 volts across the resistor.

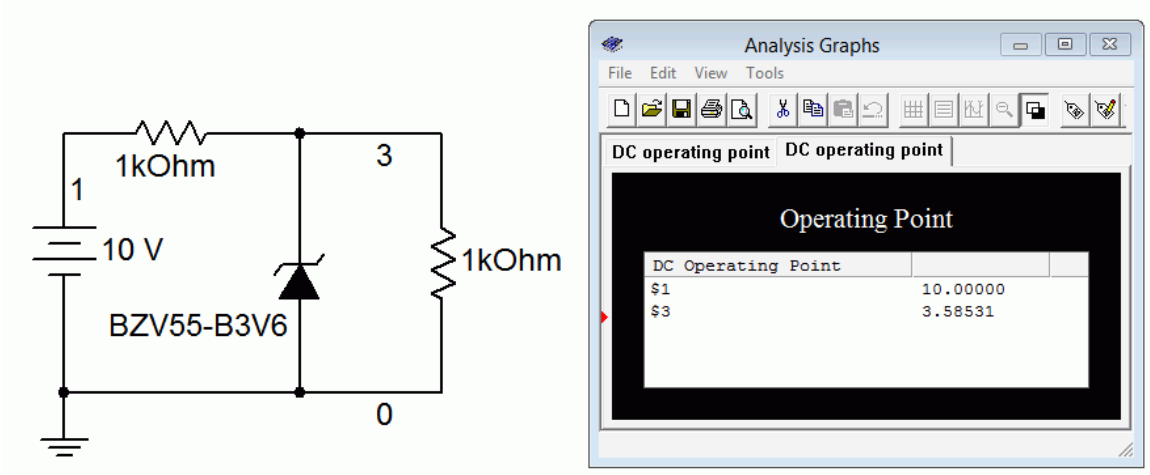
Figure 2.23
Schematic for Example 2.6.



Computer Simulation

A Zener diode circuit is simulated as shown in Figure 2.24.

Figure 2.24
Zener diode circuit simulation.



Without the Zener, the two resistors would simply split the supply voltage equally, each receiving 5 volts. If an ordinary diode was used it would be reverse-biased and act as an open. The resulting voltages would be the same. In this case, however, the Zener is activated at 3.6 volts (it is typical to include the Zener voltage as part of the model number with the letter "V" replacing the decimal point when needed). Therefore we see roughly 3.6 volts across the Zener and the parallel second resistor (node 3 to ground). The precise value of voltage will depend on the magnitude of the diode current. If the simulation is rerun with a higher voltage source, the increased current will produce a slightly higher voltage at node 3. This is because the breakdown curve is not infinitely steep once it goes past the rated Zener voltage. The effect is similar to that of R_{bulk} in a forward-biased diode. On a data sheet this value is referred to as the *differential resistance*, or R_{dif}.

An excellent use of the Zener is to limit or regulate a voltage. When a Zener is placed in parallel with other components we can ensure that those components will not see a potential higher than the rated Zener voltage. We will take a much closer look at this in the next chapter.

Light Emitting Diode (LED) and Photodiode

The light emitting diode (LED) and photodiode are complements. While the LED produces light with an electrical input, the photodiode produces a current when exposed to light. Both devices can operate within the human visible spectrum and can also be designed to operate at wavelengths outside this range, in the infrared (IR) and ultraviolet (UV). In fact, most TV remote controls rely on IR emitter/detector pairs for communication[8].

Figure 2.25
Light emitting diode (LED) schematic symbol.

The LED has displaced traditional incandescent (filament-based) light sources in many applications due to its high efficiency in turning an electrical energy input into a light output. They are small, physically robust, operate relatively cool and are available in a number of different colors. The schematic symbol is shown in Figure 2.25. The basic idea behind its operation is fairly simple. In a forward-biased PN junction, when free electrons recombine and "fall" into lower energy valence holes they must give up this energy differential in some manner. In most diodes, this energy is emitted as heat. In LEDs, the energy transition is designed such that it is radiated at shorter wavelengths (i.e., visible light). In order to achieve this, LEDs are not formed just using silicon as in a typical switching diode. Instead, somewhat more exotic materials are used. From an analysis or design standpoint, the important thing to remember is that the forward voltage tends to be noticeably higher than silicon's 0.7 volt drop. The precise value will depend on the material, which in turn effects the color. A generic red LED will likely exhibit a forward drop of around 1.8 volts or so. Other colors tend to be somewhat higher as we move through the

8 There are advantages to using the infrared over the visible spectrum for this application. It tends to be less sensitive to room lighting conditions and there are no potentially annoying visible flashes of light coming from the remote.

rainbow, ending with blue and UV LEDs (and also high brightness versions) up around 3 to 4 volts. In a lab it is easy to determine the approximate forward drop of a given diode by connecting it in series with a voltage source and current limiting resistor. The supply is increased until the desired brightness is achieved and then the diode drop can be measured with a DMM. When reverse-biased the LED behaves like a switching diode, that is, it looks like an open switch. Unlike switching and rectifying diodes, LED maximum reverse potentials tend to be relatively low, perhaps just a few volts.

A datasheet for the Cree C566D series LED is presented in Figure 2.26. Notice that the colors are specified in terms of wavelength (in nanometers) and luminous intensity (brightness) is given in millicandella (mcd).

Figure 2.26a
LED datasheet.
Courtesy of Cree, Inc.

CREE®

PRODUCT FAMILY DATA SHEET

Cree® Screen Master® 5-mm Oval LED
C566D-RFF/GFF/BFF/AFF
C566D-RFE/GFE/BFE/AFE

PRODUCT DESCRIPTION

The oval LED is specifically designed for variable-message signs and passenger-information signs.The oval-shaped radiation pattern and high luminous intensity ensure that these devices are excellent for wide-field-of -view outdoor applications where a wide viewing angle and readability in sunlight are essential.

These lamps are tinted and diffused. The encapsulation resin contains anti-UV material in order to reduce the effects of long-term exposure to direct sunlight.

FEATURES

- Size (mm): 5
- Color and Typical Dominant Wavelength:
  Red (621nm)
  Green(527nm)
  Blue(470nm)
  Amber(591nm)
- Luminous Intensity (mcd)
  C566D-RFF/RFE:(2130-5860)
  C566D-GFF/GFE:(4180-12000)
  C566D-BFF/BFE:(1100-3000)
  C566D-AFF/AFE:(2130-5860)
- Lead - Free
- RoHS Compliant

APPLICATIONS

- Electronic Signs & Signals (ESS)
- Full Color video screen
- Motorway Signs
- Variable Message Sign (VMS)
- Advertising signs
- Petrol Signs

Cree, Inc.
4600 Silicon Drive
Durham, NC 27703
USA Tel: +1.919.313.5300

These devices are *RoHS compliant* which stands for Restriction of Hazardous Substances. It is an EU directive that limits usage of toxic materials such as lead, cadmium and mercury. Continuing, we find the device maximum ratings:

Figure 2.26b
LED datasheet (continued).

CREE

Cree 5-mm Oval LED
C566D-RFF/GFF/BFF/AFF
C566D-RFE/GFE/BFE/AFE

ABSOLUTE MAXIMUM RATINGS (T_A = 25°C)

Items	Symbol	Absolute Maximum Rating		Unit
		Red and Amber	Blue and Green	
Forward Current	I_F	50 Note1	35	mA
Peak Forward Current Note2	I_{FP}	200	100	mA
Reverse Voltage	V_R	5	5	V
Power Dissipation	P_D	130	140	mW
Operation Temperature	T_{opr}	-40 ~ +95		°C
Storage Temperature	T_{stg}	-40 ~ +100		°C
Lead Soldering Temperature	T_{sol}	Max. 260°C for 3 sec. max. (3 mm from the base of the epoxy bulb)		
Electrostatic Discharge Classification (MIL-STD-883E)	ESD	Class 2		

Note:
1. For long term performance the drive currents between 10mA and 30mA are recommended. Please contact CREE sales representative for more information on recommended drive conditions.
2. Pulse width ≤0.1 msec, duty ≤1/10.

TYPICAL ELECTRICAL & OPTICAL CHARACTERISTICS (T_A = 25°C)

Characteristics	Color	Symbol	Condition	Unit	Minimum	Typical	Maximum
Forward Voltage	Red/Amber	V_F	I_F = 20 mA	V		2.1	2.6
	Blue/Green	V_F	I_F = 20 mA	V		3.4	4.0
Reverse Current	Red/Amber	I_R	V_R = 5 V	μA			100
	Blue/Green	I_R	V_R = 5 V	μA			100
Dominant Wavelength	Red	λ_D	I_F = 20 mA	nm	619	621	624
	Green	λ_D	I_F = 20 mA	nm	520	527	535
	Blue	λ_D	I_F = 20 mA	nm	460	470	475
	Amber	λ_D	I_F = 20 mA	nm	584	591	596
Luminous Intensity	Red	I_V	I_F = 20 mA	mcd	2130	3000	
	Green	I_V	I_F = 20 mA	mcd	4180	7000	
	Blue	I_V	I_F = 20 mA	mcd	1100	2000	
	Amber	I_V	I_F = 20 mA	mcd	2130	3000	

Note the different values between the various colors. The forward current is specified as 50 mA for red/amber with 35 mA for green/blue. Nominal operating currents are between 10 and 30 mA. Reverse voltage is 5 volts, typical for many LEDs although much lower than the average switching diode. Forward voltage is typically 2.1 volts for the red end of the spectrum and, as expected, 3.4 volts for the green/blue end. The expected luminous intensities also vary with color. Further, it

Figure 2.26c
LED datasheet (continued).

should be noted that LEDs do not produce "pure color" light in the manner of a laser. Rather, they produce a range of wavelengths clustered in a specific area. The wavelength that produces the highest output in this area is referred to as the peak or dominant wavelength. Human vision covers the range of roughly 400 nanometers (violet) to 700 nanometers (red)[9].

CREE

Cree 5-mm Oval LED
C566D-RFF/GFF/BFF/AFF
C566D-RFE/GFE/BFE/AFE

GRAPHS

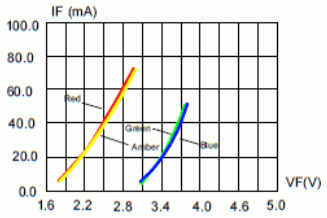


FIG.1 FORWARD CURRENT VS. FORWARD VOLTAGE.

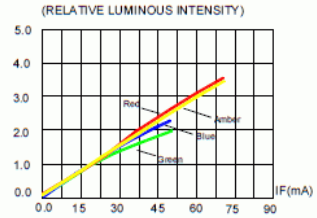


FIG.2 RELATIVE LUMINOUS INTENSITY VS. FORWARD CURRENT

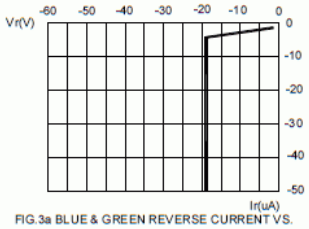


FIG.3a BLUE & GREEN REVERSE CURRENT VS. REVERSE VOLTAGE.

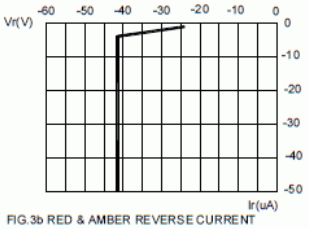


FIG.3b RED & AMBER REVERSE CURRENT VS. REVERSE VOLTAGE.

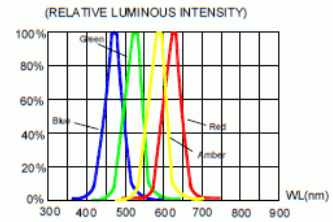


FIG.5 RELATIVE LUMINOUS INTENSITY VS. WAVELENGTH.

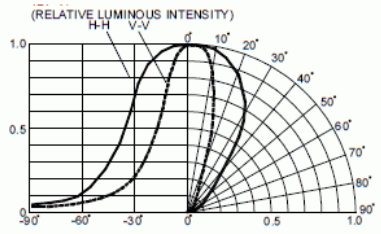


FIG.6 FAR FIELD PATTERN

The above data are collected from statistical figures that do not necessarily correspond to the actual parameters of each single LED. Hence, these data will be changed without further notice.

9 It is interesting to observe that the human visual system operates over a frequency range of less than 2:1 while the human auditory system operates over a frequency range of about 1000:1 (20 hertz to 20,000 hertz). If human hearing had a range equivalent to that of our sight, we'd hear less than a full octave of pitches in total. In other words, *do-re-mi-fa-sol-la-ti-do* would end at *ti* and anything beyond would be inaudible. In such a case one thing is certain: piano keyboards would be much shorter.

Figure 2.26c presents pertinent graphical data. We observe a roughly linear increase in luminous intensity with increasing current. Also, note the difference in the reverse voltage/current plots between blue/green and red/amber. Of particular interest is the final graph which shows the beam pattern or beam angle. You can think of this in terms of how narrow or broad the illumination pattern is. When comparing different model LEDs it is useful to remember that on-axis brightness can be increased by narrowing the angle. This graph is split in half using two different ways of showing the data. On the left side we have a linear graph depicting the relative brightness as we move off of the center axis (zero degrees). On the right side we see a polar plot version of the same data.

Example 2.7

Determine the circulating current for the circuit in Figure 2.27. Assume the power supply is 5 volts, the LED forward voltage is 2.1 volts and the resistor is 330 Ω.

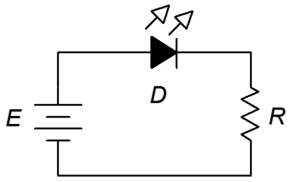


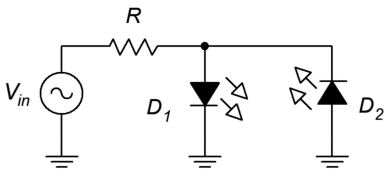
Figure 2.27
Schematic for Example 2.7.

The LED is forward-biased and as the source is greater than the LED potential, it should light. Using KVL,

$$I = \frac{E - V_{LED}}{R}$$
$$I = \frac{5\,\text{V} - 2.1\,\text{V}}{330\,\Omega}$$
$$I = 8.788\,\text{mA}$$

This should result in a relatively bright LED. The resistor can be used to effectively program the brightness by changing the current level (a smaller resistance yields a higher current and therefore a brighter LED). Given the 2.1 volt forward potential, it is likely that this is an amber or yellow LED. If a different color had been used, say a 1.6 volt red or 3.2 volt blue, there would be a change in current and most likely a change in brightness. The change in brightness might not perfectly echo the change in current because the conversion efficiency for the two diodes may not be the same (refer to Figure 2.27b to compare the luminous intensities at 20 mA for different wavelengths).

Figure 2.28
Dual LED application.



An interesting circuit using two differently colored LEDs is shown in Figure 2.28. An AC source is used to drive the LEDs. Only one of the two will be forward-biased at any given time. For positive source voltages D_1 will be on while D_2 will be off. For negative source voltages the opposite will be true. Resistor R serves to limit current for both of them. Assume D_1 is red and D_2 is blue. Further, suppose the source frequency is relative low, say 1 hertz. For the positive half cycle (.5 seconds)

the red LED will light and for the negative half cycle the blue LED will light. This alternating pattern continues for as long as the source is applied but a curious thing happens as we increase the frequency. At first, the blink rate will increase with the red and blue flickering back and forth faster and faster. At some point, perhaps around 30 hertz or so, it will appear as though both LEDs are lit continuously. This is because the human visual sense will tend to integrate the rapid motion and we effectively see the "average" intensity. In fact, this "on-off" trick is often used in digital circuits to control the brightness of LEDs or the speed of motors. Bi-color LEDs are available in a single package. Using a common lead and two control leads (one for each color), it is possible to achieve color mixing.

The logical inverse of the LED is the photodiode, the schematic symbol of which is illustrated in Figure 2.29. The photodiode includes some manner of port that allows light to hit the junction. A sufficiently energetic photon of light can knock lose an electron. This creates an electron-hole pair which results in current flow. As more light energy is added to the system, an increasing current or voltage will result[10].

Figure 2.29
Photodiode schematic symbol.

Photodiodes can operate in one of two modes. The first mode is *photovoltaic* mode. It uses zero-bias (that is, no external bias potential). In this mode, the photodiode operates as a voltage source. This is the mode used by photovoltaic solar cells. They can be thought of as very large photodiodes. The second mode of operation is *photoconductive* mode. This mode requires reverse-biasing the diode with an external potential. In this mode the diode acts more like a current source. The advantage is that the response is faster than photovoltaic mode. The downside is that noise and *dark current* are worse. Dark current is the current produced even when no light is shining on the photodiode. Ideally this would be zero. A large dark current reduces the effective dynamic range of the device.

Schottky and Varactor Diodes

The Schottky diode is a special purpose device. It is named after Walter Schottky, a German physicist. Unlike other diodes that rely on a semiconductor-to-semiconductor junction, the schottky diode is comprised of a semiconductor-to-metal contact. The Schottky diode exhibits two major advantages over traditional diodes. First, they have very fast switching times, perhaps orders of magnitude of improvement. Second, they exhibit relatively low turn on voltages. Instead of the 0.6 to 0.7 volts seen with a silicon junction diode, a Schottky diode may turn on with as little as 0.2 or 0.3 volts. Consequently, Schottky diodes are used when very fast switching speed and/or minimizing forward voltage drops are important. Examples include shunting diodes in switch mode power supplies and RF detector circuits. Its schematic symbol is shown in Figure 2.30.

Figure 2.30
Schottky diode schematic symbol.

10 As a side note, depending on their construction some LEDs can be used as crude photodiodes. Although they are not optimized for this use it can be entertaining to shine a light on an LED and watch it produce a voltage.

The varactor diode is another special purpose device. Its schematic symbol is illustrated in Figure 2.31. It is used as an electrically controlled capacitance (note that the schematic symbol appears as a hybrid of normal diode and capacitor symbols).

Figure 2.31
Varactor diode schematic symbol.

Varactors are used in reverse-bias mode. The key to understanding their operation is to consider the structure of a diode, comparing it to the construction of a capacitor. Consider the depletion region to be the dielectric of a capacitor with the anode and cathode being the capacitor plates. Consequently, all junction diodes exhibit some capacitance. Normally, designers try to minimize this effect but it is exploited with varactors. As noted in our earlier discussion, increasing the reverse-bias potential on a diode causes its depletion region to widen. All else being equal, increasing the plate separation of a capacitor decreases its capacitance. Thus, by increasing the reverse-bias potential, we increase the effective plate spacing and decrease the diode junction capacitance. We now have a capacitance the value of which is determined by a DC bias voltage. This capacitance can be used as part of electronic tuning circuits for applications such as oscillators and filters. Compared to fixed capacitors the values tend to be small, in the tens to hundreds of picofarads, but it is sufficient for much radio frequency work. The advantages over mechanically adjustable capacitors are manifold, including small size, high reliability, low cost and the ability to rapidly change the capacitance[11].

Summary

In this chapter we have examined the structure and functioning of the PN junction. A PN junction produces a depletion region which is an area devoid of free charges. This leads to an energy hill or barrier voltage, the precise value of which depends on the material used as well as other factors such as temperature. The PN junction is the basis for most diodes. Its current-voltage characteristic is described by the Shockley equation and shows a logarithmic characteristic (i.e., the voltage is proportional to the log of the current).

The terminals of a diode are identified as the anode (P material) and the cathode (N material). If a positive potential which is greater than the barrier voltage is applied from anode to cathode, the diode will conduct current. If the polarity is reversed, the diode will not conduct. Therefore a simple model of the diode is a polarity sensitive switch. Improved models include the forward barrier voltage and the bulk resistance of the diode. Another refinement includes the effect of reverse breakdown, that is, the tendency of a diode to suddenly begin conducting if the reverse-bias potential is

11 The mechanical version would require a rotary-style adjustable capacitor connected to some form of small motor or solenoid to move the capacitor plates. While this can work at lower frequencies, if rapid changes are needed the resulting friction-generated heat may cause this contraption to burst into flames. Generally speaking, this is not something we want our circuits to do.

large enough. For ordinary diodes, the reverse potential should not be allowed to reach breakdown.

Besides the common switching and rectifying diodes, other types are also available. These include the Zener which is normally used in reverse-bias mode. It is commonly used to set or limit a specific voltage. In forward-bias, a Zener behaves like an ordinary diode. LEDs produce light from an electrical input. Their forward potentials tend to be in the neighborhood of a few volts. The photodiode is the complement of the LED and produces a current or voltage that scales with incident light. The Schottky diode is notable for its fast switching speeds and low barrier potential. Finally, the varactor is used as an electrically controlled capacitance. It is used in reverse-bias mode.

Review Questions

1. What is a depletion region?
2. Draw and explain the energy diagram for a PN junction, including the Fermi level.
3. Describe and compare the three diode models.
4. Explain the difference between the effective DC resistance and AC resistance of a diode.
5. List some of the practical differences between switching diodes, Zener diodes and LEDs.

Problems

(Assume diodes are silicon unless stated otherwise)

Analysis Problems

1. For the circuit of Figure 2.32 determine the circulating current if the supply is 6 volts and the resistor is 10 kΩ.

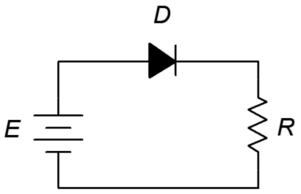


Figure 2.32

2. Repeat Problem 1 if the diode is inserted in the opposite orientation.

3. Given the circuit of Figure 2.33, determine the voltage drops across the resistors. The source is 12 volts, R_1 = 4.7 k and R_2 = 3.3 k.

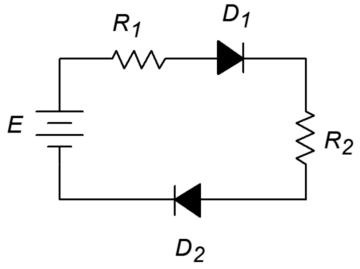


Figure 2.33

4. In Figure 2.34 determine the voltage drops across the resistors.

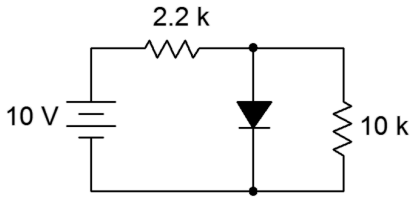


Figure 2.34

5. Determine the LED current in Figure 2.35. Assume the LED barrier is 2.1 volts, the source is 5 volts and the resistor is 330 Ω.

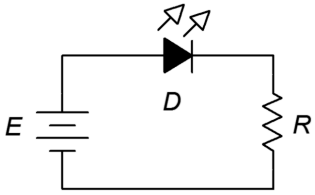


Figure 2.35

6. Repeat Problem 5 if the LED is inserted in reverse orientation.

7. Determine the resistor currents in Figure 2.36. The source is 15 volts, R_1 = 8.2 k and R_2 = 3.9 k.

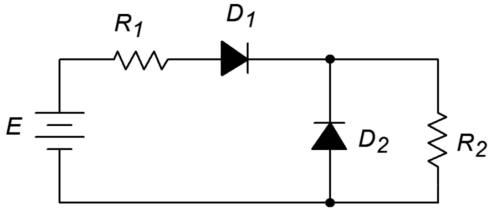


Figure 2.36

8. For the circuit of Figure 2.37, determine the resistor voltage. The source is 9 volts, the Zener potential is 5.1 volts and the resistor is 1 k.

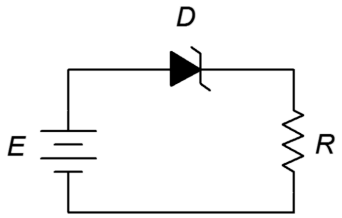


Figure 2.37

9. For the circuit of Figure 2.38, determine the resistor voltage. The source is 8 volts, the Zener potential is 3.3 volts and the resistor is 10 k.

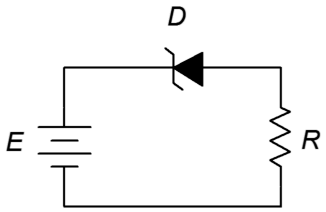


Figure 2.38

10. Determine the voltage across R_2 in Figure 2.39 if the source is 9 volts, the Zener is 6.8 volts, R_1 = 5.1 k and R_2 = 33 k.

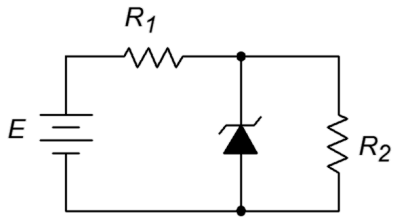


Figure 2.39

Challenge Problems

11. Determine the resistor voltage in Figure 2.40 if E_1 = 5 volts, E_2 = 9 volts and R = 1 k.

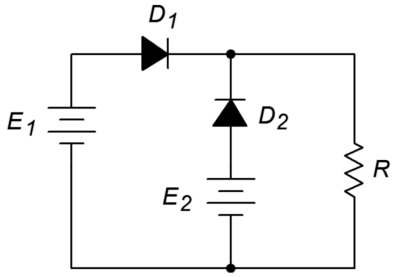


Figure 2.40

12. Determine the voltage across R_2 in Figure 2.39 if the source is 9 volts, the Zener is 5.6 volts, R_1 = 5.1 k and R_2 = 3.9 k.

Design Problems

13. Determine a value for R in Figure 2.32 to set the current to 10 mA if the source is 5 volts.

14. Determine a value for R in Figure 2.35 that will set the LED current to approximately 20 mA if the source is 9 volts and the LED is a standard red type. Use a standard resistor value.

Computer Simulation Problems

15. Simulate Problem 9.

16. Simulate the circuit designed in Problem 14 for verification.

And now for something completely different...

One of the great things about the Internet is that you can find almost anything on it. In contrast, one of the terrifying things about the Internet is that you can find almost anything on it. The following is presented in keeping with the dictum that *"It must be true because I saw it on the Internet"*.

An Alternate Hypothesis Regarding PN Junctions

In order to learn how to design circuits and systems using transistors and other solid state devices, students of electronics are told in their courses how semiconductors function. The atomic structure of crystalline silicon is examined in its intrinsic and doped states. Discussion of energy levels, conduction band electrons and hole production quickly follow. Soon, the student encounters the PN junction, a basic building block of modern electronics, and learns about majority and minority carriers, depletion regions, barrier potentials, leakage current and other exotica. This information is intended to explain just how solid state devices really work and it can end up sounding quite obtuse. It may sound so complicated, in fact, that the student assumes that only a genius could design such devices and relegates him or herself to a lower paying engineering or technician job. This should not be the case!

Recent investigations by Farcebook authors have uncovered some startling facts:

1. Semiconductors don't really work the way we've all been told. In fact, the fundamental theory is much simpler.

2. This lie has been fabricated and perpetuated by the economic and political elite, a group of people with cushy, high paying jobs; jobs so easy in the light of the real theory that even CEOs, hedge fund managers, televangelists, TV psychics and other folk of nil capability could do it in their sleep (well, OK, the executives would still need an army of assistants and the televangelists would continue to be outraged by your private life, but you get the general idea). By making their jobs sound difficult these people get to sit around all day eating eclairs and reading Esquire for an eight figure annual income.

It is high time that the truth be told and this farcical sham be torn down! As an example, we shall see how a simple diode really works.

So, you think a diode is composed of semiconducting material? Think again! One of the chief researchers at Bell labs in the 1940s and 50s was a certain Doctor Schlocking. After several experiments involving solid state diodes, Dr. Schlocking wrote in his diary:

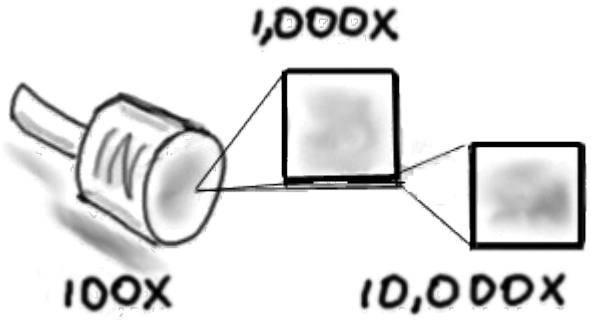
> *"This stuff don't work at all. Better go back to tubes before they can me. Ooops, must let the dog out."*

Schlocking was often pestered by his dog Melvin who reminded him of a small self-propelled dust mop, and who was approximately as clean. The diary continued:

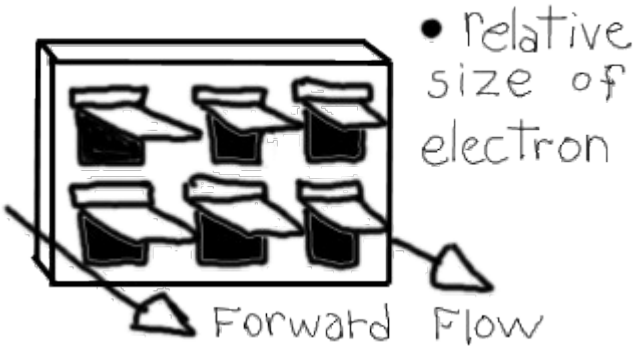
> *"If only there was a way in which Melvin could let himself out. Even better...if he couldn't get back in!"*

This ominous tone led to Schlocking's invention of the now infamous one-way *doggie door* that can be seen mounted to the bottom of normal doors across the country. Schlocking knew how to milk an idea and took the form to its extremes by developing the *cat door*, the *mouse door*, the *grasshopper door*, the *flea door* (an early attempt at a flea and tick collar) and even the *amoeba door*. This last unit when properly designed could force microscopic parasites and bacteria out of the human body and not allow them to re-enter. This was instrumental in the development of the polio vaccine, in spite of the fact that the vaccine was produced some years earlier. Schlocking's greatest achievement, however, came when he shrunk the doggie door still further to produce the *electron door*. This is the fundamental unit of modern electronics.

In the figure below, we see a cross section of a diode and close-ups.



Even at 10,000X magnification we can still see nothing of the PN junction. If we go a bit further, something interesting comes into focus (see the second figure).



Yes! A PN junction is nothing more than a huge array of real tiny one-way doggie doors! Here's how it works: Electrons are a lot like marbles. When one hits a doggie door from behind, the door flips open allowing the marble through (i.e., allowing current to flow). If the electron hits the doggie door from the front, the flap closes and the electron can't get through (i.e., no current flow). Now obviously, if we hang the diode vertically, gravity should open all the doors and we'll get lots of electrons (i.e., current flow) in either direction. In truth, a real diode doesn't do this. Its operation will not matter on how the diode is oriented in space. This feature is

accomplished by simply adding a small coil spring to the doggie door's hinge, forcing it to stay shut in the face of gravity. This has the negative side effect of requiring somewhat higher energy levels from the electrons to force the door open. This force happens to be the barrier potential of the diode! It has nothing to do with so-called *depletion regions*. If you were an electron, would you want to go through a place called a depletion region? Of course not! Neither would electrons. They're not stupid, you know. In any case, the stronger the spring, the greater the barrier potential. Presently diodes are made of either silicon or germanium with barrier potentials of approximately 0.7 volts or 0.3 volts, respectively. In fact, *silicon* and *germanium* are really code words meaning *strong spring* and *weak spring*! It took a while to develop small, strong springs and this is why germanium diodes were the first ones built.

Note that spring strength also plays a role in how tightly the flap can shut thus indicating the reverse leakage current. Here again we see strong spring “silicon” units having lower leakage. Theory also indicates that leakage should increase with temperature. This effect can be seen clearly in the doggie door model. At present it is impossible to create both the frame and the door out of precisely the same material and thus two different expansion coefficients exist. Because the flap is smaller than the frame it will tend to curl away at higher temperatures allowing more electrons to sneak through the gaps. At very low temperatures the flap tends to stick to the frame in much the same fashion that your tongue or lips will stick to a metal flag pole in freezing weather (also known as the *Christmas Story phenomenon*).

At very high forward energy levels the flaps may be literally torn off their hinges. This high volume of electrons at high energy will yield the *maximum forward current*. Also, note that if the energy level is high enough in the reverse direction, either the flaps will be bent and pushed through the frames or they will start to bounce violently at resonance, allowing electrons through. These two modes are referred to as *avalanche* and *Zener conduction*, respectively. The required energy level indicates the *reverse breakdown voltage*.

Other fine points can be explained equally by the doggie door model, as well as bipolar and field effect transistors, IGBTs and just about everything else in the field of solid state electronics with the exception of the original 7400 series TTL logic gates which utilized an array of small, edible fungi and miniature harvester ants.

More on that in a future exposé.

3 Diode Applications

3.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Solve basic AC rectifier circuits for resulting waveforms.
- Detail the differences between half-wave, full-wave and full-wave bridge diode rectifier configurations.
- Solve basic regulator circuits employing Zener diodes.
- Outline a complete AC-to-DC power supply with regulation, describing the function of each component, including power transformer.
- Solve AC clipper circuits for output waveforms.
- Solve AC clamper circuits for output waveforms.

3.1 Introduction

The preceding chapter was concerned primarily with introducing the practical considerations of diodes while presenting them in DC circuits. This chapter will extend the discussion by focusing on AC circuit applications. A prime example is AC to DC conversion, the concept behind most electronic power supplies. It also includes the basics behind regulation and limiting/level shifting circuits such as clippers and clampers. The inherent asymmetry in the conductance of diodes, that is, their sensitivity to the direction of current flow, is what makes these circuits possible. Non-ideal effects such as a diode's forward voltage drop might be ignored in some instances but may be quite important in others.

3.2 Rectification

Rectification is the process of turning an alternating current waveform into a direct current waveform, i.e., creating a new signal that has only a single polarity. In this respect it's reminiscent of the common definition of the word, for example where “to rectify the situation” means “to set something straight”. Before continuing, remember that a DC voltage or current does not have to exhibit a constant value (like a battery). All it means is that the polarity of the signal never changes. To distinguish between a fixed DC value and one that varies in amplitude in a regular fashion, the latter is sometimes referred to as *pulsating DC*.

The concept of rectification is crucial to the operation of modern electronic circuits. Most electronic devices such as a TV or computer require a fixed, unchanging DC voltage to power their internal circuitry. In contrast,

residential and commercial power distribution is normally AC. Consequently, some form of AC to DC conversion is required[12]. This is where the asymmetry of the diode comes in.

Half-wave Rectification

To understand the operation of a single diode in an AC circuit, consider the diagram of Figure 3.1. This is a simple series loop consisting of a sine wave source, a diode and a resistor that serves as the load. That is, primarily we will be interested in the voltage developed across the resistor.

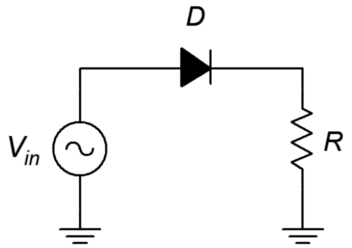


Figure 3.1
Basic AC diode-resistor circuit.

For positive portions of the input wave, the diode will be forward-biased. To a first approximation it will appear as a closed switch. Consequently, all of the input signal will drop across the resistor. In contrast, when the input signal switches to a negative polarity on the other half of the waveform, the diode will be reverse-biased. Therefore, the diode acts as an open switch. The circulating current drops to zero thereby producing no voltage across the resistor. All of the applied potential drops across the diode, as indicated by Kirchhoff's voltage law (KVL). The input and load resistor's voltage waveforms can be seen in Figure 3.2.

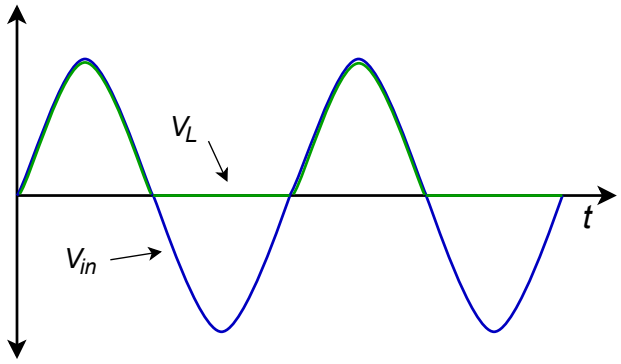


Figure 3.2
Half-wave rectification waveforms.

12 If you're wondering why we don't just use DC distribution instead in order to “cut out the middle man”, the reasons are manifold. First, it is generally more efficient to distribute power via AC rather than DC. Second, even if DC is available, it may not be at the amplitude the circuitry requires. Therefore some form of DC-to-DC conversion would be needed. Depending on the application, this can turn out to be more expensive than AC-to-DC conversion.

The resulting signal seen across the load resistor is a pulsating DC waveform. We have effectively removed the negative half of the waveform leaving just the positive portion. Because only half of the input waveform makes it to the load, this is referred to as *half-wave rectification*.

It is worth noting that if the AC peak input voltage is not particularly large, there can be an obvious discrepancy between the peak levels of the input and load signals. For example, if the peak input voltage is in the range of three or four volts and a silicon diode is used, the resulting waveforms would look more like Figure 3.3.

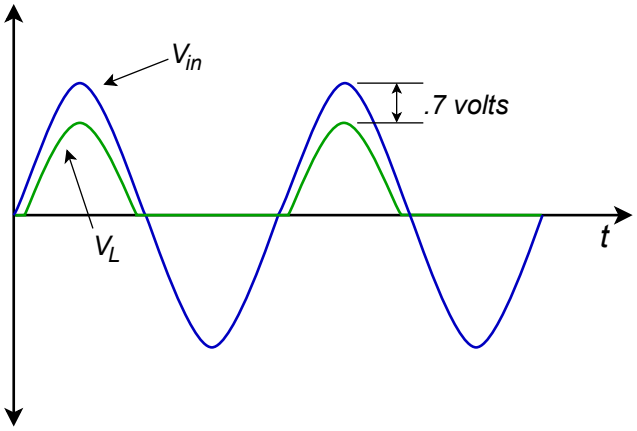


Figure 3.3
Half-wave rectification waveforms including forward diode drop.

In this case the 0.7 volt forward drop cannot be ignored as it represents a sizable percentage of the input peak. The positive pulses are also slightly narrowed as current will not begin to flow at reasonable levels until the input voltage reaches 0.6 to 0.7 volts.

If the diode was oriented in reverse, it would block the positive portion of the input and allow only the negative portion through. In this instance the load waveform would appear flipped top to bottom compared to Figures 3.2 and 3.3.

Computer Simulation

A simulation schematic for a simple half-wave rectifier is shown in Figure 3.4. A sine wave source of 10 volts peak is used to feed a popular 1N4000 series rectifier diode connected to a 100 Ω load. The source frequency is 60 hertz, the North American standard for power distribution.

A transient analysis is run resulting in the waveforms shown in Figure 3.5. The source voltage waveform is shown in red while the load voltage waveform is depicted in blue. While the half-wave rectification is obvious, the loss due to the forward voltage drop of the diode is clearly evident. Based on the vertical scale, a value just under one volt would be a reasonable estimate. The simulation agrees

nicely with the expected result as drawn in Figure 3.3, although not as extreme due to the increased source voltage.

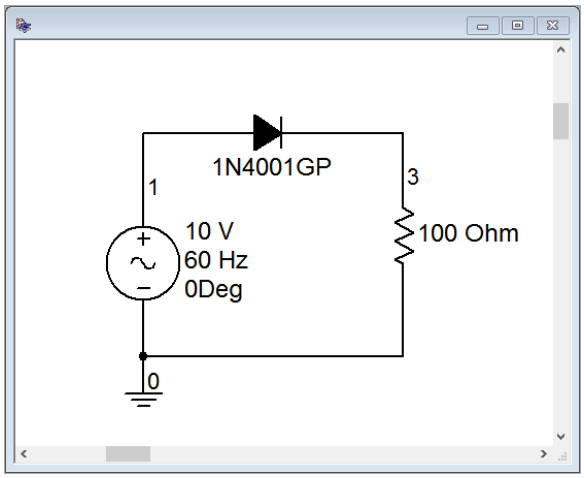


Figure 3.4
Simulation schematic for half-wave rectifier.

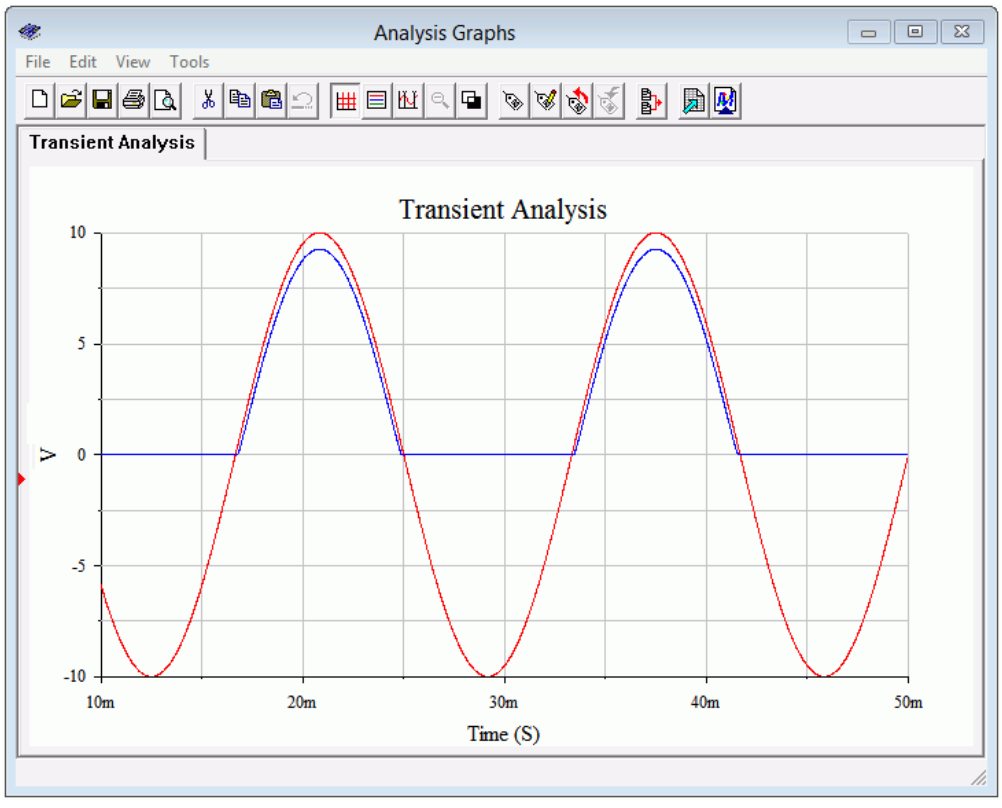


Figure 3.5
Transient analysis for half-wave rectifier.

On a practical note, there are still two items to consider when it comes to converting AC to DC. The first item is the issue of scaling the 120 VAC RMS outlet voltage to a more useful level. In many cases this means lowering the voltage although there are some applications such as high power amplifiers where the voltage will need to be increased. The second item involves smoothing the pulsating DC to produce a constant value, much like a battery.

A Note Regarding Transformers

The aforementioned voltage scaling issue can be addressed through the use of a transformer. While a complete exploration of transformers is beyond the scope of this chapter, we can present the basics. In simple terms, a transformer has an input side, or *primary*, and an output side, or *secondary*. Each side is made up of a coil of wire and these coils are wound around a common magnetic core. The current in the primary-side coil creates a magnetic flux in the core. This flux induces a current in the secondary coil. Ideally, the voltage is decreased and the current is increased by the ratio of the number of loops between these coils. For example, if the secondary-side coil has half as many turns as the primary-side coil then the secondary voltage will be half of the primary voltage and its current will be twice as large as the primary current. This implies that in the ideal case there is no power lost within the transformer. It simply transforms the power from high-voltage/low-current to low-voltage/high-current (or vice versa), hence the name. In reality, transformers do have voltage and current limits, and they are specified in terms of a volt-amp or VA rating which is simply the product of the nominal secondary voltage and maximum allowed secondary current. Transformers that decrease the voltage are referred to as *step-down* while those that increase the voltage are referred to as *step-up*. Finally, it is possible to create transformers with multiple primaries and secondaries (via either separate coils or multi-tapped coils). The resulting series and parallel coil configurations make them much more flexible.

Smoothing (Filtering) the Output

The second issue we have is smoothing and leveling the pulsating DC. The most straightforward method to achieve this is to add a capacitor in parallel with the load. The capacitor will charge up during the conduction phase, thus storing energy. When the diode turns off, the capacitor will begin to discharge, thus transferring its stored energy into the load. The larger the capacitor, the greater its storage capacity and the smoother the load voltage will be. It turns out that there is a down side to large capacitors, as we shall see. Consequently, the goal will not be to use as large of a capacitor as possible but rather to use an optimal size for a given application. A half-wave rectifier with transformer and capacitor is shown in Figure 3.6.

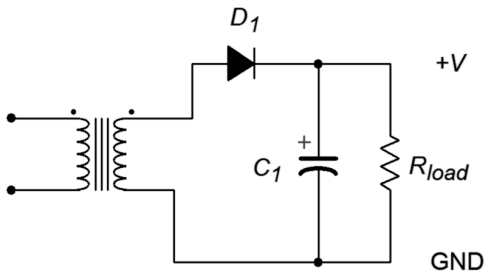


Figure 3.6
Half-wave rectifier with transformer and filter capacitor.

One way of looking at the inclusion of the smoothing capacitor is to consider that it, along with the load resistance, make up an *RC* discharge network. To achieve a smooth load voltage the discharge time constant should be much longer than the gap produced when the diode turns off. For 60 hertz operation, this gap is half of the period, or roughly 8.3 milliseconds. The time constant equation is

$$\tau = RC$$

Recalling that in one time constant the capacitor voltage will fall to well below half of the starting value (roughly 37%), we will need a time constant several times larger than 8.3 milliseconds. For example, suppose our effective load resistance is 100 Ω. If we use a 1000 µF capacitor, the resulting time constant would be 100 milliseconds, or over ten times the gap duration. A much smaller capacitor, say around 50 µF, would not be nearly so effective at keeping the voltage constant.

The variation in output voltage due to capacitor discharge is referred to as *ripple*. It can be modeled as an AC voltage riding on a larger DC output. The magnitude of the ripple worsens as the load current increases. Under light load conditions, the output will tend to float to the peak voltage of the secondary with very little ripple. As load current demand goes up, the ripple magnitude increases and the nominal output voltage begins to drop.

Computer Simulation

Two variations on a filtered half-wave rectifier are simulated below. Both versions use a 100 Ω load with a 10 volt source, similar to the prior simulation. The first version uses a 50 µF filter capacitor while the second ups this to 1000 µF. In both cases a 1 Ω resistor is added in series with the capacitor to serve as a current sensor. The first version is shown in Figure 3.7.

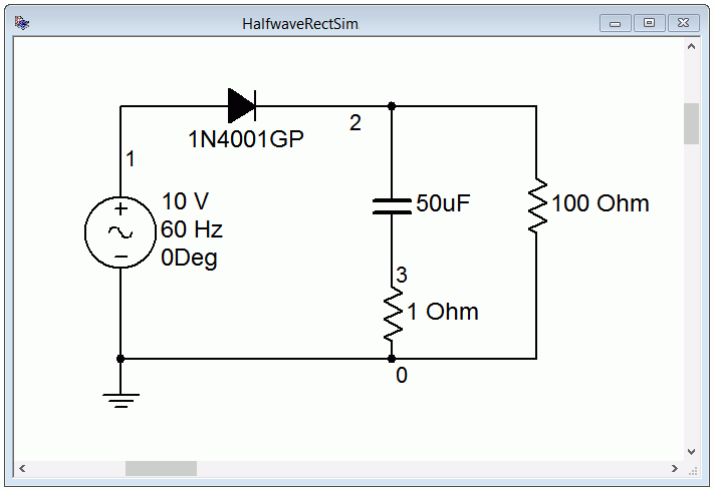


Figure 3.7
Simulation schematic for half-wave rectifier with 50 µF filter capacitor.

A transient analysis simulation graph is shown in Figure 3.8. The input waveform is colored blue while the load voltage is red. Comparing this waveform to that depicted in Figure 3.5 shows the effect of the capacitor stretching out the pulse and partially filling in the gap. It is obvious that this capacitor is too small given the load resistance and the resulting current demand. Indeed, by the time the next pulse arrives the capacitor is nearly depleted and the output voltage has dropped to around one volt.

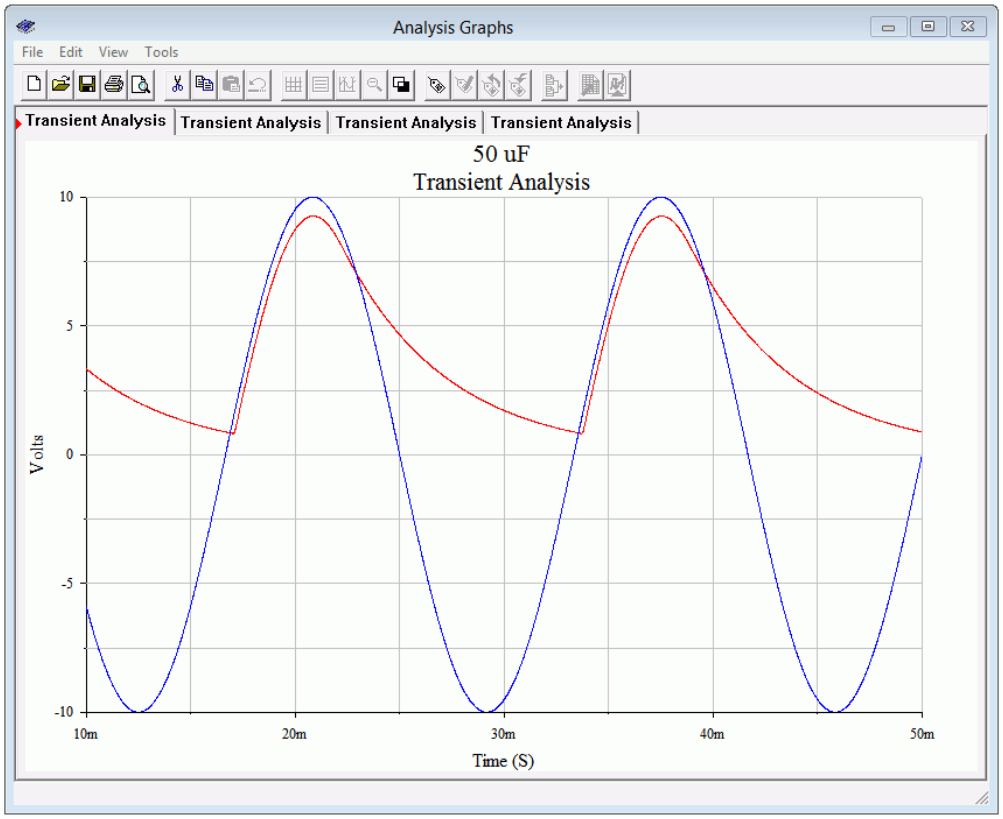


Figure 3.8
Transient analysis simulation for half-wave rectifier using a 50 μF filter capacitor.

In Figure 3.9 the simulation is rerun, but this time using a 1000 μF capacitor in place of the 50 μF. As expected, the increased *RC* time constant results in a much more stable load voltage. In this version the output has dropped from a little over nine volts to about eight volts yielding a peak-to-peak ripple of a volt and a half or so. The peak voltage of just over nine volts versus the applied ten volts is largely due to the voltage drop across the rectifying diode.

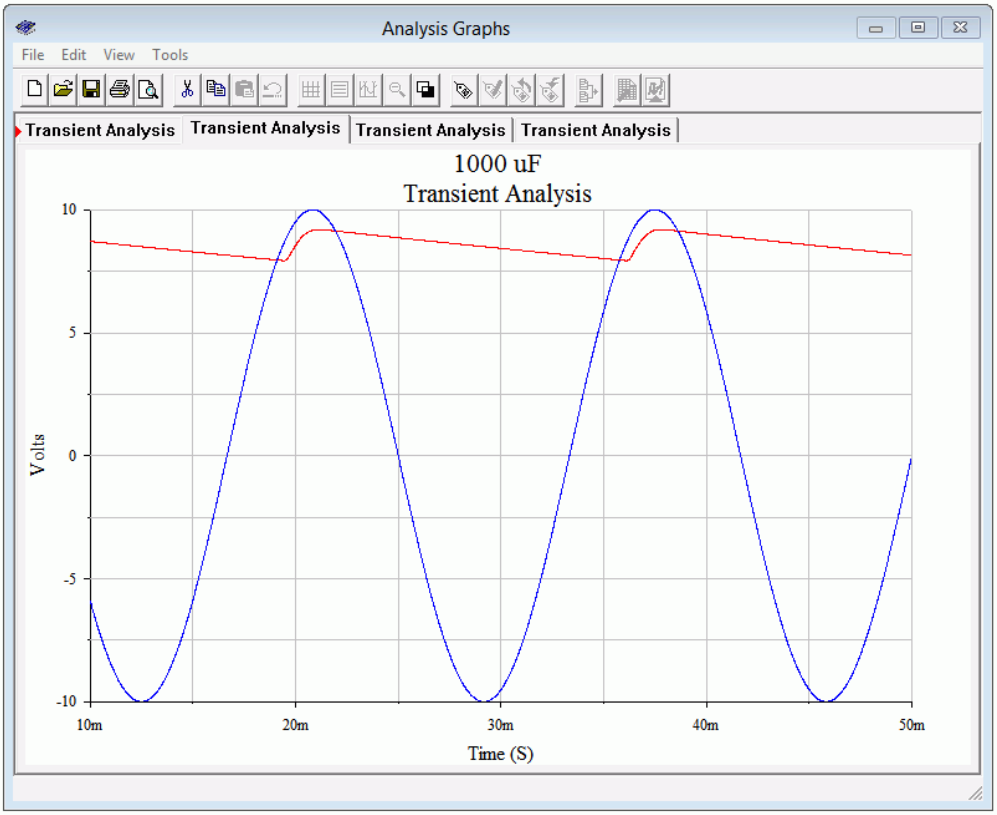


Figure 3.9
Transient analysis simulation for half-wave rectifier using a 1000 µF filter capacitor.

One thing that may not be apparent immediately is that the charge time for the larger capacitor is much shorter than for the smaller unit. This is perhaps counterintuitive. With a larger capacitor, the diode turns on for a shorter time because its cathode is held at a high voltage due to the capacitor. That is, it will only turn on when the input voltage exceeds the capacitor voltage by roughly 0.7 volts. It is only during this time that the capacitor will be replenished, and this can lead to very large current spikes.

To investigate this effect, the simulations are rerun, but this time adding the voltage across the 1 Ω sensing resistor. This relatively small value will have only a modest effect on the charging and discharging, and conveniently scales to the current value (i.e., 100 millivolts signifies 100 milliamps). First, examine the transient simulation of Figure 3.10 using the 50 µF capacitor.

The red sweep is the output voltage while the blue sweep represents the capacitor current. The output voltage plot uses the left vertical axis while the current plot uses the right vertical axis. As the load voltage begins to rise, we see an abrupt spike in the capacitor current. This is current charging the capacitor and it peaks at about 180 milliamps. The total time for the charge phase is around 4 milliseconds. Once the output voltage peaks, the capacitor starts to discharge into the load. During the discharge phase note that the capacitor current's polarity has reversed. It is negative, peaking at roughly −80 milliamps, and delivering current to the load.

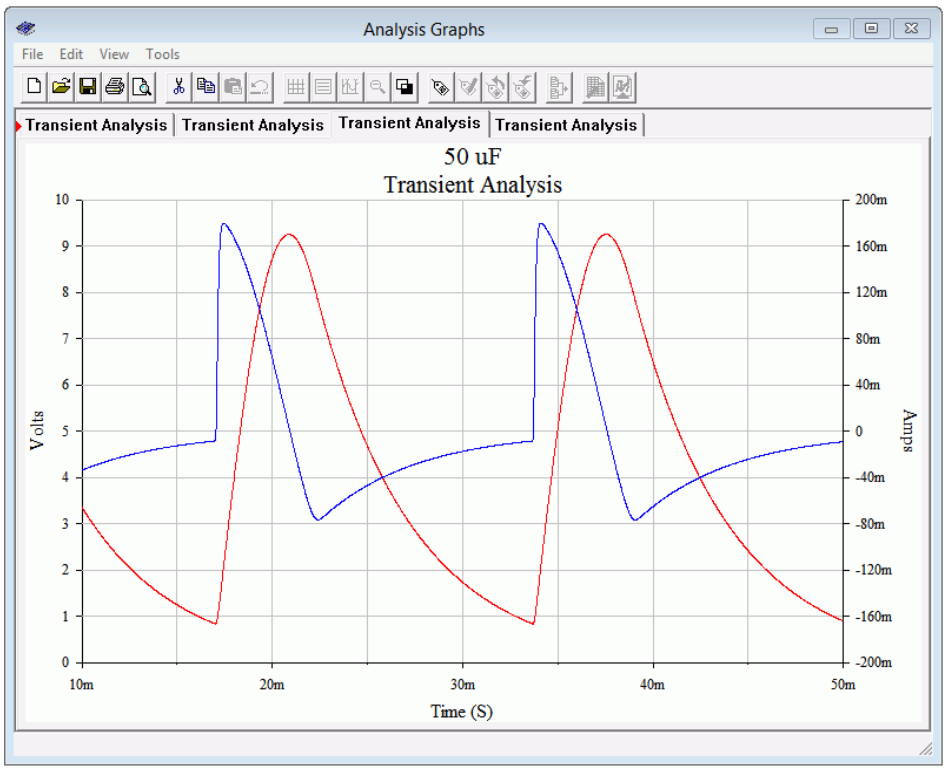


Figure 3.10
Transient analysis current waveform using a 50 µF filter capacitor.

This simulation is repeated using the 1000 µF capacitor. The results are shown in Figure 3.11.

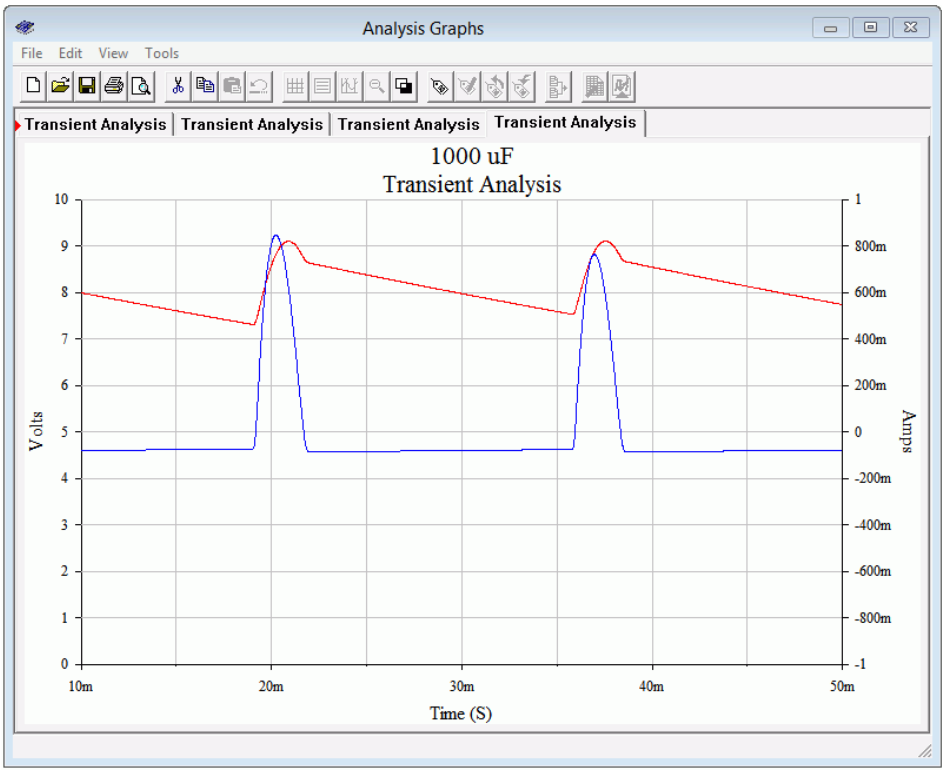


Figure 3.11
Transient analysis current waveform using a 1000 µF filter capacitor.

The blue current waveform peaks at approximately 800 milliamps, or over four times the value compared to using the smaller capacitor. Also, the width of the positive pulse has decreased to about 2.5 milliseconds. The discharge phase is nearly flat, implying that the output voltage must be more stable as this capacitor is the only source for load current during this phase.

Full-wave Rectification

An improvement on half-wave rectification is *full-wave rectification*. Half-wave rectification is inefficient because it essentially throws away the negative portion of the input. In contrast, full-wave rectification makes use of the negative portion by inverting or flipping its polarity. The resulting circuit is modestly larger and more complicated but results in large performance improvements. For example, filter capacitor size is greatly reduced.

There are two popular methods to achieve full-wave rectification. The first method uses a pair of diodes with a center-tapped (i.e., split) secondary. The second method uses a four diode bridge network. The diode bridge form is also capable of producing a bipolar output (i.e., a positive output along with a negative output, typically of the same magnitude).

The two diode center-tapped secondary circuit is shown in Figure 3.12. This schematic also includes the filter capacitor.

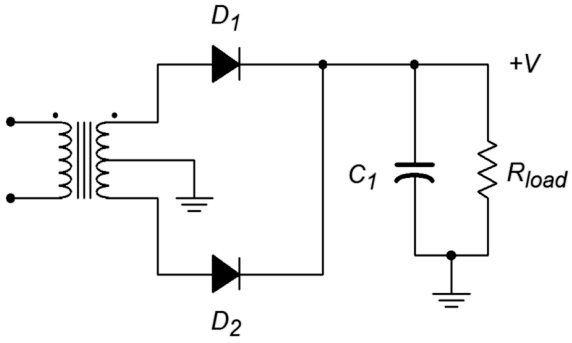


Figure 3.12
Full-wave center-tapped rectifier with capacitor.

The operation is as follows. During the positive half of the source voltage diode D_1 is forward-biased while D_2 is reverse-biased. Therefore the upper half of the secondary behaves like a simple half-wave rectifier allowing current to flow through D_1 and into the load. Due to the reverse-bias on D_2, the lower half presents an open circuit and is effectively removed. In mirror fashion, when the applied potential switches polarity D_1 will be reverse-biased while D_2 becomes forward-biased. Current is now free to flow through D_2 into the load. Thus, both halves of the input waveform are used. The resulting waveforms are illustrated in Figure 3.13. For clarity, the filtering effect of the capacitor is not shown and V_{in} represents one half of the total secondary voltage.

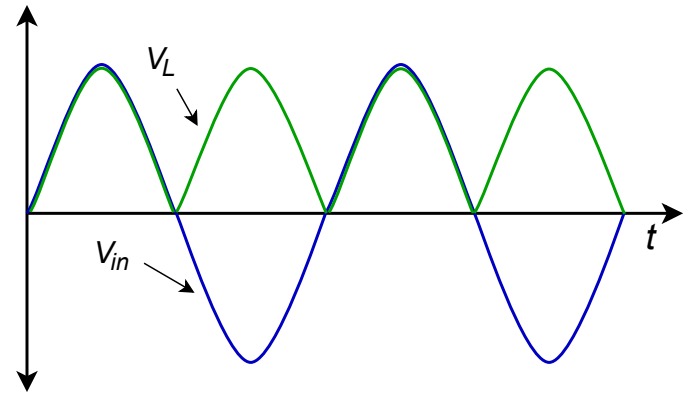


Figure 3.13
Full-wave rectifier waveforms.

An important point to remember about this configuration is that the load only "sees" half of the secondary at any given time. Therefore, the load voltage will only be half of the total secondary voltage (minus one forward diode drop). For example, if the transformer has a 10:1 turns ratio and is being fed from a standard 120 volt source, the secondary will produce 12 volts RMS. Ignoring the diode drop, the load would see half of this, or 6 volts RMS (about 8.5 volts peak). Typically, transformers are rated by their total secondary voltage so this transformer would be referred to as having a "12 volt center-tapped secondary".

A four diode bridge rectifier is shown in Figure 3.14. A filter capacitor is included. Also, note the usage of a standard, non center-tapped secondary. As this is a very common configuration, the four diode bridge is available as a single four-lead part in a variety of sizes and current capacities.

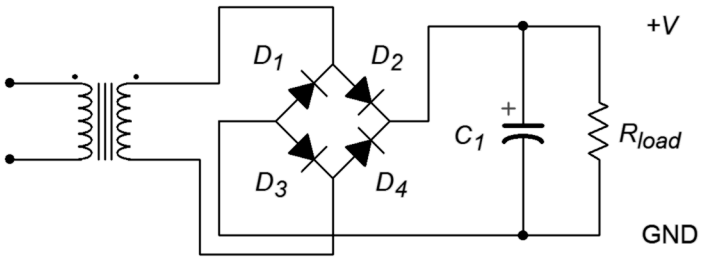


Figure 3.14
Full-wave bridge rectifier with capacitor.

The operation of this circuit is illustrated in Figure 3.15 for the positive portion of the input. First, current flows from the top of the secondary to the D_1/D_2 junction. Only D_2 offers a forward-bias path so current flows through D_2 to the junction with D_4 and the load. As D_4 presents a reverse-bias path, current must flow down through the load. From ground, current continues to the D_1/D_3 junction. Although at first glance it appears that current could flow through either diode, remember that the cathode of D_1 is tied to the high side of the secondary. Therefore, its potential must be higher than the anode side, making it reverse-biased. Consequently, the current flows down through D_3. A similar situation occurs at D_4 and current is directed back to the low side of the secondary. In short, D_2 and D_3 are forward-biased while D_1 and

D_4 are reverse-biased. The load sees the entire secondary voltage minus two forward diode drops.

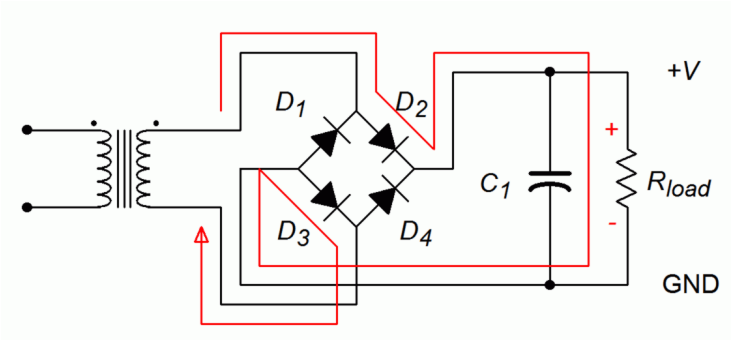


Figure 3.15
Full-wave bridge rectifier analysis, positive input.

During the negative polarity portion of the input the situation is reversed as illustrated in Figure 3.16. Current will flow from the bottom of the secondary through D_4, down through the load, and finally back to the top of the secondary via D_1. Thus, D_1 and D_4 are forward-biased while D_2 and D_3 are reverse-biased. The important thing is that in both cases, the current flows down through the load, top to bottom, resulting in a positive output voltage.

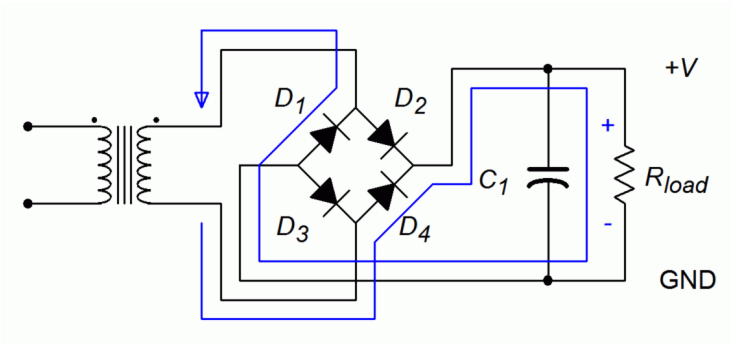


Figure 3.16
Full-wave bridge rectifier analysis, negative input.

Example 3.1

Design a rectifier/filter that will produce an output voltage of approximately 30 volts with a maximum current draw of 300 milliamps. It is to be fed from a 120 VAC RMS source. The ripple voltage should be less than 10% of the nominal output voltage at full load.

For this design we shall focus on using common off-the-shelf parts. As we have seen, the full-wave rectifiers are more efficient at converting AC to DC so we shall go that route, specifically, a four diode bridge arrangement. We will use the circuit of Figure 3.14 as a guide.

The first item to consider is the size of the transformer. A 30 volt output would require a peak secondary voltage of at least 32 volts as we must add in two forward diode drops. The equivalent RMS value is 32/√2 or 22.6

volts. At full load the filtered output voltage will droop somewhat so a somewhat larger value is called for. A standard 24 volt secondary should suffice. Given the 300 milliamp load current rating, the transformer must be at least 0.3 amps · 24 volts or 7.2 VA.

As far as the capacitor is concerned, it must be rated for the peak voltage. The peak equivalent is 24 VAC RMS · √2 or 34 volts. Although a 35 volt rated capacitor might be tried, a standard 50 volt rating would leave a generous safety margin and increase reliability. To find the capacitance value we must first find the effective worst case load impedance.

$$R = \frac{V_{out}}{I_{max}}$$
$$R = \frac{30\,\text{V}}{0.3\,\text{A}}$$
$$R = 100\,\Omega$$

It will be useful to compare this back to the simulation depicted in Figure 3.9. Our ripple specification is somewhat tighter than that achieved in the prior simulation. This is apparent by noting how far the output voltage has dropped by midway through the off portion of the cycle. Consequently, we will need a larger time constant, perhaps by a factor of two. That puts us at 200 milliseconds.

$$\tau = RC$$
$$C = \frac{\tau}{R}$$
$$C = \frac{0.2\,\text{s}}{100\,\Omega}$$
$$C = 2000\,\mu\text{F}$$

A 2200 μF standard value should be sufficient.

Computer Simulation

To verify our results, the design from Example 3.1 is simulated. The schematic is shown in Figure 3.17. To simplify the simulation, a 24 volt RMS source is used in place of the transformer. The worst case load is simulated via a 100 Ω resistor. For the initial test the filter capacitor is omitted so that we can ensure the proper peak voltage and waveforms are created. The results of a transient analysis are shown in Figure 3.18. The secondary voltage is shown in red while the load voltage is shown in blue. The full-wave waveform is exactly as expected, including a slight reduction in the peak voltage value due to two forward diode drops. The output peak is just above 30 volts, as desired.

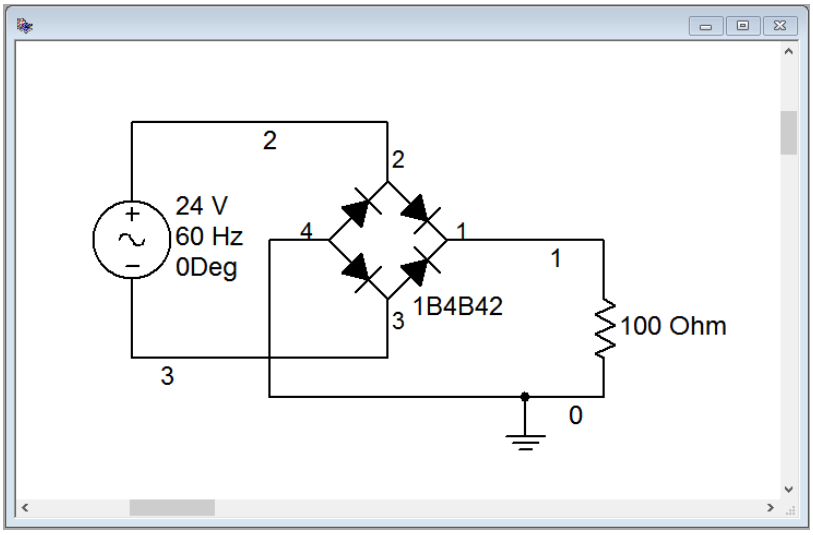


Figure 3.17
Simulation schematic for the design of Example 3.1 without capacitor.

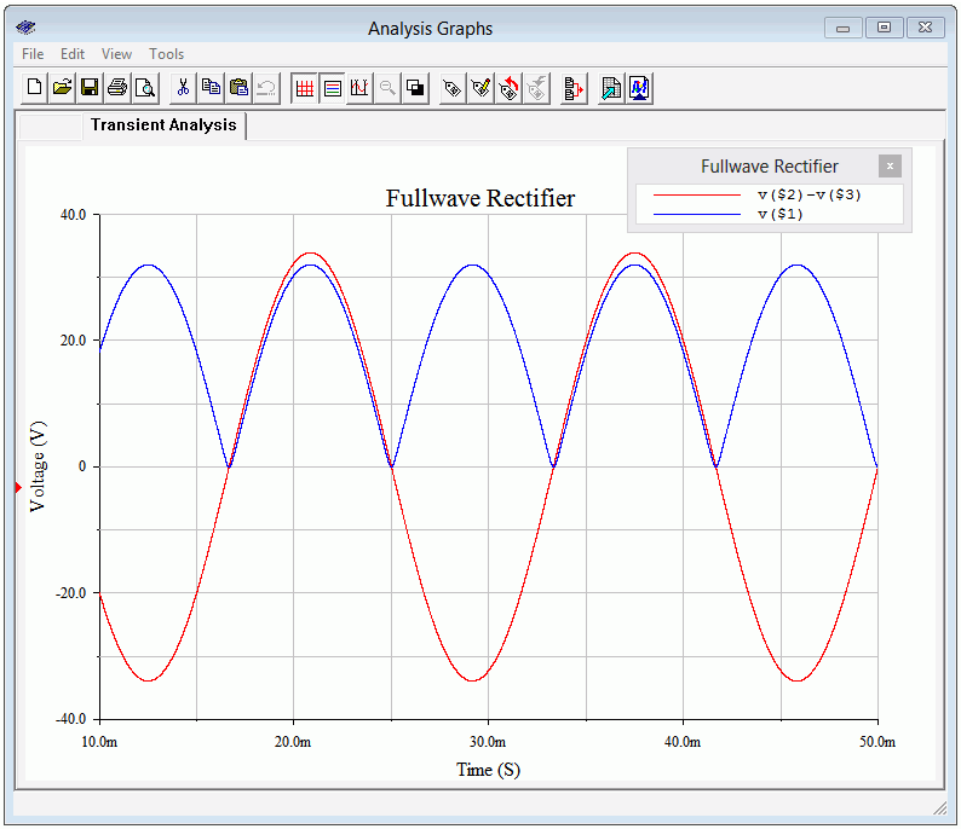


Figure 3.18
Transient analysis of the design of Example 3.1 without capacitor.

Now that we have confidence in the voltage level and waveform, the output filter capacitor is added as shown in Figure 3.19. A transient analysis is run again with the resulting input and load voltage waveforms depicted in Figure 3.20. The load voltage is shown in red. The average value is just over 30 volts and the peak-to-peak ripple is less than two volts, as desired. Note that the full-load peak voltage with the capacitor is slightly less than what was seen in the capacitor-less version. If the load current demand were to increase, both droop and ripple would get worse.

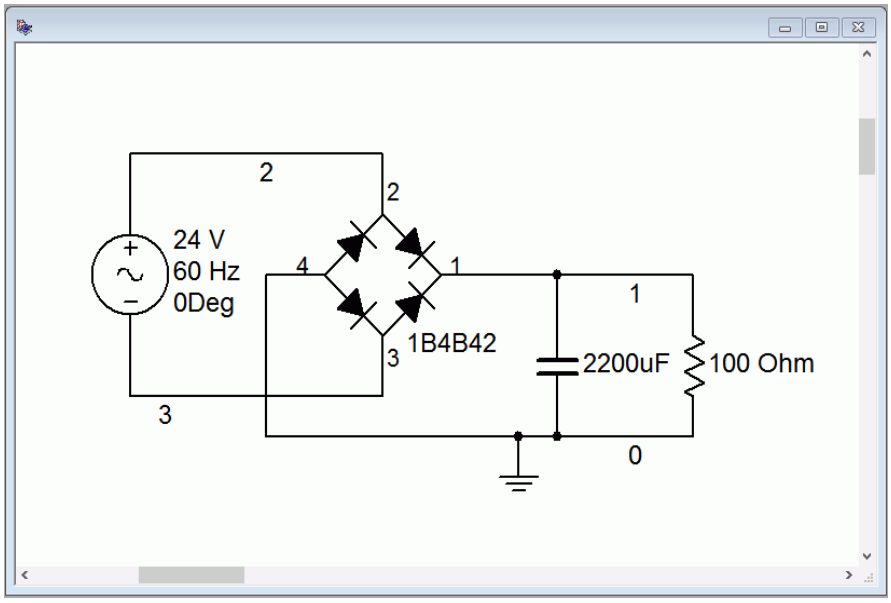


Figure 3.19
Simulation schematic for the design of Example 3.1 with capacitor.

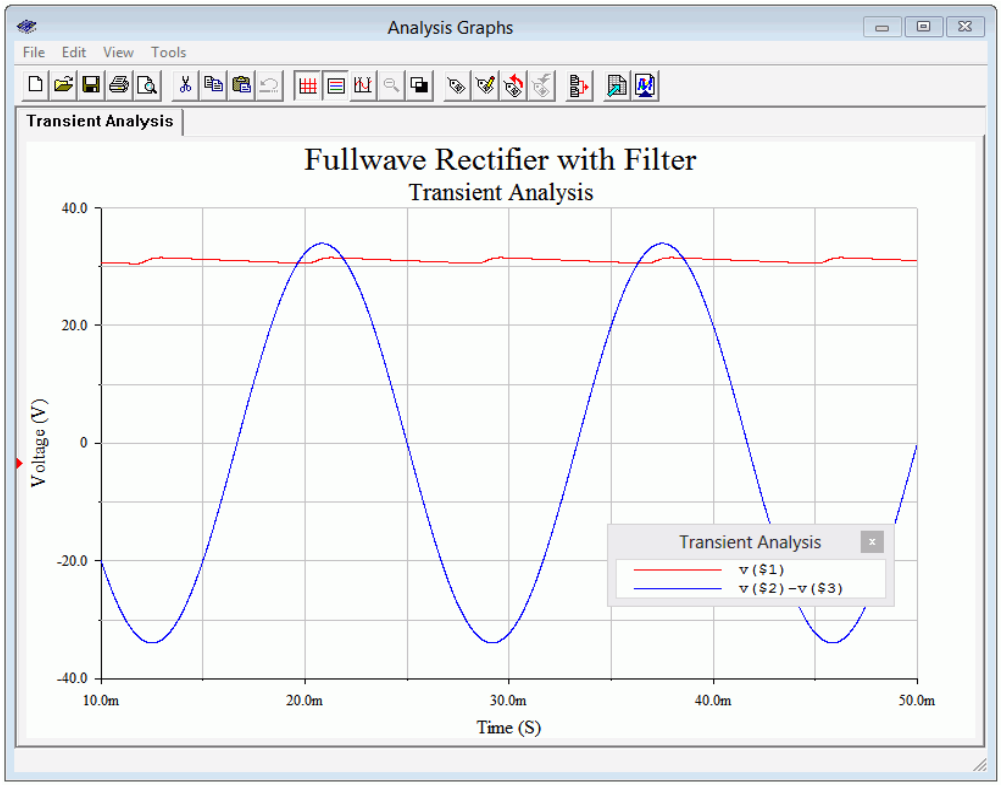


Figure 3.20
Transient analysis of the design of Example 3.1 with capacitor.

Full-wave Bridge With Dual Outputs

As mentioned, the full-wave bridge can be configured to create a dual output bipolar supply. This is shown in Figure 3.21. Note the inclusion of the center tap on the secondary of the transformer and the location of the ground connection between the two loads and their associated capacitors.

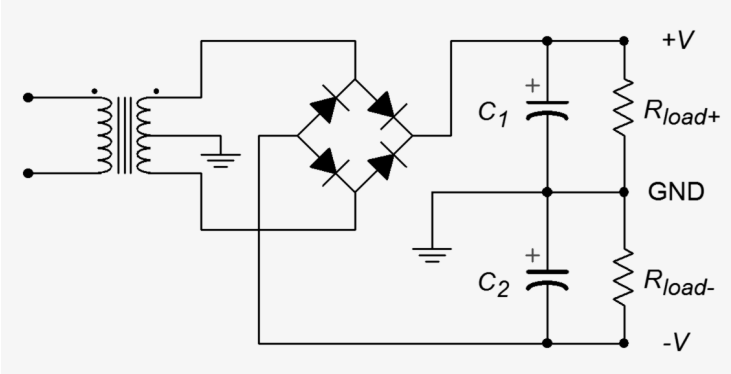


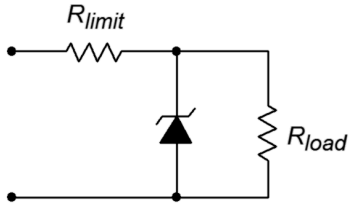
Figure 3.21
Dual output full-wave bridge rectifier.

One way of thinking of this is that we have simply created a new reference point, splitting in half the total output potential of the circuit presented in Figure 3.14. Alternately, it can be thought of as the upper half of the secondary driving R_{load+} while the bottom half drives R_{load-}, as if the bridge and two-diode versions were somehow combined in a transporter accident, as in the 1958 movie The Fly, although it doesn't scream "Help me! Help me!" in a tiny little voice at the end.

Zener Regulation

Adding a large capacitor to a rectifier is necessary to store and transfer energy so that a smooth, ideally non-varying voltage results. As noted previously, under heavy load the ripple would increase in amplitude and the average voltage would drop. This issue can be greatly reduced by adding a Zener diode and current limiting resistor to the output, following the capacitor. This is called a *Zener regulator* and is shown in Figure 3.22.

Figure 3.22
Simple Zener regulator.



The operation of the Zener regulator is fairly straightforward. Recall that when reverse-biased with a sufficiently large potential, the normal reverse diode behavior of an open switch abruptly changes to maintain a fixed voltage; the Zener potential. The current through the diode begins to increase dramatically once this potential is reached. If we place a Zener diode across the output of our filtered rectifier, the Zener will attempt to limit the output voltage to the Zener potential. To prevent excessive and possibly destructive current draw by the Zener diode, the voltage difference between the capacitor voltage and the Zener potential is dropped across a series current limiting resistor. This limiting resistor will set the maximum amount of output current. This current is then split between the Zener diode and the load. Under light load conditions, most of this current will flow through the Zener diode. Under heavy load conditions, most of the current will be drawn by the load with little flowing through the Zener diode. If the load current demand is too heavy, no current is available for the Zener diode and it stops conducting. Regulation is lost and the limiting resistor forms a voltage divider with the load.

A complete rectifier/filter/Zener regulator circuit is show in Figure 3.22. Let's examine how R_{limit} interacts with the load.

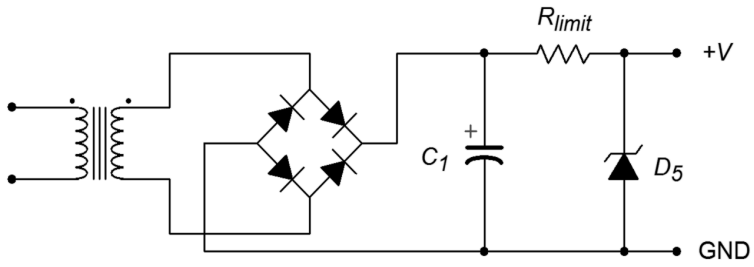


Figure 3.22
Full-wave bridge rectifier and filter with Zener regulator.

For proper operation, the Zener potential (D_5) is the desired DC output voltage and the peak secondary voltage is set somewhat higher. We wish to guarantee that under full load conditions the lowest capacitor voltage due to ripple is still greater than the desired DC output voltage. The difference between the capacitor voltage and the Zener potential drops across R_{limit}. Therefore

$$I = \frac{V_{cap} - V_Z}{R_{limit}}$$

Under no-load conditions all of this current flows down through the Zener diode. The maximum load current is equal to this value (at which point no current flows through the Zener diode).

Example 3.2

Determine the maximum load current for a DC supply such as that found in Figure 3.22. The capacitor voltage is 15 volts average with ±1 volt of ripple (i.e., 16 volts dropping to 14 volts). The Zener potential is 12 volts and R_{limit} is 4.7 Ω.

The highest possible continuous load current is the current through R_{limit} (ignoring I_{ZT}). The limiting case for continuous draw will occur when the capacitor voltage is at its lowest value, or 14 volts.

$$I = \frac{V_{cap} - V_Z}{R_{limit}}$$
$$I = \frac{14\,\text{V} - 12\,\text{V}}{4.7\,\Omega}$$
$$I = 426\,\text{mA} \quad \text{(actually a few mA less due to } I_{ZT}\text{)}$$

The highest peak current through the Zener diode is found at the maximum capacitor voltage and assumes no current is drawn by the load.

$$I = \frac{V_{cap} - V_Z}{R_{limit}}$$
$$I = \frac{16\,\text{V} - 12\,\text{V}}{4.7\,\Omega}$$
$$I = 851\,\text{mA}$$

Note that this worst case current times the Zener potential results in a power dissipation of about 10 watts. Of course, during normal operation with a load drawing current, the diode dissipation is much reduced. It is interesting to note that the Zener dissipates maximal power when the load current is zero. Consequently, we can think of this circuit as shifting current from the Zener diode to the load as the load demands more current[13].

3.3 Clippers

Sometimes it is useful to limit the maximum amplitude of a signal. This might be done for protection, for example when too large of an input signal might damage the following circuit. It might also be employed as a means of wave shaping, that is, morphing a signal into another shape. A good example is the purely aesthetic desire to emulate the sound of "fuzz" guitar. In the early days of rock music it was discovered that over-driving a guitar amplifier in an attempt to make it louder created considerable distortion and this produced a new and interesting sound quality. Technically, this is largely caused by the power stage of the amplifier reaching its maximum output level. Any portion of the waveform above this level is simply cut off or *clipped*[14]. The practical problem here is that the only way to achieve this sound is to crank up the guitar amplifier's volume to ten[15] and live with the attendant high loudness level. Not too popular with the neighbors, that's for sure. In contrast, if the signal could be limited before the power amplifier stage in an attempt to mimic the clipping, the effect could be achieved without the resulting loudness. This proved to be so popular among guitarists that by the 1970s numerous companies were making "fuzz boxes" and "distortion pedals", each with their own twist on the concept.

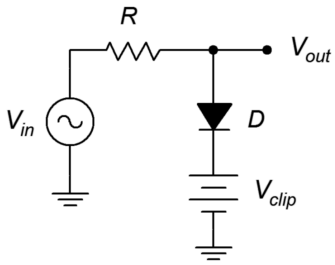
13 As you might guess, this is not particularly efficient because even when the load demand is nil, the Zener diode is still drawing current from the transformer. An improved circuit may include a bipolar transistor, as examined in Chapter 4. For details on more sophisticated techniques to regulate voltage, see Fiore, J, *Operational Amplifiers and Linear Integrated Circuits: Theory and Application*, another free OER text.

14 We will take a closer look at amplifier clipping in the chapters that cover power amplifiers.

15 Or eleven, if you have a custom Spinal Tap amplifier.

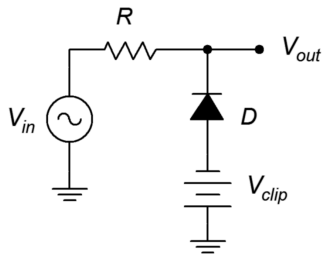
The simplest form of clipper places a diode (or two parallel diodes of opposing polarity) in parallel with the load. The diode will limit the output voltage swing to its forward turn-on potential; 0.7 volts for a silicon device. This circuit is somewhat limiting (pun intended) as you are stuck with a 0.7 volt limit value. What if we need to limit at some other potential, say 12 volts? While it is possible to stack a bunch of diodes in series to increase the limit point, a more flexible and practical approach involves biasing the diode with a DC source. This is called a biased diode clipper.

Figure 3.23
Biased positive clipper.



A biased diode clipper used to limit positive signals is shown in Figure 3.23. The operation is as follows. For any input signal that is less positive than the bias potential V_{clip}, the diode will be in reverse-bias. This means that the diode branch presents a high resistance and is effectively removed from the circuit. Therefore, V_{in} flows through R to the output unimpeded. If the input signal exceeds by V_{clip} by approximately 0.7 volts, the diode turns on resulting in a very low internal resistance. As the internal resistance of the DC source is also very low, this creates a low impedance path to ground and results in a voltage divider with R. As R is a much greater resistance value, virtually all of the input signal above the turn-on voltage will be dropped across R, never reaching the output. Therefore, we can control or program the clip point by adjusting the bias voltage. Clipping will occur at approximately V_{clip} plus 0.7 volts, assuming a silicon diode is used.

Figure 3.24
Biased negative clipper.



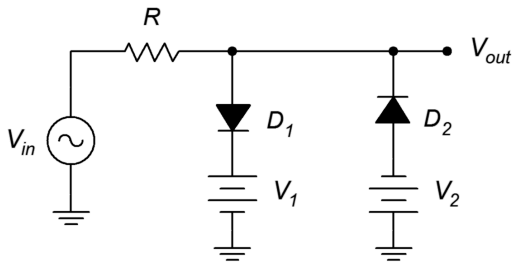
A negative biased clipper is shown in Figure 3.24. Both the diode and DC bias voltages have been flipped to the opposite polarity. The operation of this circuit is similar to the positive clipper. In this variant, the negative bias potential ensures that the diode is reverse-biased as long as the input level is more positive than V_{clip} minus 0.7 volts. Once the input signal goes below this voltage, the diode turns on creating the shorting path and limiting the output voltage.

A bipolar or dual-polarity clipper can be created by combining the positive and negative clippers. It is possible to limit the positive and negative swings independently, as illustrated in the following Example.

Example 3.3

Determine the output signal for the circuit of Figure 3.25. The input voltage is a 12 volt peak sine wave at 100 Hz. D_1 and D_2 are silicon switching diodes. R = 10 kΩ, V_1 = 4 volts and V_2 = 8 volts.

Figure 3.25
Dual clipper for Example 3.3.



First off, note that the precise value of R is unimportant here. It simply needs to be significantly larger than the on-resistance of the diodes. Based on our prior study of diode resistance in Chapter 2 (see the discussion around Figure 2.14) it is likely that the dynamic diode resistance will be under 100 Ω once full turn-on is reached. This resistance ratio is more than sufficient to create an effective voltage divider.

The positive clip level will be set by V_1. Adding in the forward potential of D_1 we arrive at 4.7 volts. The negative clip level will be set by V_2. Including in the forward potential of D_2 we arrive at −8.7 volts.

Thus, we expect to see a sine wave that is clipped at +4.7 volts and −8.7 volts. It should appear as a sort of lopsided cross between a sine wave and a square wave.

Computer Simulation

To verify and visualize our computations, the circuit of Example 3.3 is simulated with a transient analysis. The circuit schematic is shown in Figure 3.26. For the diodes, common 1N914 switching diodes are used.

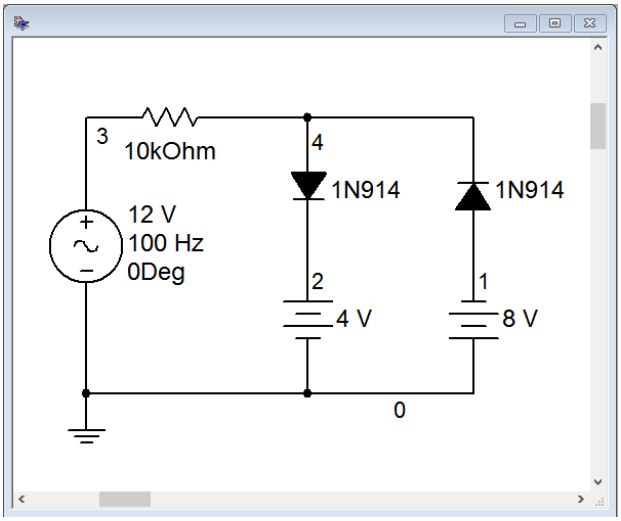


Figure 3.26
Simulation schematic for dual clipper of Example 3.3.

The results of the transient analysis are shown in Figure 3.27. The input waveform is shown in red while the output waveform is shown in blue. The input sine wave appears as expected with the specified 12 volt peak. The output follows the input perfectly for all values that are within the clip points. Beyond the clip points, the output voltage flattens. That is, it is limited to just below +5 volts (the programmed +4.7) and to just above −9 volts (the programmed −8.7 volts).

Careful inspection of the output waveform plot reveals that it is not perfectly flat at the voltage limits. In fact, there is a slight rounding that is most noticeable toward the transitions. This is due to the fact that the diodes do not turn on immediately. The dynamic resistance of the diodes change with the size of the signal. That is, the greater the input signal is above the clip point, the more current that will flow, and thus the dynamic resistance decreases, strengthening the effect of the voltage divider.

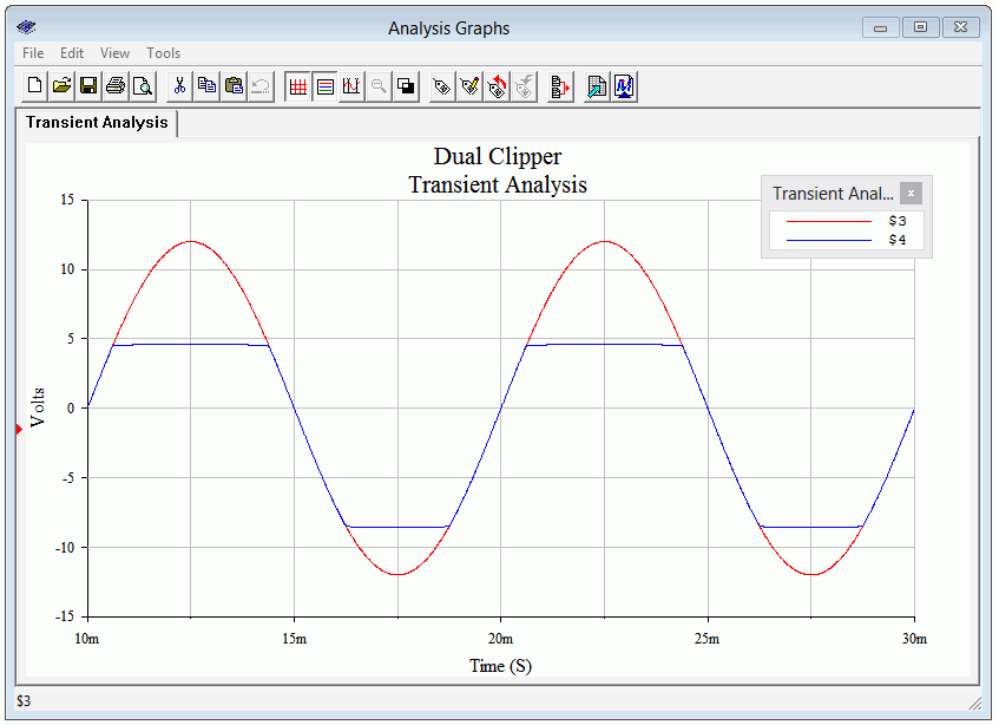


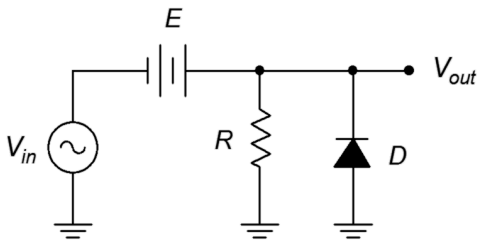
Figure 3.26
Transient analysis for dual clipper of Example 3.3.

3.4 Clampers

A *clamper* is circuit that adds a DC offset to an AC signal in such a way that the resulting voltage is uni-polar. A positive clamper adds a positive offset such that the former negative peak now sits at zero volts. In like fashion, a negative clamper adds in a negative offset such that the former positive peak now sits at zero volts. Clampers are also referred to as *DC restorers*. Clampers can also be biased so that the new peak point is something other than zero volts.

The concept of a clamper is fairly simple; we just add a DC voltage to the existing AC signal. The trick is in getting the circuit to automatically determine what the DC shift needs to be. This way, if the amplitude of the input signal changes, the offset can track with it.

Figure 3.27
Prototype DC offset circuit.

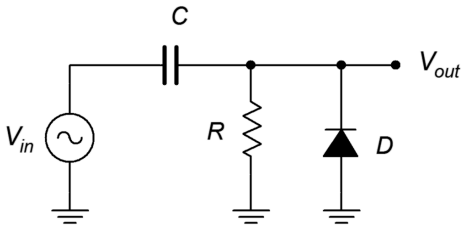


First, let's consider the prototype circuit in Figure 3.27. This is a fixed DC offset circuit. The DC source *E* adds a positive offset to the input signal. If the offset is equal to the peak value of the input, the negative peak will rise up to zero volts and the diode will never turn on (meaning that it will not load the input and change the wave shape).

In Figure 3.28 the fixed DC source has been replaced with a capacitor. This capacitor is used to create the DC offset. Unlike the fixed source, the capacitor voltage will vary with the peak value of the input and therefore precisely compensate to produce an ideally clamped output signal. As long as the

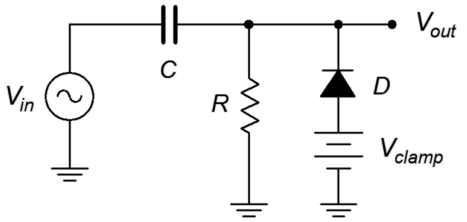
time constant for this capacitor and the surrounding resistance is much longer than the period of the input waveform, it will achieve proper clamping action.

Figure 3.28
Positive clamper.



Here is how the circuit operates. For the initial positive cycle, the capacitor is uncharged and the diode is reverse-biased. As the *RC* time constant is much longer than the input period, the output voltage merely follows the input voltage. Once the input signal swings negative, the diode turns on. This bypasses the parallel resistor and drastically reduces the charge time constant. This means that the capacitor voltage will begin to track the negative portion of the input signal while the output stays near zero volts. Note that the capacitor voltage will have a polarity of minus-to-plus from left to right, in keeping with Kirchhoff's voltage law. The capacitor voltage will track the negative input voltage all the way down to the negative peak. Once the input begins to reverse slope and rise toward zero, the diode will be turned off due to the potential now held on the capacitor. At this point, the capacitor has a voltage across it that is equivalent to the negative peak value of the input signal and it will behave just like the fixed DC voltage source in the prototype. The input is just now starting to track in the positive direction from its negative peak while the capacitor holds this same magnitude of voltage. The result is that the output is at zero volts and as the input continues to swing positive, the output will track it, thus producing the desired level shift.

Figure 3.29
Biased positive clamper.



Of course, circuits are never perfect. First, the forward voltage drop of the diode will result in a negative peak that's not precisely at zero volts but is instead about −0.7 volts. Second, it may take more than one cycle of the input to "grab" the peak value, all depending on the period and the precise charge and discharge time constants. As you might guess, flipping the polarity of the diode will result in a negative clamper instead of a positive clamper. Also, if we add a DC source in series with the diode, like we did with the biased clipper, we can create a biased clamper. This is shown in Figure 3.29.

Example 3.4

Determine the output signal for the circuit of Figure 3.29. The input voltage is a 10 volt peak sine wave at 1 kHz. $C = 10\ \mu F$, $R = 10\ k\Omega$, $V_{clamp} = 5.7$ volts and *D* is a silicon switching diode.

The configuration is a positive biased clamper. First, we need to ensure that the discharge time constant is much longer than the period. The period is $1/f$, or 1 millisecond. The discharge time constant is

$$\tau = RC$$
$$\tau = 10\,k\Omega \times 10\mu F$$
$$\tau = 100\,\text{milliseconds}$$

The DC clamping source will produce a positive offset of 5 volts (5.7 volts minus the 0.7 volt forward diode drop). This means that we should see a 20 volt peak-to-peak sine wave that swings between +5 volts and +25 volts.

Computer Simulation

To verify the analysis of Example 3.4, the circuit is captured as shown in Figure 3.30. A common 1N914 switching diode is used.

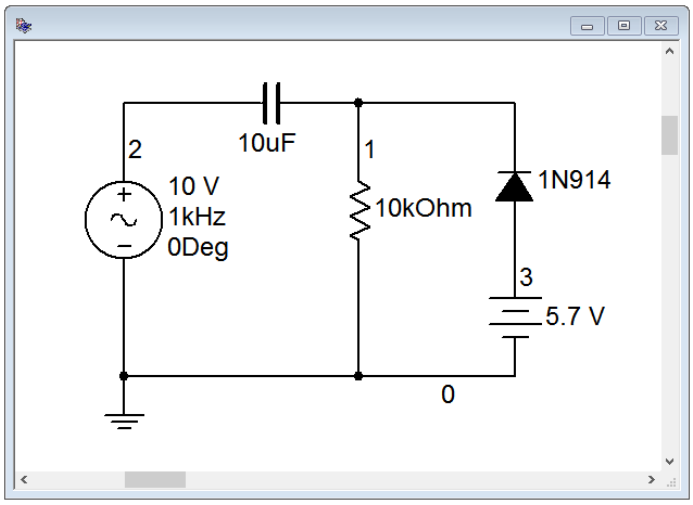


Figure 3.30
Simulation schematic for biased clamper of Example 3.4.

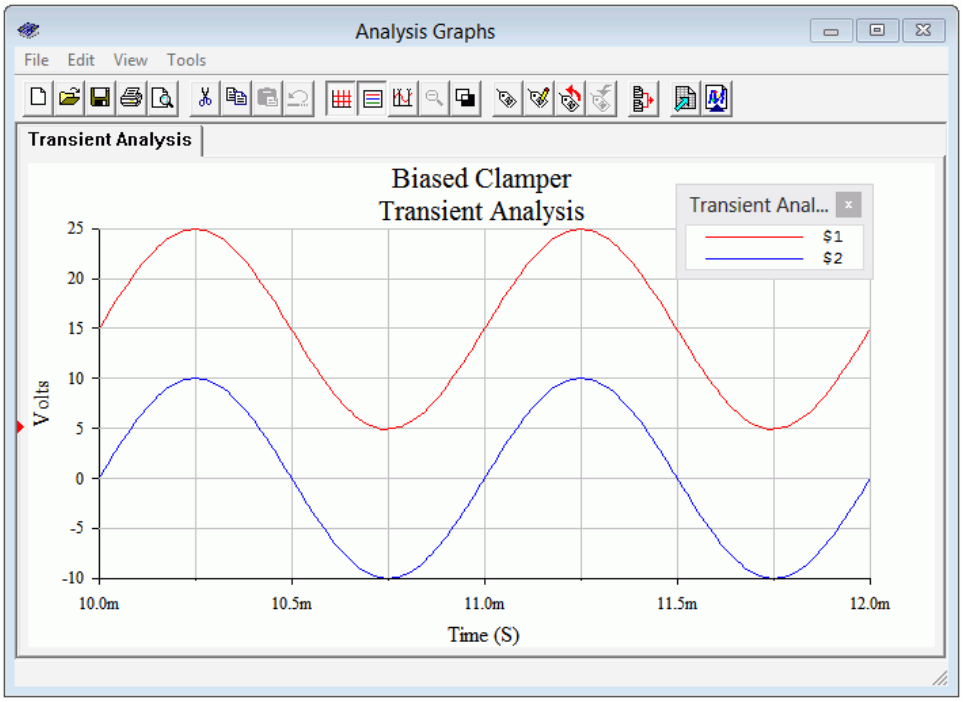


Figure 3.31
Transient analysis for biased clamper of Example 3.4.

The transient analysis is shown in Figure 3.31 and is precisely as predicted. The input waveform is blue and the clamped output is shown in red. The analysis was delayed ten milliseconds in order to get past the initial charge phase and observe the steady state operation.

Summary

Numerous AC diode applications have been examined in this chapter. A single diode may be used to create half-wave rectification, producing pulsating DC from AC. This is achieved by simply blocking one of the two polarities with a diode. A more efficient form of rectification is full-wave rectification. In this scheme, one of the two polarities is effectively flipped. This may be achieved via a two diode circuit that employs a split secondary transformer or via a four diode bridge circuit using a non-tapped secondary. The addition of a split secondary to the bridge circuit enables a dual polarity output.

In order to smooth the pulsating DC into a relatively constant level, a filter capacitor is added in parallel with the load. The larger the capacitor, the greater the filtering and smoothing action, however, this will also increase peak charging current. A Zener diode may be employed to further stabilize the output voltage.

Clippers are used to limit the range of an input signal. They may be designed to clip the positive portion, the negative portion or both polarities of the input waveform. The positive and negative clip levels may be adjusted independently.

Clampers are used to create a DC level shift that is dependent on the peak level of the input waveform. The shift may be positive or negative, and may also include an optional bias. The operation of the clamper hinges upon the charge versus discharge time constants for a series capacitor and associated diode.

Review Questions

1. List the advantages and disadvantages of half-wave versus full-wave rectifiers.
2. Discuss the advantages and disadvantages of a full-wave bridge rectifier versus a two diode center-tapped rectifier.
3. What is the purpose of the capacitor in a rectifier/power supply circuit?
4. Under what load conditions will a Zener regulator fail to maintain regulation of the output voltage?
5. What is *ripple*? How might it be reduced?
6. What is the function of the DC source(s) in a biased clipper?
7. What is the function of the capacitor in a clamper circuit?

Problems

(Assume diodes are silicon unless stated otherwise)

Analysis Problems

1. For the circuit of Figure 3.32, determine the peak output voltage. V_{sec} = 12 volts RMS, R_{load} = 50 Ω, C_1 = 1500 μF.

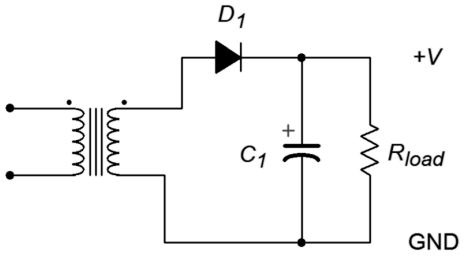


Figure 3.32

2. Sketch the output voltage waveform for the circuit of Problem 1, Figure 3.32, with and without the capacitor.

3. Determine the peak output voltage for the circuit of Figure 3.33. V_{sec} = 18 volts RMS, R_{load} = 75 Ω, C_1 = 470 μF.

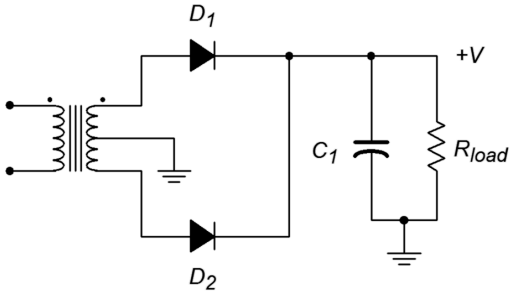


Figure 3.33

4. Sketch the output voltage waveform for the circuit of Problem 3, Figure 3.33, with and without the capacitor.

5. For the circuit of Figure 3.34, determine the peak output voltage. V_{sec} = 18 volts RMS, R_{load} = 40 Ω, C_1 = 1000 μF.

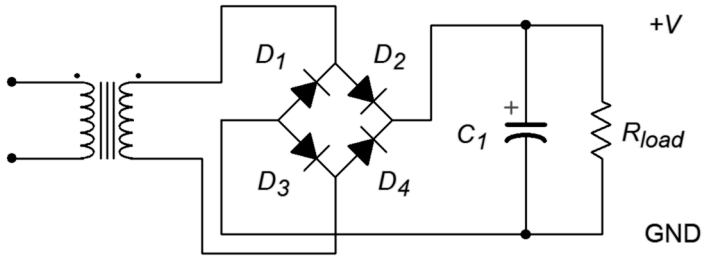


Figure 3.34

6. Sketch the output voltage waveform for the circuit of Problem 5, Figure 3.34, with and without the capacitor.

7. Determine the output voltage waveform and its amplitude for the circuit of Figure 3.35. V_{in} = 10 sin 2π100t, V_{clip} = 8 volts, R = 10 kΩ.

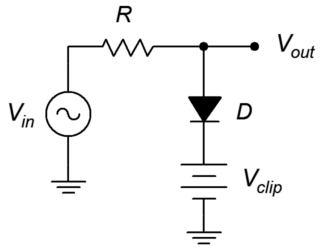


Figure 3.35

8. Draw the output waveform with its amplitudes for the circuit of Figure 3.36. V_{in} = 10 sin 2π100t, V_{clip} = 5 volts, R = 10 kΩ.

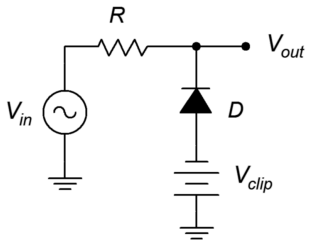


Figure 3.36

9. Draw the output waveform with its amplitudes for the circuit of Figure 3.37. V_{in} = 12 sin 2π200t, V_1 = 6 volts, V_2 = 4 volts, R = 10 kΩ.

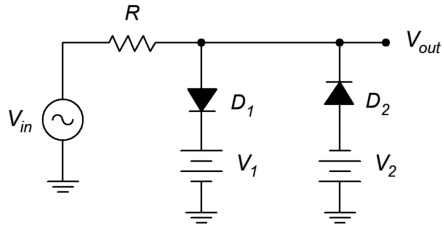


Figure 3.37

10. Draw the output waveform with its amplitudes for the circuit of Figure 3.38. V_{in} = 5 sin 2π2000t, C = 10 μF, R = 4.7 kΩ.

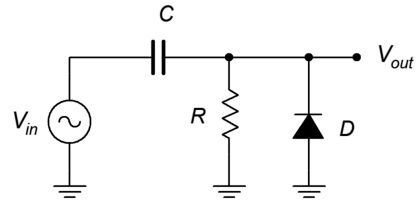


Figure 3.38

11. Draw the output waveform with its amplitudes for the circuit of Figure 3.39. V_{in} = 8 sin 2π500t, V_{clamp} = 2 volts, C = 4.7 μF, R = 33 kΩ.

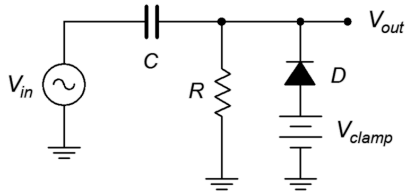


Figure 3.39

Design Problems

12. Design a 15 volt AC to DC power supply capable of drawing 200 mA.

13. Design a circuit that will limit its output voltage to a range of −5 volts to +10 volts.

14. Design a circuit that will shift its output voltage so that it is always positive. The input frequency is 2 kHz.

Challenge Problems

15. Design a circuit that will shift its output voltage so that its negative peak is at +3 volts. The input frequency range is from 100 Hz to 1 kHz.

Computer Simulation Problems

16. Run a transient analysis of the circuit in Figure 3.32, Problem 1.

17. Run a transient analysis of the circuit in Figure 3.33, Problem 3.

18. Run a transient analysis of the circuit in Figure 3.34, Problem 5.

19. Run two transient analyses on the clamper circuit of Example 3.4, first using a capacitor 100 times larger than specified, and second using a capacitor 100 times smaller than specified. Discuss the resulting waveforms.

4 Bipolar Junction Transistors (BJTs)

4.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Draw and explain the energy diagram for a biased bipolar junction transistor (BJT).
- Describe the differences between NPN and PNP BJTs.
- Explain forward-reverse bias operation.
- Define the transistor parameters *alpha* and *beta*, and determine them from device curves and/or circuit currents.
- Draw and explain the various regions of a BJT collector curve, along with *Early voltage*.
- Describe and utilize the Ebers-Moll BJT model.
- Describe the concept of the DC load line.
- Solve and design basic switching and driver circuits utilizing BJTs.

4.1 Introduction

The bipolar junction transistor, or BJT, is a foundational electronic component. It serves as the basis for a variety of applications ranging from simple amplifiers to device control to complex digital computing circuitry. Variations exist for applications spanning very low to very high frequency work; low, medium and high power; inexpensive general purpose through highly specialized niche items; and so forth. No matter what a BJT has been optimized for, all BJTs can be considered to be current boosting devices. Of course, if you can boost current, then you can also boost voltage and power, depending on the associated impedances. Further, all BJTs share the same basic structure: three alternating layers of N-type and P-type material with one external lead attached to each layer. In this manner, the BJT can be thought of as an extension of the basic diode: just add another segment of oppositely doped material to one end of the diode creating a second PN junction. The configuration could be either PNP or NPN. There are uses for both types and circuits often work best when the two types are used together.

4.2 The Bipolar Junction Transistor

In prior work we discovered that the PN junction is the foundation of the basic diode. Under normal operating conditions the interface between the N-type and P-type materials is devoid of free charges and is referred to as a depletion region. The dissimilar Fermi levels of N-type and P-type materials lead to an “energy hill” between them, and without an external potential of the proper polarity, the junction will not allow current to flow. The required magnitude is a function of the material used but it is always the case that the P material (anode) must be positive with respect to the N material (anode). We extend this idea by adding a second portion of N material to

the other side of the P material, creating an N-P-N "sandwich" of sorts. This is shown in Figure 4.1.

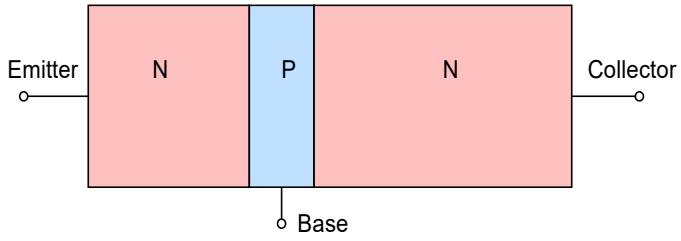


Figure 4.1
Basic configuration of NPN bipolar junction transistor.

This diagram is drawn to ease the understanding of the operation of the device, extending our earlier diode work. In contrast, real BJTs are built in more of a "layer cake" fashion, N-P-N bottom to top[16]. Of course, the spatial orientation of the device has no bearing on its operation so this is not a major issue for our purposes. The three terminals are named the *emitter*, *base* and *collector*. The collector is the largest of the three regions while the base is relatively thin and lightly doped.

Above absolute zero there will be recombination and two depletion regions will form as shown in Figure 4.2. Compare this figure to the basic PN junction drawing found at the beginning of Chapter 2, Figure 2.1.

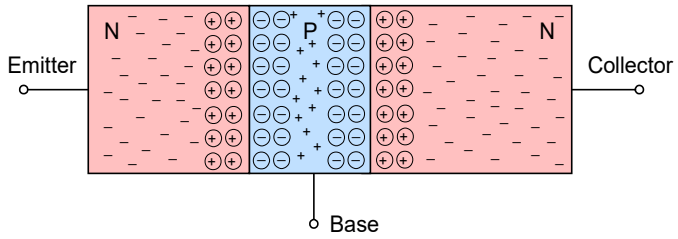


Figure 4.2
Charges in NPN BJT (base region widened to show detail).

A Simple Two-Diode Model

Because this device contains two depletion regions, a much simplified model can be created using two diodes as shown in Figure 4.3. Please keep in mind that this is a very limited model (as we shall soon see).

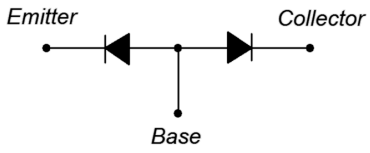


Figure 4.3
Diode model of NPN BJT.

If you were to test an NPN BJT with an ohmmeter, two leads at time, this model would successfully predict the results. If the red (positive) lead of the ohmmeter is connected to the base and the black (negative) lead is connected to either the emitter

16 Homer says, "Mmm, NPN layer cake sandwich..."

or collector, a low resistance will be indicated. This is because the ohmmeter will modestly forward-bias the base-emitter or base-collector junction. Similarly, if the leads are reversed, the meter will indicate high resistance because the junction under consideration will be reverse-biased. If the two leads are connected to the emitter and collector, a high reading will result regardless of the polarity. This is because one of the two junctions will be reverse-biased which results in no current flow through either of them due to the series connection.

Biasing the BJT

Now let's consider adding external sources to bias the transistor. We begin by adding two DC sources with associated current limiting resistors as shown in Figure 4.4.

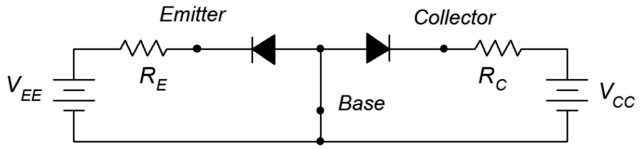


Figure 4.4
Double reverse-bias.

This circuit is comprised of two loops, one between the base-emitter and the second between the base-collector. In the B-E loop, the emitter supply V_{EE} reverse-biases the base-emitter diode. A similar situation occurs in B-C loop where the collector supply reverse-biases the base-collector diode. The result is that virtually no current flows anywhere in the circuit. If the two supplies are reversed in polarity then both diodes become forward-biased and we see currents flowing in both loops dependent on the precise values of the supplies and associated resistors. No surprises so far. Now consider if we forward-bias the base-emitter diode while simultaneously reverse-biasing the base-collector diode, as shown in Figure 4.5.

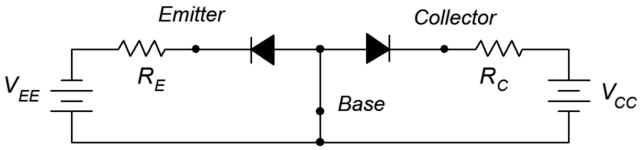


Figure 4.5
Forward-reverse bias.

With a simple pair of diodes we'd expect the B-E loop to show a high current and the B-C loop show negligible current. With a BJT, this is **not** what happens. Instead, what we see is a high current in both loops, and those currents are very nearly equal in magnitude. How does this come about?

The key to understanding this situation is that the base of the BJT is thin and lightly doped. In contrast, the dual diode model splits the base into two separate pieces of material and that makes all the difference. To get a better handle on what's happening here, let's take a closer look at this forward-reverse bias circuit but this time substituting the transistor diagram of Figure 4.2. Refer to Figure 4.6.

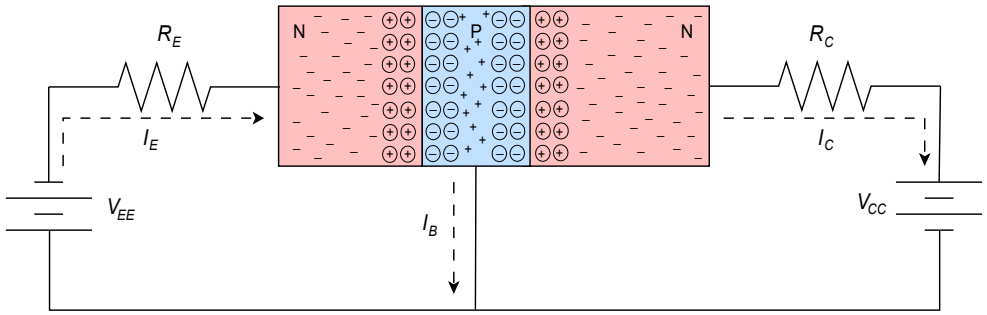


Figure 4.6
Forward-reverse bias, electron flow.

Electron flow will facilitate this explanation so we'll draw the current directions using dashed lines. From the left side of the diagram, electrons exit the emitter supply and enter the N emitter. Here they are the majority carrier. The base-emitter depletion creates an energy hill just as it did with a single PN junction. As long as there is sufficient potential from the emitter supply, the electrons will be pushed into the base. These electrons will attempt to recombine with the majority base holes, however, because the base is physically thin and lightly doped, only a small percentage of the injected electrons will recombine with base holes and exit the base terminal back to ground. This current is called the base current current or the recombination current. Meanwhile, the vast majority of the remaining electrons (95% to over 99%) will find their way to the base-collector depletion region and then to the collector. Once in the collector, the electrons are again the majority carrier and flow back to the positive terminal of the collector power supply. The energy diagram of the transistor is depicted in Figure 4.7. Compare this to the single PN junction energy diagram found at the beginning of Chapter 2, Figure 2.2.

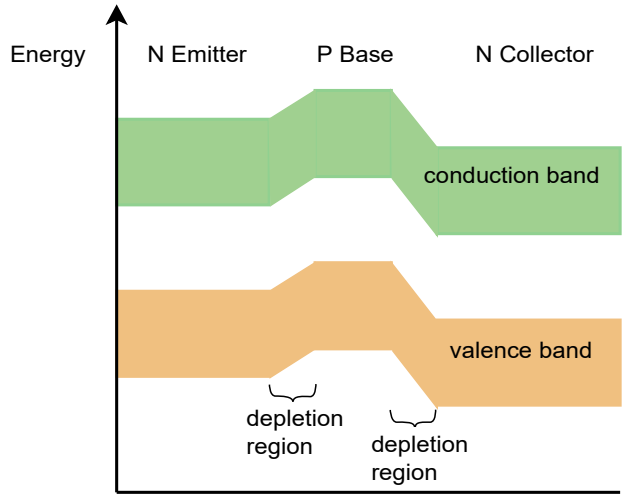


Figure 4.7
Energy diagram of BJT.

At first glance, it might appear as though the emitter and collector leads can be swapped with no change in operation. With real-world devices this is not possible generally because the emitter and collector regions are optimized and not physically identical. Thus, placing transistors into a circuit backwards, with emitter and collector leads swapped, will usually result in unpredictable behavior.

Based on the foregoing discussion and what we already know about PN junctions, we can summarize transistor performance as follows:

- From KCL, $I_E = I_C + I_B$.
- $I_C >> I_B$, therefore $I_E \approx I_C$.
- The base-emitter junction is forward-biased, therefore $V_{BE} \approx 0.7$ V (silicon).
- The base-collector junction is reverse-biased, therefore V_{CB} is large.
- Conventional current flows into the collector and base, and out of the emitter.

We can also define a couple of transistor performance parameters. The ratio of collector current to emitter current is called α (alpha). α typically is greater than 0.95. A somewhat more useful parameter is the ratio of collector current to base current. This is called β (beta) and can also be found on transistor spec sheets as h_{FE} (h_{FE} is one of four *hybrid parameters*). It is also referred to generically as *current gain* (if I_B is in the input signal and I_C is the output signal then β represents the amount of signal boost or gain). For small signal transistors β typically is in the range of 100 to 200, although it can be larger. For power transistors, β tends to be smaller, more like 25 to 50. Presented as formulas we have:

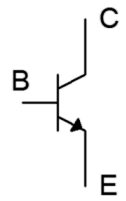
$$\alpha = I_C / I_E \quad (4.1)$$
$$\beta = I_C / I_B \quad (4.2)$$

And with a little math,

$$\alpha = \beta / (\beta+1)$$
$$\beta = \alpha / (1-\alpha)$$
$$I_C = \beta I_B$$

Finally, we come to the schematic symbol of the NPN BJT, as shown in Figure 4.8. A common variation places the body of the device within a circle. Following the standard, the arrow points to N material and in the direction of easy conventional current flow.

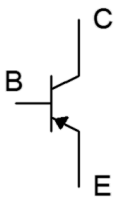
Figure 4.8
NPN Schematic symbol



The PNP Bipolar Junction Transistor

The PNP version of the BJT is created by swapping the material for each layer. The outcome is the logical inverse of the NPN regarding current directions and voltage polarities. That is, conventional current flows into the emitter, and out of the collector and base (echoing the electron flow of the NPN). Further, voltages across the device have reversed polarity, for example, $V_{BE} \approx -0.7$ V. All of the other characteristics remain unchanged so equations such 4.1 and 4.2 are still applicable. Just about any NPN-based circuit has its PNP counterpart. The schematic symbol of the PNP reverses the emitter arrow. As the base is now the N material, the arrow points toward the base. This is illustrated in Figure 4.9.

Figure 4.9
PNP Schematic symbol



4.3 BJT Collector Curves

One of the more useful BJT device plots is the family of collector curves. This is a series of plots of collector current, I_C, versus collector-emitter voltage, V_{CE}, at varying levels of base current, I_B. To generate these curves we drive the base terminal with a fixed current source establishing I_B. A DC power supply is attached from the collector to emitter and then swept from zero volts to some upper value. This establishes V_{CE}. Simultaneously, we track the resulting collector current and plot the result. This results in one trace. The base current is then increased and the DC supply swept again for a second trace. This process is repeated to create a family of curves. An example is shown in Figure 4.10.

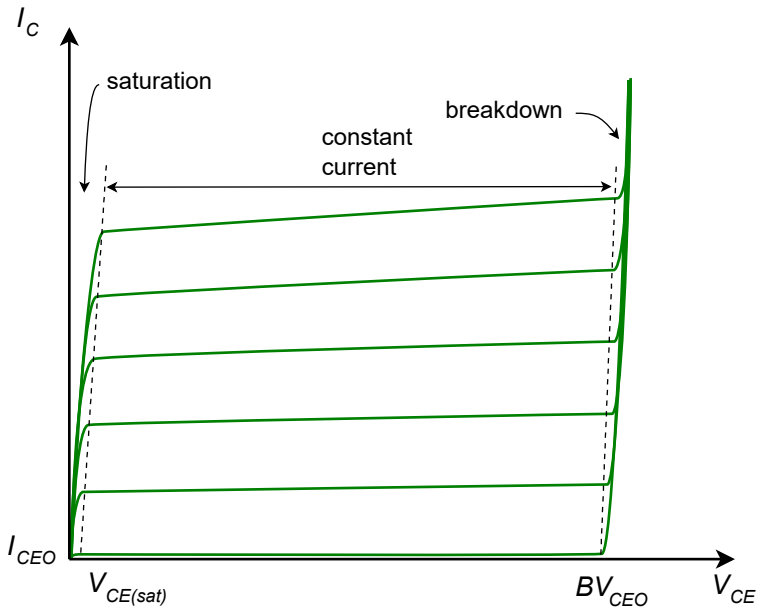


Figure 4.10
A family of collector curves.

The bottom curve results when $I_B = 0$. Ideally, the corresponding collector current would be 0 but a small leakage current occurs. This is usually referred to as I_{CEO}, meaning the *C*ollector-*E*mitter current with the base terminal *O*pen (i.e., no base current). The curves above this correspond to increasing levels of base current; each new base current stepping up a fixed amount for each subsequent trace (e.g., 0 µA, 10 µA, 20 µA, 30 µA, etc.).

The most striking thing about this set of curves is that there are three distinct regions or zones of operation. To the extreme left of the curve the current rises rapidly. This is known as the *saturation region*. The break-over point is fairly small at just a few tenths of a volt. This can be found on a data sheet as $V_{CE(sat)}$. The saturation region is used in transistor switching applications.

At the extreme right is another region where the collector current rises rapidly. This is called the *breakdown region*. This is the same effect we saw with individual diodes. We do not wish to operate devices in this region as damage may result. The breakdown voltage is denoted on most data sheets as BV_{CEO} (*C*ollector to *E*mitter voltage with an *O*pen base). For general purpose devices this will be in the range of 30 to 60 volts or so, but it can be much higher.

In between these two extremes is a region where the collector current is relatively constant, showing only a modest positive slope. This is the *constant current region*. This is where we want the transistor to operate for applications such as linear amplifiers.

A device called a *curve tracer* can be used to generate this family of curves in the lab. A very good approximation for β can be determined using these curves. First, we determine the approximate circuit values for I_C and V_{CE}, and locate this point on the graph. We then find the nearest plot line to that point. From the intersection of V_{CE} and and this plot line we track back to the vertical axis to find the precise value of I_C for that trace. We count the number of traces and multiply by the base current step size to determine the corresponding base current. We then divide the two values and arrive at β.

Example 4.1

Assume we have a BJT operating at V_{CE} = 30 V and I_C = 4 mA. If the device is placed in a curve tracer and the resulting family of curves appears as in Figure 4.11, determine the value of β. Assume the base current is increased 10 μA per trace.

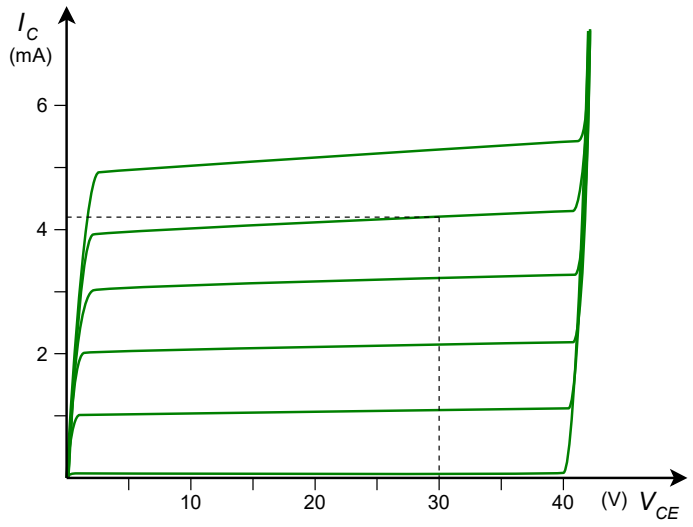


Figure 4.11
Curve tracer display for Example 4.1.

First, find a trace close to the operating point of 30 volts and 4 mA. Draw a vertical line at 30 volts and stop when that line intersects the trace nearest to

4 mA. In this example that's the second trace from the top. From that intersection point, track back to the vertical axis to determine the precise collector current. That's roughly 4.2 mA here. To determine the base current count up the number of traces to the selected trace. The selected trace is the fourth one up (do not include the bottom trace where I_B is 0). The base current is increased by 10 μA per trace so that leaves us with $I_B = 40$ μA.

$$\beta = \frac{I_C}{I_B}$$
$$\beta = \frac{4.2\,\text{mA}}{40\,\mu\,\text{A}}$$
$$\beta = 105$$

Rise in β and Early Voltage

When looking at the collector curves, a good question we might ask is why the collector current rises as V_{CE} increases. This is due to the fact that the increased collector-emitter voltage is responsible for an increase in collector-base voltage (by definition, $V_{CE} = V_{CB} + V_{BE}$). V_{CB} is the reverse-bias potential on the collector-base PN junction. As this reverse potential increases, the collector-base depletion region widens. As it widens, it penetrates further into the base layer. Because the base is effectively narrowed, the chances for recombination are reduced, thus reducing base current and effectively increasing β.

If we extend the constant current region traces back into the second quadrant they intersect at a point called the Early Voltage, named after James Early, and denoted as V_A. This is illustrated in Figure 4.12.

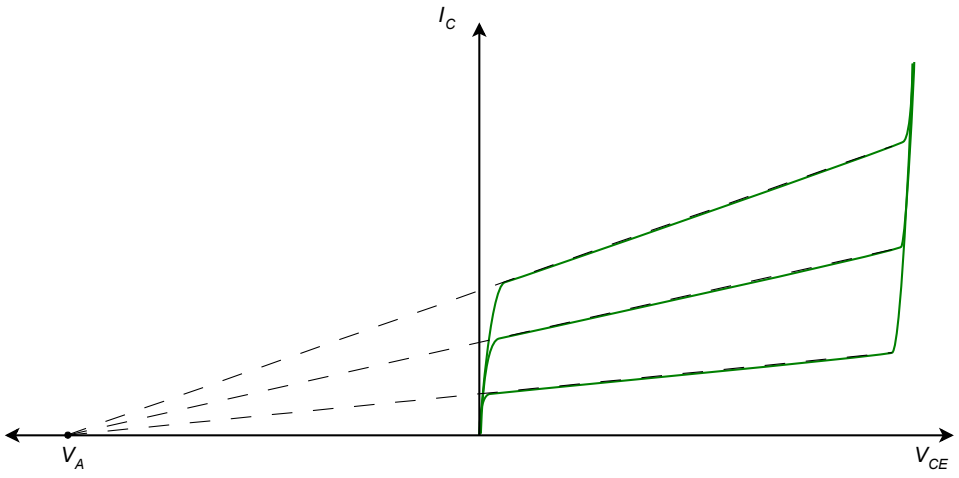


Figure 4.12
Early Voltage.

4.4 BJT Data Sheet Interpretation

The data sheet for a common NPN transistor, the 2N3904, is shown in Figure 4.13. This model is available from several different manufacturers. First off, note the case style. This a TO-92 plastic case for through-hole mounting and is commonly used for small signal transistors. Under the maximums we find the device has a maximum power dissipation of 625 mW in free air (ambient temperature of 25°C), a maximum collector current of 200 mA and a maximum collector-emitter voltage of 40 V. Obviously, the device cannot withstand maximum current and voltage simultaneously.

Figure 4.13a
2N3904 data sheet.
Used with permission from SCILLC dba ON Semiconductor.

2N3903, 2N3904

General Purpose Transistors

NPN Silicon

Features

- Pb–Free Packages are Available*

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Collector – Emitter Voltage	V_{CEO}	40	Vdc
Collector – Base Voltage	V_{CBO}	60	Vdc
Emitter – Base Voltage	V_{EBO}	6.0	Vdc
Collector Current – Continuous	I_C	200	mAdc
Total Device Dissipation @ T_A = 25°C Derate above 25°C	P_D	625 5.0	mW mW/°C
Total Device Dissipation @ T_C = 25°C Derate above 25°C	P_D	1.5 12	W mW/°C
Operating and Storage Junction Temperature Range	T_J, T_{stg}	−55 to +150	°C

THERMAL CHARACTERISTICS (Note 1)

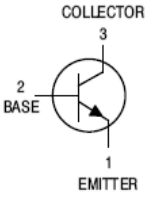
Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction–to–Ambient	$R_{\theta JA}$	200	°C/W
Thermal Resistance, Junction–to–Case	$R_{\theta JC}$	83.3	°C/W

Stresses exceeding Maximum Ratings may damage the device. Maximum Ratings are stress ratings only. Functional operation above the Recommended Operating Conditions is not implied. Extended exposure to stresses above the Recommended Operating Conditions may affect device reliability.
1. Indicates Data in addition to JEDEC Requirements.

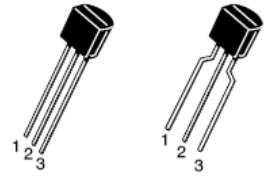
August, 2012 – Rev. 8

ON Semiconductor®

http://onsemi.com



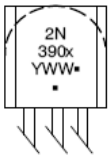
TO–92
CASE 29
STYLE 1



STRAIGHT LEAD
BULK PACK

BENT LEAD
TAPE & REEL
AMMO PACK

MARKING DIAGRAMS



x = 3 or 4
Y = Year
WW = Work Week
▪ = Pb–Free Package
(Note: Microdot may be in either location)

In Figure 4.13b we find a variety of characteristics including nominal values for β (listed here as h_{FE}) under various conditions. At particularly small or large collector currents β tends to drop off. Also, note the wide 3:1 variance at 10 mA. Perhaps more illustrative are the graphs from the third page, Figure 4.13c.

Figure 4.13b
2N3904 data sheet (cont).

ELECTRICAL CHARACTERISTICS (T_A = 25°C unless otherwise noted)

Characteristic		Symbol	Min	Max	Unit
OFF CHARACTERISTICS					
Collector – Emitter Breakdown Voltage (Note 2) (I_C = 1.0 mAdc, I_B = 0)		$V_{(BR)CEO}$	40	–	Vdc
Collector – Base Breakdown Voltage (I_C = 10 μAdc, I_E = 0)		$V_{(BR)CBO}$	60	–	Vdc
Emitter – Base Breakdown Voltage (I_E = 10 μAdc, I_C = 0)		$V_{(BR)EBO}$	6.0	–	Vdc
Base Cutoff Current (V_{CE} = 30 Vdc, V_{EB} = 3.0 Vdc)		I_{BL}	–	50	nAdc
Collector Cutoff Current (V_{CE} = 30 Vdc, V_{EB} = 3.0 Vdc)		I_{CEX}	–	50	nAdc
ON CHARACTERISTICS					
DC Current Gain (Note 2)		h_{FE}			–
(I_C = 0.1 mAdc, V_{CE} = 1.0 Vdc)	2N3903		20	–	
	2N3904		40	–	
(I_C = 1.0 mAdc, V_{CE} = 1.0 Vdc)	2N3903		35	–	
	2N3904		70	–	
(I_C = 10 mAdc, V_{CE} = 1.0 Vdc)	2N3903		50	150	
	2N3904		100	300	
(I_C = 50 mAdc, V_{CE} = 1.0 Vdc)	2N3903		30	–	
	2N3904		60	–	
(I_C = 100 mAdc, V_{CE} = 1.0 Vdc)	2N3903		15	–	
	2N3904		30	–	
Collector – Emitter Saturation Voltage (Note 2)		$V_{CE(sat)}$			Vdc
(I_C = 10 mAdc, I_B = 1.0 mAdc)			–	0.2	
(I_C = 50 mAdc, I_B = 5.0 mAdc			–	0.3	
Base – Emitter Saturation Voltage (Note 2)		$V_{BE(sat)}$			Vdc
(I_C = 10 mAdc, I_B = 1.0 mAdc)			0.65	0.85	
(I_C = 50 mAdc, I_B = 5.0 mAdc)			–	0.95	
SMALL – SIGNAL CHARACTERISTICS					
Current – Gain – Bandwidth Product		f_T			MHz
(I_C = 10 mAdc, V_{CE} = 20 Vdc, f = 100 MHz)	2N3903		250	–	
	2N3904		300	–	
Output Capacitance (V_{CB} = 5.0 Vdc, I_E = 0, f = 1.0 MHz)		C_{obo}	–	4.0	pF
Input Capacitance (V_{EB} = 0.5 Vdc, I_C = 0, f = 1.0 MHz)		C_{ibo}	–	8.0	pF
Input Impedance		h_{ie}			k Ω
(I_C = 1.0 mAdc, V_{CE} = 10 Vdc, f = 1.0 kHz)	2N3903		1.0	8.0	
	2N3904		1.0	10	
Voltage Feedback Ratio		h_{re}			X 10^{-4}
(I_C = 1.0 mAdc, V_{CE} = 10 Vdc, f = 1.0 kHz)	2N3903		0.1	5.0	
	2N3904		0.5	8.0	
Small – Signal Current Gain		h_{fe}			–
(I_C = 1.0 mAdc, V_{CE} = 10 Vdc, f = 1.0 kHz)	2N3903		50	200	
	2N3904		100	400	
Output Admittance (I_C = 1.0 mAdc, V_{CE} = 10 Vdc, f = 1.0 kHz)		h_{oe}	1.0	40	μmhos
Noise Figure		NF			dB
(I_C = 100 μAdc, V_{CE} = 5.0 Vdc, R_S = 1.0 k Ω, f = 1.0 kHz)	2N3903		–	6.0	
	2N3904		–	5.0	
SWITCHING CHARACTERISTICS					
Delay Time	(V_{CC} = 3.0 Vdc, V_{BE} = 0.5 Vdc, I_C = 10 mAdc, I_{B1} = 1.0 mAdc)	t_d	–	35	ns
Rise Time		t_r	–	35	ns
Storage Time	(V_{CC} = 3.0 Vdc, I_C = 10 mAdc, I_{B1} = I_{B2} = 1.0 mAdc) 2N3903	t_s	–	175	ns
	2N3904		–	200	
Fall Time		t_f	–	50	ns

2. Pulse Test: Pulse Width ≤ 300 μs; Duty Cycle ≤ 2%.

The upper-most graph depicts the variation of β with both collector current and temperature. The *normalized* β is plotted on the vertical axis. That is, this is not the expected value but is a ratio used to compare β under varying conditions.

Figure 4.13c
2N3904 data sheet (cont).

2N3903, 2N3904

TYPICAL STATIC CHARACTERISTICS

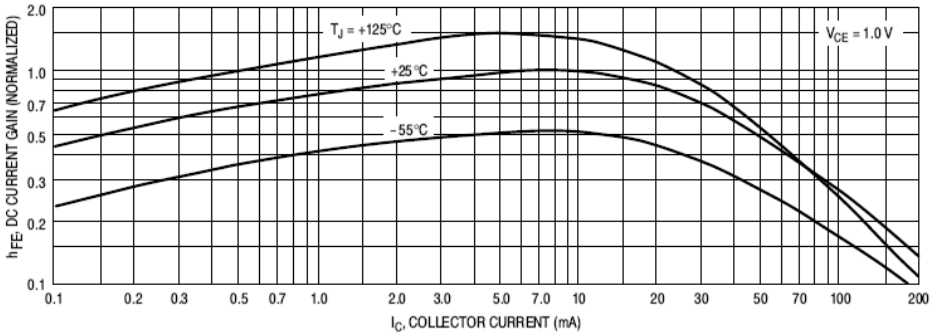


Figure 15. DC Current Gain

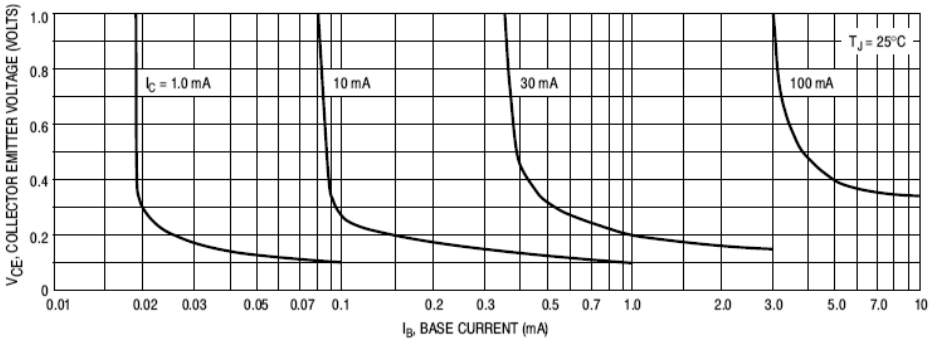


Figure 16. Collector Saturation Region

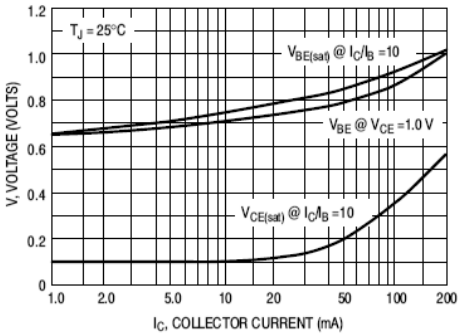


Figure 17. "ON" Voltages

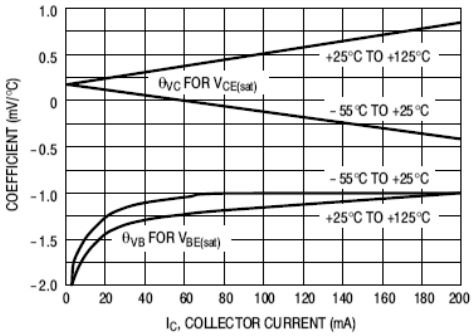


Figure 18. Temperature Coefficients

http://onsemi.com

For example, at room temperature and 10 mA, the normalized value is 1.0. The second page indicated a range of 100 to 300 for the 2N3904's β under these conditions. Let's say we measure one particular transistor to have a β of 200. If we were to operate this transistor at a lower current, say 0.2 mA, the β would drop. From the graph, the normalized β value at 0.2 mA and 25° C is 0.7. Therefore, the β

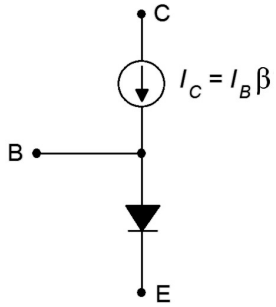
under these conditions would be 0.7/1.0 • 200, or 140. The graph also shows that, generally speaking, β tends to increase with increasing temperature.

The middle graph plots the collector-emitter saturation voltage, or $V_{CE(sat)}$, for various current conditions. This is an important parameter when dealing with transistor switching circuits. We shall refer back to this graph a little later in this chapter.

4.5 Ebers-Moll Model

A good, functional model of the BJT is the simplified Ebers-Moll model shown in Figure 4.14. This utilizes an ideal diode to model the base-emitter junction and a current-controlled current source located at the collector-base. This model is sufficient to achieve good analysis results with a variety of DC and low frequency circuits. It is important to remember, though, that β varies not only from device to device, but also varies with changes in temperature, collector current and collector-emitter voltage.

Figure 4.14
Ebers-Moll model of the NPN BJT.



We can put the Ebers-Moll model to use in basic DC biasing circuits. To properly bias the BJT we need to make the collector-base reverse-biased and the base-emitter forward-biased. In other words, $V_C > V_B > V_E$. There are many ways to achieve this. One method places the emitter at ground, a modest DC source in the base-emitter loop, and a somewhat higher DC source at the collector. An example is shown in Figure 4.15.

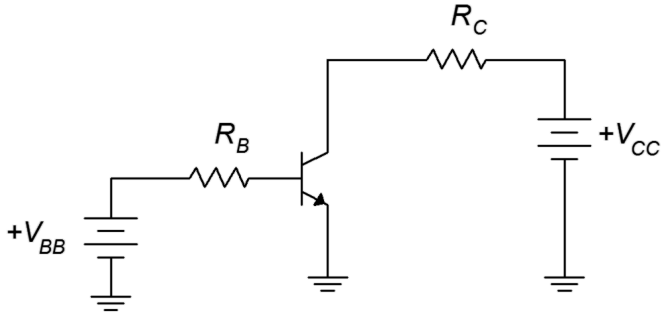


Figure 4.15
Simple base bias circuit.

The two resistors serve to limit the transistor's currents and voltages. Because the emitter is at ground, the common point, this circuit is classified as having a *common emitter configuration*. There are many possible common emitter circuits. We shall refer to this one specifically as *base bias*.

Now, let's replace the transistor with the Ebers-Moll model. The result, with added voltage polarities and current directions, is shown in Figure 4.16.

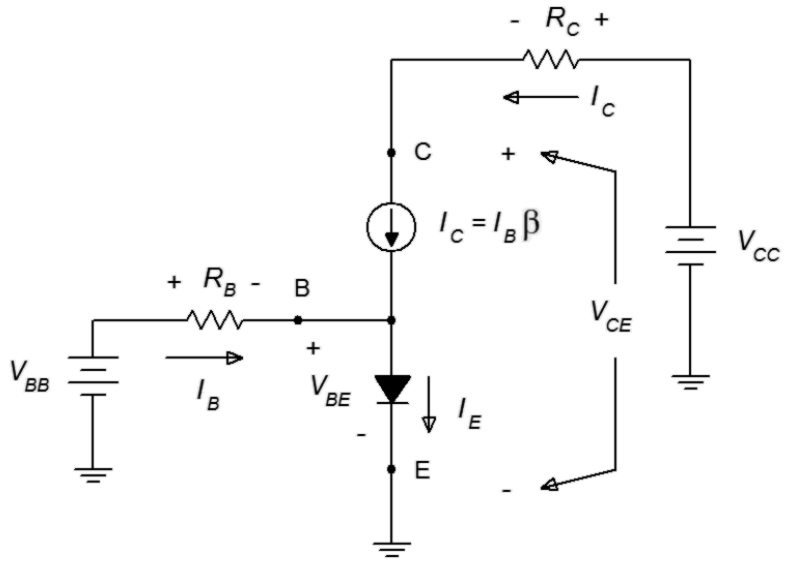


Figure 4.16
Base bias circuit with Ebers-Moll model.

Given values for the resistors, power supplies and β, all of the currents and voltages may be determined using fundamental circuit analysis techniques. The basic idea is to create KVL equations for the two loops and then expand using Ohm's law. We begin with the base-emitter loop.

$$\begin{aligned} V_{BB} &= V_{R_B} + V_{BE} \\ V_{BB} &= I_B R_B + V_{BE} \\ I_B &= \frac{V_{BB} - V_{BE}}{R_B} \end{aligned} \tag{4.3}$$

And for the collector-emitter loop:

$$\begin{aligned} V_{CC} &= V_{R_C} + V_{CE} \\ V_{CC} &= I_C R_C + V_{CE} \\ V_{CE} &= V_{CC} - I_C R_C \end{aligned} \tag{4.4}$$

To obtain I_C, recall that $I_C = \beta\, I_B$.

Example 4.2

Determine the circulating currents and device voltages for the circuit of Figure 4.15 if V_{BB} = 10 V, V_{CC} = 15 V, R_B = 200 kΩ, R_C = 1 kΩ and β = 100. Assume that the transistor is silicon.

First, find the base current. KVL dictates that the voltage across R_B is 9.3 volts.

$$I_B = \frac{V_{BB} - V_{BE}}{R_B}$$
$$I_B = \frac{10\,\text{V} - 0.7\,\text{V}}{200\,\text{k}\Omega}$$
$$I_B = 46.5\,\mu\text{A}$$

Now find the collector current and follow with Ohm's law and KVL.

$$I_C = \beta I_B$$
$$I_C = 100 \times 46.5\,\mu\text{A}$$
$$I_C = 4.65\,\text{mA}$$

$$V_{CE} = V_{CC} - I_C R_C$$
$$V_{CE} = 15\,\text{V} - 4.65\,\text{mA} \times 1\,\text{k}\Omega$$
$$V_{CE} = 10.35\,\text{V}$$

For the sake of completeness, the voltage across R_C is 4.65 volts, V_{CB} is 9.65 volts and I_E is 4.6965 mA.

The preceding example illustrates that the place to start the analysis is in the base-emitter loop instead of the collector-emitter loop. This is because in the base-emitter loop we have the forward-biased base-emitter junction which has a known potential (approximately 0.7 volts). In contrast, the collector-emitter voltage is an unknown as it includes the reverse-biased collector-base junction. That voltage will depend on other circuit elements, most notably the collector resistor and associated supply.

An improvement on the circuit of Example 4.2 would be to redesign it for a single power supply rather than two supplies. This is easy to do. All that is needed is to keep the base current unchanged. If that remains at its original value then the collector current won't change and consequently nothing in the collector-emitter loop will change either. Using the 15 volt source for V_{BB} means that the voltage across R_B will increase to 14.3 volts. Ohm's law then indicates that R_B must be 14.3 volts divided by 46.5 μA, or 307.5 kΩ.

β Variation Issues

There is a major problem with the circuit of Figure 4.15, namely, it lacks stability of collector current and collector-emitter voltage. As we shall see in upcoming chapters, it is important to keep these parameters stable in order to ensure consistent performance for many kinds of circuits. As we noted from our inspection of the 2N3904 data sheet, the variation of β can be quite large at a given operating point. If we also add in the variance due to temperature and other factors, we may be looking at a 10:1 range. If we repeat Example 4.2 with a doubled β of 200, the base-emitter loop is unchanged but the collector current doubles to 9.3 mA. This increases the

voltage drop across R_C to 9.3 volts which then forces V_{CE} to drop to 5.7 volts. Given a typical production run of transistors, this circuit might exhibit collector currents from less than 4 mA to more than 10 mA. In some applications this variation in current might be tolerable but not in all of them. For example, suppose an LED is placed in series with R_C. Because the brightness of an LED depends on its current level, the brightness will now depend on the β of the specific BJT used. If this is one LED in a larger display made up of similar circuits, then the illumination will be uneven between them causing the entire display to appear off kilter.

In fact, if this circuit was built in the lab, it is quite likely that after turning on the power, you could watch I_C slowly rise on your ammeter. This is because the BJT will begin to warm up as it dissipates power. As noted from the data sheet, β increases with increasing temperature. Because I_B is a fixed value, any rise in β means that I_C must also rise. This increased current will tend to cause a further rise in power dissipation and temperature which causes a further increase in β, and the process cycles. We have created an inadvertent thermal positive feedback loop. Left unchecked, devices could overheat and be destroyed. We will examine biasing circuits that achieve high stability in the next chapter.

There is another interesting aspect to this circuit. As noted, if we substituted the original BJT with another unit that had a higher β, the collector current would rise. What if we continued this to higher and higher β values? For example, if we increased β to 400 (admittedly, rather high) the new collector current would seem to jump up to 46.5 μA• 400, or 18.6 mA. There is a “small” problem with this value. Ohm's law indicates that this current would develop a drop of 18.6 volts across the 1 kΩ R_C but that's impossible because V_{CC} is only 15 volts. The only way that “works” is if somehow the BJT is transformed magically into a 3.6 volt battery. No amount of prayer or letters to Santa will make that happen[17].

4.6 DC Load Lines

So how do we determine the range of possible values of collector current and collector-emitter voltage in any given DC BJT circuit? One answer is to employ the concept of the *DC load line*. In general, a load line is a plot of all possible coordinate pairs of I_C and V_{CE} for a transistor in a given circuit. Referring back to Figure 4.16, we pick up with Equation 4.4 and solve it for I_C:

$$\begin{aligned} V_{CE} &= V_{CC} - I_C R_C \\ I_C &= \frac{1}{R_C}(V_{CC} - V_{CE}) \\ I_C &= -\frac{1}{R_C} V_{CE} + \frac{V_{CC}}{R_C} \end{aligned} \qquad (4.5)$$

17 Both being equally effective.

Equation 4.5 is a linear equation of the form $y = mx + b$. The y intercept (the value of I_C when $V_{CE} = 0$) is V_{CC}/R_C. This is the maximum collector current that can be achieved. At this point the transistor is saturated and this maximum is referred to as $I_{C(sat)}$. The x intercept (the value of V_{CE} when $I_C = 0$) is V_{CC}. This represents the largest possible voltage across the transistor's collector-emitter. At this point the current is cut off, and therefore this voltage is called $V_{CE(cutoff)}$. Lastly, the slope of the line is $-1/R_C$. A plot is shown in Figure 4.17.

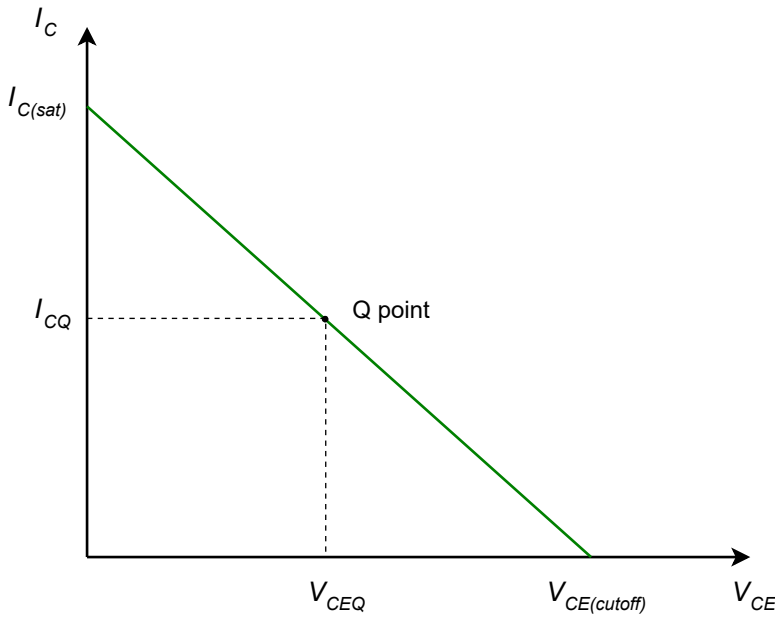


Figure 4.17
Generic DC load line.

To complete the graph, we also include the operating point for some specific transistor. This is called the quiescent point, or *Q point*, and the associated device current and voltage are called I_{CQ} and V_{CEQ}. All possible Q points lay on this line.

Referring back to Example 4.2 and using Equation 4.5, we can summarize the circuit as follows:

$I_{C(sat)} = 15$ mA
$V_{CE(cutoff)} = 15$ V

Q Point for $\beta = 100$:
$I_C = 4.65$ mA
$V_{CE} = 10.35$ V

Q Point for $\beta = 200$:
$I_C = 9.3$ mA
$V_{CE} = 5.7$ V

This is plotted in Figure 4.18.

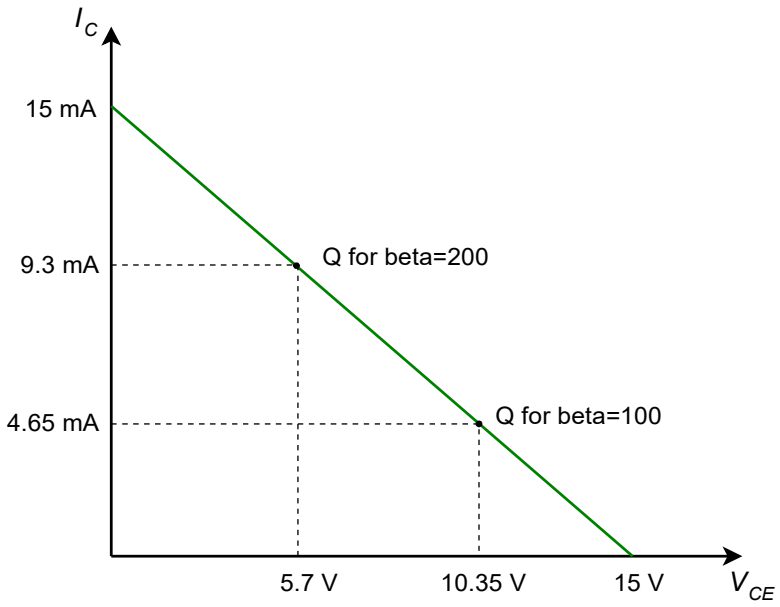


Figure 4.18
Load line for the variations on Example 4.2.

If we calculate a collector current that is greater than the saturation current, then we know that the actual current will be the saturation current maximum. For this circuit, any calculated value greater than 15 mA indicates that the transistor would produce only 15 mA (our earlier example using β = 400, for instance). In reality, the true value will be very slightly less. This is because the collector-emitter voltage does not go all the way down to zero volts when the device is saturated. Typically, $V_{CE(sat)}$ will be a tenth of a volt or so for small signal devices. Precise values can be determined from device graphs such as the middle graph of Figure 4.13c, labeled "Collector Saturation Region". As an example, if I_C = 10 mA and I_B = 0.3 mA, then $V_{CE(SAT)}$ is approximately 0.15 V. It turns out that we can use saturation to our advantage in switching circuits, as we are about see.

4.7 BJT Switching and Driver Applications

As mentioned, variation in β can cause changes in collector current. This can cause performance issues. For example, when driving an LED, this can lead to variance in brightness. But what if we purposely put the transistor into saturation? Saturation is a fixed value. It is inherently stable and β no longer matters. Effectively, when a BJT saturates, β is forced to drop to whatever value is needed to produce $I_{C(sat)}$. We just need to make sure that even the smallest β is large enough to cause saturation.

The Saturating Switch

A good example of this is the saturating LED driver circuit shown in Figure 4.19. To begin with, the whole point of the driver is to offload the current demand from the prior circuit. For example, we may wish to light an LED from the output of a logic

gate or microcontroller chip. The problem is that those circuits might only be able to deliver, say, 5 mA when we might need well over 10 mA to achieve the desired brightness. The LED driver circuit is used to overcome this limitation.

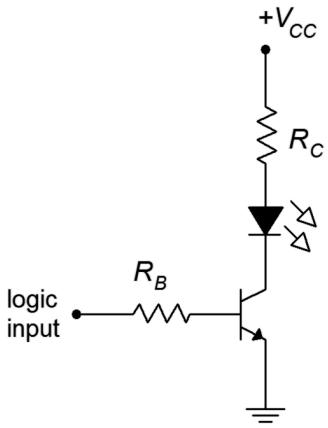


Figure 4.19
Saturating LED driver circuit (positive logic).

Note: The negative terminal of V_{CC} is connected to ground (not shown).

With the driver, the logic circuit will only need to supply base current, not LED current. Here is how it works: If the logic input voltage is zero, there will be no base current. This means that there will be no collector current and therefore the LED will be off. At this point the BJT is in cutoff. In contrast, when the logic level goes high, all of the logic voltage drops across R_B, with the exception of V_{BE}. This creates I_B. If properly designed, this current will be sufficient to put the BJT into saturation. The BJT acts as a switch, completing the circuit between the DC supply, the LED and the current limiting resistor, R_C. For this to work reliably, we have to make sure that the ratio of saturation current to base current is much less than β. A value of 10 or so would guarantee hard saturation.

If we would like to invert the logic, that is, have a logic low turn on the LED and a logic high turn it off, we can achieve that with a PNP version of the circuit as shown in Figure 4.20.

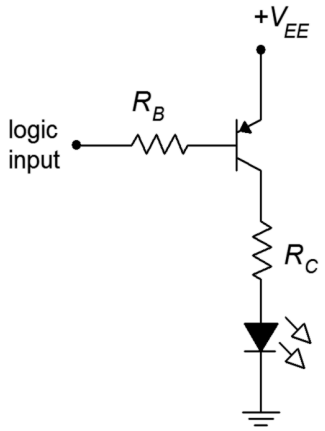
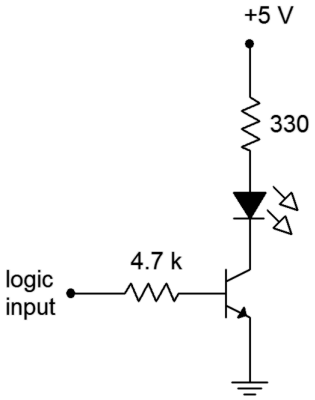


Figure 4.20
Saturating LED driver circuit (negative logic).

Example 4.3

Determine the LED "on" current for the circuit of Figure 4.21. Assume the logic "on" voltage is 5 volts, V_{LED} = 1.8 volts and $V_{CE(sat)}$ = 0.

Figure 4.21
Circuit for Example 4.3.



First, find the base current.

$$I_B = \frac{V_{logic} - V_{BE}}{R_B}$$
$$I_B = \frac{5\,\text{V} - 0.7\,\text{V}}{4.7\,\text{k}\Omega}$$
$$I_B = 915\,\mu\text{A}$$

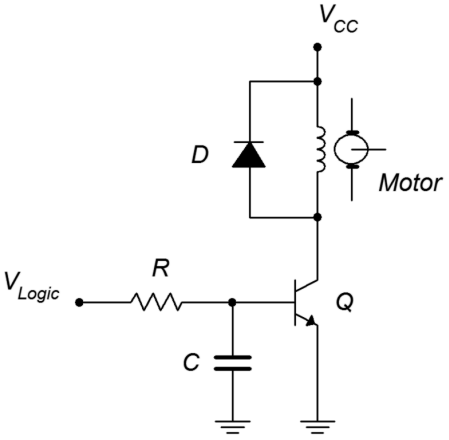
Now find $I_{C(sat)}$, making sure the BJT is in saturation. This will be the LED current.

$$I_{C(sat)} = \frac{V_{CC} - V_{LED}}{R_C}$$
$$I_{C(sat)} = \frac{5\,\text{V} - 1.8\,\text{V}}{330\,\Omega}$$
$$I_{C(sat)} = 9.7\,\text{mA}$$

The ratio of these two currents is just over 10:1. This will guarantee hard saturation.

There are many different applications for saturating switches. Just about anywhere you can imagine a relay being used, you can consider a transistor switch. The transistor switch has the advantages of small size, no moving parts to wear out and very fast switching speeds. Relays have the advantage for very high currents. Figure 4.22 shows an example of direct motor drive using a saturating BJT switch.

Figure 4.22
Direct DC motor drive.



This circuit is used to control the speed of a DC motor through a technique called *pulse width modulation*. The speed of the motor will depend on the average voltage applied to it. The trick here is that instead of applying a continuously variable voltage to the motor, we instead apply a series of pulses of varying width. These pulses are sufficient to saturate the BJT, causing it to behave as a switch. These pulses are so fast that the motor does not start and stop, but rather inertia keeps it going. Instead, the motor responds to the averaged value of these pulses. If the pulses are narrow and widely spaced, the average value will be low and the motor speed will be slow. If the pulses are wide and closely spaced, the average will high and the motor speed will be fast.

The resistor and capacitor at the base are used to shape the incoming pulse to improve performance. The diode across the motor winding is particularly important. It is referred to as a *snubbing diode*[18]. Without it, the switching transistor might experience large and damaging transient spikes. Here's why: Let's assume the BJT is on and conducting fully. This current is the same current flowing through the motor's armature, which is little more than a huge coil of wire. That means it exhibits a lot of inductance. When we turn off the transistor, we are attempting to turn off the armature current, but the current through an inductor cannot change instantaneously. The result is that the winding now generates a large flyback voltage (also called an "inductive kick") directly across the BJT. That is, the winding momentarily appears as a high voltage source of opposite polarity and, via KVL, this potential appears from collector to emitter. This could damage the BJT. The snubbing diode effectively short-circuits the winding when it reverses voltage polarity, preventing the large spike. The remainder of the time the diode is reverse-biased and effectively out of the circuit.

The Non-Saturating Driver

It is also possible to create a switch or driver that is non-saturating. An example of a non-saturating LED driver is shown in Figure 4.23.

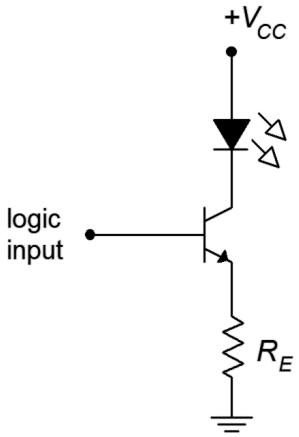


Figure 4.23
Non-saturating LED driver circuit (positive logic).

The advantage of this circuit is that it requires less current from the logic circuit. Unfortunately, it also exhibits higher transistor power dissipation and requires a DC source that is higher than the logic level. The operation is as follows: As in the saturating driver, if the logic level is zero, there is no rise in the base-emitter loop and the collector current will also be zero. With a high logic voltage, via KVL around the base-emitter loop, all of the logic input voltage drops across R_E, with the exception of V_{BE}. This creates I_E which is virtually the same as I_C (which is I_{LED}). This

18 It is also known as a commutating diode, clamp diode, flyback diode and by a host of other names. But as Shakespeare said, "A snubbing diode by any other name would clamp a flyback voltage as well". Or something like that.

circuit "programs" the emitter current via the resistor and logic voltage. Therefore it is fixed and stable. This process is sometimes referred to as *bootstrapping*. It might be said that the emitter voltage is "bootstrapped" to within 0.7 volts of the logic input level, keeping it stable[19]. In any case, if β varies, this will cause an inverse change in I_B with no change in I_C. A negative logic PNP version is also possible and left as an exercise.

Example 4.4

Determine the LED "on" current for the circuit of Figure 4.24. Assume the logic "on" voltage is 5 volts, V_{LED} = 1.8 volts and β = 100.

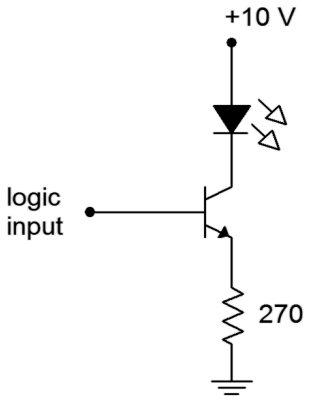


Figure 4.24
Circuit for Example 4.4.

We can find I_C directly because $I_C \approx I_E$. This will be the LED current.

$$I_C = \frac{V_{logic} - V_{BE}}{R_E}$$
$$I_C = \frac{5\,\text{V} - 0.7\,\text{V}}{270\,\Omega}$$
$$I_C = 15.9\,\text{mA}$$

Note that β was not used. All it tells us is that I_B = 15.9 mA/100, or 159 μA. A higher β would simply lead to a lower base current.

For the sake of completeness, we might also note that

$$V_{CE} = V_{CC} - V_{LED} - V_{RE}$$
$$V_{CE} = 10\,\text{V} - 1.8\,\text{V} - 4.3\,\text{V}$$
$$V_{CE} = 3.9\,\text{V}$$

Clearly, if V_{CE} is 3.9 volts, the transistor is not in saturation.

The Zener Follower

In the prior chapter we examined a method of regulating the output voltage of a filtered full-wave rectifier through the use of a Zener diode. The downfall of that specific circuit is that it was not particularly efficient because it drew a fair amount of current even when the demand for load current was light. Using the concept of locking one voltage to another, as in the non-saturating switch, we can create a nice improvement, the Zener Follower.

19 This is in reference to the old phrase "pulling yourself up by your bootstraps". To be honest, that saying never made sense to this author and all that ever happened when I tried to do it was that my arms got tired.

A Zener Follower is shown in Figure 4.25. The input signal is the positive rectified and filtered output of the AC-to-DC power supply.

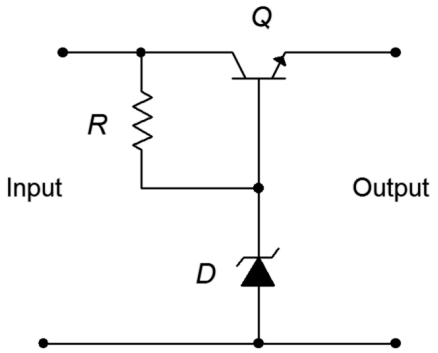


Figure 4.25
Zener Follower.

The primary thing to notice is that the Zener diode is reverse-biased via resistor *R*. That is, a current will flow down through *R* and into the Zener. The Zener presents a fixed potential, V_Z. Consequently, the difference between the input voltage and V_Z must drop across *R*, and by extension, V_{CB}. Further, the final output voltage is the voltage at the BJT's emitter which must be $V_Z - V_{BE}$. As these are both fixed, stable potentials, the output must likewise be a fixed, stable voltage. Lastly, because $V_{CE} = V_{CB} + V_{BE}$, it is apparent that any variation between the input voltage and the desired output (for example, due to ripple) must be dropped across the BJT.

The diode current is kept low in the Zener Follower and thus its power dissipation is also modest. Further, current draw from the input circuit is a direct reflection of load current demand. If the load current requirement is low, very little current will flow through the transistor, and ultimately, from the input circuit. This makes for a more efficient system.

Summary

A bipolar junction transistor may be thought of as an extension of a simple diode or PN junction. Another layer of doped material is added, resulting in either an NPN or PNP configuration, both with two depletion regions. The two depletion regions create two hills in the energy diagram. The three terminals of the device are called the emitter, base (middle) and collector. BJTs are not normally constructed symmetrically and swapping the collector and emitter can result in unpredictable behavior.

For proper operation, the base-emitter junction is forward-biased while the collector-base junction is reverse-biased. This results in the emitter and collector currents being very nearly equal and much, much larger than the base current. The ratio of collector current to base current is called β (beta) while the ratio of collector current to emitter current is called α (alpha). β in particular is subject to wide variations and

it can have a major impact on circuit parameters. A plot of collector current versus collector-emitter voltage reveals the three main regions of the BJT circuit: saturation, constant current and breakdown.

The Ebers-Moll model consists of a diode from the base to emitter and a controlled current source from the collector to base. This simple model of the BJT can be used to solve a variety of transistor circuits, particularly when used in conjunction with a DC load line. The DC load line is a plot of all possible operating points for a given transistor circuit.

Finally, it is possible to create switching and driver circuits using BJTs that produce stable output currents. These may utilize saturating or non-saturating configurations with NPN or PNP devices.

Review Questions

1. Describe the energy diagram for a forward-reverse biased BJT.
2. Define α.
3. Define β.
4. Define Early voltage. What is its significance?
5. What is a family of collector curves? What information can we derive from it?
6. Describe the Ebers-Moll BJT model.
7. Explain some of the issues involving variation of β.
8. What is a DC load line?
9. How is a saturating switch different from a non-saturating driver? What are the advantages and disadvantages of each?
10. What is a Zener Follower?

Problems

Analysis Problems

1. Determine β if α = 0.99.

2. Determine α if β = 200.

3. Determine the currents for the circuit of Figure 4.26 if V_{BB} = 5 V, V_{CC} = 20 V, R_B = 200 kΩ, R_C = 2 kΩ, β = 100.

4. Determine the transistor voltages for the circuit of Figure 4.26 if V_{BB} = 5 V, V_{CC} = 20 V, R_B = 200 kΩ, R_C = 2 kΩ, β = 200.

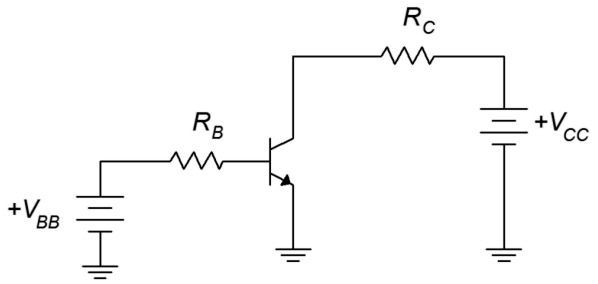


Figure 4.26

5. Determine the LED current in the circuit of Figure 4.27 if V_{logic} = 5 V, V_{CC} = 5 V, V_{LED} = 2.1 V, R_B = 3.6 kΩ, R_C = 270 Ω, β = 100.

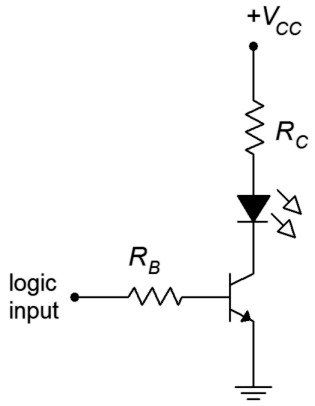


Figure 4.27

6. Determine the LED current in the circuit of Figure 4.27 if V_{logic} = 0 V, V_{CC} = 5 V, V_{LED} = 2.1 V, R_B = 3.6 kΩ, R_C = 270 Ω, β = 100.

7. Determine the LED current in the circuit of Figure 4.28 if V_{logic} = 5 V, V_{EE} = 5 V, V_{LED} = 2.2 V, R_B = 2.7 kΩ, R_C = 220 Ω, β = 100.

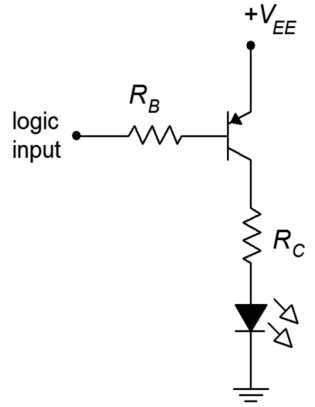


Figure 4.28

8. Determine the LED current in the circuit of Figure 4.28 if $V_{logic} = 0$ V, $V_{EE} = 5$ V, $V_{LED} = 2.2$ V, $R_B = 2.7$ kΩ, $R_C = 220$ Ω, β = 100.

9. Determine the LED current in the circuit of Figure 4.29 if $V_{logic} = 3.6$ V, $V_{CC} = 10$ V, $V_{LED} = 2.3$ V, $R_E = 270$ Ω, β = 200.

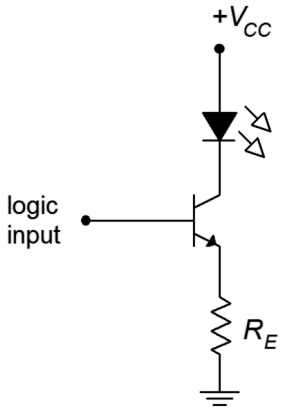


Figure 4.29

10. Determine the LED current in the circuit of Figure 4.29. $V_{logic} = 0$ V, $V_{CC} = 10$ V, $V_{LED} = 2.3$ V, $R_E = 270$ Ω, β = 200.

11. Using the 2N3904 data sheet, determine $V_{CE(sat)}$ if $I_C = 30$ mA and $I_B = 1$ mA.

12. Using the 2N3904 data sheet, determine the percent change in β if $I_C = 10$ mA and the temperature rises from 25°C to 125°C.

13. Using the 2N3904 data sheet, determine the percent change in β if $I_C = 40$ mA and the temperature drops from 25°C to −55°C.

Design Problems

14. Using Figure 4.26, determine a value for R_B to set I_C to 5 mA if $V_{BB} = 5$ V, $V_{CC} = 25$ V, $R_C = 2$ kΩ, β = 100.

15. Using Figure 4.26, determine a value for R_C to set V_{CE} to 6 V if $V_{BB} = 10$ V, $V_{CC} = 25$ V, $R_B = 330$ kΩ, β = 200.

16. For the circuit of Figure 4.27, determine a value for R_C to set the LED current to 15 mA. $V_{logic} = 5$ V, $V_{CC} = 5$ V, $V_{LED} = 1.6$ V, $R_B = 3.3$ kΩ.

17. For the circuit of Figure 4.28, determine a value for R_C to set the LED current to 20 mA. $V_{logic} = 0$ V, $V_{EE} = 5$ V, $V_{LED} = 2.0$ V, $R_B = 2.7$ kΩ.

18. For the circuit of Figure 4.29, determine a value for R_E to set the LED current to 25 mA. V_{logic} = 5 V, V_{CC} = 9 V, V_{LED} = 2.8 V.

Challenge Problems

19. Determine the maximum and minimum values for I_C in the circuit of Figure 4.26 if all resistors have a 10% tolerance and β = ranges from 100 to 200. V_{BB} = 5 V, V_{CC} = 20 V, R_B = 200 kΩ, R_C = 2 kΩ.

20. Derive and draw a PNP non-saturating LED driver circuit.

Computer Simulation Problems

21. Simulate the circuit of Problem 3.

22. Simulate the circuit of Problem 5.

23. Simulate the circuit of Problem 7.

24. Verify the design of Problem 14 using a simulator.

"If I have seen farther than others,
it was because I stood on the shoulders of giants."

- Sir Isaac Newton

5 BJT Biasing

5.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Explain the need for DC biasing of BJT amplifiers.
- Solve various BJT biasing circuits for device currents and voltages.
- Plot DC load lines for a variety of BJT biasing circuits.
- Discuss methods to increase circuit stability with regard to transistor parameter variation.

5.1 Introduction

As we saw in the preceding chapter, a bipolar junction transistor requires a forward-bias of the base-emitter junction and a reverse-bias of the collector-base junction in order to operate properly. One of the prime BJT parameters is the current gain, β. It can have a considerable impact on the operation of the circuit. Unfortunately, β also varies with changes in temperature, collector-emitter voltage, etc., and this can lead to circuit instability. In this chapter we shall investigate a variety of circuit topologies to bias the BJT, always with an eye toward stability.

5.2 The Need For Biasing

Why bias a transistor in the first place? After all, if the device exhibits current gain (i.e., β), why not just apply an AC signal at the base and obtain an amplified version of it at the collector? The first thing to remember is that current gain is an outgrowth of forward-reverse bias. Given that fact, and without an additional source of energy, amplification cannot be produced. Also, remember the magnitude of the energy hill required for forward-biasing the base-emitter. In order to achieve that, V_{BE} needs to be around 0.7 volts. If we simply applied an AC signal to the base, we could only hope to forward-bias the base-emitter when that signal exceeded 0.7 volts. The entire negative half of the AC signal would be ignored along with everything positive that's below 0.7 volts. Seeing that the voltage generated from many input devices such as microphones and sensors may only be a few hundred millivolts, the entire signal could be ignored! The solution to these problems is to apply a DC bias to the transistor and then superimpose the AC signal on top of that. In other words, if the AC voltage is riding on a much larger DC voltage, then even the negative peak of the AC signal will be a net positive voltage, and we can maintain proper transistor function.

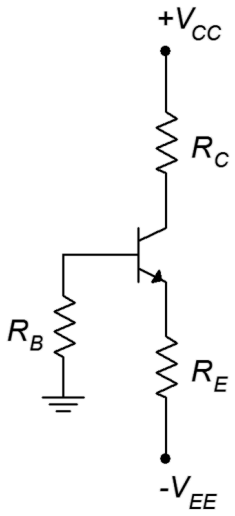
There are numerous ways to establish a proper polarity DC bias on a transistor. The trick is to find ways to make a stable bias, that is, to establish a Q point that doesn't move in spite of parameter changes such as changes in β. As we shall see in following chapters, an unstable Q point can have negative effects on the AC performance of

an amplifier. For example, it could make the gain unstable, increase distortion or reduce output power. This lack of stability is a major problem with the base bias configuration examined in the prior chapter. What we would like is a circuit that will establish a collector current that does not shift even when β changes.

5.3 Two-Supply Emitter Bias

For proper functioning, the collector-base junction needs to be reverse-biased and the base-emitter junction needs to be forward-biased. For an NPN transistor that means that the collector must be at the highest potential, the base somewhat lower and the emitter at the lowest potential of the three. One way of doing this is to apply the usual positive supply to the collector, but instead of using a second potential at the base, the base is tied to ground through a resistor. The requisite forward-bias on the base-emitter is then achieved by connecting the emitter to a negative power supply. This circuit configuration is shown in Figure 5.1 using an NPN device. We shall refer to this as *two-supply emitter bias*.

Figure 5.1
Two-supply emitter bias, NPN version.



We can derive an equation for the collector current by applying KVL to the base-emitter loop:

$$V_{EE} = V_{R_B} + V_{BE} + V_{R_E}$$
$$V_{EE} = I_B R_B + V_{BE} + I_E R_E$$

Recalling that $I_B = I_C/\beta$ and $I_E \approx I_C$,

$$V_{EE} = (I_C/\beta) R_B + V_{BE} + I_C R_E$$

Solving for I_C we arrive at

$$I_C = \frac{|V_{EE}| - V_{BE}}{R_E + R_B/\beta} \qquad (5.1)$$

The absolute value has been added to the emitter supply voltage so there is no confusion regarding the sign of this potential in the equation.

The thing to notice about equation 5.1 is that β only partly determines the collector current. In fact, if we can make $R_E >> R_B/\beta$, then the equation reduces to

$$I_C \approx \frac{|V_{EE}| - V_{BE}}{R_E} \qquad (5.2)$$

It is relatively easy to achieve the $R_E >> R_B/\beta$ stipulation. Given typical values of β, this will be the case if R_E is approximately equal to or larger than R_B. What we find in this instance is that almost all of the emitter supply drops across R_E to establish a

stable I_C with β playing virtually no role. If β changes, the result will be an inverse change in I_B with I_C remaining largely unchanged.

Now that we have the collector current, any other current or voltage in the circuit may be derived by applying Ohm's law, KVL and the like. For example, to find V_C, the voltage from the collector to ground,

$$V_C = V_{CC} - V_{R_C}$$
$$V_C = V_{CC} - I_C R_C$$

And to find the transistor's collector-emitter voltage, V_{CE},

$$V_{CE} = V_{CC} + |V_{EE}| - V_{R_C} - V_{R_E}$$
$$V_{CE} = V_{CC} + |V_{EE}| - I_C R_C - I_C R_E$$
$$V_{CE} = V_{CC} + |V_{EE}| - I_C (R_C + R_E) \qquad (5.3)$$

Note that V_{CE} can also be found via $V_{CE} = V_C - V_E$. Dropping voltages along the base-emitter loop yields

$$V_E = -V_{R_B} - V_{BE}$$
$$V_E = -I_B R_B - V_{BE}$$

Also, it is to our advantage to develop the DC load line for this configuration. The load line can serve as a "sanity check" for our computations. To find the endpoints, $I_{C(sat)}$ is the maximum current and will occur when $V_{CE} = 0$. If we imagine the current rising as V_{CE} collapses, eventually all of the available supply voltage will have dropped across R_C and R_E. Thus

$$I_{C(sat)} = \frac{V_{CC} + |V_{EE}|}{R_C + R_E} \qquad (5.4)$$

Similarly, $V_{CE(cutoff)}$ occurs when $I_C = 0$. That means that there will be no potentials across R_C and R_E. Therefore, V_{CE} "absorbs" the entire available source voltage.

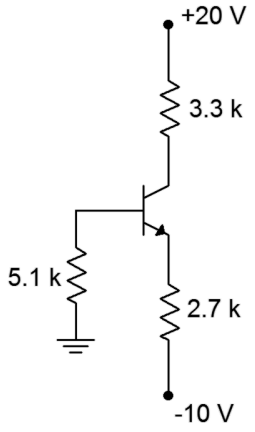
$$V_{CE(cutoff)} = V_{CC} + |V_{EE}| \qquad (5.5)$$

Do not attempt to memorize all of the myriad equations presented. There are simply too many variations on the theme and it will only get worse when other biasing configurations are introduced. Instead, remember how to find the collector current and then get in the habit of applying Ohm's law and KVL to derive whatever else you may need.

At this point a comprehensive example is called for.

Assuming β = 100, plot the Q point (I_C and V_{CE}) on the load line for the circuit of Figure 5.2.

Figure 5.2
Circuit for Example 5.1.



Using Equation 5.1:

$$I_C = \frac{|V_{EE}| - V_{BE}}{R_E + R_B/\beta}$$
$$I_C = \frac{10 - 0.7}{2.7\,\text{k}\Omega + 5.1\,\text{k}\Omega/100}$$
$$I_C = 3.38\,\text{mA}$$

Noting the relative sizes of R_E and R_B, the approximation should be close.

$$I_C = \frac{|V_{EE}| - V_{BE}}{R_E}$$
$$I_C = \frac{10 - 0.7}{2.7\,\text{k}\Omega}$$
$$I_C = 3.44\,\text{mA}$$

To find V_{CE} we can use the equation derived above (Equation 5.3).

$$V_{CE} = V_{CC} + |V_{EE}| - I_C(R_C + R_E)$$
$$V_{CE} = 20\,\text{V} + 10\,\text{V} - 3.38\,\text{mA}(3.3\,\text{k}\Omega + 2.7\,\text{k}\Omega)$$
$$V_{CE} = 9.72\,\text{V}$$

Now calculate the load line endpoints:

$$I_{C(sat)} = \frac{V_{CC} + |V_{EE}|}{R_C + R_E}$$
$$I_{C(sat)} = \frac{20\,\text{V} + 10\,\text{V}}{3.3\,\text{K}\Omega + 2.7\,\text{K}\Omega}$$
$$I_{C(sat)} = 5\,\text{mA}$$

$$V_{CE(cutoff)} = V_{CC} + |V_{EE}|$$
$$V_{CE(cutoff)} = 20\,\text{V} + -10\,\text{V}$$
$$V_{CE(cutoff)} = 30\,\text{V}$$

The load line for the circuit in Example 5.1 is shown in Figure 5.3.

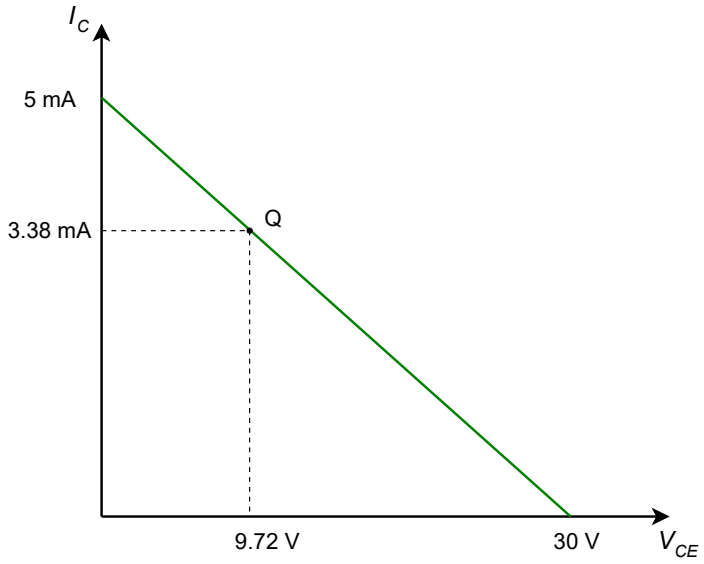


Figure 5.3
DC load line for the circuit of Figure 5.2.

Note the proportions between voltage and current for the Q point. The voltage is a little less than one-third of the maximum while the current is a little more than two-thirds of its maximum.

Verification of Stability

The claim was made that two-supply emitter bias circuits like the one Figure 5.1 potentially have a stable Q point. If we were to plot a second Q point for a large change in β, it should hardly move, thus indicating very high stability. For example, doubling β to 200 results in $I_C =$ 3.41 mA and $V_{CE} = 9.53$ V. The new Q point has edged just slightly closer to saturation, producing about a 1% change in current for a 100% change in β. Clearly, this configuration can produce very small changes in the Q point in spite of very large changes in β.

Computer Simulation

The two-supply emitter bias circuit of Figure 5.4 is simulated using the DC Bias function. A quick estimation shows that we expect about 2 mA of collector current (9.3 V/4.7 kΩ) and a collector voltage of about 8 volts (15 V − 2 mA · 3.6 kΩ). We also expect a small negative potential at the base $-I_B R_B$). Given typical β values for the 2N3904 (200-ish at this current, refer back to the data sheet), we expect a base current of around 10 to 15 μA, leaving us with a V_B of approximately −0.1 volts. The emitter voltage would be about 0.7 volts less than that, perhaps −0.8 volts or so.

In short, for a properly designed circuit of this type we expect V_B to be pretty close to 0 V and V_E to be about −0.7 volts.

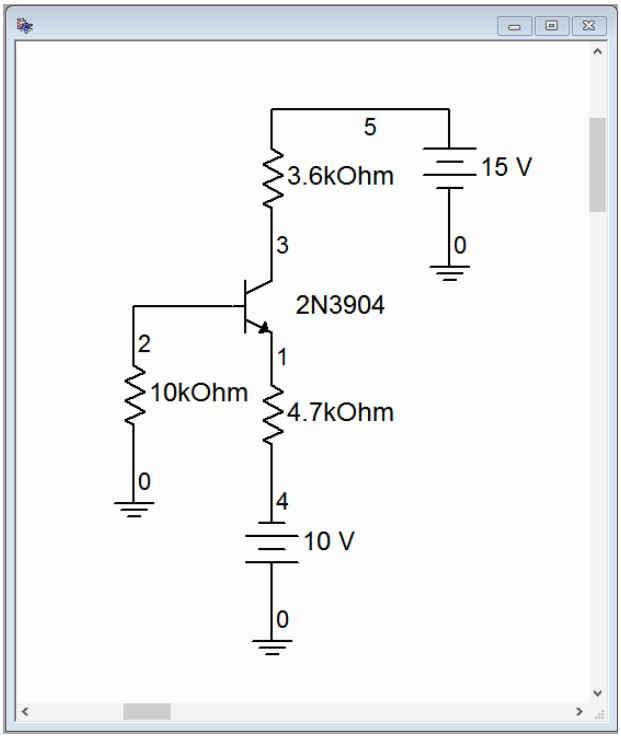


Figure 5.4
Schematic for two-supply emitter bias simulation.

The results are shown in Figure 5.5. The node voltages agree with our estimations. Node 3 is the collector voltage, very close to the estimation. The results for the base voltage (node 2) and the emitter voltage (node 1) are also in line with the estimates.

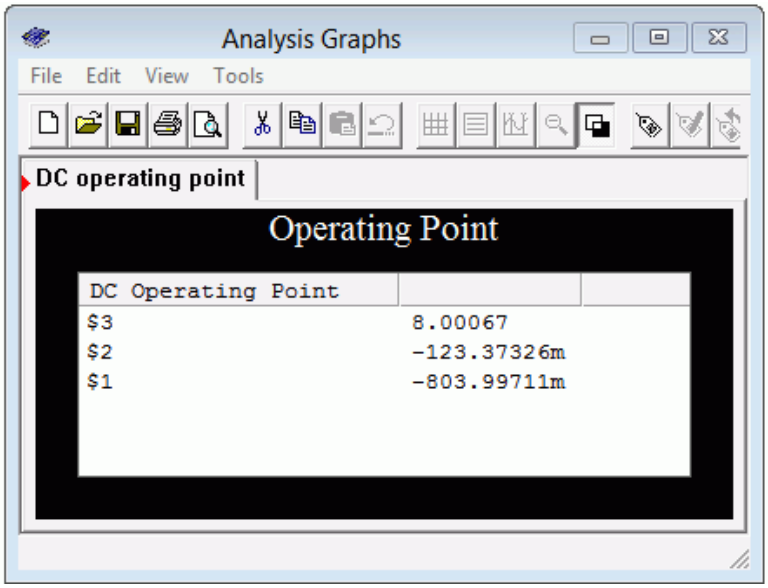
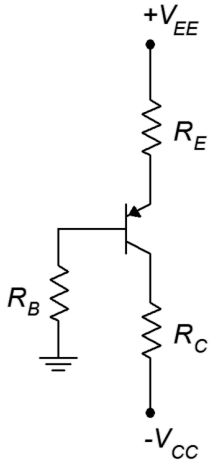


Figure 5.5
Simulation results for two-supply emitter bias circuit.

PNP Two-Supply Emitter Bias

While it is possible to create a PNP version of bias circuits by simply swapping out the device and then changing the signs of the power supplies, it is common to "flip" the entire circuit from top to bottom so that the emitter winds up on top and the collector on the bottom. One advantage of this is that, in a multi-transistor circuit schematic, all of the DC bias currents "run down the page", that is, the collector currents flow from the top of the page to the bottom of the page. Figure 5.6 shows a PNP two-supply emitter bias circuit.

Figure 5.6
PNP two-supply emitter bias circuit.

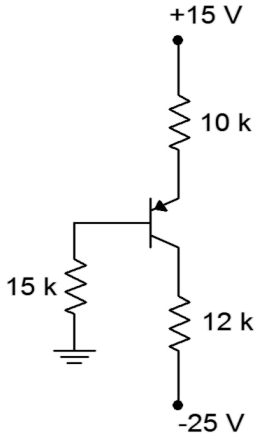


All of the device current equations and component voltage equations derived for the NPN version will hold for the PNP version. The differences to remember are that the voltage polarities will be reversed (what was positive in the NPN is negative in the PNP) and the current directions will be reversed (e.g., conventional current flows into the NPN's collector but out of the PNP's collector. For example, in the NPN we expect the base current to flow into the base terminal. This creates a small negative voltage at the base and a somewhat more negative voltage (by 0.7 V) at the emitter. In the PNP, the base current flows out of the base. This creates a small positive voltage at the base and results in the emitter being slightly more positive (by 0.7 V). This is perhaps best illustrated with an example.

Example 5.2

Assuming β = 100, determine the Q point and load line endpoints of the circuit of Figure 5.7.

Figure 5.7
Circuit for Example 5.2.



First, note that this is a PNP drawn upside down so the emitter is at the top. Using Equation 5.1:

$$I_C = \frac{|V_{EE}| - V_{BE}}{R_E + R_B/\beta}$$
$$I_C = \frac{15 - 0.7}{10\,\text{k}\Omega + 15\,\text{k}\Omega/100}$$
$$I_C = 1.409\,\text{mA}$$

As a cross check, noting the relative sizes of R_E and R_B, the approximation should be close.

$$I_C = \frac{|V_{EE}| - V_{BE}}{R_E}$$
$$I_C = \frac{15 - 0.7}{10\,\text{k}\Omega}$$
$$I_C = 1.43\,\text{mA}$$

To find V_{CE} we can use Equation 5.3 with a slight modification.

$$V_{CE} = V_{EE} + |V_{CC}| - I_C(R_C + R_E)$$
$$V_{CE} = 15\,\text{V} + 25\,\text{V} - 1.409\,\text{mA}(12\,\text{k}\Omega + 10\,\text{k}\Omega)$$
$$V_{CE} = 9\,\text{V}$$

We complete the picture by determining the endpoints of the load line.

$$I_{C(sat)} = \frac{V_{EE} + |V_{CC}|}{R_C + R_E}$$
$$I_{C(sat)} = \frac{10\,\text{V} + 25\,\text{V}}{12\,\text{K}\Omega + 10\,\text{K}\Omega}$$
$$I_{C(sat)} = 1.818\,\text{mA}$$

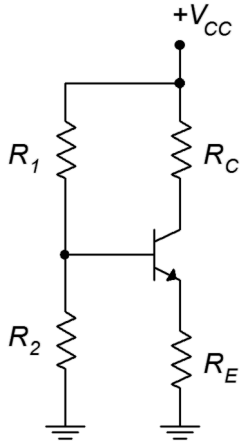
$$V_{CE(cutoff)} = V_{EE} + |V_{CC}|$$
$$V_{CE(cutoff)} = 15\,\text{V} + 25\,\text{V}$$
$$V_{CE(cutoff)} = 40\,\text{V}$$

The Q point is about 3/4$^{\text{ths}}$ of the maximum current and 1/4$^{\text{th}}$ of the maximum voltage.

5.4 Voltage Divider Bias

Another configuration that can provide high bias stability is *voltage divider bias*. Instead of using a negative supply off of the emitter resistor, like two-supply emitter bias, this configuration returns the emitter resistor to ground and raises the base voltage. So as to avoid issues with a second power supply, this base voltage is derived from the collector power supply via a voltage divider. The bias template is shown in Figure 5.8.

Figure 5.8
Voltage divider bias.



Let's derive the equations for the load line. First, let's consider the saturation and cutoff endpoints. For saturation, assume V_{CE} goes to 0. What resistances are left to limit the current?

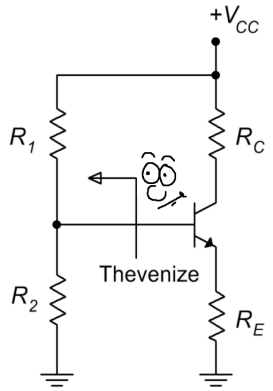
$$I_{C(sat)} = \frac{V_{CC}}{R_C + R_E} \tag{5.6}$$

$V_{CE(cutoff)}$ occurs when $I_C = 0$ and that means that there will be no potentials across R_C and R_E. Therefore, V_{CE} takes on the entire available source voltage.

$$V_{CE(cutoff)} = V_{CC} \tag{5.7}$$

The key to finding the Q point (and pretty much any other current or voltage in the circuit) is to find I_C. To simplify the process, Thevenize the voltage divider as shown in Figure 5.9.

Figure 5.9
Thevenizing the voltage divider.



(a)

By inspection of Figure 5.9b,

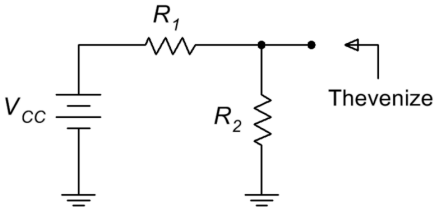
$$V_{TH} = V_{CC}\frac{R_2}{R_1+R_2}$$
$$R_{TH} = R_1 \parallel R_2 = \frac{R_1 R_2}{R_1+R_2}$$

Now we can derive an equation for the collector current by applying KVL to the base-emitter loop of Figure 5.9c:

$$V_{TH} = V_{R_{TH}} + V_{BE} + V_{R_E}$$
$$V_{TH} = I_B R_{TH} + V_{BE} + I_E R_E$$

Recalling that $I_B = I_C/\beta$ and $I_E \approx I_C$,

$$V_{TH} = (I_C/\beta) R_{TH} + V_{BE} + I_C R_E$$

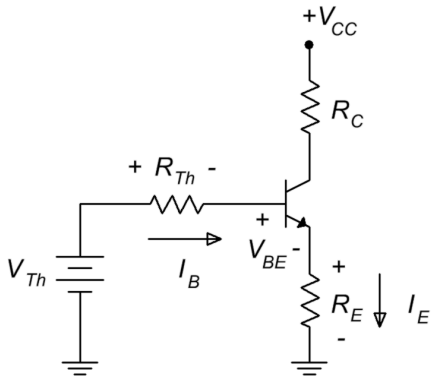


(b)

Solving for I_C we arrive at

$$I_C = \frac{V_{TH} - V_{BE}}{R_E + R_{TH}/\beta} \qquad (5.8)$$

Can we find a quick approximation for I_C as well? If we assume that the voltage divider of R_1 and R_2 is lightly loaded, in other words, that the base current is much, much less than the divider current, finding I_C is easy. The divider voltage yields the base voltage. We then subtract the 0.7 volt drop on the base-emitter and what's left drops across R_E. From there it's one short application of Ohm's law to get I_E, which is approximately equal to I_C. But how do we know if the divider is lightly loaded in the first place without going through the Thevenin equivalent? Looking at Equation 5.8, as long as $R_E >> R_{TH}/\beta$, we can ignore the second term in the denominator, leaving us with our quick approximation. Given typical values for β, as long as R_2 is not much larger than R_E, the approximation will be reasonably accurate.



(c)

Once I_C is obtained we can find the transistor's collector-emitter voltage, V_{CE},

$$V_{CE} = V_{CC} - V_{R_C} - V_{R_E}$$
$$V_{CE} = V_{CC} - I_C R_C - I_C R_E$$
$$V_{CE} = V_{CC} - I_C(R_C + R_E) \qquad (5.9)$$

Time for yet another thrilling illustrative example.

Example 5.3

Assuming β = 200, plot the Q point (I_C and V_{CE}) on the load line for the circuit of Figure 5.10. Also determine the value of V_B.

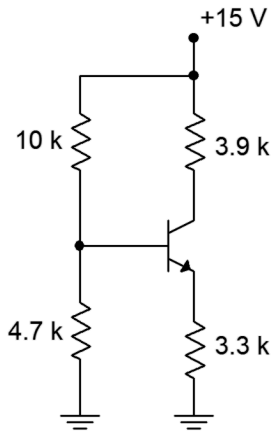


Figure 5.10
Circuit for Example 5.3.

Calculate the load line endpoints so we know the maximums.

$$I_{C(sat)} = \frac{V_{CC}}{R_C + R_E}$$
$$I_{C(sat)} = \frac{15\,\text{V}}{3.9\,\text{K}\,\Omega + 3.3\,\text{K}\,\Omega}$$
$$I_{C(sat)} = 2.08\,\text{mA}$$

$$V_{CE(cutoff)} = V_{CC}$$
$$V_{CE(cutoff)} = 15\,\text{V}$$

To obtain the Q point, first find the Thevenin values.

$$V_{TH} = V_{CC}\frac{R_2}{R_1 + R_2}$$
$$V_{TH} = 15\,\text{V}\frac{4.7\,\text{k}\,\Omega}{10\,\text{k}\,\Omega + 4.7\,\text{k}\,\Omega}$$
$$V_{TH} = 4.8\,\text{V}$$

$$R_{TH} = R_1 \| R_2 = \frac{R_1 R_2}{R_1 + R_2}$$
$$R_{TH} = \frac{10\,\text{k}\,\Omega \times 4.7\,\text{k}\,\Omega}{10\,\text{k}\,\Omega + 4.7\,\text{k}\,\Omega}$$
$$R_{TH} = 3.2\,\text{k}\,\Omega$$

Using Equation 5.8:

$$I_C = \frac{V_{TH} - V_{BE}}{R_E + R_{TH}/\beta}$$
$$I_C = \frac{4.8\,\text{V} - 0.7\,\text{V}}{3.3\,\text{k}\,\Omega + 3.2\,\text{k}\,\Omega/200}$$
$$I_C = 1.236\,\text{mA}$$

Noting the relative sizes of R_E and R_2, the approximation should be fairly accurate.

$$I_C = \frac{V_{TH} - V_{BE}}{R_E}$$
$$I_C = \frac{4.8\,\text{V} - 0.7\,\text{V}}{3.3\,\text{k}\,\Omega}$$
$$I_C = 1.242\,\text{mA}$$

To find V_{CE} we can use Equation 5.9.

$$V_{CE} = V_{CC} - I_C(R_C + R_E)$$
$$V_{CE} = 15\,\text{V} - 1.236\,\text{mA}(3.9\,\text{k}\,\Omega + 3.3\,\text{k}\,\Omega)$$
$$V_{CE} = 6.1\,\text{V}$$

As far as finding V_B is concerned, a decent approximation would be the value of V_{TH} because we have determined that the divider is lightly loaded. In a more general sense, we could also find the drop across R_E and then add V_{BE}. The approximation yields 4.8 volts and the more accurate method yields

$$V_B = V_{BE} + I_C R_E$$
$$V_B = 0.7\,\text{V} + 1.236\,\text{mA} \times 3.3\,\text{k}\,\Omega$$
$$V_B = 4.78\,\text{V}$$

The load line for the circuit in Example 5.3 is shown in Figure 5.11.

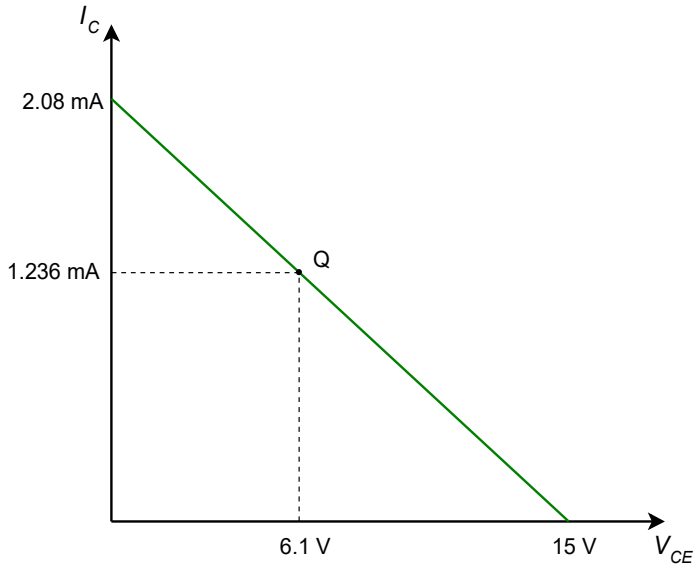


Figure 5.11
DC load line for the circuit of Figure 5.10.

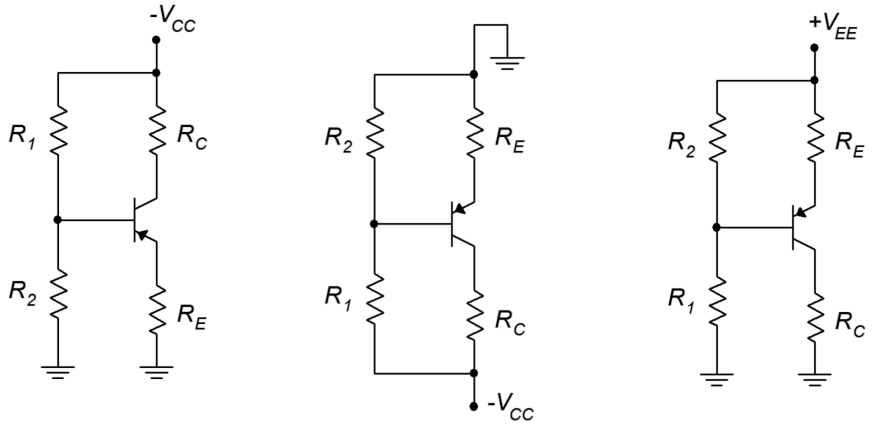
Once again the proportions between voltage and current for the Q point appear to be proper when compared against the endpoints.

Verification of Stability

How much does the Q point move if β were to get cut half? Recalculating with a β of 100 yields I_C = 1.23 mA and V_{CE} = 6.14 V. This represents a shift in both current and voltage of less than 1%. This will, of course, cause a near doubling of I_B but this will be hardly noticed here as the divider current is so much larger; approximately 15V/(10 k + 4.7 k) or 1 mA versus about 1.23 mA/100 or 12.3 μA.

PNP Voltage Divider Bias

To create the PNP version of the voltage divider bias, we replace the NPN with a PNP and then change the sign of the power supply. As mentioned with the two-supply emitter bias, these circuits are usually flipped top to bottom resulting in the flow of DC current going down the page. All of the currents and component voltages are unchanged except that their directions and polarities are reversed. The current equations and so forth remain valid. Something a little odd-looking happens with the voltage divider bias, though: we end up with ground being the most positive potential and a negative supply at the bottom of the schematic. It works, but it's an issue if we're using a traditional positive supply elsewhere in the circuit. After all, why have two supplies where one will do? It turns out that we can make a positive supply version fairly easily. All we need to do is add the magnitude of the negative source voltage to the ground and power connections. This progression is shown in Figure 5.12.



(a) (b) (c)

Figure 5.12
Progression of PNP voltage divider bias circuit.
a. Direct conversion from NPN.
b. Top-to-bottom flip.
c. DC supply offset added to achieve a positive supply.

There is nothing magic about this procedure. In essence, all we've really done is renamed the reference point. All of the individual component voltages remain unchanged. For example, looking at Figure 5.12c versus 5.12b, it is still the case that the top connection to R_E is more positive than the bottom connection to R_C by the voltage V_{CC} (although we did rename the supply to V_{EE} to be consistent with where it's connected). What *has* happened is that all ground-referenced (i.e., single

subscript) voltages have changed. For example, V_B in Figures 5.12a and 5.12b is the voltage across R_2. In contrast, V_B in Figure 5.12c is the voltage across R_1. That makes sense. If we move the reference then any voltage that is measured against the reference will change.

When analyzing the PNP voltage divider, we could simply parrot the collector current formula developed for the NPN, but there are other techniques. Two methods are illustrated in the following example.

Example 5.4

Assuming β = 200, determine the Q point (I_C and V_{CE}) for the circuit of Figure 5.13. Also determine the values of V_C and V_B.

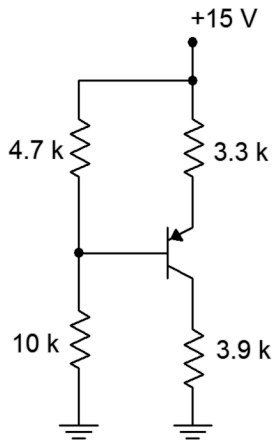


Figure 5.13
Circuit for Example 5.4.

First off, R_2 (now on top) is around the same size as R_E so the approximation method should be accurate and we can assume the divider is lightly loaded.

Method One

We will focus on the base-emitter loop as usual because V_{BE} is a known potential. Our immediate goal is to find the voltage across R_E so that we can use Ohm's law to find I_C. First we note that the voltage drop across R_2 is equal to the combined drops across R_E and V_{BE}. The drop across R_2 is found via the voltage divider rule.

$$V_{R_2} = V_{EE}\frac{R_2}{R_1 + R_2}$$
$$V_{R_2} = 15\,\text{V}\frac{4.7\,\text{k}\,\Omega}{10\,\text{k}\,\Omega + 4.7\,\text{k}\,\Omega}$$
$$V_{R_2} = 4.8\,\text{V}$$

And

$$V_{R_E} = V_{R_2} - V_{BE}$$
$$V_{R_E} = 4.8\,\text{V} - 0.7\,\text{V}$$
$$V_{R_E} = 4.1\,\text{V}$$

Therefore

$$I_C = \frac{V_{R_E}}{R_E}$$
$$I_C = \frac{4.1\,\text{V}}{3.3\,\text{k}\,\Omega}$$
$$I_C = 1.24\,\text{mA}$$

Method Two

Here we will determine all voltages with respect to ground.

$$V_B = V_{EE}\frac{R_1}{R_1+R_2}$$
$$V_B = 15\,\text{V}\frac{10\,\text{k}\Omega}{10\,\text{k}\Omega+4.7\,\text{k}\Omega}$$
$$V_B = 10.2\,\text{V}$$

The voltage from base to emitter has a − to + polarity, meaning it is a rise of 0.7 volts. Therefore

$$V_E = V_B+V_{BE}$$
$$V_E = 10.2\,\text{V}+0.7\,\text{V}$$
$$V_E = 10.9\,\text{V}$$

The voltage across R_E is the difference between V_{EE} and V_E.

$$V_{R_E} = V_{EE}-V_E$$
$$V_{R_E} = 15\,\text{V}-10.9\,\text{V}$$
$$V_{R_E} = 4.1\,\text{V}$$

This is the same value we arrived at using method one, so the collector current must be the same at 1.24 mA.

To find V_{CE} we also have options. One path is to use a slightly modified Equation 5.9.

$$V_{CE} = -\left(V_{EE}-I_C\left(R_C+R_E\right)\right)$$
$$V_{CE} = -15\,\text{V}+1.24\,\text{mA}\left(3.9\,\text{k}\Omega+3.3\,\text{k}\Omega\right)$$
$$V_{CE} = -6.07\,\text{V}$$

The collector is negative relative to the emitter, hence the negative sign. To avoid this, we could just swap the leads and refer to V_{EC} instead.

Alternately, we could find V_{CE} by determining V_C and then subtracting V_E from it.

$$V_C = I_C R_C$$
$$V_C = 1.24\,\text{mA}\times 3.9\,\text{k}\Omega$$
$$V_C = 4.84\,\text{V}$$

$$V_{CE} = V_C - V_E$$
$$V_{CE} = 4.84\,\text{V} - 10.9\,\text{V}$$
$$V_{CE} = -6.06\,\text{V}$$

We see a very slight difference here due to carried rounding errors.

It is instructive to compare the results of Example 5.4 back to Example 5.3. These circuits are otherwise identical except for the fact that one is NPN and the other is PNP. We find the same results for device currents (I_C) and component voltage magnitudes (V_{CE} or the voltage across R_E); only the signs and directions are reversed. On the other hand, we find that ground referenced potentials such as V_B, V_C and V_E are decidedly different between the two circuits.

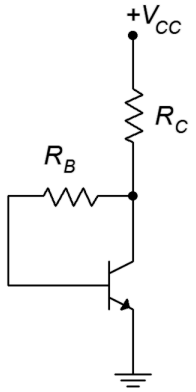
5.5 Feedback Biasing

While two-supply emitter bias and voltage divider bias can produce very high stability, there are other bias configurations available. Their stability tends not to be quite as good, but they are superior to simple base bias. They also tend to use fewer components than their high stability cousins. As a group, we refer to these as *feedback biasing* configurations. They use the concept of *negative feedback*. This is a technique where a change in the output can be reflected back to the input in such a way that it tends to partially offset the output change.

Collector Feedback Bias

With a simple move of R_B in the basic base bias configuration, we arrive at collector feedback bias. The NPN template is shown in Figure 5.14. Compared to base bias, all that has changed is that R_B is connected to the lower part of R_C rather than to the power supply. That small change can have a noticeable effect on stability.

Figure 5.14
Collector feedback bias.



To understand how feedback works, assume that a current is flowing from the supply, through R_C, into the collector and finally, out of the emitter to ground. Via KVL, $V_{CE} = V_C = V_{CC} - I_C \cdot R_C$. Now suppose for some reason, a temperature change perhaps, β increases. This should cause an increase in I_C. An increase in I_C, though, would cause an increase in the drop across R_C due to Ohm's law. This, in turn, would force V_C to drop. Here is the key: V_C is also equal to the drop across R_B plus the voltage V_{BE}. The base-emitter potential is fixed at approximately 0.7 volts so any decrease in V_C is reflected as a decrease in voltage across R_B. By Ohm's law, that means that I_B must decrease by a similar proportion. This decrease tends to offset the initial tendency of the collector current to increase.

To derive an equation for the collector current, we can use KVL.

$$V_{CC} = V_{R_C} + V_{R_B} + V_{BE}$$
$$V_{CC} = I_E R_C + I_B R_B + V_{BE}$$
$$V_{CC} = I_C R_C + \frac{I_C}{\beta} R_B + V_{BE}$$
$$I_C = \frac{V_{CC} - V_{BE}}{R_C + R_B/\beta} \qquad (5.10)$$

This equation is very similar to the current derivations for the two-supply emitter bias (Eq 5.1) and voltage divider bias (Eq 5.8). Again, if we can set $R_C >> R_B/\beta$ then I_C will be relatively immune from Q point shifts due to β. The problem here is that it's not nearly so easy to meet that stipulation in this circuit. Consequently, collector feedback tends to have only modest stability.

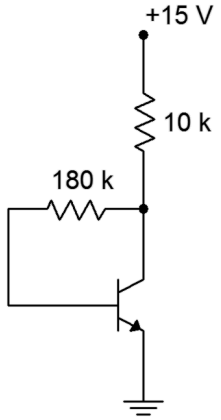
Concerning the cutoff and saturation endpoints on the DC load line, once again, cutoff is determined by the DC power supply while saturation is determined by the amount of resistance in the collector-emitter to limit said power supply's current.

$$I_{C(sat)} = \frac{V_{CC}}{R_C} \qquad (5.11)$$
$$V_{CE(cutoff)} = V_{CC} \qquad (5.12)$$

Example 5.5

Assuming β = 100, determine the Q point (I_C and V_{CE}) for the circuit of Figure 5.15. How much does the Q point change if β is halved?

Figure 5.15
Circuit for Example 5.5.



Using Equation 5.10

$$I_C = \frac{V_{CC} - V_{BE}}{R_C + R_B/\beta}$$
$$I_C = \frac{15\,\text{V} - 0.7\,\text{V}}{10\,\text{k}\,\Omega + 180\,\text{k}\,\Omega/100}$$
$$I_C = 1.21\,\text{mA}$$

Using KVL

$$V_{CE} = V_{CC} - V_{R_C}$$
$$V_{CE} = V_{CC} - I_C R_C$$
$$V_{CE} = 15\,\text{V} - 1.21\,\text{mA} \times 10\,\text{k}\,\Omega$$
$$V_{CE} = 2.9\,\text{V}$$

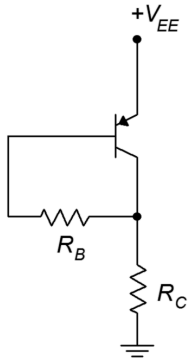
If β is halved to 50

$$I_C = \frac{V_{CC} - V_{BE}}{R_C + R_B/\beta}$$
$$I_C = \frac{15\,\text{V} - 0.7\,\text{V}}{10\,\text{k}\,\Omega + 180\,\text{k}\,\Omega/50}$$
$$I_C = 1.05\,\text{mA}$$

$$V_{CE} = V_{CC} - I_C R_C$$
$$V_{CE} = 15\,\text{V} - 1.05\,\text{mA} \times 10\,\text{k}\,\Omega$$
$$V_{CE} = 4.5\,\text{V}$$

For a 2:1 drop in β we see about a 13% reduction in I_C with a somewhat larger change in V_{CE}. This circuit is clearly not as stable as the two-supply emitter bias or the voltage divider bias but it is superior to base bias.

Figure 5.16
PNP Collector feedback bias.

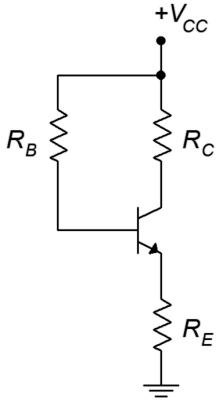


The PNP version of the collector feedback bias configuration should come as no surprise. The template is shown in Figure 5.16. Here, we use the same technique of power supply shifting that was used with the PNP voltage divider in order to wind up with a positive power supply. As with the PNP voltage divider, because we have changed the reference point, all ground referenced voltages will be different from their NPN counterparts. All currents and component voltages will have the same magnitudes but with opposite directions and polarities.

Emitter Feedback Bias

The emitter feedback bias uses the same overall idea as the collector feedback circuit, namely, that changes at the output will be reflected back to the input and thus help mitigate the initial change. While collector feedback focuses on collector current establishing V_C via R_C, emitter feedback uses the fact that emitter current establishes V_E via R_E. In both cases, these voltages are used to change the voltage across R_B, which results in a change in I_B that opposes the original collector current change.

Figure 5.17
Emitter feedback bias.



A basic emitter feedback bias circuit is shown in Figure 5.17. We shall use KVL to develop an equation for collector current.

$$V_{CC} = V_{R_B} + V_{BE} + V_{R_E}$$
$$V_{CC} = I_B R_B + V_{BE} + I_E R_E$$
$$V_{CC} = \frac{I_C}{\beta} R_B + I_C R_E + V_{BE}$$
$$I_C = \frac{V_{CC} - V_{BE}}{R_E + R_B/\beta} \qquad (5.13)$$

If we can set $R_E >> R_B/\beta$ then the Q point will be stable in spite of changes in β. The problem here is the same as was the case in collector feedback, namely that this stipulation is not easy to achieve. Consequently, the emitter feedback configuration tends to have only modest stability. In any event, once the collector current is known, V_{CE} can be found using the techniques illustrated with the voltage divider configuration. The endpoints for the DC load line are found in the usual manner.

$$I_{C(sat)} = \frac{V_{CC}}{R_C + R_E} \tag{5.14}$$

$$V_{CE(cutoff)} = V_{CC} \tag{5.15}$$

Example 5.6

Assuming β = 100, determine the Q point (I_C and V_{CE}) for the circuit of Figure 5.18.

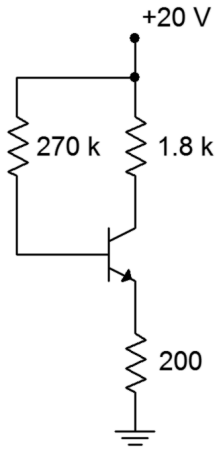


Figure 5.18
Circuit for Example 5.6.

Using Equation 5.13

$$I_C = \frac{V_{CC} - V_{BE}}{R_E + R_B/\beta}$$

$$I_C = \frac{20\,\text{V} - 0.7\,\text{V}}{200\,\Omega + 270\,\text{k}\,\Omega/100}$$

$$I_C = 6.66\,\text{mA}$$

Using KVL

$$V_{CE} = V_{CC} - V_{R_C} - V_{R_E}$$

$$V_{CE} = V_{CC} - I_C(R_C + R_E)$$

$$V_{CE} = 20\,\text{V} - 6.66\,\text{mA}\,(1.8\,\text{k}\,\Omega + 200\,\Omega)$$

$$V_{CE} = 6.68\,\text{V}$$

To complete the load line, we find

$$I_{C(sat)} = 10\ \text{mA}$$

$$V_{CE(cutoff)} = 20\ \text{V}$$

Dropping β to 50 will result in

$$I_C = 3.45\ \text{mA}$$

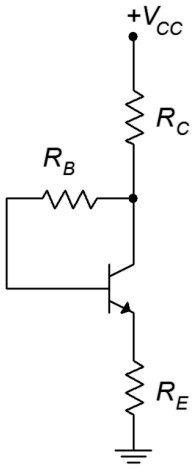
$$V_{CE} = 13.1\ \text{V}$$

We see less than a 2:1 change in I_C and V_{CE} but the stability is not dramatic.

Combination Feedback Bias

The final feedback bias configuration combines both collector feedback and emitter feedback to arrive at the circuit depicted in Figure 5.19. We shall bestow upon it the highly original name of *combination feedback bias*.

Figure 5.19
Combination feedback bias.



This circuit applies feedback to R_B from both ends, so to speak, so it tends to have slightly better stability than either collector feedback or emitter feedback bias. Of course, it is now only one resistor shy from the voltage divider circuit which is considerably more stable.

The equations for the load line are listed below. The derivations are left as an exercise.

$$I_C = \frac{V_{CC} - V_{BE}}{R_C + R_E + R_B/\beta} \tag{5.16}$$

$$V_{CE} = V_{CC} - I_C(R_C + R_E) \tag{5.17}$$

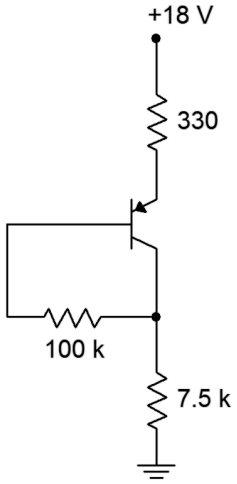
$$I_{C(sat)} = \frac{V_{CC}}{R_C + R_E} \tag{5.18}$$

$$V_{CE(cutoff)} = V_{CC} \tag{5.19}$$

Example 5.7

Assuming β = 125, determine the Q point (I_C and V_{CE}) for the circuit of Figure 5.20.

Figure 5.20
Circuit for Example 5.7.



Note that this is the upside down PNP version. Using Equation 5.16

$$I_C = \frac{V_{CC} - V_{BE}}{R_C + R_E + R_B/\beta}$$

$$I_C = \frac{18\,\text{V} - 0.7\,\text{V}}{7.5\,\text{k}\Omega + 330\,\Omega + 100\,\text{k}\Omega/125}$$

$$I_C = 2\,\text{mA}$$

Using KVL

$$V_{CE} = V_{CC} - V_{R_C} - V_{R_E}$$

$$V_{CE} = V_{CC} - I_C(R_C + R_E)$$

$$V_{CE} = 18\,\text{V} - 2\,\text{mA}(7.5\,\text{k}\Omega + 330\,\Omega)$$

$$V_{CE} = 2.34\,\text{V}$$

Summary

DC biasing is required in order to maintain the proper junction potentials and operation of the BJT. Several different circuit configurations are available to establish a DC bias on both NPN and PNP transistors. These circuits vary in complexity and their ability to maintain a constant operating point, or Q point, in the face of variations of β.

The two-supply emitter bias topology offers very high Q point stability. It achieves this through the use of two powers supplies; one connected through a resistor to the emitter and a second unit connected through a resistor to the collector. It is unique in that the supplies are bipolar; one being positive and the other being negative.

The voltage divider bias circuit offers similar stability performance to the two-supply emitter bias circuit. It uses a single supply and a resistive voltage divider to establish a second, lower potential at the base terminal.

The three feedback bias configurations offer only modest enhancements in stability but use the least amount of parts. They all rely on a single DC power source.

A DC load line is a plot of all possible collector current and corresponding collector-emitter voltage operating points. No matter what the β for a circuit happens to be, the transistor's operating point must lie on this line. It is a valuable DC analysis tool.

Review Questions

1. Explain the need for DC biasing. Why can't we just apply an AC signal to the base of a BJT and expect proper amplification of the signal?
2. What is a Q point?
3. What are the four values found on a DC load line?
4. Rank the bias configurations presented in this chapter in terms of their Q point stability relative to β.
5. Rank the bias configurations presented in this chapter in terms of their circuit complexity.
6. Describe the process of making a PNP version of an NPN bias circuit.

Problems

Unless otherwise specified, use β = 100.

Analysis Problems

1. Plot the load line for the circuit of Figure 5.21. V_{CC} = 20 V, V_{EE} = −8 V, R_B = 7.5 kΩ, R_E = 10 kΩ, R_C = 12 kΩ.

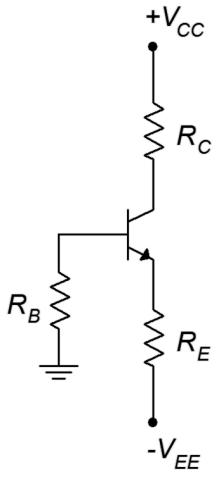


Figure 5.21

2. Determine the new Q point for Problem 1 if β = 250.

3. Plot the load line for the circuit of Figure 5.22. V_{EE} = 5 V, V_{CC} = −18 V, R_B = 22 kΩ, R_E = 1.2 kΩ, R_C = 1.5 kΩ.

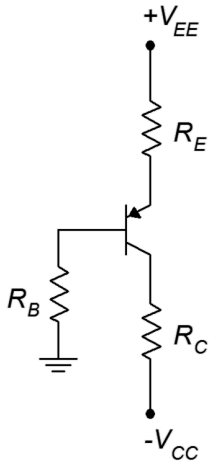


Figure 5.22

4. Determine the new Q point for Problem 3 if β = 50.

5. Plot the load line for the circuit of Figure 5.23. V_{CC} = 20 V, R_1 = 15 kΩ, R_2 = 5 kΩ, R_E = 4.3 kΩ, R_C = 9.1 kΩ.

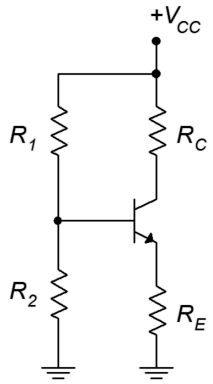


Figure 5.23

6. Determine the new Q point for Problem 5 if β = 150.

7. Plot the load line for the circuit of Figure 5.24. V_{EE} = 16 V, R_1 = 12 kΩ, R_2 = 4.7 kΩ, R_E = 6.2 kΩ, R_C = 10 kΩ.

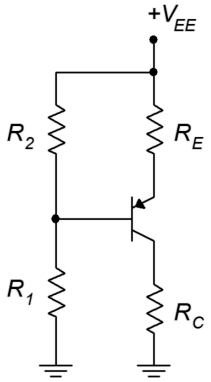


Figure 5.24

8. Determine the new Q point for Problem 7 if β = 200.

9. Plot the load line for the circuit of Figure 5.25. V_{CC} = 12 V, R_B = 560 kΩ, R_C = 3.3 kΩ.

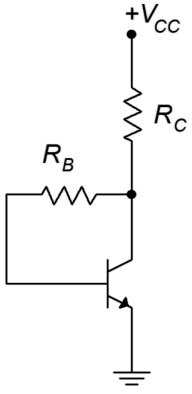


Figure 5.25

10. Determine the new Q point for Problem 9 if β = 75.

11. Plot the load line for the circuit of Figure 5.26.

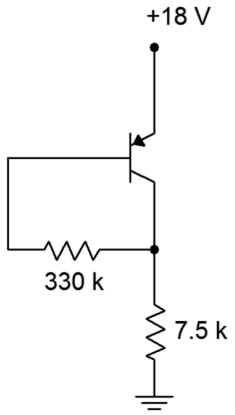


Figure 5.26

12. Determine the new Q point for Problem 11 if β = 200.

13. Plot the load line for the circuit of Figure 5.27. V_{CC} = 15 V, R_B = 470 kΩ, R_E = 560 Ω, R_C = 3.3 kΩ.

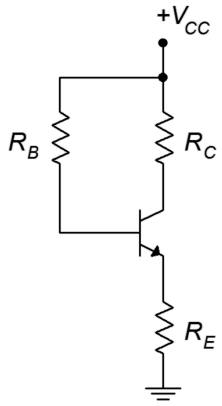


Figure 5.27

14. Determine the new Q point for Problem 13 if β = 170.

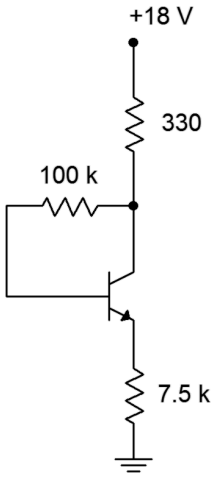
15. Plot the load line for the circuit of Figure 5.28.

16. Determine the new Q point for Problem 15 if β = 75.

17. Plot the load line for the circuit of Figure 5.29. V_{EE} = 18 V, R_B = 680 kΩ, R_E = 270 Ω, R_C = 3.9 kΩ.

18. Determine the new Q point for Problem 17 if β = 200.

Figure 5.28



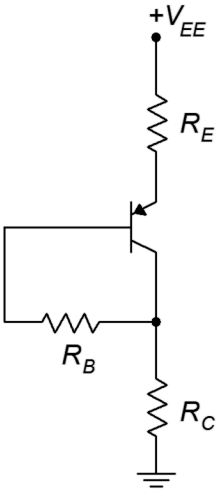


Figure 5.29

Design Problems

19. Determine a value for R_E in the circuit of Figure 5.21 to set $I_C = 2$ mA. Use $V_{CC} = 20$ V, $V_{EE} = -8$ V, $R_B = 10$ kΩ, $R_C = 5.6$ kΩ.

20. Determine a value for R_C in the circuit of Figure 5.22 to set $V_{CE} = 10$ V. Use $V_{CC} = -25$ V, $V_{EE} = 6$ V, $R_B = 15$ kΩ, $R_E = 6.8$ kΩ.

21. Determine a value for R_C in the circuit of Figure 5.23 to set $V_{CE} = 8$ V. Use $V_{CC} = 24$ V, $R_1 = 22$ kΩ, $R_2 = 10$ kΩ, $R_E = 5.6$ kΩ.

22. Determine new values for R_1 and R_2 in the circuit of Figure 5.24 in order to set $I_C = 500$ μA. $V_{EE} = 22$ V, $R_E = 15$ kΩ, $R_C = 6.8$ kΩ.

Challenge Problems

23. Determine the maximum and minimum values for I_C in Problem 1 if every resistor has a 10% tolerance.

24. Determine the maximum and minimum values for V_{CE} in Problem 3 if every resistor has a 5% tolerance.

25. Determine a value for R_E in the circuit of Figure 5.23 to set $V_{CE} = 10$ V. $V_{CC} = 30$ V, $R_1 = 12$ kΩ, $R_2 = 3$ kΩ, $R_C = 8.2$ kΩ.

26. Derive Equation 5.16.

27. Determine the the power drawn from the supply for the circuit of Problem 5.

28. Using a 15 volt power supply, design a bias circuit to create a very stable Q point of 2 mA and 5 volts.

Computer Simulation Problems

29. Perform a series of DC simulations to test the Q point stability versus β of the circuit of Problem 1.

30. Perform a Monte Carlo simulation to investigate the Q point stability of the circuit of Problem 5 if the emitter resistor has a 10% tolerance.

6 Amplifier Concepts

6.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Explain the differences between voltage gain, current gain and power gain.
- Describe a basic voltage amplifier model using voltage gain, input impedance and output impedance.
- Determine the effects of source and load impedance on system gain and explain how they interact with an amplifier's input and output impedance.
- Describe and distinguish the concepts of *noise* and *waveform distortion*.
- Define the concept of *output compliance*.
- Discuss the frequency limits of an amplifier in general terms.
- Define *Miller's Theorem*.

6.1 Introduction

The concept of signal amplification finds numerous uses in the field of electronics. This includes applications such as boosting the signal level from a sensor or driving loads like loudspeakers or antennas. Reduced to its most simple terms, amplification is just multiplication. The ideal amplifier multiplies the amplitude of the input signal by a constant. It should not change the frequency of the signal, alter its shape, add noise or in any other way warp or distort the signal.

Amplifiers can be designed to be voltage sensing or current sensing and can be modeled as either controlled voltage sources or controlled current sources. As a functional block, we are primarily interested in describing an amplifier in terms of its amplification factor, input impedance and output impedance. The amplification factor is also referred to as the *gain* and may be expressed in terms of voltage gain, current gain or power gain, depending on the application. Other items of interest include the maximum output level or *compliance*, useful frequency range, noise and distortion characteristics.

6.2 Amplifier Model

The ideal amplifier does nothing except increase the amplitude of the input signal. The factor of increase is defined as the ratio of the output signal to the input signal. It is a unit-less quantity. For example, if the input signal has a power of 10 milliwatts and the circuit boosts the signal up to 50 milliwatts, we say it has a power gain of 50 milliwatts /10 milliwatts, or 5. Similarly, if the input signal is 2 volts and the output signal is 16 volts, we say it has a voltage gain of 16 volts/2 volts, or 8. Historically, power gain is denoted as G. For voltage gain

and current gain we use A_v and A_i, where A stands for *Amplification factor.* Some amplifiers invert the signal from input to output. Basically, they flip the wave shape upside down. For a simple sine wave this is equivalent to shifting the phase of the signal by 180°, and for a sine wave input the amplifier produces a −sine output. To reflect this effect, the amplification factor is denoted as negative. For example, an A_v of −10 indicates an amplification factor of 10 with a signal inversion.

The size and complexity of an amplifier circuit can vary considerably, ranging from a single transistor to dozens of transistors. To ease system design it is helpful to use simplified functional models. Typically, these models use a resistor to represent the impedance seen looking into the amplifier along with a controlled source and its associated internal resistance. An example is shown in Figure 6.1.

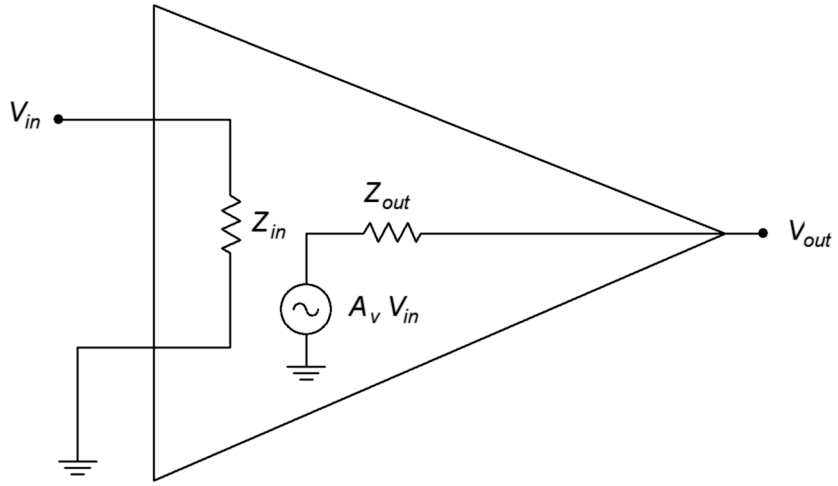


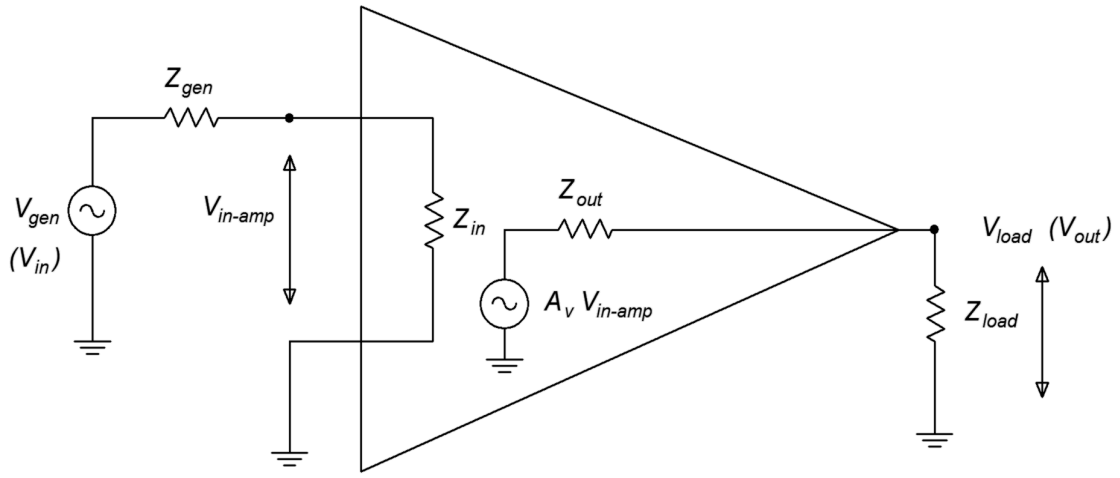
Figure 6.1
A simple voltage amplifier model.

This is a model of a voltage amplifier (note that V_{in} and V_{out} are specified along with a controlled voltage source within the model). The controlled voltage source and its series Z_{out} is the Thevenin equivalent of the output when viewed from the load (i.e., the V_{out} pin). Likewise, Z_{in} is the equivalent impedance seen by the driving source. Because a voltage amplifier is designed to maximize voltage transfer, the input impedance tends to be high to minimize loading (think of a voltmeter). Similarly, the output impedance would tend to be low (think of an ideal voltage source). In contrast, a circuit designed for maximum current transfer would tend to have a low Z_{in} and a high Z_{out}. Precisely *how* the circuit creates the signal boost is not a concern of this model, we only care that it *does*.

Loading Effects

Once the model is established it is relatively easy to recognize and compute loading effects. Loading effects are signal losses caused by interactions between the amplifier's impedances and those of the circuits and loads connected to it. A generic model including loading effects is shown in Figure 6.2.

Figure 6.2
Voltage amplifier model with loading effects.



A cursory examination of Figure 6.2 shows that there is a voltage divider between Z_{gen} and Z_{in} along with a second divider between Z_{out} and Z_{load}. Each of these dividers causes signal loss, that is, they reduce the final output voltage. The input voltage to the amplifier is reduced as follows

$$V_{in-amp} = \frac{Z_{in}}{Z_{in}+Z_{gen}} \times V_{gen}$$

The load voltage is reduced as follows

$$V_{load} = \frac{Z_{load}}{Z_{load}+Z_{out}} \times A_v \times V_{in-amp}$$

Therefore the combined gain is

$$A_{system} = \frac{V_{load}}{V_{gen}} = \frac{Z_{in}}{Z_{in}+Z_{gen}} \times A_v \times \frac{Z_{load}}{Z_{load}+Z_{out}}$$

To minimize these losses we'd like $Z_{in} >> Z_{gen}$ and $Z_{load} >> Z_{out}$.

Example 6.1

A voltage amplifier has the following specifications: A_v=20, Z_{in}=10 kΩ, Z_{out}=200 Ω. It is driven by a 30 millivolt source with a 600 Ω internal impedance and drives a 1 kΩ load. Determine the load voltage.

The voltage that appears at the amplifier's input is

$$V_{in-amp} = \frac{Z_{in}}{Z_{in}+Z_{gen}} \times V_{gen}$$
$$V_{in-amp} = \frac{10\,\text{k}\Omega}{10\,\text{k}\Omega+600\,\Omega} \times 30\,\text{mV}$$
$$V_{in-amp} = 28.3\,\text{mV}$$

This is multiplied by the voltage gain of 20 and then reduced by the output divider.

$$V_{load} = \frac{Z_{load}}{Z_{load}+Z_{out}} \times A_v \times V_{in-amp}$$
$$V_{load} = \frac{1\,\text{k}\Omega}{1\,\text{k}\Omega+200\,\Omega} \times 20 \times 28.3\,\text{mV}$$
$$V_{load} = 471.7\,\text{mV}$$

Without the loading effects the output signal would be simply 30 millivolts times the voltage gain of 20, or 600 millivolts. Further, note that if the source is replaced with a typical laboratory grade function generator exhibiting an internal impedance of 50 Ω and the load is removed, being replaced by an oscilloscope exhibiting a typical 1 MΩ input impedance, the loading effects would be minimal and we would measure just a few millivolts shy of the ideal 600 millivolts.

6.3 Compliance and Distortion

At some point the idealization that the output signal is merely the input signal times the gain fails. All amplifiers have a limit on just how large the output signal can be. This is set by the DC power supply and the amplifier design. The maximum output signal (typically, the maximum output voltage) is referred to as the *compliance*. Any attempt to produce an output signal that swings beyond the compliance will result in waveform distortion. In the simplest case, the output signal is strictly and abruptly limited to the compliance level and any portion of the output waveform that would otherwise lay above that will be removed. It is as if some form of electronic scissors clipped off the top of the waveform. Hence, this is often referred to as *clipping*. An example of clipping is illustrated in Figure 6.3. The ideal output waveform is shown in brown and the clipped waveform is shown in blue. The clipping is so severe here that the clipped waveform now looks less like a sine wave and more like a square wave. This is extreme waveform distortion and has important consequences.

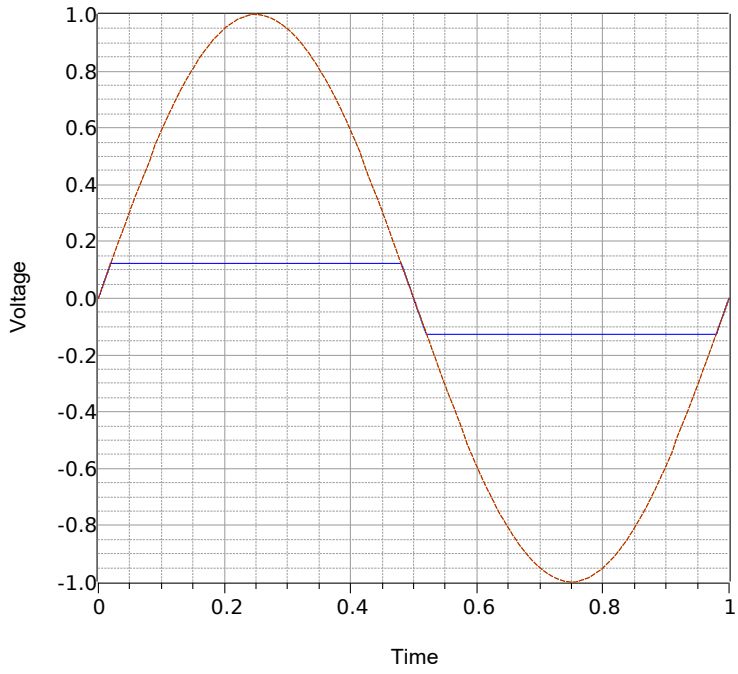


Figure 6.3
Clipped waveform.

Whenever a signal is altered in the time domain, there will be an alteration of its frequency content. Depending on what we started with and the manner in which the wave shape is altered, new frequency components may be added to the signal and the levels of existing components may be changed or even deleted. The extreme clipping that created a near-square wave added a large number of new frequency components to the signal.

To see how the wave shape and frequency content are connected, consider the waveform depicted in Figure 6.4. We start with a simple sine wave shown in green. We refer to this as the *fundamental* or base frequency. We then add a *harmonic*. A harmonic is another sine wave that is an integer multiple of the fundamental frequency. It may be larger or smaller in amplitude and the phase may be shifted. In this example we have a single sine at three times the fundamental frequency (blue). When we add this harmonic to the fundamental we arrive at a new waveform shown in red. This new waveform looks something like a square wave but with a "lumpy" top and bottom. If we add more harmonics, these variations will begin to smooth out, as shown in Figure 6.5. This waveform appears to be fairly close to a square wave and not too distant from our earlier clipped waveform. Based on this, we can conclude that the clipped sine wave has new frequency components added to it. We can also conclude that if our amplifier clips a more complicated waveform such as a snippet of music, that process will add new harmonics as well. Further, it is likely that these harmonics will be audible and can change our perception of the music, perhaps subtly but maybe drastically.

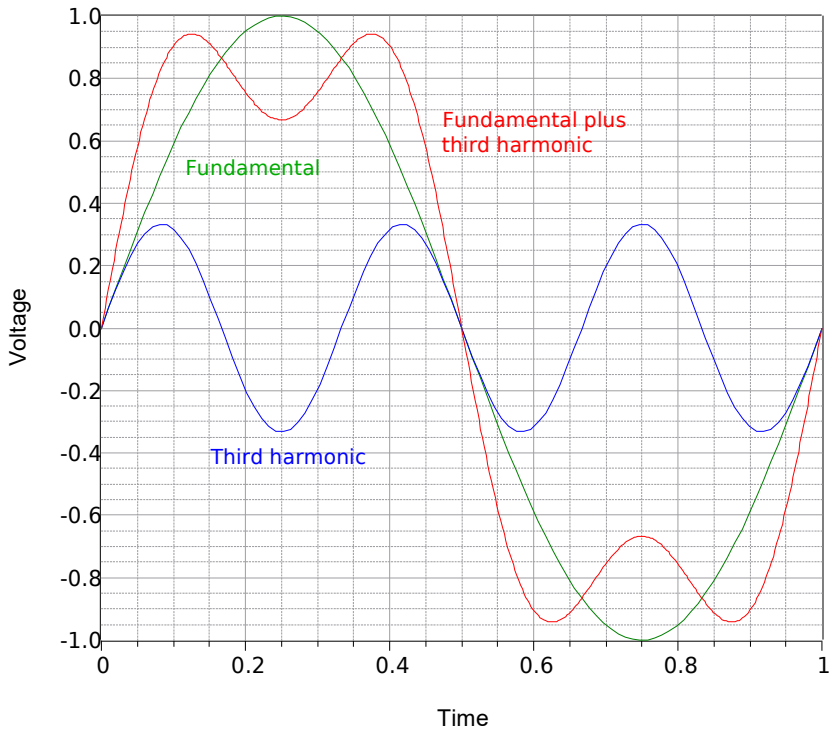


Figure 6.4
Sine wave with third harmonic.

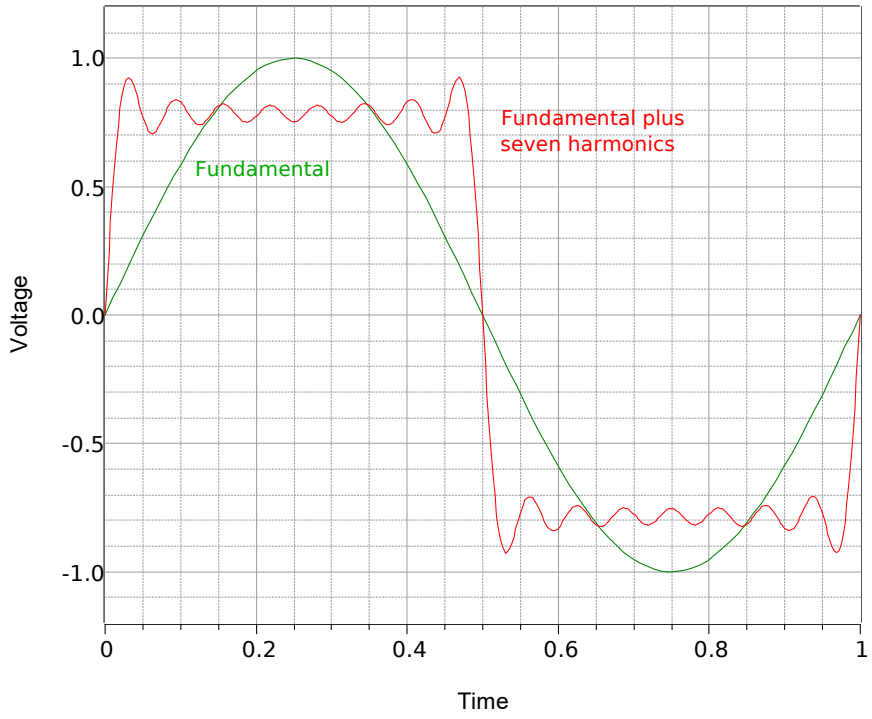


Figure 6.5
Sine wave with seven harmonics, approaching a square wave.

Along with clipping, amplifiers can exhibit more subtle forms of distortion due to internal nonlinearity. For example, it is possible for the gain to vary slightly as the signal swings from low to high or from negative to positive. An example is shown in Figure 6.6 with the distorted wave shown in red.

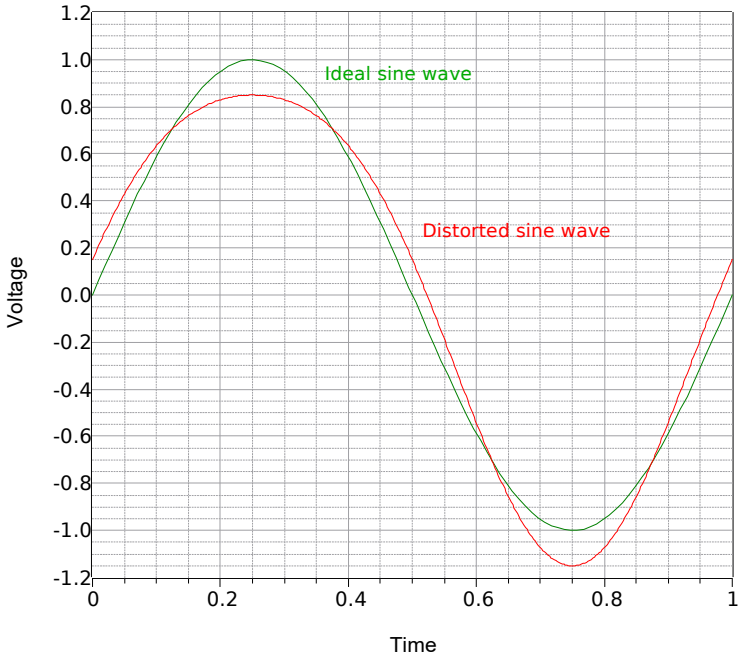


Figure 6.6
Sine wave with distortion.

At first glance it may appear as though the wave is merely offset negatively. This is not the case. If we shift the wave vertically, as in Figure 6.7, it becomes apparent that the wave is truly distorted and is no longer a pure sine wave.

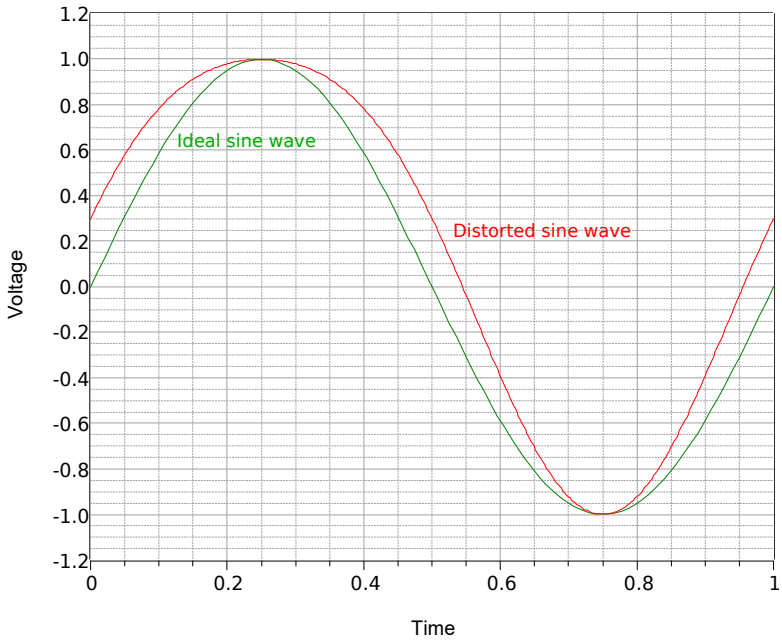


Figure 6.7
Sine wave with distortion, level shifted.

Unlike the clipped wave, the distorted wave in Figure 6.7 exhibits an asymmetry; the negative portion does not appear to be a mirror image of the positive portion. In other words, this wave lacks *half-wave symmetry*. Waves that exhibit half-wave symmetry contain only odd harmonic distortion (harmonics that are odd integer multiples of the fundamental). In contrast, waves that lack half-symmetry have *at least one* even harmonic. Here is how to test for half-wave symmetry. First, consider the sawtooth wave shown in Figure 6.8.

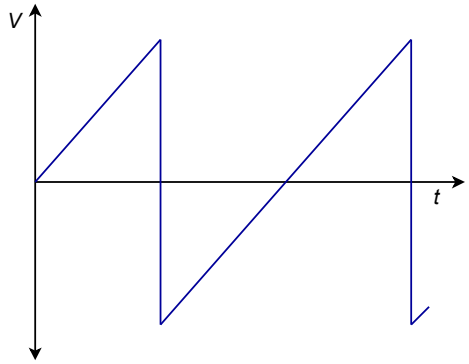


Figure 6.8
Half-wave symmetry test: Sawtooth wave.

Rotate the negative portion of the wave around the time axis as shown in Figure 6.9.

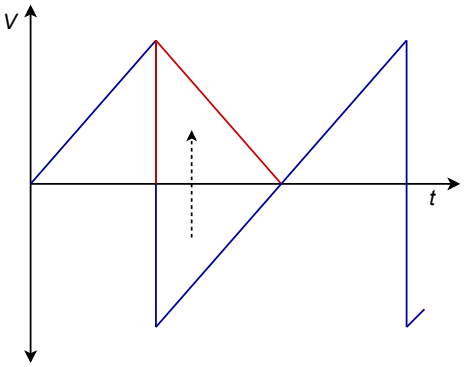


Figure 6.9
Half-wave symmetry test: Sawtooth wave, negative portion rotated.

Finally, slide the negative portion over the positive portion and see if they're identical, as in Figure 6.10.

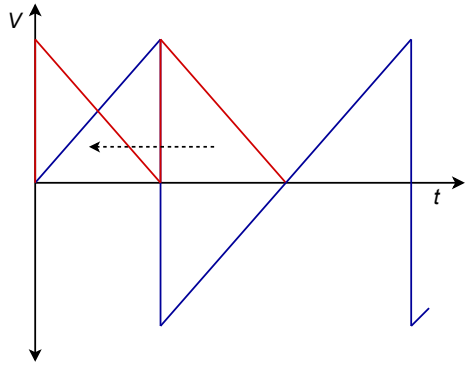


Figure 6.10
Half-wave symmetry test: Sawtooth wave, negative portion rotated and slid.

If the two halves are identical then the wave has half-wave symmetry. The sawtooth wave does not exhibit half-wave symmetry, therefore it must contain at least one even harmonic.

An amplifier's linearity is often quantified through a Total Harmonic Distortion, or THD, measurement. The measurement is carried out by applying a very pure, low distortion sine wave to the amplifier. This is the fundamental. At the output of the amplifier, a very selective filter is used to remove the fundamental. This leaves behind just the added distortion harmonics.[20] These harmonics are then treated as a lumped value and presented as a percentage of the total signal. On an oscilloscope, it is relatively easy for a person to discern THD levels in the double digits. On the other hand, it is very difficult, if not impossible, for an individual to discern THD levels much below 1% by eye. Of course, what matters is what we can hear, not how the waveform looks. To put this in perspective, many high fidelity audio amplifiers exhibit THD levels below 0.1% while an over-driven guitar amplifier might be running over 20%. THD is not the final word on distortion though. It has its limits. For example, all of the distortion products are lumped together. It says nothing about which harmonics are particularly strong or their distribution. It also doesn't say much about what happens when multiple frequencies interact. One method of trying to quantify that is to apply two sine waves at different frequencies to the amplifier simultaneously. The result is called an Intermodulation Distortion rating, or IMD. This is also expressed as a percentage.

6.4 Frequency Response and Noise

Like compliance and distortion, two other practical limits on amplifier performance are its frequency response and output noise. First, let's discuss frequency response.

Although we describe an amplifier as having a specific gain or amplification factor, this is true only for a certain range of frequencies. All amplifiers are limited in terms of the range of frequencies over which they can operate. If we examine an amplifier's performance at extreme frequencies, the gain may be much less than the nominal value. In fact, if we go far enough, the gain may even be fractional, meaning that the “amplifier” is actually reducing the signal level.

The region where the nominal gain is accurate is referred to as the *mid-band*. This range is defined by one or two *corner* or *break frequencies*. The lower limit is referred to as f_1 while the upper limit is referred to as f_2. At these frequencies, the output level has dropped to half the power exhibited by a mid-band frequency of the same input level. A gain versus frequency response plot that encapsulates this concept is shown in Figure 6.11.

20 To be strictly accurate, the residual consists of the harmonics plus any noise produced by the amplifier. Therefore it is more accurate to refer to this as a THD+noise spec.

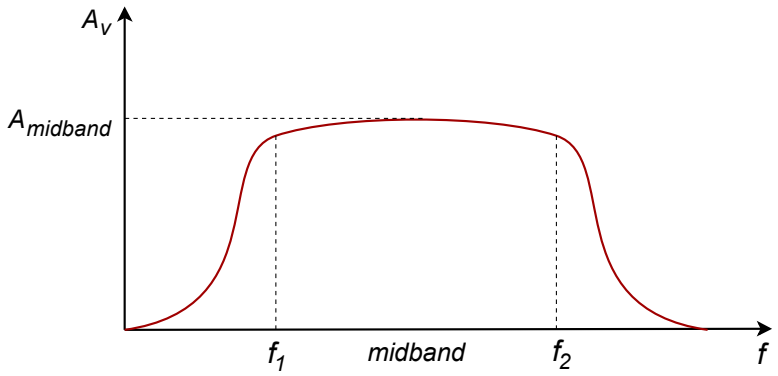


Figure 6.11
A generic gain versus frequency plot.

In this representative plot it is apparent that only input signals whose frequencies lay between f_1 and f_2 will receive full amplification. As the input frequency moves to either side of the middle band, the gain begins to drop off. The drop-off increases as the signal frequency moves farther and farther away. Eventually the gain will fall to practically zero and virtually no trace of the input signal will appear at the output.

Precise values of the corner frequencies will depend on the application. For example, a high fidelity audio amplifier will most likely have an f_1 below 20 Hz and an f_2 above 20 kHz[21] while an amplifier used for telephone systems might range from 300 Hz to 4 kHz[22]. In contrast, a radio frequency amplifier may be operating at frequencies orders of magnitude higher than these.

Without exception, all amplifiers have an upper limit frequency, f_2, but not all of them have a lower frequency limit, f_1. Amplifiers without a lower limit can amplify signals with frequencies all the way down to DC. They are referred to as *direct coupled* or *DC amplifiers*. The lower frequency limit is usually caused by in-line coupling capacitors, and in some cases, transformers. Among other uses, these components are added to purposely block DC. There are good reasons to do this, as we shall see in upcoming work, however, it is possible to design amplifiers without them. The resulting amplifier will then have no limit on how low of a frequency it can amplify.

The upper limit frequency is another story. While components are often added to tailor the upper frequency response of an amplifier, even if no tailoring was desired the amplifier would still have an upper limit frequency. This would be due to small and unavoidable capacitances and inductances that exist in the circuit, for example stray wiring capacitance. Ultimately, the corresponding reactances will cause a signal level reduction that worsens as frequency increases. As you might guess, these reactances also cause varying phase shifts between the input signal and the output signal. Frequency response is examined in great detail in later chapters.

21 20 Hz – 20 kHz is the range of frequencies heard by a typical healthy young human.

22 Decidedly not hi-fi, but do we really need high fidelity to call-in a take-out order?

Amplifier performance is also limited by its internal noise. Noise is an undesired signal that appears at the output of an amplifier. Unlike distortion, noise is usually not correlated with the input signal level. Generally, noise is broad-band, meaning that it contains a very wide range of frequencies. As such, it does not have a discernible pitch. Examples in nature include the sound of leaves rustling in the wind or the sound of a waterfall. Noise is best thought of as a truly random signal. As such, it cannot be accurately predicted and therefore there is no easy way to remove it once it has been added to a desired signal. There are many potential sources of noise in an amplifier. They range from process issues in semiconductors to thermal effects in resistive elements. In general, noise gets worse as temperature, resistance and frequency range increase. Noise is unavoidable in absolute terms but ultimately what we care about is whether or not it is low enough for a given application. In other words, is the noise level significantly lower than the signal level, to the point where it is no longer a problem? This is quantified by simply creating a ratio between the nominal output signal level and the output noise level. This ratio is given the very creative name *signal-to-noise ratio*, or *S/N* for short. All other factors being equal, the higher the *S/N*, the better.

6.5 Miller's Theorem

Some inverting voltage amplifier designs employ an impedance bridged between the input and output, as shown in Figure 6.12.

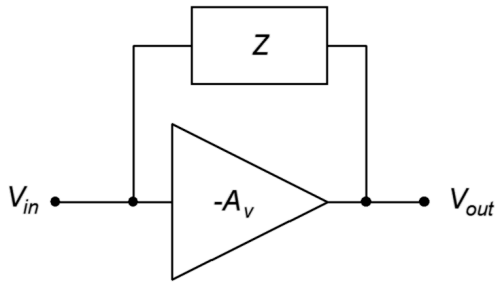


Figure 6.12
An input/output bridged impedance.

There are different reasons for doing this, a prime example involving shaping the frequency response of the amplifier. To simplify the analysis of such a configuration, we may employ *Miller's Theorem*, named after American engineer John Milton Miller.

The goal here is to determine equivalent impedances that lie in parallel with the input and output of the amplifier. These equivalents simply become part of the input and output networks around the amplifier.

First, let's consider the equivalent impedance at the input, $Z_{in\text{-}miller}$. By definition, this is the impedance in parallel with the input of the amplifier that would draw the same amount of current as the original bridging Miller impedance. The current through the Miller impedance is simply the the voltage across it divided by the Miller impedance, *Z*. The voltage across it is the difference between the input and output voltages.

$$i_{in-miller} = \frac{V_{in} - V_{out}}{Z}$$

$$i_{in-miller} = \frac{V_{in} - A_v V_{in}}{Z}, \text{ the gain is negative so}$$

$$i_{in-miller} = \frac{V_{in}(|A_v|+1)}{Z}$$

Dividing this current into the input voltage yields the equivalent impedance.

$$Z_{in-miller} = \frac{Z}{|A_v|+1}$$

A similar derivation yields the output equivalent.

$$Z_{out-miller} = \frac{Z|A_v|}{|A_v|+1}$$

The Miller equivalent circuit for a general impedance *Z* is shown in Figure 6.13.

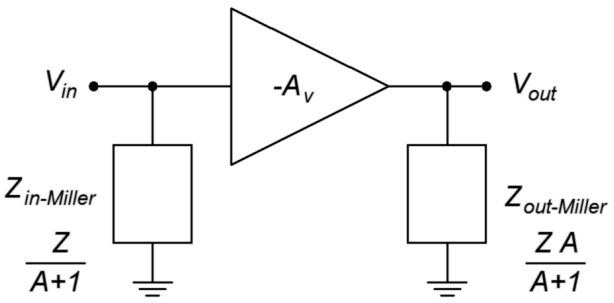


Figure 6.13
Miller general equivalent.

The general rule to remember is that the Miller equivalent presents equivalent impedances that are less than the original bridging impedance. In the case of the input section, the reduction effect is very large at higher gains.

Two typical cases for the impedance *Z* are a resistance, *R*, and a capacitance, *C*. For a pure resistance, we can perform a direct substitution for *Z*. The original and Miller equivalents are shown in Figure 6.14.

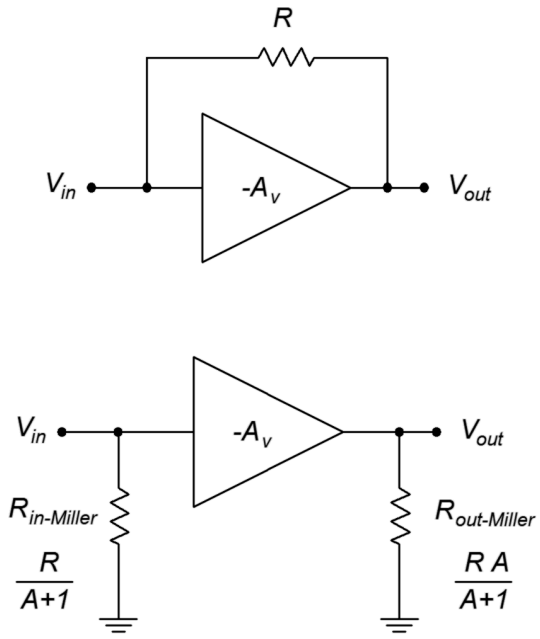


Figure 6.14
Original and Miller equivalent for a resistor.

For a capacitor, the situation is similar, however, we will substitute X_c for Z and recall that $C=1/(2\pi fX_c)$.

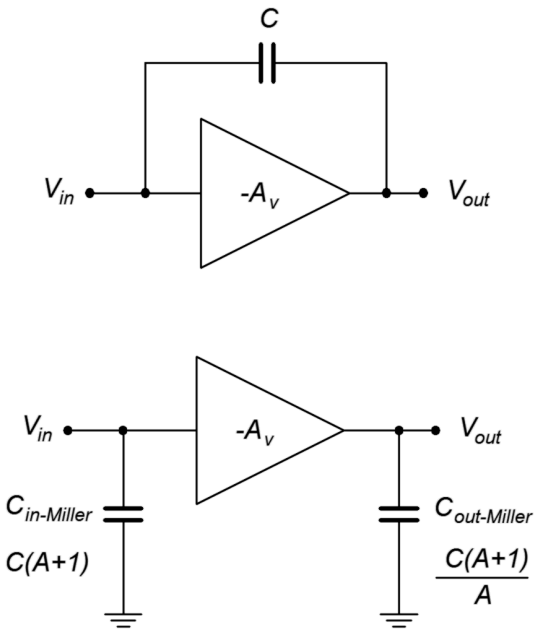


Figure 6.15
Original and Miller equivalent for a capacitor.

Note that for the capacitor, there is a multiplicative effect. That is, the effect of the original bridging capacitor on a high gain amplifier is equivalent to a much larger input shunt capacitor.

Summary

Complex amplifier circuits can be modeled with a functional block. The ideal model includes the input and output impedances along with a controlled source. This source would exhibit a signal gain or amplification factor. Usually, G stands for power gain while A_v and A_i represent voltage and current amplification, respectively. This amplification factor may be negative which indicates that the amplifier inverts the phase of the input, that is, the waveform is flipped upside down. The impedances allow calculation of loading effects while the gain determines the the size of the output signal.

If the input signal is too large, the output signal may be limited in amplitude or clipped. The maximum output amplitude is referred to as the compliance. Clipping is a gross form of distortion but more subtle forms exist as well. In general, distortion creates new frequency components. If these new components are integer multiples of the original input frequency, which they are typically, they are referred to as harmonics. One method of quantifying distortion performance is to sum all of the harmonics and compare that to the original signal. This is called THD or total harmonic distortion. Along with distortion, the amplifier might also add undesirable noise to the output signal. Noise is a random signal that contains many different frequencies. Typically, this is measured via a signal-to-noise ratio, or S/N, at the output.

An amplifier also operates over a given range of frequencies, from a lower limit, f_1, to a high limit, f_2. Some amplifiers are able to amplify down to 0 Hz (DC) and effectively do not have an f_1 but all amplifiers do have an upper limit.

Finally, Miller's Theorem is an analysis technique that allows an impedance that bridges from the input of an inverting voltage amplifier to its output to be split into equivalent input and output parallel impedances. These impedances will be smaller than the original bridging impedance and are a function of the gain of the amplifier.

Review Questions

1. Explain how an amplifier's input impedance might react with a source to produce a signal loss.
2. Explain how an amplifier's output impedance might react with a load to produce a signal loss.
3. What is compliance?
4. Describe clipping.
5. Describe half-wave symmetry. What does it have to do with amplifier distortion?
6. What is noise? How does it differ from distortion?
7. Draw a generic frequency response plot for an amplifier.
8. Detail the purpose and use of Miller's Theorem.

Problems

Analysis Problems

1. Determine the load voltage for the model of Figure 6.16 if V_{gen} = 10 mV, Z_{gen} = 50 Ω, Z_{in} = 1 MΩ, Z_{out} = 75 Ω, Z_{load} = 1 k Ω and A_v = 50.

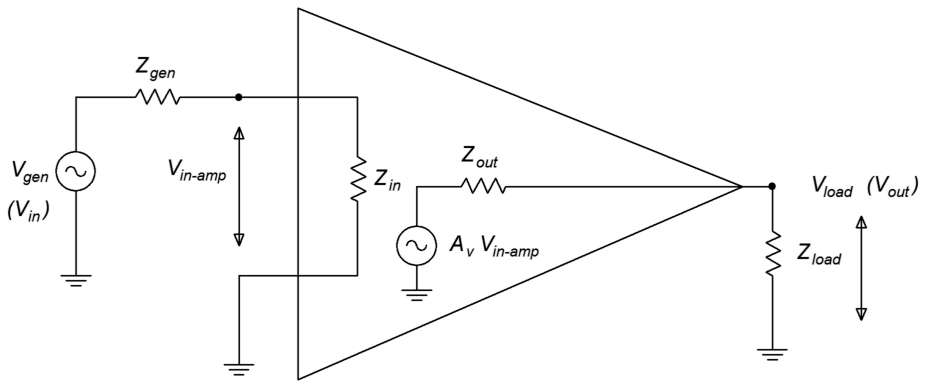


Figure 6.16

2. Determine the load voltage for the model of Figure 6.16 given V_{gen} = 8 mV, Z_{gen} = 1 k Ω, Z_{in} = 6 kΩ, Z_{out} = 500 Ω, Z_{load} = 2 k Ω and A_v = 100.

3. If the circuit of Problem 1 has a compliance of 2 volts, will the output clip? What if the input is increased to 100 mV?

4. If the circuit of Problem 2 has a compliance of 5 volts, will the output clip? What if the input is increased to 200 mV?

5. If an amplifier has $A_v = 25$, $V_{in} = 20$ mV and there is no appreciable loading, determine the output signal-to-noise ratio if the amplifier generates an output noise voltage of 10 μV.

6. Determine which waveforms from Figures 6.17 through 6.21 exhibit half-wave symmetry.

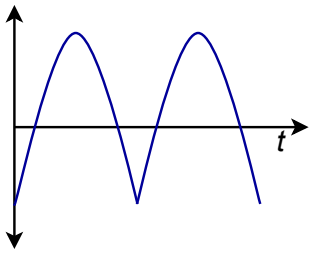


Figure 6.17

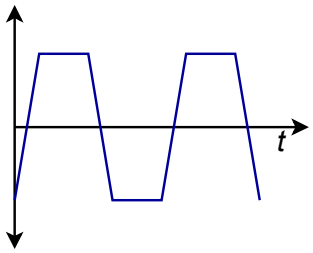


Figure 6.18

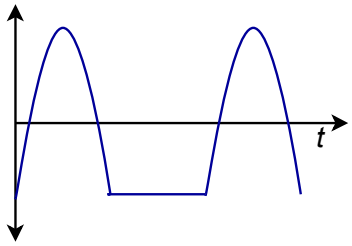


Figure 6.19

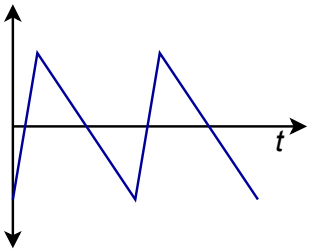


Figure 6.20

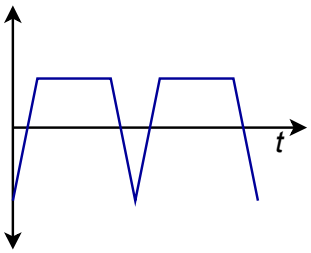


Figure 6.21

7. Determine the Miller equivalent resistances for the circuit of Figure 6.22 if $A_v = -20$ and $R = 60$ kΩ.

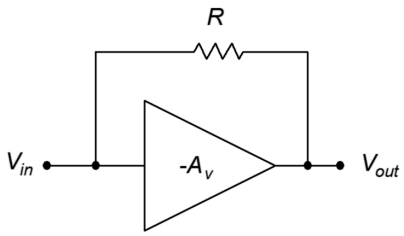


Figure 6.22

8. Determine the Miller equivalent capacitances for the circuit of Figure 6.23 assuming $A_v = -30$ and $C = 200$ pF.

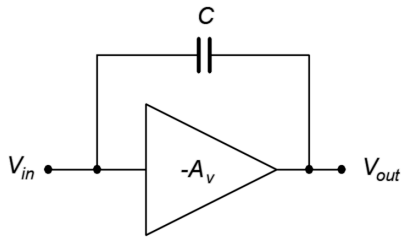


Figure 6.23

Challenge Problems

9. If the circuit of Problem 1 has a compliance of 20 volts, how large can the input signal be before the load voltage is clipped?

10. If the circuit of Problem 2 has a compliance of 10 volts, how large can the input signal be before the load voltage is clipped?

11. Using Figure 6.22 as a guide and assuming that $R = 100$ kΩ, how large would the gain have to be such that the input equivalent resistance is 4 kΩ?

12. Using Figure 6.23 as a guide and assuming that $A_v = -35$, determine a value for C such that the input equivalent capacitance is 1.2 nF.

Computer Simulation Problems

13. Simulate the circuit of Problem 1 and verify the load voltage.

14. Simulate the circuit of Problem 2 and verify the load voltage.

7 BJT Small Signal Amplifiers

7.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Determine the voltage gain, input impedance and output impedance of simple BJT amplifiers.
- Detail the functional differences between voltage amplifiers and voltage followers.
- Explain the advantages and disadvantages of using localized feedback (swamping).
- Determine the combined characteristics of multistage BJT amplifiers.
- Detail the advantages and disadvantages of using direct coupling versus capacitor coupling in multistage amplifiers.
- Explain the operation of the *Darlington pair.*

7.1 Introduction

In the prior chapter we discussed the general operational characteristics of amplifiers including voltage gain, input and output impedance, compliance, distortion and so forth. In this chapter we shall focus on the analysis of small signal amplifiers, specifically, their voltage gain and input/output impedances. As we will be performing a *small signal analysis*, we will not be concerned with compliance, maximum load power, device dissipation or the like. There is no specific definition of small signal versus large signal but for our purposes we shall define small signal as output signals that are well below the clipping limit and with power dissipation of no more than a few hundred milliwatts for either the load or transistor.

There are two popular techniques used to analyze BJT amplifier circuits. One is through the use of *hybrid parameters*. There are four different hybrid parameters. We have already seen one of them, the forward current gain, h_{fe}. We simply call it β. The other three are h_{ie}, the input impedance; h_{oe}, the output admittance; and h_{re}, the reverse voltage gain. The second letter of the subscript (the "*e*" in h_{fe}) indicates it is for the common emitter configuration (that is, input applied to the base, output taken at the collector and the emitter at the common ground).

The second approach uses *r′ parameters* (pronounced "r prime"). The *r′* approach is sufficient for all of our analyses and, given an understanding of Ohm's law, KVL and KCL, produces straightforward equations for circuit gain, input impedance and the like. As a consequence, we shall focus on the *r′* system.

7.2 Simplified AC Model of the BJT

Just as we created a DC model to ease the analysis of DC bias circuits, we shall make use of an AC BJT model for our AC analyses. In fact, our AC model is based on the DC model. The collector-base region is still represented with a current-controlled current source although it's AC instead of DC: $i_C = \beta\, i_B$. The base-emitter junction is a bit trickier. Although a simple 0.7 volt junction worked fine for DC, we now have to consider the AC resistance of the diode.

To find the dynamic resistance of the junction, first recall that the AC signal is riding on the DC bias current, as plotted in Figure 7.1.[23] We can imagine that the AC signal is causing this point to trace back and forth along the curve. Of course, as this is a small signal analysis, this sweep will be very small, perhaps only a few percent of the quiescent current and can be approximated as a straight line segment. The slope of this line segment represents its conductance.[24] The reciprocal of conductance is resistance; therefore, the reciprocal of the slope represents the resistance of the device. Consequently, we can approximate the dynamic resistance of the device as the reciprocal of the slope of the line tangent to the operating point (that is, the reciprocal of the slope of the line tangent to the quiescent bias current I_C).

In reality, this slope is changing slightly as the signal swings back and forth along the base-emitter I-V curve. As the signal swings positive and goes above the quiescent point the slope is a little steeper producing a slight reduction in dynamic resistance. In contrast, as it swings negative, going below the quiescent point, the slope becomes a bit more shallow and produces a slightly higher resistance. As a result, we are effectively computing an average value for the dynamic resistance by assuming this is a straight line segment. The variance in this resistance will be a source of asymmetrical distortion in the amplifier of the type shown in Chapter 6, Figure 6.6. We shall see more on this later.

23 The values plotted along the current axis are typical of a generic device and do not represent the current values for all BJTs.

24 Technically, this value is called the device's *transconductance* and is denoted as g_m. We shall be seeing this again in upcoming work.

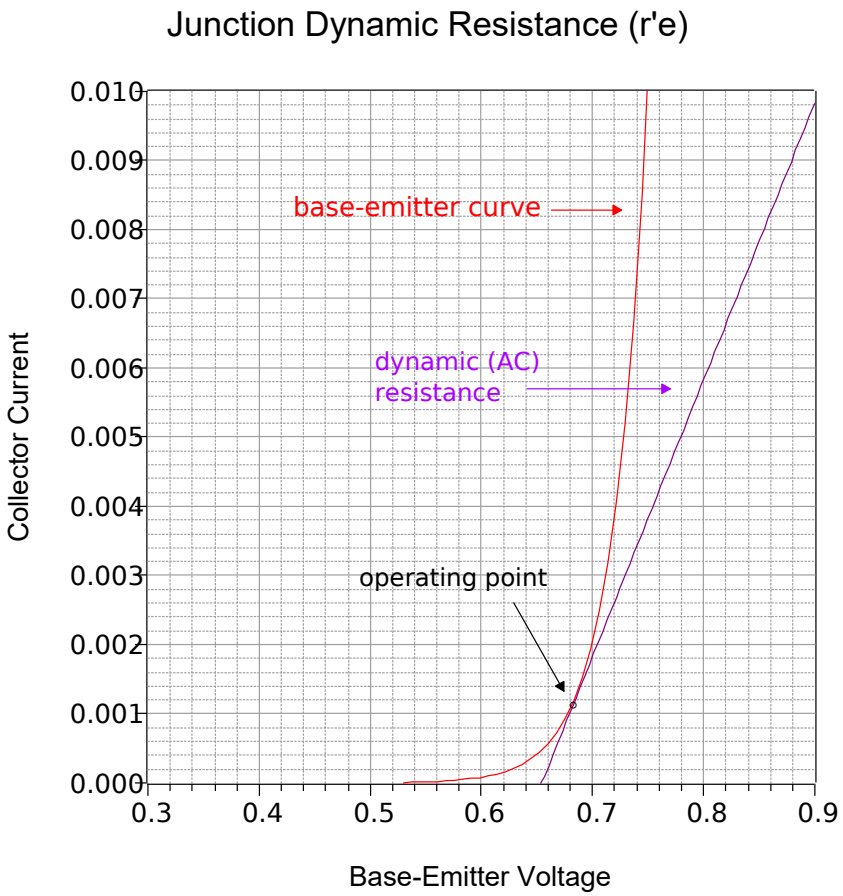


Figure 7.1
Base-emitter junction plot and dynamic resistance.

In order to derive an equation for the dynamic resistance, we begin with the Shockley equation from Chapter 2, Equation 2.1, slightly modified to reflect the terminal names of a BJT.

$$I_C = I_S\left(e^{\frac{V_{BE}\,q}{n\,k\,T}} - 1\right)$$

Where
I_C is the junction (collector) current,
I_S is the reverse saturation current,
V_{BE} is the voltage across the junction (base-emitter),
q is the charge on an electron, 1.6E−19 coulombs,
n is the quality factor (typically between 1 and 2),
k is the Boltzmann constant, 1.38E−23 joules/kelvin,
T is the temperature in kelvin.

At 300 kelvin (about 80° F), q/kT is approximately 38.6, thus for any reasonable value of V_{BE} the "−1" term is small enough to ignore. Also, we shall take n as 1.

The equation then reduces to

$$I_C = I_S e^{38.6 V_{BE}} \tag{7.1}$$

To find the slope we take the first derivative of Equation 7.1 with respect to V_{BE}.

$$\frac{dI_C}{dV_{BE}} = 38.6\, I_S e^{38.6 V_{BE}} \tag{7.2}$$

Substituting Equation 7.1 into Equation 7.2 yields

$$\frac{dI_C}{dV_{BE}} = 38.6\, I_C$$

By definition, the dynamic junction resistance is the reciprocal of the slope.

$$\frac{dV_{BE}}{dI_C} = \frac{25.9\,\text{mV}}{I_C}$$

We call this r'_e. This value is slightly low as it doesn't include bulk resistance so a good approximation is

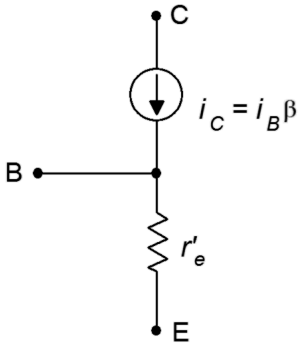
$$r'_e = \frac{26\,\text{mV}}{I_C} \tag{7.3}$$

It is important to note that I_S in Equation 7.2 varies with temperature. Therefore r'_e varies with temperature as well, decreasing with increasing temperature. This carries important ramifications with the thermal stability of higher power amplifiers as we shall see in subsequent work.

One of the most important things to remember here is that the DC collector current sets up the resistance of the AC model. In other words, the stability of the AC circuit will depend in part on the stability of the DC bias (hence our emphasis on stable bias circuits in Chapter 5).

We now have our AC model, as shown in Figure 7.2. This is a simplified model in that it does not include junction capacitance effects, lead inductance and the like. It is appropriate, therefore, as a low to mid-band frequency model. To summarize, the AC collector current, i_C, is determined by the AC input current, i_B; which is in turn a function of the the size of the applied input signal. In contrast, r'_e is set by the DC bias current, I_C. The AC input can produce small variations in r'_e which are manifested as waveform distortion.

Figure 7.2
Simplified AC model of BJT.



Given the model, there are three ways to configure the transistor as an amplifier:

- *Common Emitter.* The input is applied to the base and the output is taken at the collector. The emitter terminal is at the common or ground point. This configuration exhibits both voltage gain and current gain. It also inverts the phase of the signal.

- *Common Collector.* The input is applied to the base and the output is taken at the emitter. The collector terminal is at the common or ground point. This configuration offers a voltage gain of about unity but does exhibit current gain. It maintains the phase of the input signal. It is also referred to as an emitter follower or voltage follower.

- *Common Base.* The input is applied to the emitter and the output is taken at the collector. The base terminal is at the common or ground point. This configuration exhibits voltage gain but the current gain is unity at best. It also maintains the phase of the input signal.

We shall examine each of these topologies in turn. Each of these can be made using a variety of DC bias techniques. For example, a two-supply emitter bias or voltage divider bias could be used for any of the three AC topologies, and further, they could utilize either an NPN or PNP transistor.

7.3 The Common Emitter Amplifier

The common emitter configuration finds wide use as a general purpose voltage amplifier. We begin with a basic DC biasing circuit and then add a few other components. For example, refer to Figure 7.3.

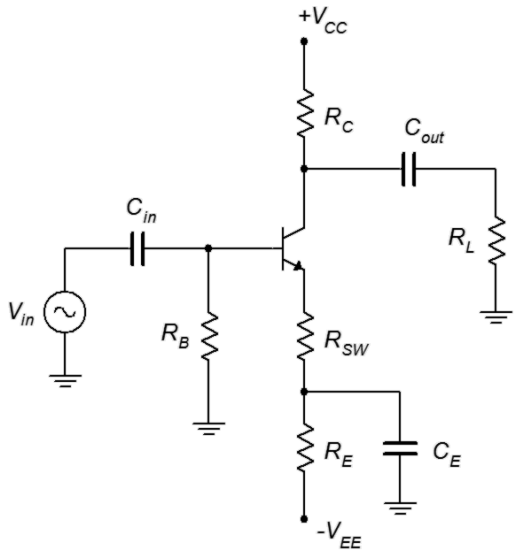


Figure 7.3
Common emitter amplifier using two-supply emitter bias.

This amplifier is based on a two-supply emitter bias circuit. The notable changes are the inclusion of an input signal voltage, V_{in}, and a load, R_L. So that these components do not alter the bias, we isolate the input and load through the use of *coupling capacitors* C_{in} and C_{out}. These capacitors will act as opens to DC creating the desired isolation. As for the AC signal, the capacitances will be chosen such that their reactances will be much smaller than the surrounding resistors at the frequency of the input. Consequently, the capacitors will appear as shorts and allow the AC signal to pass through the amplifier.

The final alteration involves the emitter resistor. The single resistor of the bias network is replaced by a pair of resistors, R_E and R_{SW}, along with a *bypass capacitor*, C_E. For DC, the capacitor is open and the effective emitter bias resistance is $R_E + R_{SW}$. For AC, the capacitor will behave ideally as a short so the AC emitter resistance will fall to just R_{SW}. This resistor is called a *swamping* or *emitter degeneration* resistor. It is used primarily to help control the voltage gain of the amplifier.

We can use our AC transistor model along with the Superposition Theorem to arrive at an equivalent AC circuit of the amplifier, as shown in Figure 7.4.

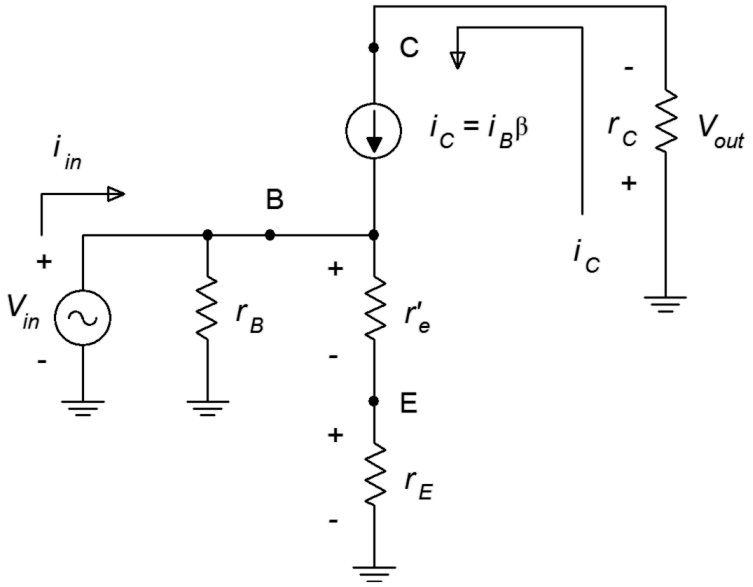


Figure 7.4
AC equivalent of common emitter amplifier.

First, we have shorted all of the capacitors. Second, we have replaced the DC sources with their ideal internal resistance (a short) which places those points at AC ground. Third, we swapped out the transistor for the model. Lastly, we have combined and/or renamed resistances where needed. Because this is an AC circuit, we use the convention of lower case *r* for resistance to avoid confusion with the DC resistance (which are upper case). Thus, r_E is the AC resistance from the emitter to AC ground. This corresponds to R_{SW} in the original schematic. Similarly, r_C represents the total resistance seen from the collector to AC ground. In the original schematic this corresponds to R_C in parallel with R_L. If this circuit was unloaded,

then r_C would just be equal to R_C. Finally, r_B corresponds to R_B but in a voltage divider bias it would be equal to R_1 in parallel with R_2.

Voltage Gain

Voltage gain, A_v, is defined as the ratio of v_{out} to v_{in}. Using Ohm's law we find

$$A_v = \frac{v_{out}}{v_{in}} = \frac{v_C}{v_B}$$
$$A_v = \frac{-i_C r_C}{i_C(r'_e + r_E)}$$
$$A_v = -\frac{r_C}{r'_e + r_E} \tag{7.4}$$

First, the negative sign indicates that this amplifier inverts the waveform, top to bottom. For a sine wave, this is equivalent to shifting the phase 180°. In some applications this can be a major issue, in others, not so much. If it is an issue, it can be resolved by using a second inverting gain amplifier in sequence with the first (inverting the inversion).

The second thing we see is that the gain is little more than a ratio of collector to emitter resistances. This is where splitting the emitter resistor into two parts comes in. In the equation, r_E is the swamping resistor R_{SW}. The larger the swamping resistor, the lower the gain. The maximum gain will be achieved when $R_{SW} = 0$. That is, when the emitter is completely bypassed. The down side of this is that the gain will now depend entirely on r'_e. This will increase the distortion. The reason is because R_{SW}, being so much larger, effectively "swamps out" the variation in r'_e and reduces distortion. The larger R_{SW} is relative to r'_e, the greater the reduction in distortion, but with the cost of reduced gain. This is why a swamping resistor is also called an emitter degeneration resistor: it degrades the voltage gain.

Input Impedance

Input impedance, Z_{in}, is defined as the ratio of v_{in} to i_{in}. In Figure 7.2 this is equal to r_B in parallel with the impedance looking into the base terminal, $Z_{in(base)}$. Using Ohm's law we find

$$Z_{in(base)} = \frac{v_B}{i_B}$$
$$Z_{in(base)} = \frac{i_C(r'_e + r_E)}{i_B}$$
$$Z_{in(base)} = \frac{i_C(r'_e + r_E)}{i_C/\beta}$$
$$Z_{in(base)} = \beta(r'_e + r_E) \tag{7.5}$$

Therefore

$$Z_{in} = r_B \,||\, Z_{in(base)} \qquad (7.6)$$

We see that both the swamping resistor and β play a role in setting the input impedance. Larger values of R_{SW} and β produce larger input impedances.

In sum, we find that while swamping decreases voltage gain, it reduces distortion and increases input impedance, the latter two generally desirable for a voltage amplifier. A non-swamped amplifier will have the largest gain but will suffer from the worst distortion and a low input impedance. This is a classic "quality versus quantity" trade-off: a large low quality gain versus a modest high quality gain[25].

Output Impedance

Output impedance, Z_{out}, is defined as the internal impedance of the equivalent source that drives the load. If we position ourselves at the load and look back into the amplifier shown in Figure 7.1, C_{out} is shorted ideally and V_{CC} is at AC ground. This leaves us with R_C in parallel with the transistor. The transistor is modeled as a current source and its ideal internal resistance would approach infinity. In reality, the effective value, r'_C, is likely in the region of 100 kΩ or so, depending on bias current. This parallel combination comprises the output impedance of the current source. We model this circuit as a voltage amplifier so to be proper, we'd convert the current source with parallel internal resistance to a voltage source with series internal resistance. Those resistance values are identical, though, and we arrive at

$$Z_{out} = r'_C \,||\, R_C$$

In many circuits, R_C is considerably smaller than r'_C, therefore

$$Z_{out} \approx R_C \qquad (7.7)$$

Example 7.1

Determine the input and output impedances of the amplifier shown in Figure 7.5. Also compute the voltage gain. Assume β = 150.

First, the easy bit. We can determine the output impedance by inspection. It is approximately equal to R_C, or 22 kΩ.

25 The obvious question is, "How do we get both high gain and low distortion?" One answer is to use multiple low gain stages in cascade.

In order to find Z_{in} and A_v, we will need to determine r'_e. To obtain r'_e we need to find I_C. Using KVL around the base-emitter loop, if we approximate the DC base voltage to be near zero, then all of the emitter supply drops across the DC emitter resistance, with the exception of V_{BE}.

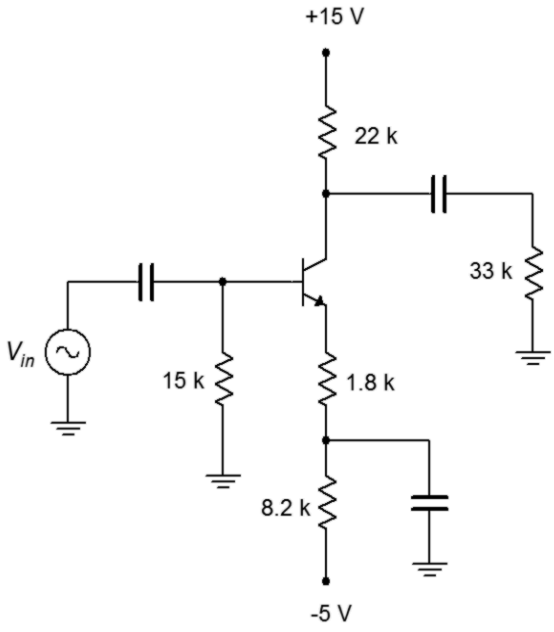


Figure 7.5
Schematic for Example 7.1.

$$I_C = \frac{|V_{EE}| - V_{BE}}{R_E + R_{SW}}$$
$$I_C = \frac{5\,\text{V} - 0.7\,\text{V}}{8.2\,\text{k}\Omega + 1.8\,\text{k}\Omega}$$
$$I_C = 0.43\,\text{mA}$$

$$r'_e = \frac{26\,\text{mV}}{I_C}$$
$$r'_e = \frac{26\,\text{mV}}{0.43\,\text{mA}}$$
$$r'_e = 60.5\,\Omega$$

$$Z_{in-base} = \beta(r'_e + r_E)$$
$$Z_{in-base} = 150(60.5\,\Omega + 1.8\,\text{k}\Omega)$$
$$Z_{in-base} = 279\,\text{k}\Omega$$

This value in parallel with the base biasing resistor creates the input impedance.

$$Z_{in} = R_B \parallel Z_{in(base)}$$
$$Z_{in} = 15\,\text{k}\Omega \parallel 279\,\text{k}\Omega$$
$$Z_{in} = 14.2\,\text{k}\Omega$$

Now for the voltage gain. We have two options. The first option is to find the loaded gain directly, meaning we include the load resistor as part of r_C. This would be the fastest method for this particular problem. The second option is to find the unloaded gain (i.e., with the load resistor removed) and then include the voltage divider effect caused by the load interacting with the output impedance. This would be the preferred method if we were swapping out different load values. We'll do both just for funsies.[26]

First, method one.

$$r_C = R_C \parallel R_L$$
$$r_C = 22\,\text{k}\Omega \parallel 33\,\text{k}\Omega$$
$$r_C = 13.2\,\text{k}\Omega$$

26 Is "funsies" a *real* word? It is if we all agree that it is. Besides, if it was an *imaginary* word, we'd spell it "*j* funsies".

$$A_v = -\frac{r_C}{r'_e + r_E}$$
$$A_v = -\frac{13.2\,\text{k}\Omega}{60.5\,\Omega + 1.8\,\text{k}\Omega}$$
$$A_v = -7.1$$

And now method two; first the unloaded gain, then the divider effect and finally, the composite gain.

$$A_{v(unloaded)} = -\frac{r_C}{r'_e + r_E}$$
$$A_{v(unloaded)} = -\frac{22\,\text{k}\Omega}{60.5\,\Omega + 1.8\,\text{k}\Omega}$$
$$A_{v(unloaded)} = -11.82$$

$$A_{divider} = \frac{R_L}{R_L + R_C}$$
$$A_{divider} = \frac{33\,\text{k}\Omega}{33\,\text{k}\Omega + 22\,\text{k}\Omega}$$
$$A_{divider} = 0.6$$

$$A_v = A_{v(unloaded)} \times A_{divider}$$
$$A_v = -11.82 \times 0.6$$
$$A_v = -7.1$$

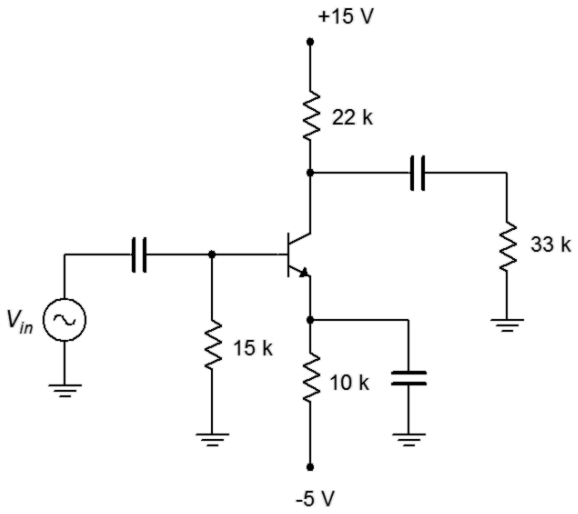
We shall repeat the prior example using the same circuit but with one change: the emitter resistor will be completely bypassed. This will show the effect that swamping has on voltage gain and input impedance.

Example 7.2

Determine the voltage gain and input impedance of the amplifier shown in Figure 7.6. Assume β = 150.

The DC equivalent of this circuit is identical to that of the circuit shown in Figure 7.5. In both cases, the DC emitter resistance is 10 kΩ. Therefore, I_C and r'_e are unchanged. The bypass capacitor shorts this entire value for the AC equivalent because there is no swamping resistor. Consequently, $r_E = 0$. We can simply use 0 for r_E in the equations previously derived.

Figure 7.6
Schematic for Example 7.2.



We begin with the input impedance.

$$Z_{in(base)} = \beta(r'_e + r_E)$$
$$Z_{in(base)} = 150(60.5\,\Omega + 0)$$
$$Z_{in(base)} = 9075\,\Omega$$

This value is considerably smaller than the value obtained from the swamped circuit. Continuing,

$$Z_{in} = R_B \parallel Z_{in(base)}$$
$$Z_{in} = 15\,\text{k}\,\Omega \parallel 9075\,\Omega$$
$$Z_{in} = 5654\,\Omega$$

$$A_v = -\frac{r_C}{r'_e + r_E}$$
$$A_v = -\frac{13.2\,\text{k}\,\Omega}{60.5\,\Omega + 0}$$
$$A_v = -218.2$$

The end result is an input impedance less than half of the swamped case and a voltage gain over 30 times greater. What these calculations do not show is the increase in distortion that will be created by this change. More on that in a moment.

Let's consider something slightly different: a voltage divider bias PNP amplifier.

Example 7.3

Determine the input impedance and voltage gain for the circuit shown in Figure 7.7. Also determine v_{load} if v_{in} = 20 mV peak. Assume β = 100.

We need to first determine r'_e which means that we need to find the collector current. If we assume a lightly loaded divider, the base voltage will be approximately 15 volts and the emitter will be 0.7 volts higher, or 15.7 volts. This leaves 20 volts − 15.7 volts, or 4.3 volts, across the DC equivalent emitter resistance. That's 4.1 k Ω + 200 Ω, or 4.3 kΩ , yielding 1 mA for I_C. This will produce r'_e = 26 Ω.

$$Z_{in(base)} = \beta(r'_e + r_E)$$
$$Z_{in(base)} = 100(26\,\Omega + 200\,\Omega)$$
$$Z_{in(base)} = 22.6\,\text{k}\,\Omega$$

This value is in parallel with the voltage divider biasing resistors, creating the input impedance.

$$Z_{in} = R_1 \parallel R_2 \parallel Z_{in(base)}$$
$$Z_{in} = 15\,\text{k}\Omega \parallel 5\,\text{k}\Omega \parallel 22.6\,\text{k}\Omega$$
$$Z_{in} = 3.22\,\text{k}\Omega$$

$$A_v = -\frac{r_C}{r'_e + r_E}$$
$$A_v = -\frac{7.5\,\text{k}\Omega \parallel 10\,\text{k}\Omega}{26\,\Omega + 200\,\Omega}$$
$$A_v = -19$$

We also need to include the effect of the 600 Ω source impedance. This will create a voltage divider with the input impedance.

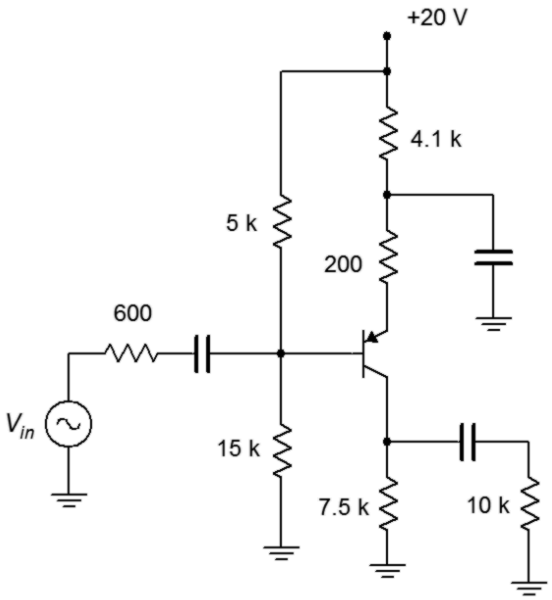
$$A_{divider} = \frac{Z_{in}}{Z_{in} + Z_{source}}$$
$$A_{divider} = \frac{3.22\,\text{k}\Omega}{3.22\,\text{k}\Omega + 600\,\Omega}$$
$$A_{divider} = 0.843$$

$$A_{v(system)} = A_v \times A_{divider}$$
$$A_{v(system)} = -19 \times 0.843$$
$$A_{v(system)} = -16$$

Finally, we get to the load voltage.

$$V_{load} = A_{v(system)} \times V_{in}$$
$$V_{load} = -16 \times 20\,\text{mV}$$
$$V_{load} = 320\,\text{mV peak, inverted}$$

Figure 7.7
Schematic for Example 7.3.



If we were to inspect the circuit of Figure 7.7 using a direct coupled oscilloscope, we would see the superposition of the AC and DC components. In other words, we'd see the AC signal riding on a DC offset. In some cases, the AC signal would be too small to notice compared to the DC portion. In proper scale it might be no thicker than the trace itself. In order to measure it accurately, we'd have to AC couple the oscilloscope.

The voltages at the source and load would be just AC as the coupling capacitors serve to block DC. At the base we'd have 15 volts DC with an AC signal riding on top

of it. The AC would be the 20 mV input times the input impedance/source impedance divider of 0.843, or 16.86 mV. Recalling that I_C is 1 mA, the DC drop across R_C must be 7.5 volts. This is, of course, V_C. Therefore, at the the collector we'd see an inverted 320 mV signal riding on 7.5 volts DC.

Computer Simulation

In order to get some insight into the swamping-versus-distortion issue, we shall take a look at a more involved circuit simulation. This will echo Examples 7.1 and 7.2 in that we will simulate two circuits with the same DC equivalents. The only circuit change will be that one version will have a fully bypassed emitter while the other version will utilize a swamping resistor. In order to keep the comparison fair, we will increase the input signal voltage of the lower gain swamped amplifier so that both versions have a similar load voltage. In this way we guarantee that they are both using a similar percentage of the junction curve.

The unswamped circuit is shown in Figure 7.8. This utilizes a straightforward two-supply emitter bias.

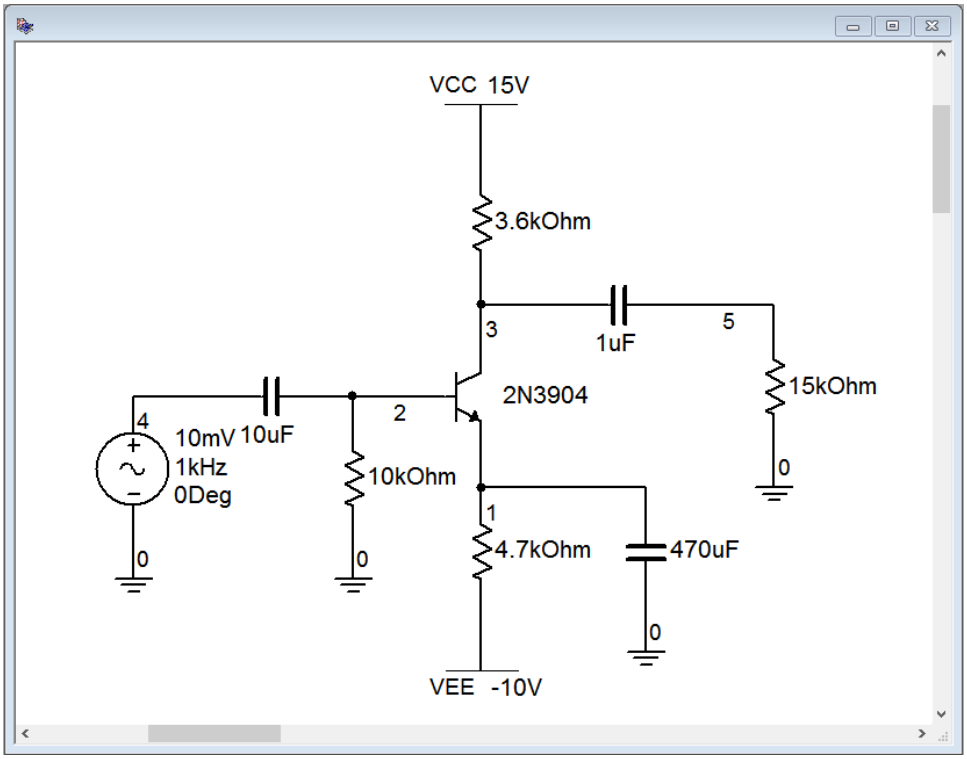


Figure 7.8
Unswamped CE amplifier in simulator.

A quick "back-of-an-envelope" estimate gives $I_C \approx 2$ mA, yielding $r'_e \approx 13\ \Omega$. The load will be around 3 kΩ which gives a gain in the low 200s. Thus, we expect the load voltage to be around 2 volts.

The transient analysis graph is depicted in Figure 7.9. Several traces are shown.

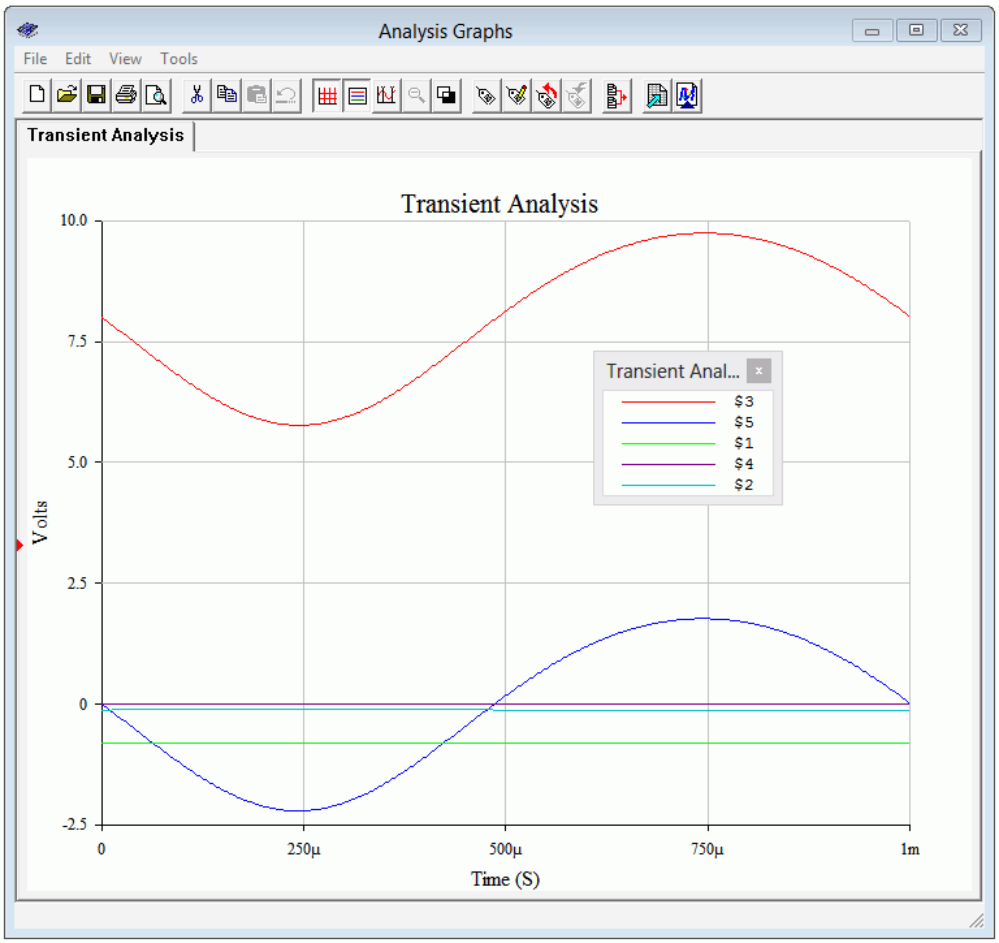


Figure 7.9
Unswamped CE amplifier, Transient Analysis.

At this scale, the AC signal at the input (node 4, purple) and the base (node 2, aqua) cannot be seen. As expected, we see a small negative DC value at the base and at the emitter, around −0.7 VDC. The DC offset at the collector is around 8 volts, as expected. Finally, the load voltage (node 5, blue) is sitting right around 2 volts.

What might not be visible immediately in the load voltage plot is some waveform asymmetry distortion. This can be quantified through a THD simulation, the output of which is shown in Figure 7.10. THD is nearly 8%. Not so good.

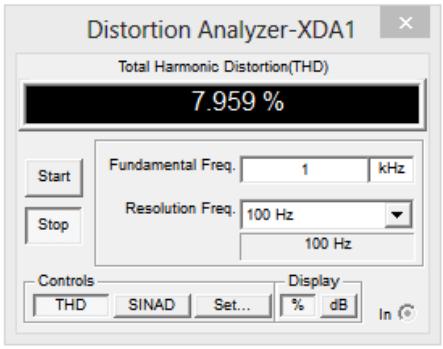


Figure 7.10
Non-swamped CE amplifier, THD Analysis.

For the second pass, the circuit is modified to include a swamping resistor, as illustrated in Figure 7.11. The original 4.7 kΩ emitter resistor has been split into a 4.5 kΩ and a 200 Ω swamping resistor. The bias in this circuit is identical to the first, therefore r'_e is unchanged. This will lower our expected gain to around 13, decreasing by a factor of 15. The input signal is raised by a factor of 15 to compensate so that our load voltage will still be around 2 volts.

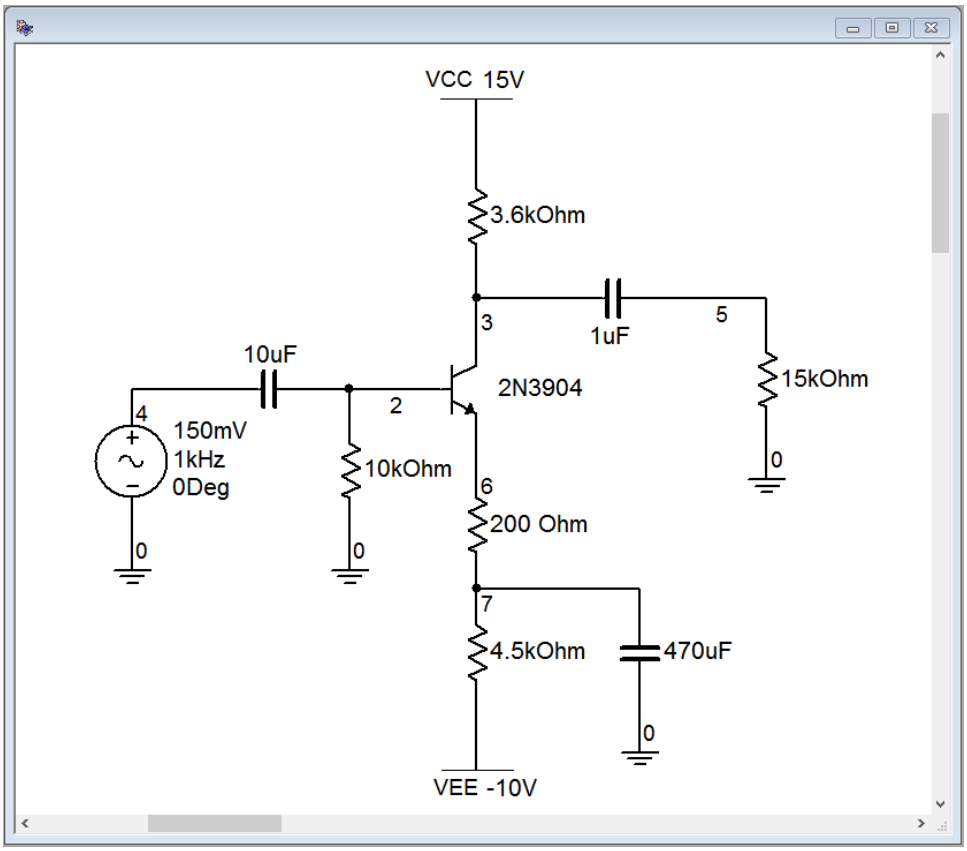


Figure 7.11
Swamped CE amplifier in simulator.

Once again, we run a transient analysis. The results are shown in Figure 7.12. In this case we have done something a little different. By zooming in, we can now confirm the signal inversion. The input signal is the purple trace at node 4. We can also see this signal at the base, riding on the small negative DC bias voltage (aqua trace, node 2). The DC offset is about −0.1 volts. Looking at the emitter we see the expected 0.7 volt DC base-emitter drop below this, or about −0.8 volts DC. Notice that there is no AC signal at the emitter whatsoever. This is expected as the emitter bypass capacitor forces this point to an AC ground.

The load voltage is the blue trace, node 5. While much of it is not visible at this zoom level, clearly it is an inverted waveform when compared to the input signal.

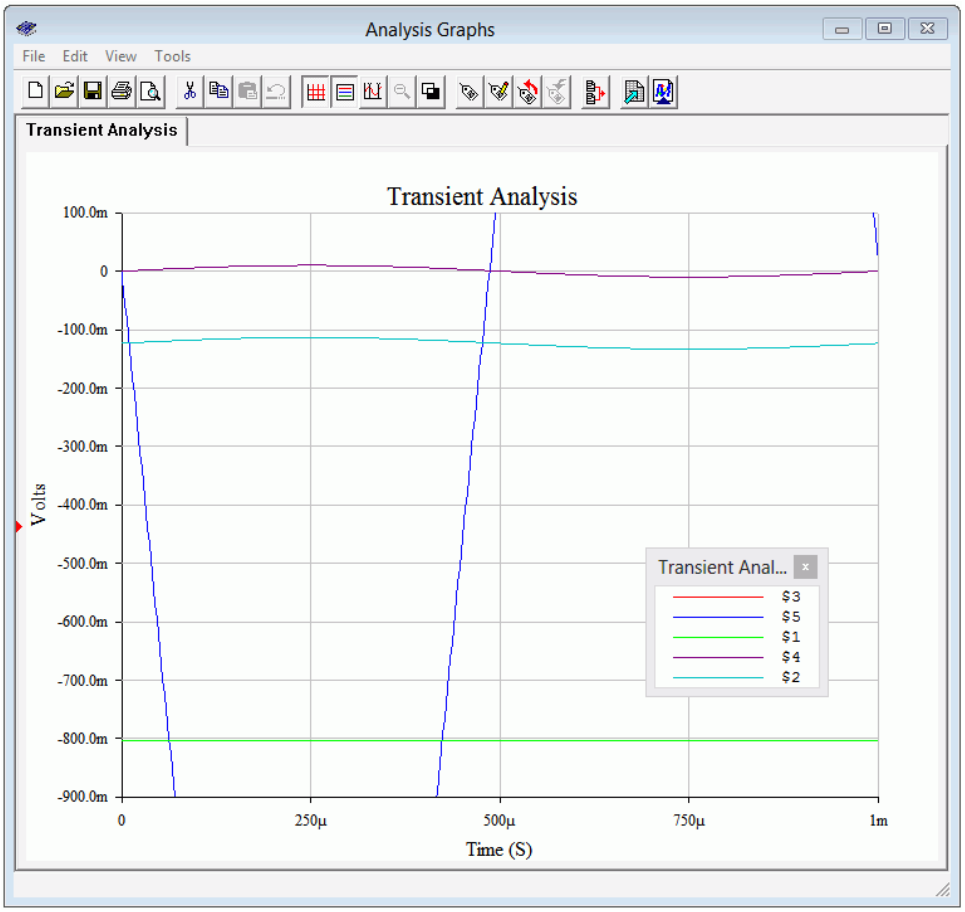


Figure 7.12
Swamped CE amplifier, Transient Analysis.

But what about the load voltage distortion? A THD simulation is performed on the swamped amplifier with the results shown in Figure 7.13. The THD is now under 0.6 %, a considerable improvement, even if not audiophile quality. Interestingly, as a ratio, the reduction in distortion is roughly equal to the reduction in gain. The more you give up, the more you get.

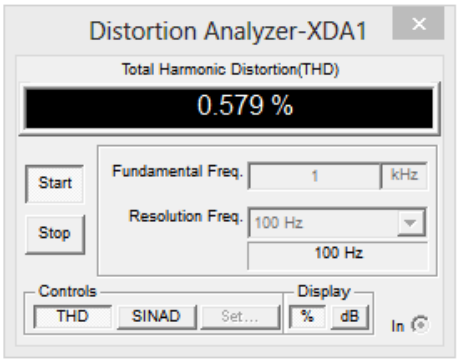


Figure 7.13
Swamped CE amplifier, THD Analysis.

Finally, the change in signal quality can be seen readily by plotting both load voltages concurrently, as shown in Figure 7.14. The non-swamped output (in blue)

exhibits telltale asymmetry. Notice that the positive peak does not quite reach 2 volts but the negative peak exceeds −2 volts. The positive peak is also broadened and flattened, while the negative peak is sharper. In contrast, the swamped output (in red) has virtually identical positive and negative peak values with no apparent shape changes on them. Compare this simulation to the waveform distortion discussion from Chapter 6. In particular, compare Figure 7.14 to Figure 6.6.

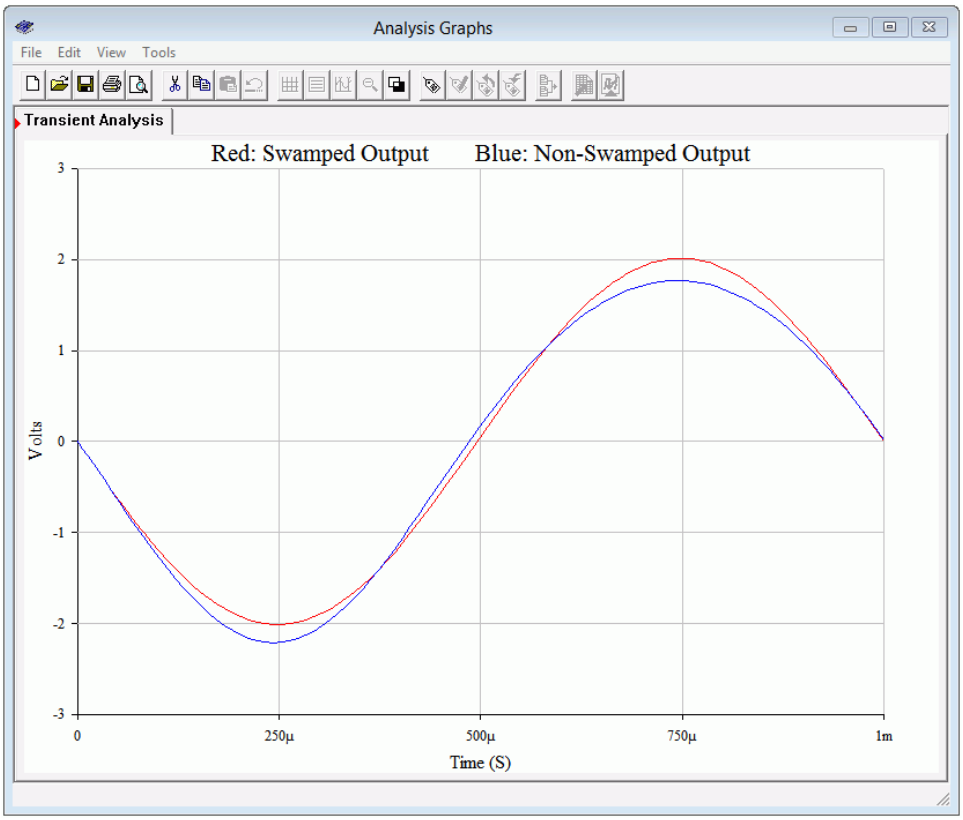


Figure 7.14
Swamped versus non-swamped CE amplifiers, Transient Analysis.

Power Supply Bypass and Decoupling

In the prior analyses we have assumed ideal behavior from the DC power sources. First, we assumed that they present a perfect AC ground and second, that they exhibit no ripple or noise. In reality, this may not be the case and non-ideal behavior may lead to a number of problems that diminish the quality of the amplified output signal, including hum and oscillations.

To battle the first issue, power supply bypass capacitors may be used. These capacitors are usually modest in size, perhaps 1 µF or so, although they can be much larger, particularly with high output power amplifiers. Power supply bypass capacitors are located physically close to the active devices. This location minimizes the resistive and inductive effects of power supply circuit board traces and wiring that could result in the power supply not being a good AC ground.

The second issue involves the noise and ripple from the power supply finding its way into the input signal and becoming part of the output signal. A classic example of this is amplifiers that use voltage divider biasing such as the one shown in Figure 7.7. Not only does the divider create the needed DC potential at the base terminal, but it also couples in any noise or ripple that might be riding on the DC voltage. This is particularly nasty because this undesirable signal is being applied to the base where it will get amplified.

The obvious solution to this problem is to create a very high quality, regulated DC supply, but this is not always practical given cost constraints. A relatively simple solution is to decouple the undesirable AC components through an *RC* network as shown in Figure 7.15.

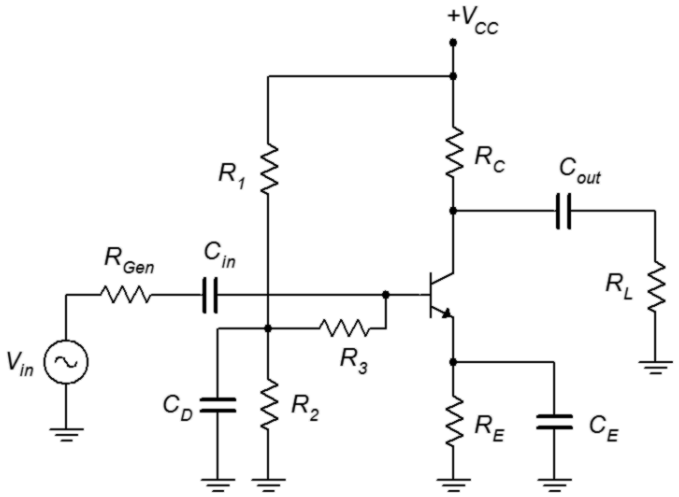


Figure 7.15
Decoupled voltage divider.

The capacitor C_D is used to create a low impedance at the divider junction, thus shunting any noise or ripple to ground. Unfortunately, this would also short out the input signal so R_3 is added to prevent this. R_3 is in parallel with $Z_{in(base)}$ to create the input impedance.

7.4 The Common Collector Amplifier

The common collector amplifier is often referred to as an *emitter follower*, or more generically, as a *voltage follower*. The key characteristics of a voltage follower are a high input impedance, a low output impedance and a non-inverting voltage gain of approximately one. The name comes from the fact that output voltage *follows* the input, that is, it's at the same voltage level and is in phase with the input. While this configuration does not produce voltage gain, it does produce current gain, and therefore, power gain. It's primary purpose is to reduce impedance loading effects, for example, to match a high impedance source to a low impedance load. Consequently, they are used as high-Z input buffer stages or as drivers for low

impedance loads such as loudspeakers.

A common collector amplifier using two-supply emitter bias is shown in Figure 7.16. The input is coupled into the base like the common emitter amplifier, however, the output signal is taken at the emitter instead of at the collector. Because the collector is at the AC common, there is no need for a collector resistor.

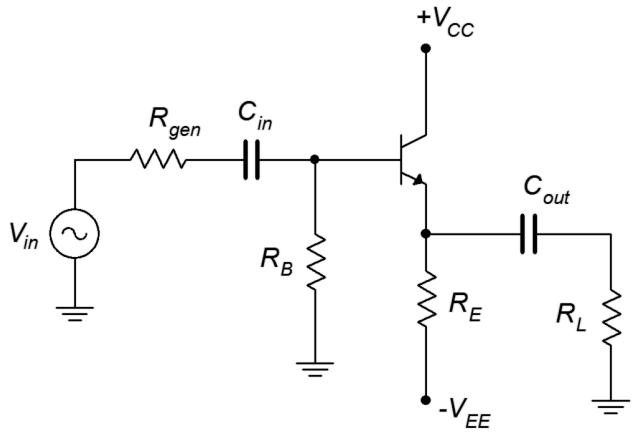


Figure 7.16
Common collector amplifier.

Perhaps the best way to think about the follower is not that it gives a voltage gain of one, but that it will prevent signal loss. The analysis follows, using Figure 7.17.

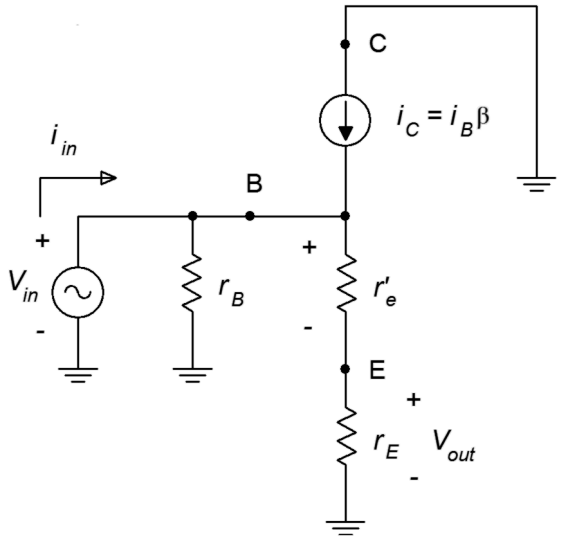


Figure 7.17
AC equivalent of common collector amplifier.

First, the AC emitter resistance, r_E, is either the emitter bias resistor, R_E, or the parallel combination of R_E and the load resistance, R_L. We'll use the former in order to determine the unloaded gain and the latter to determine the loaded gain, similar to what we did with the common emitter amplifier concerning R_C and R_L. The AC base resistance, r_B, typically boils down to the base biasing resistor just as we saw with

the common emitter amplifier (R_B in a two-supply emitter bias or $R_1 \| R_2$ for a voltage divider bias).

Voltage Gain

The derivation for the emitter follower's voltage gain equation is similar to that shown for the common emitter amplifier. We begin with the basic definition of voltage gain and then expand using Ohm's law.

$$A_v = \frac{v_{out}}{v_{in}} = \frac{v_E}{v_B}$$
$$A_v = \frac{i_C r_E}{i_C (r'_e + r_E)}$$
$$A_v = \frac{r_E}{r'_e + r_E} \qquad (7.8)$$

This equation is very similar to that of Equation 7.4. Here we see that the output signal is in phase with the input and that if $r_E >> r'_e$, the gain approaches unity. Signal distortion tends to be low in followers because a gain of one is a desired goal.

Input Impedance

The derivation for Z_{in} and $Z_{in(base)}$ are unchanged compared to the common emitter configuration. The formulas are repeated below for convenience.

$$Z_{in(base)} = \beta(r'_e + r_E)$$

$$Z_{in} = r_B \| Z_{in(base)}$$

Output Impedance

The derivation for common collector output impedance varies considerably from that of the common emitter. We shall use Figure 7.18 for the analysis.

First, note that this diagram splits the AC emitter resistance into its two components, R_L and the biasing resistor R_E. This is because we want to find the effective resistance of the source that drives the load, so logically we can't include the load in that value. We begin by looking back into the emitter from the perspective of the load. We see the emitter bias resistor in parallel with whatever the impedance is looking back into the emitter terminal.

$$Z_{out} = R_E \| Z_{out(emitter)} \qquad (7.9)$$

$Z_{out(emitter)}$ is equal to r'_e in series with the equivalent resistance of the network above it and to the left. The internal resistance of the current source is high enough to ignore so we're left with the equivalent resistance looking back off the base. We'll call this $Z_{B(equivalent)}$. At first glance this might appear to be the parallel combination of r_{gen} and r_B, but this ignores the effect of the collector current source. What we really want is the effective resistance as seen from the perspective of r'_e, not as seen from the base terminal.

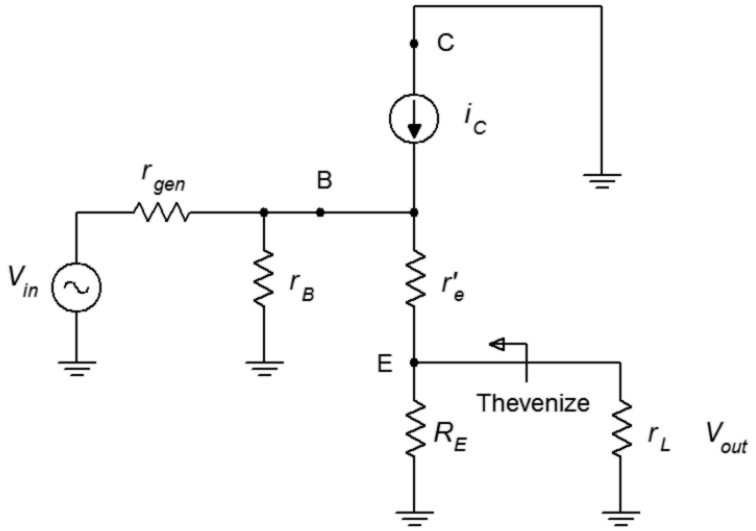


Figure 7.18
Common collector output impedance analysis.

$$Z_{out\,(emitter)} = r'_e + Z_{B(equivalent)} \tag{7.10}$$

$$Z_{B(equivalent)} = \frac{v_B}{i_C}$$

$$Z_{B(equivalent)} = \frac{i_B(r_B \parallel r_{gen})}{\beta i_B}$$

$$Z_{B(equivalent)} = \frac{r_B \parallel r_{gen}}{\beta} \tag{7.11}$$

Combining Equations 7.9, 7.10 and 7.11 yields

$$Z_{out} = R_E \parallel \left(r'_e + \frac{r_B \parallel r_{gen}}{\beta} \right) \tag{7.12}$$

In many instances the emitter bias resistor is large enough to ignore.

Example 7.4

For the follower shown in Figure 7.19, determine the input impedance, output impedance and load voltage. Assume $\beta = 100$ and $V_{in} = 100$ mV.

First, find I_C in order to find r'_e. Assuming an unloaded divider, V_B will equal half of the DC supply, or 10 volts. We lose 0.7 volts across the base-emitter junction leaving 9.3 volts across the 10 kΩ. This results in a collector current of 930 μA and an r'_e of 28 Ω.

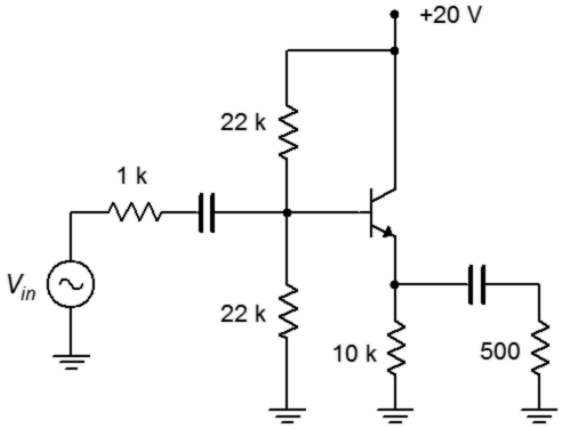


Figure 7.19
Schematic for Example 7.4.

To find Z_{in}

$$Z_{in(base)} = \beta(r'_e + r_E)$$
$$Z_{in(base)} = 100(28\,\Omega + 10\,\text{k}\Omega \parallel 500\,\Omega)$$
$$Z_{in(base)} = 50.4\,\text{k}\Omega$$

$$Z_{in} = R_1 \parallel R_2 \parallel Z_{in(base)}$$
$$Z_{in} = 22\,\text{k}\Omega \parallel 22\,\text{k}\Omega \parallel 50.4\,\text{k}\Omega$$
$$Z_{in} = 9.03\,\text{k}\Omega$$

This value is not particularly high when compared with the rather large source resistance of 1 kΩ. There will be some signal loss here due to the voltage divider effect between the two impedances. And now for Z_{out}

$$Z_{out} = R_E \parallel \left(r'_e + \frac{r_B \parallel r_{gen}}{\beta} \right)$$
$$Z_{out} = 10\,\text{k}\Omega \parallel \left(28\,\Omega + \frac{22\,\text{k}\Omega \parallel 22\,\text{k}\Omega \parallel 1\,\text{k}\Omega}{100} \right)$$
$$Z_{out} = 37\,\Omega$$

This value is much, much lower than anything we saw with the common emitter amplifiers. Therefore this circuit can drive much lower impedance loads with minimal signal loss. The loaded gain from base to emitter is

$$A_v = \frac{r_E}{r'_e + r_E}$$
$$A_v = \frac{500\,\Omega \parallel 10\,\text{k}\Omega}{28\,\Omega + 500\,\Omega \parallel 10\,\text{k}\Omega}$$
$$A_v = 0.9444$$

As mentioned, we need to include the effect of the 1 kΩ source impedance. This will create a voltage divider with the input impedance.

$$A_{divider} = \frac{Z_{in}}{Z_{in} + Z_{source}}$$
$$A_{divider} = \frac{9.03\,\text{k}\Omega}{9.03\,\text{k}\Omega + 1\,\text{k}\Omega}$$
$$A_{divider} = 0.9$$

$$A_{v(system)} = A_v \times A_{divider}$$
$$A_{v(system)} = 0.9444 \times 0.9$$
$$A_{v(system)} = 0.85$$

Finally, we get to the load voltage.

$$V_{load} = A_{v(system)} \times V_{in}$$
$$V_{load} = 0.85 \times 100\,\text{mV}$$
$$V_{load} = 85\,\text{mV}$$

At this point the question might be, "Why did we go to the trouble of building this circuit when we lost 15% of the input signal?" Well, consider what would have happened without the circuit. If we had connected the source directly to the load, the resulting 1 kΩ/500 Ω voltage divider would have dropped the load voltage to 33 mV. This circuit prevented that loss.

A High Impedance Source: The Guitar Pickup

In Example 7.4, the source had an internal impedance of 1 kΩ, much higher than we would see with, say, a laboratory function generator (probably 50 Ω). Things could be much worse. Consider the electric guitar pickup. The job of a pickup is to transform the vibrations of the guitar strings into an electrical signal so that it can be amplified. It is commonly thought that a pickup is some form of microphone but this is not true[27].

A guitar pickup is little more than a magnet surrounded by numerous turns of fine wire, as shown in Figure 7.20. This particular pickup is for a bass guitar but the construction is similar for all types of guitars and basses.

Figure 7.20
Electric bass pickup (cover removed).

27 Don't believe it? Just try screaming into one and listen to what comes out of the guitar amp.

Here is it how works: The magnet creates a field around the guitar strings. Because the strings are steel, their reluctance is much less than the surrounding air, therefore, they distort or bend the magnetic field. When a string is plucked, the field moves back and forth along with it. As the field moves, the flux lines cut across the coil's wire and this action induces a current in the conductor in accordance with Faraday's Law of Induction. This current is then fed to the amplifier.

A typical guitar pickup consists of perhaps 5000 turns of very fine wire, 42 AWG being typical. 42 gauge copper wire has a resistance of around 1.6 Ω per foot so the DC resistance of the coil can be over 5 kΩ. Further, that many turns of wire around a magnet can produce a very large inductance, perhaps several henries, that is in series with this resistance. There is also distributed capacitance and cable capacitance in parallel that could be upwards of 1 nF. The result is a complex impedance with resonance effects, regions of which can be tens of kΩ in magnitude. What makes this more challenging is that because the impedance is a function of frequency, the voltage divider effect with the amplifier's input impedance also becomes a function of frequency. For example, the increasing impedance due to X_L will result in increasing attenuation with frequency. This is akin to turning down the treble on the amplifier. Generally, not a good result. How do we limit this effect? Simple. We make a circuit with a very, very high input impedance. How do we do that? Well, there are several ways, including the use of field effect transistors and operational amplifiers, but we can also obtain high input impedances through the use of a dual BJT configuration called the *Darlington pair*.

The Darlington Pair

The Darlington pair was invented by Sidney Darlington, an American engineer. The configuration leads to a compound device with a very high β. Used properly, this can lead to amplifier circuits with very high input impedance. A Darlington pair is shown in Figure 7.21.

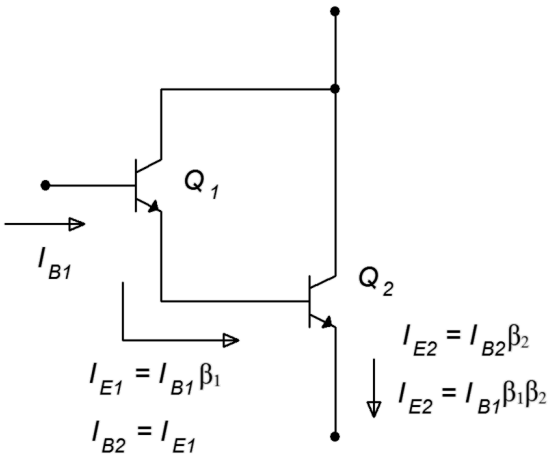


Figure 7.21
Darlington pair.

The operation is as follows. The base current of the first transistor, Q_1, is multiplied by the β of Q_1 resulting in Q_1's emitter current. This current is fed into the base of the second transistor, Q_2, where it is multiplied by the β of Q_2 resulting in Q_2's emitter current. If we treat the pair as a single device, then the effective β of the pair is $\beta_1\beta_2$. Given typical values for β, the compound value can be in the vicinity of 5000 to 10,000. The functional downside to this arrangement is that V_{BE} is now doubled to 1.4 volts (for silicon) and the effective r'_e of the pair is doubled as well. These issues are minor when compared to the advantage of the huge current gain that can be obtained.

The bottom line when using a Darlington pair is to treat it like an ordinary transistor except that it has a very large β and both V_{BE} and r'_e are doubled compared to the ordinary values.

Example 7.5

Determine the output voltage for the follower shown in Figure 7.22. Assume the input is 100 mV peak and the β for the Darlington pair is 10,000.

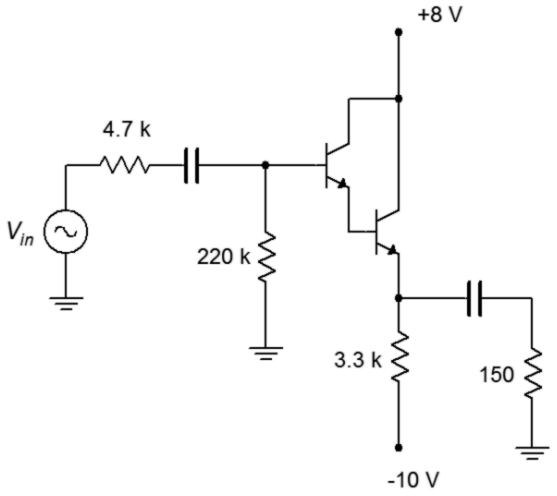


Figure 7.22
Schematic for Example 7.5.

The first thing that might look a little odd, at least compared to previous bias circuits, is that the base biasing resistor is so much larger than the emitter biasing resistor. Normally, this would lead to an unstable Q point but that's not a problem here. Because β is so large, R_B can be much larger than normal and we'll still achieve good stability. In fact, we can still use the approximation that the base is at DC ground. This being true, the analysis proceeds as follows

$$I_C = \frac{|V_{EE}| - V_{BE}}{R_E}$$
$$I_C = \frac{10\,\text{V} - 1.4\,\text{V}}{3.3\,\text{k}\Omega}$$
$$I_C = 2.61\,\text{mA}$$

$$r'_e = 2 \times \frac{26\,\text{mV}}{I_C}$$
$$r'_e = \frac{52\,\text{mV}}{2.61\,\text{mA}}$$
$$r'_e = 20\,\Omega$$

$$Z_{in(base)} = \beta(r'_e + r_E)$$
$$Z_{in(base)} = 10{,}000(20\,\Omega + 3.3\,\text{k}\Omega \parallel 150\,\Omega)$$
$$Z_{in(base)} = 1.63\,\text{M}\Omega$$

This value is in parallel with the base biasing resistor, creating the input impedance.

$$Z_{in} = R_B \parallel Z_{in(base)}$$
$$Z_{in} = 220\,\text{k}\Omega \parallel 1.63\,\text{M}\Omega$$
$$Z_{in} = 194\,\text{k}\Omega$$

This is much higher than we have seen in previous circuits. The loaded gain from base to emitter is

$$A_v = \frac{r_E}{r'_e + r_E}$$
$$A_v = \frac{150\,\Omega \parallel 3.3\,\text{k}\Omega}{20\,\Omega + 150\,\Omega \parallel 3.3\,\text{k}\Omega}$$
$$A_v = 0.88$$

Now to include the effect of the 4.7 kΩ source impedance. This will create a voltage divider with the input impedance, minimal as it turns out.

$$A_{divider} = \frac{Z_{in}}{Z_{in} + Z_{source}}$$
$$A_{divider} = \frac{194\,\text{k}\Omega}{194\,\text{k}\Omega + 4.7\,\text{k}\Omega}$$
$$A_{divider} = 0.976$$

$$A_{v(system)} = A_v \times A_{divider}$$
$$A_{v(system)} = 0.88 \times 0.976$$
$$A_{v(system)} = 0.86$$

The load voltage is

$$V_{load} = A_{v(system)} \times V_{in}$$
$$V_{load} = 0.86 \times 100\,\text{mV}$$
$$V_{load} = 86\,\text{mV}$$

If we had connected the source directly to the load, the 4.7 kΩ/150 Ω divider would have squashed the applied signal into a shadow of its former size, leaving us with just 3 mV.

Computer Simulation

To verify the results of Example 7.5, we'll run a transient analysis. The input schematic is shown in Figure 7.23.

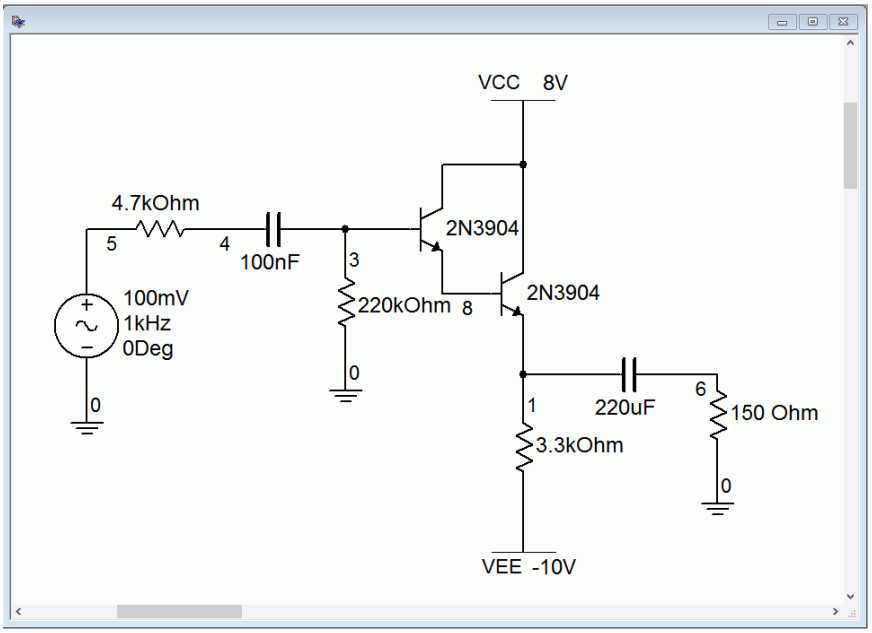
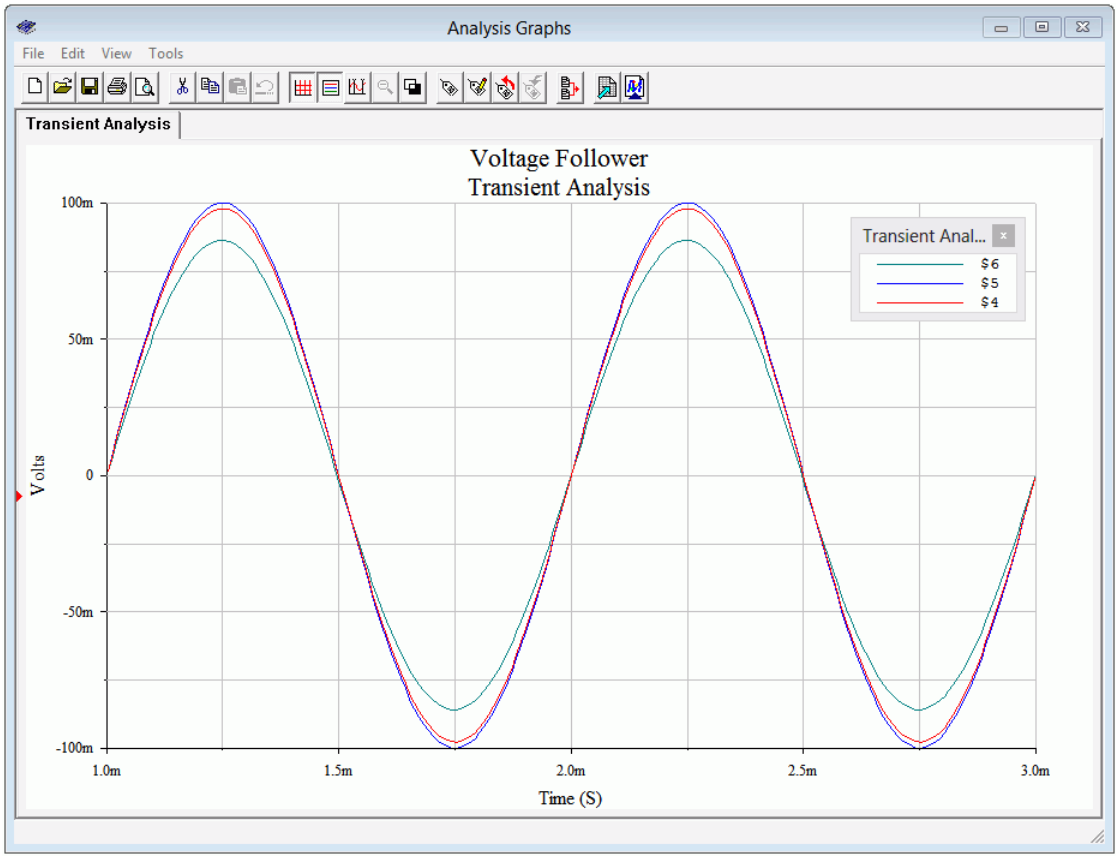


Figure 7.23
Simulation schematic for Darlington pair follower.

Of interest here will be the voltages at the source, base and load. As the input impedance/source impedance divider was 0.976, we expect 97.6 mV at node 4. At the output, node 6, we expect to see our final computed value of 86 mV. The output plot of the simulation is shown in Figure 7.24. The simulation concurs.

Figure 7.24
Transient analysis for Darlington pair follower.



The Phase Splitter

A *phase splitter* is a combination of a common emitter amplifier and a common collector follower using a single transistor. The purpose of the circuit is to produce two versions of the input signal: a buffered version identical to the input and an inverted version, both waves having the same amplitude. The circuit is used for differential line driver systems. This scheme helps to minimize outside noise and interference picked up by communications cables. There are other ways to create phase splitters, including using differential amplifiers or op amps, but this BJT-based version is a minimalist solution[28]. The basic circuit is shown in Figure 7.25.

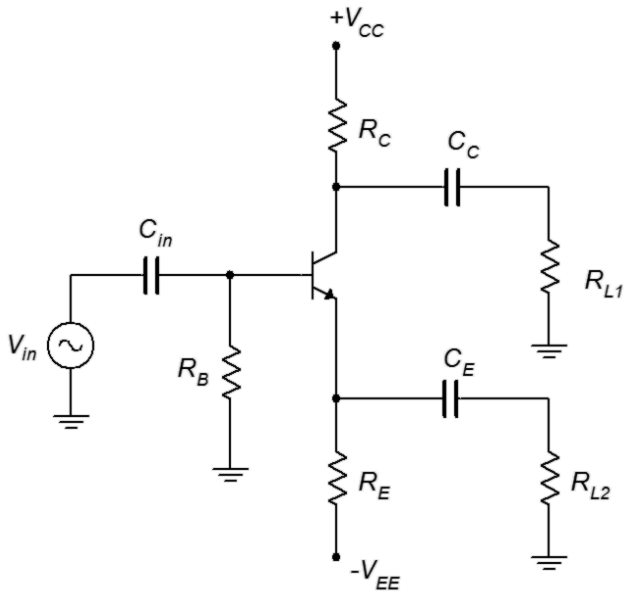


Figure 7.25
A simple phase splitter.

For proper operation, the circuit is largely symmetrical. That is, $R_{L1} = R_{L2}$, $R_E = R_C$, and $C_C = C_E$. That being the case, the AC collector and emitter resistances will be equal ($r_C = r_E$). If we then look at the basic gain equations, we find that both loads will receive the same gain magnitude (just under unity), although R_{L1} will see the signal inverted.

$$A_v = -\frac{r_C}{r'_e + r_E} \quad \text{Common emitter amplifier}$$

$$A_v = \frac{r_E}{r'_e + r_E} \quad \text{Common collector follower}$$

28 For details on alternate methods, see Fiore, J, *Operational Amplifiers and Linear Integrated Circuits: Theory and Application*, another free OER text.

7.5 The Common Base Amplifier

The third and final prototype is the *common base amplifier*. In this configuration the input signal is applied to the emitter and the output is taken from the collector. The base terminal is at the common ground point. An example, using two-supply emitter bias, is shown in Figure 7.26. Note that because neither the input nor output is connected to the base, there is no need for a base resistor. Consequently, the base terminal is connected directly to ground.

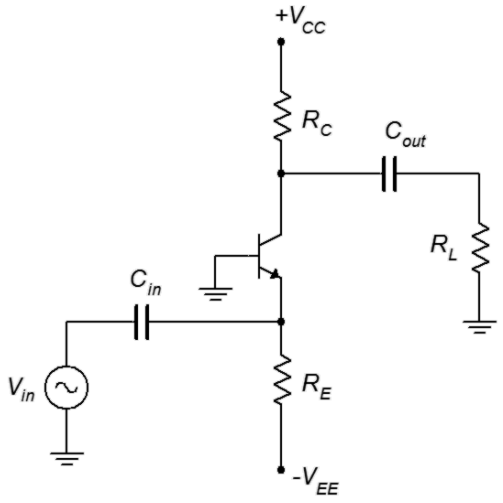


Figure 7.26
Common base amplifier.

Some people find that redrawing the schematic horizontally helps to visualize the signal flow. This version is shown in Figure 7.27.

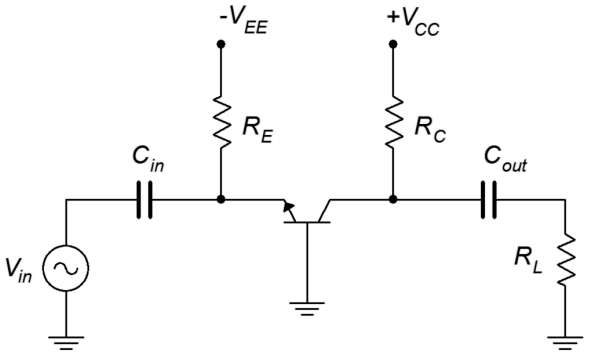


Figure 7.27
Common base amplifier redrawn.

One nice thing about the horizontal version is that when we make the AC equivalent, it becomes obvious that R_E is in parallel with the input and R_C is in parallel with the load. For the AC analysis we shall modify Figure 7.27 by substituting the BJT model for transistor, shorting the capacitors and taking the DC sources to AC ground. The result is shown in Figure 7.28.

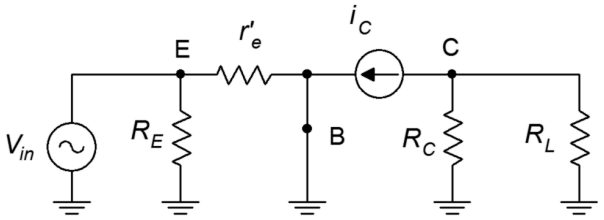


Figure 7.28
Common base amplifier with BJT model.

Voltage Gain

We begin with the basic definition of voltage gain and then expand using Ohm's law.

$$A_v = \frac{v_{out}}{v_{in}} = \frac{v_C}{v_E}$$
$$A_v = \frac{i_C r_C}{i_E r'_e}$$
$$A_v = \frac{r_C}{r'_e} \tag{7.13}$$

This equation is very similar to that of a non-swamped common emitter amplifier except that it does not invert the input signal. Therefore, gain potential is fairly high.

Input Impedance

The derivation for Z_{in} is obtained via direct inspection of the schematic.

$$Z_{in} = R_E \,||\, r'_e \tag{7.14}$$

r'_e normally dominates and thus we see that the common base configuration tends to have a low input impedance. For audio frequencies this can be an issue but it is less of a problem at higher frequencies as, generally speaking, system impedances need to be lower to avoid complications with capacitive effects.

Output Impedance

The derivation for Z_{out} unchanged compared to the common emitter configuration. The formula is repeated below for convenience.

$$Z_{out} \approx R_C$$

For the amplifier shown in Figure 7.29, determine the voltage gain and input impedance.

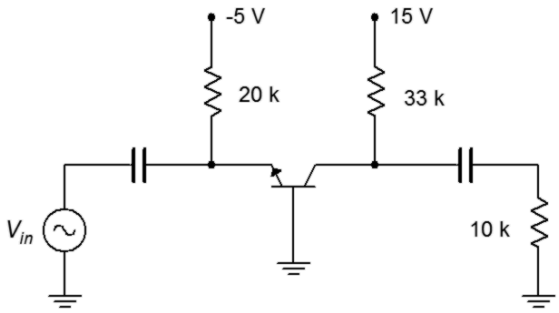


Figure 7.29
Schematic for Example 7.6.

$$I_C = \frac{|V_{EE}| - V_{BE}}{R_E}$$
$$I_C = \frac{5\,\text{V} - 0.7\,\text{V}}{20\,\text{k}\Omega}$$
$$I_C = 0.215\,\text{mA}$$

$$r'_e = \frac{26\,\text{mV}}{I_C}$$
$$r'_e = \frac{26\,\text{mV}}{0.215\,\text{mA}}$$
$$r'_e = 121\,\Omega$$

$$Z_{in} = r'_e \parallel r_E$$
$$Z_{in} = 121\,\Omega \parallel 20\,\text{k}\,\Omega$$
$$Z_{in} = 120\,\Omega$$

$$A_v = \frac{r_C}{r'_e}$$
$$A_v = \frac{33\,\text{k}\,\Omega \parallel 10\,\text{k}\,\Omega}{120\,\Omega}$$
$$A_v = 64$$

7.6 Multi-Stage Amplifiers

In order to achieve a higher gain than we can obtain from a single stage, it is possible to cascade two or more stages. Different biasing types might be used along with a mix of AC configurations such as a common collector follower for the first stage that drives a common emitter voltage amplifier. A mix of NPN and PNP devices may also be present.

In general terms, each stage serves as the load for the preceding stage. That is, the Z_{in} of one stage is the R_L of the previous stage. The gains of the individual stages are then multiplied together to arrive at the system gain. The system input impedance is the input impedance of the first stage only. The source drives the first stage alone. The first stage, in turn, drives the second stage, and so on. Therefore the source only "sees" the first stage because it is the only stage to which it delivers current. In a

similar fashion, the output impedance of the system is the Z_{out} of the last stage. An example is shown in Figure 7.30.

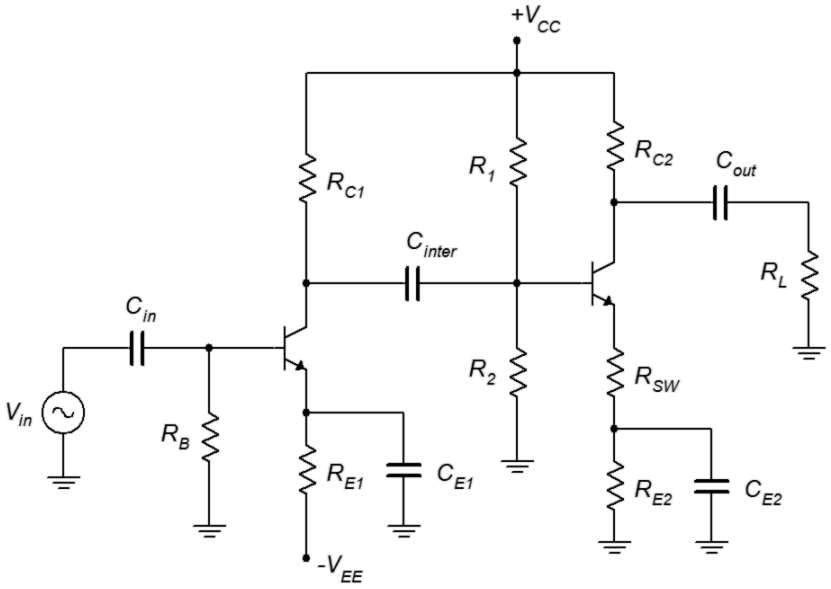


Figure 7.30
Two stage amplifier.

In this circuit, stage one is a non-swamped common emitter amplifier utilizing two-supply emitter bias. Stage two is a swamped common emitter amplifier using voltage divider bias. As far as the DC analysis is concerned, these are two separate circuits. The inter-stage coupling capacitor, C_{inter}, prevents the DC potential at the collector of the first transistor from interfering with the bias established by R_1 and R_2 for transistor number two. For the AC computation, the first stage is analyzed in normal fashion except that its load resistance is comprised of $R_1 \parallel R_2 \parallel Z_{in\text{-}base2}$ (i.e., Z_{in} of stage 2). The second stage is analyzed without changes and its gain is multiplied by the first stage's gain to arrive at the final gain for the pair. The input impedance of the system is $R_B \parallel Z_{in\text{-}base1}$ (i.e., Z_{in} of stage 1).

It should be obvious that by cascading several stages it is possible to achieve very high system gains, even if each stage is heavily swamped in order to reduce distortion. For example, three swamped common emitter stages with voltage gains of just 10 each would produce a system voltage gain of 1000.

Direct Coupling

With a little creativity, it is possible to create multi-stage designs that use fewer components but which achieve higher performance. One technique is to employ *direct coupling* of the stages. Direct coupling allows DC to flow from stage to stage. As such, it is possible to design an amplifier that has no lower frequency limit. An example is shown in Figure 7.31.

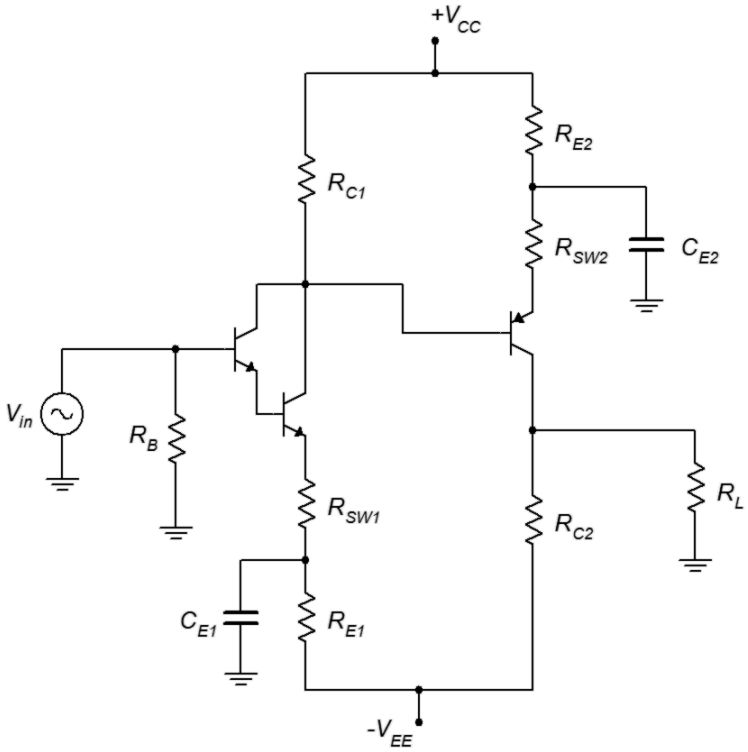


Figure 7.31
Direct coupled amplifier.

This two-stage amplifier uses no coupling capacitors nor does it rely on voltage divider resistors for the second stage[29]. Here is how it works: The first stage is a fairly ordinary swamped common emitter amplifier using two-supply emitter bias. It also uses a Darlington pair to maximize the input impedance. Because the base current is so low, the DC drop on R_B could be small enough to ignore so we may dispense with the input coupling capacitor. The DC potential at the collector of the Darlington is applied directly to the base of the second stage. This is used to set up the bias of the second stage via the stage two emitter resistors. This is precisely what we did with the circuit of Figure 7.7. The only difference is that here the base voltage is derived from the preceding stage instead of from a voltage divider. The computations for I_C, r'_e and the like would proceed unchanged. In any event, this eliminates two biasing resistors and another coupling capacitor.

Note the use of the PNP device for the second stage. By using a PNP, its collector voltage must be less than its emitter voltage. As we're also using a bipolar power supply, we can eliminate the need for the final output coupling capacitor. All we need to do is set up the resistor values such that the drop across R_{C2} is the same as V_{EE}. This will place the stage two DC collector voltage at 0 volts. If there's no DC voltage then there's nothing to block, and therefore no need for the coupling capacitor.

29 This circuit does use emitter bypass capacitors so the DC gain will be less than the AC gain. In that sense we might say that this amplifier is not fully DC coupled.

Summary

A simplified AC model of the bipolar junction transistor consists of a controlled current source in the collector and a dynamic resistance in the emitter called r'_e. This resistance is a function of the DC bias current, I_C; the higher the biasing current, the lower the resistance. Fluctuations in this resistance can lead to waveform distortion. Swamping, also known as emitter degeneration, is a technique used to reduce distortion and stabilize gain. The basic idea is to add a fixed resistor in series with the emitter so as to buffer or "swamp out" the changes of r'_e.

There are three basic AC amplifier configurations: common emitter, common collector and common base. The common emitter configuration produces a voltage amplifier with high gain and intermediate input impedance. It also inverts the signal. Because it exhibits both voltage gain and current gain, it has a potential for high power gain. The common collector configuration is known as a follower because its output follows the input. It produces a non-inverting voltage gain of one and exhibits high input impedance and low output impedance. Therefore, it is useful as either an input buffer or as a final drive stage to a low impedance load. The common base configuration exhibits high non-inverting voltage gain. It has a low input impedance and a high output impedance.

The Darlington pair is a two-transistor configuration that may be treated as a single device. As such, it exhibits a doubling of both V_{BE} and r'_e, and a very large β.

In order to achieve higher gains, multiple stages may be cascaded. Their gains multiply together to produce the combined system gain. The stages may be coupled through capacitors or via a capacitor-less direct coupling technique that can improve performance while reducing component count.

Review Questions

1. How does the AC BJT model compare with the DC model? What are the differences and similarities?
2. Explain how swamping reduces waveform distortion.
3. Compare and contrast common emitter, common collector and common base amplifiers in terms of voltage gain, power gain, input impedance and output impedance.
4. Why might biasing circuits that produce stable Q points be preferred for non-swamped amplifiers?
5. What is a phase splitter?
6. What are the advantages of direct coupling?
7. Give at least one example of a high internal impedance source.

Problems

Unless otherwise specified, use $\beta = 100$.

Analysis Problems

1. Determine the input and output impedances of the circuit of Figure 7.32.

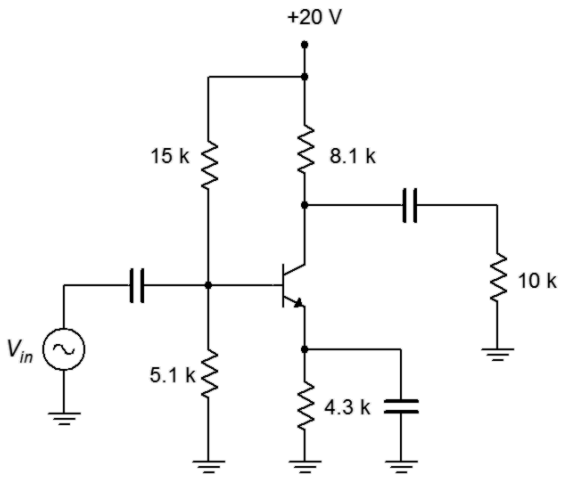


Figure 7.32

2. Determine the load voltage for the circuit of Figure 7.32 if V_{in} is 10 mV.

3. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.33 if V_{in} is 70 mV.

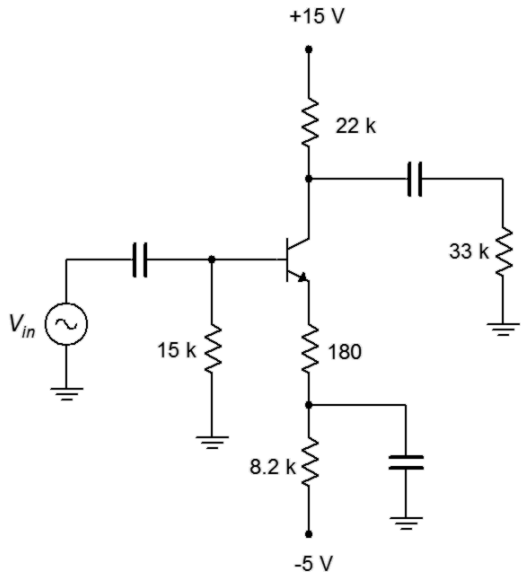


Figure 7.33

4. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.34 if V_{in} is 50 mV.

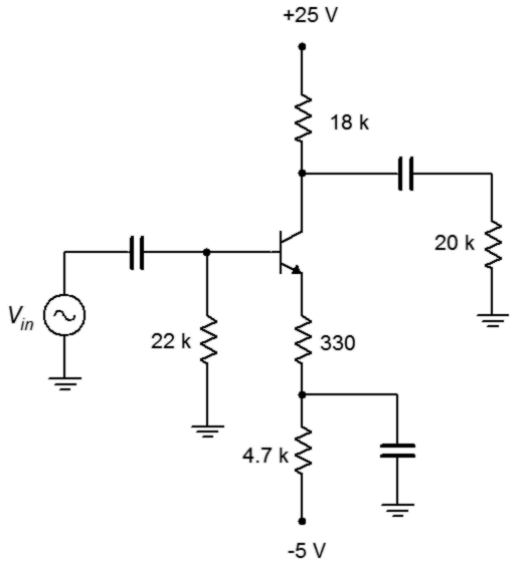


Figure 7.34

5. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.35 if V_{in} is 25 mV.

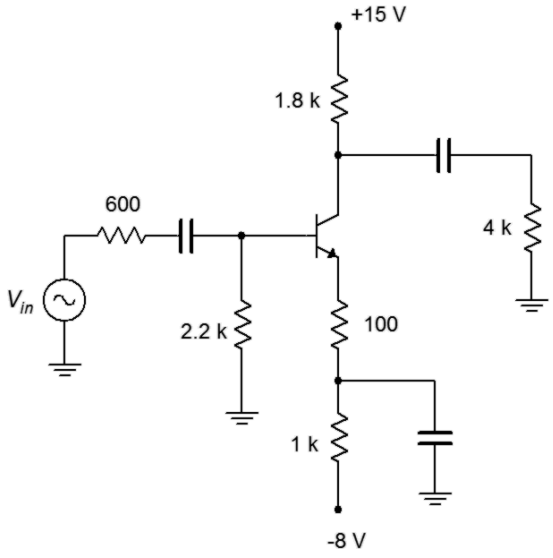


Figure 7.35

6. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.36 if V_{in} is 30 mV.

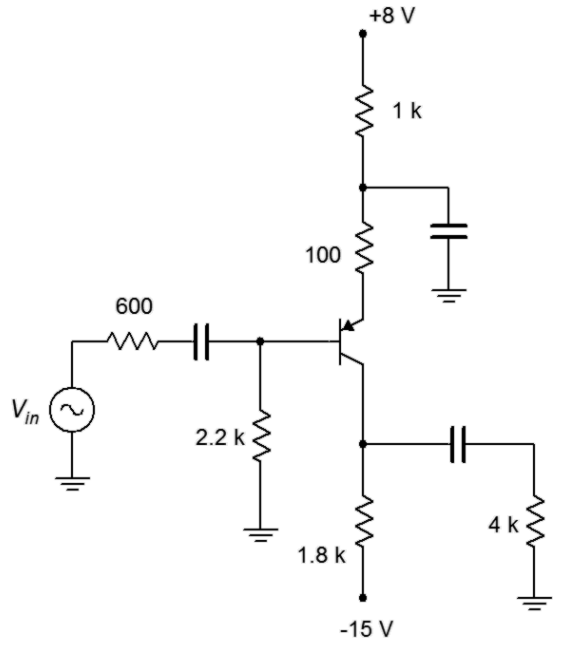


Figure 7.36

7. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.37 if V_{in} is 60 mV.

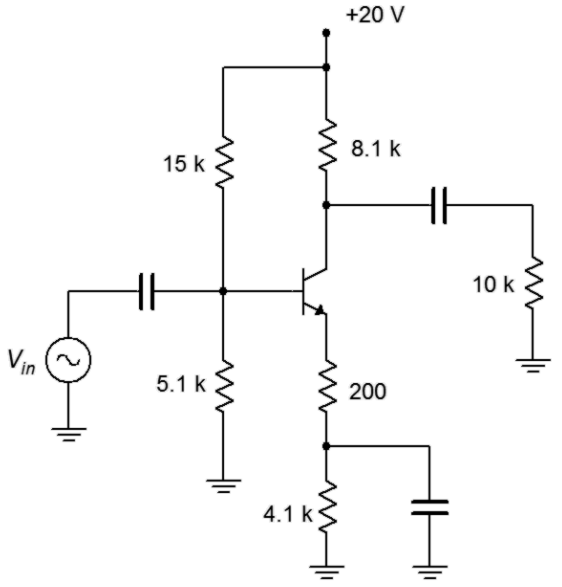


Figure 7.37

8. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.38 if V_{in} is 150 mV.

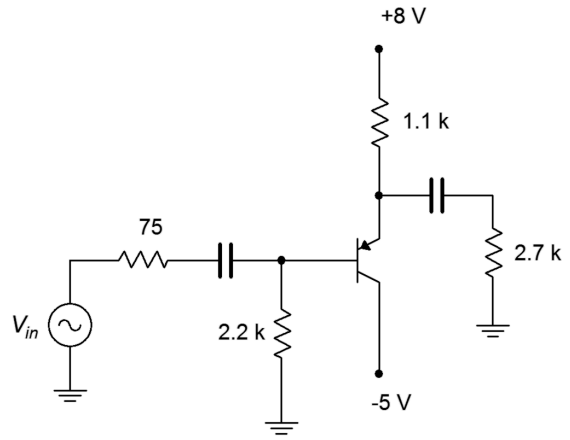


Figure 7.38

9. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.39 if V_{in} is 200 mV.

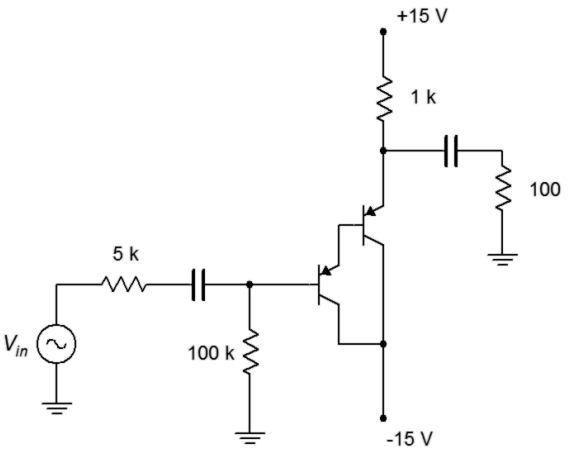


Figure 7.39

10. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.40 if V_{in} is 250 mV.

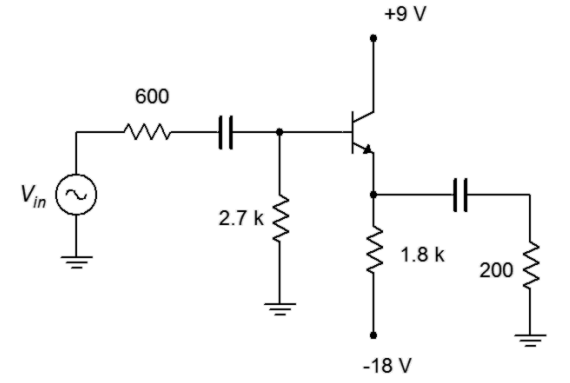


Figure 7.40

11. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.41 if V_{in} is 300 mV.

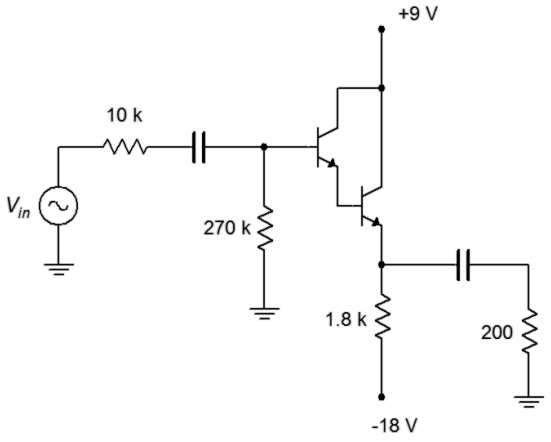


Figure 7.41

12. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.42 if V_{in} is 50 mV.

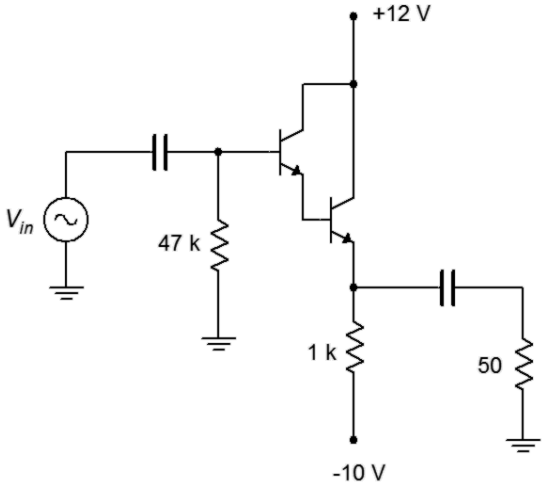


Figure 7.42

13. Determine Z_{in}, Z_{out}, and the load voltage for the circuit of Figure 7.43 if V_{in} is 2 mV.

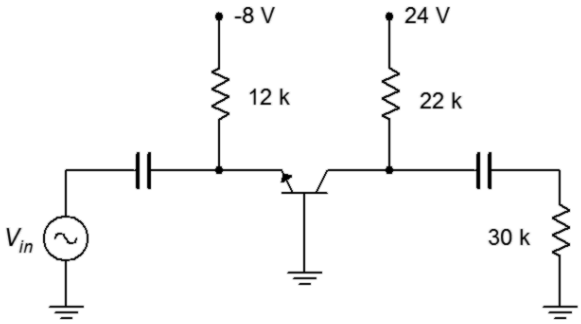


Figure 7.43

Design Problems

14. Redesign the circuit of Figure 7.33 to halve the existing gain while keeping the Q point where it is currently.

15. By using a Darlington pair, redesign the circuit of Figure 7.34 to double Z_{in}.

16. Redesign the circuit of Figure 7.34 so that it exhibits the same performance parameters but uses a PNP device.

17. Redesign the circuit of Figure 7.36 to double the existing gain while keeping the Q point where it is currently.

18. Redesign the circuit of Figure 7.38 so that it exhibits the same performance parameters but uses an NPN device.

Challenge Problems

19. Determine the gain and input impedance for the circuit of Figure 7.44. V_{CC} = 20 V, V_{EE} = −10 V, R_B = 18 kΩ, R_{E1}= 10 kΩ, R_{C1} = 12 kΩ, R_1 = 33 kΩ, R_2 = 15 kΩ, R_{E2} = 5.6 kΩ, R_{SW} = 400 kΩ, R_{C2} = 6.8 kΩ, R_L = 24 kΩ.

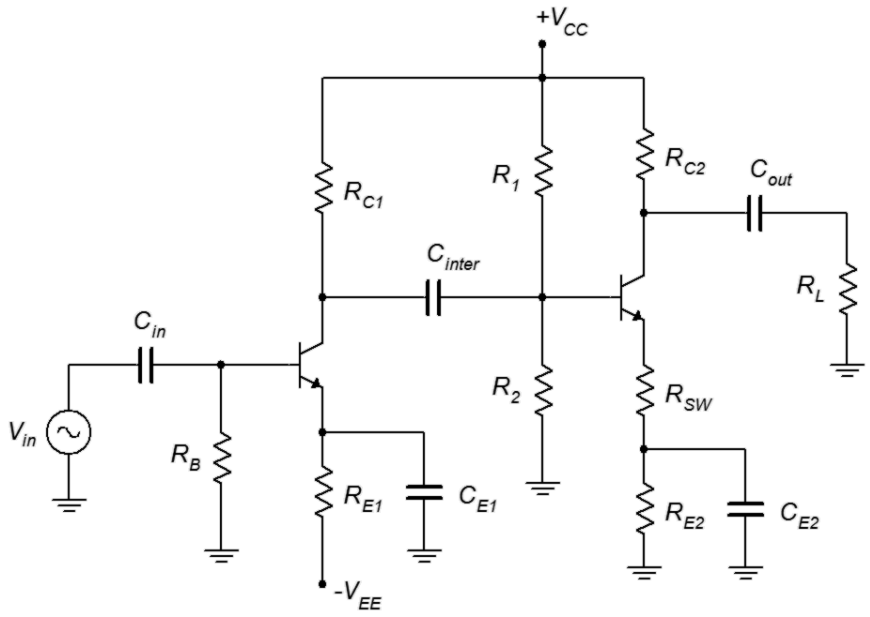


Figure 7.44

20. For the circuit of Figure 7.41, replace its load resistor with the circuit of Figure 7.37 and determine the combined gain and input impedance of the system.

Computer Simulation Problems

21. Use a transient analysis to verify the load voltage of problem 3.

22. Use a transient analysis to verify the load voltage of problem 4.

23. Use a transient analysis to verify the load voltage of problem 8.

24. Consider the amplifier of Figure 7.32. Replace the 4.3 kΩ emitter resistor with a potentiometer of the same value. Connect the wiper arm to the emitter bypass capacitor. Run several transient analyses at different pot settings (0%, 25%, 50%, etc.). What can you conclude from the results?

8 BJT Class A Power Amplifiers

8.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Define class A operation.
- Determine AC load lines for class A amplifier stages.
- Determine the compliance and maximum load power for class A amplifier circuits.
- Determine the efficiency and required device ratings for class A amplifier circuits.
- Describe the operation of a dynamic loudspeaker.
- Understand the need for heat sinks and other thermal management techniques.

8.1 Introduction

Now that we have examined BJT voltage amplifiers in terms of their gain, input impedance and output impedance, it is time that we extend the small signal analysis to larger signals. Of primary importance will be determination of the maximum output voltage swing, or compliance, along with maximum load power, device dissipation requirements and amplifier efficiency. To assist with this, we introduce the concept of the *AC load line*. In general, we will not concern ourselves with input impedance, voltage gain or even r'_e. Indeed, we shall simply consider r'_e as a source of distortion. Most power amplifiers are configured as voltage followers so we will focus largely on those.

8.2 Amplifier Classes

There are several classes of amplifier operation. The class of an amplifier has nothing to do with the fidelity or quality of the amplifier. Rather, the class indicates the fundamental operational principle of the circuit. In general, as the class letter increases, the designs become more complicated but also more efficient. For audio and other linear applications, classes A, B and D are relatively common these days. Class C is largely relegated to high power telecommunications while classes G and H are essentially variations of class B.

The definition of class A is that signal current in the collector flows 360° out of the cycle. In other words, it flows for the entire cycle without interruption. All of the amplifiers that were presented in the prior chapter are class A designs. In class B, i_C flows for just 180°, and for class D, i_C is discontinuous; the transistor is used as a switch. Class B and D designs are examined in later chapters.

8.3 Class A Operation and Load Line

The signal current in the class A amplifier flows continuously throughout the entire cycle of the waveform. Ultimately, we would like to known just how large this signal can be before it is limited and grossly distorted. To do so, we need to examine the AC equivalent of the amplifier. A generic AC equivalent is shown in Figure 8.1. This includes both AC collector and emitter resistances so it can be used for either swamped or unswamped common emitter amplifiers or for emitter followers. If one of the resistances is not used (for example, r_C in a follower), we can just substitute a value of zero for it.

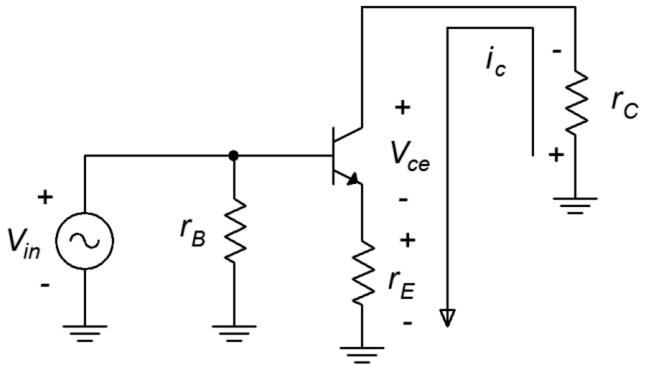


Figure 8.1
AC equivalent circuit.

The voltage polarities and current direction are shown for a positive input voltage. To determine the maximum load voltage swing (compliance), we will need to construct an *AC load line* as shown in Figure 8.2.

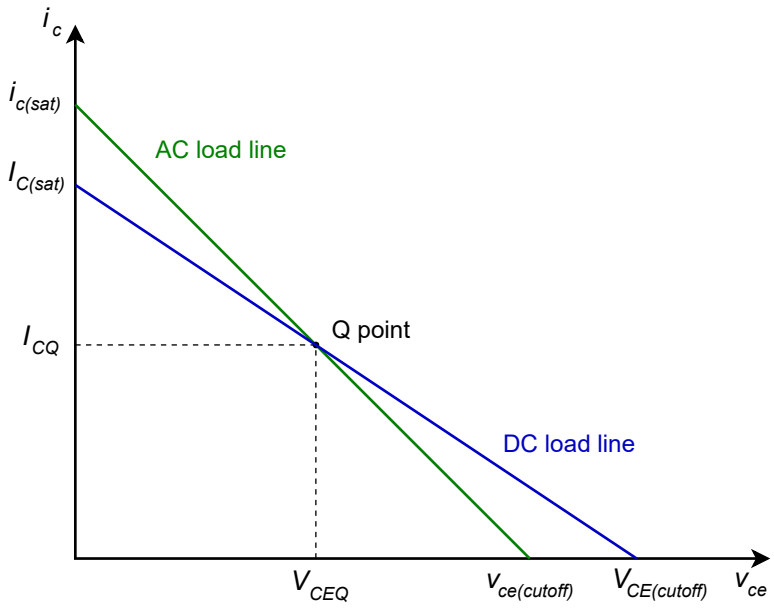


Figure 8.2
AC and DC load lines.

The AC load line is similar to the DC load line that was used for analyzing biasing circuits. As in the DC version, there will be a cutoff voltage, $v_{CE(cutoff)}$, and a saturation current, $i_{C(sat)}$. The AC and DC load lines normally are not the same, however, they must share one point in common, and that's the Q point. Usually, the slope of the AC load line is steeper than that of the DC load line. This is because the AC resistance tends to be less than the DC resistance due to loading and capacitor bypassing. Consequently, $v_{CE(cutoff)}$ tends to be smaller than $V_{CE(cutoff)}$ and $i_{C(sat)}$ tends to be larger than $I_{C(sat)}$.

To determine expressions for the AC load line endpoints, let's examine the AC equivalent circuit. Because both load lines share the Q point, we can consider the circuit of Figure 8.1 as having a no-signal current of I_{CQ} and a no-signal transistor voltage of V_{CEQ}. As the input signal grows, i_C increases. The effect of this is to increase the voltage drops across r_E and r_C due to Ohm's law. This, in turn, forces v_{CE} to decrease due to KVL. The collector current can only increase to the point where v_{CE} drops to 0 V. This is a maximum increase of $V_{CEQ}/(r_C+r_E)$. Therefore

$$i_{C(sat)} = I_{CQ} + \frac{V_{CEQ}}{r_E + r_C} \tag{8.1}$$

In terms of cutoff voltage, the transistor starts with V_{CEQ} and I_{CQ}. The largest v_{CE} increase that can occur is if the current falls to zero. Then, all of the potential originally developed across r_E and r_C by I_{CQ} must be absorbed by the transistor. Therefore

$$v_{CE(cutoff)} = V_{CEQ} + I_{CQ}(r_E + r_C) \tag{8.2}$$

There are three possible ways this can be configured: Q point closer to saturation, Q point closer to cutoff, or Q point centered on the AC load line. Let's first consider the Q point closer to saturation. This is shown in Figure 8.3.

Here we have plotted the input voltage in red and drawn the corresponding collector current and collector-emitter voltage in blue. It is apparent that as the input signal increases, eventually, the output signal is limited at zero for v_{CE} and at $i_{C(\text{sat})}$ for i_C. The two blue waveforms are severely clipped and distorted. The largest unclipped peak voltage swing is V_{CEQ} and the largest peak current swing is $i_{C(sat)} - I_{CQ}$, or more conveniently, $V_{CEQ}/(r_E+r_C)$.

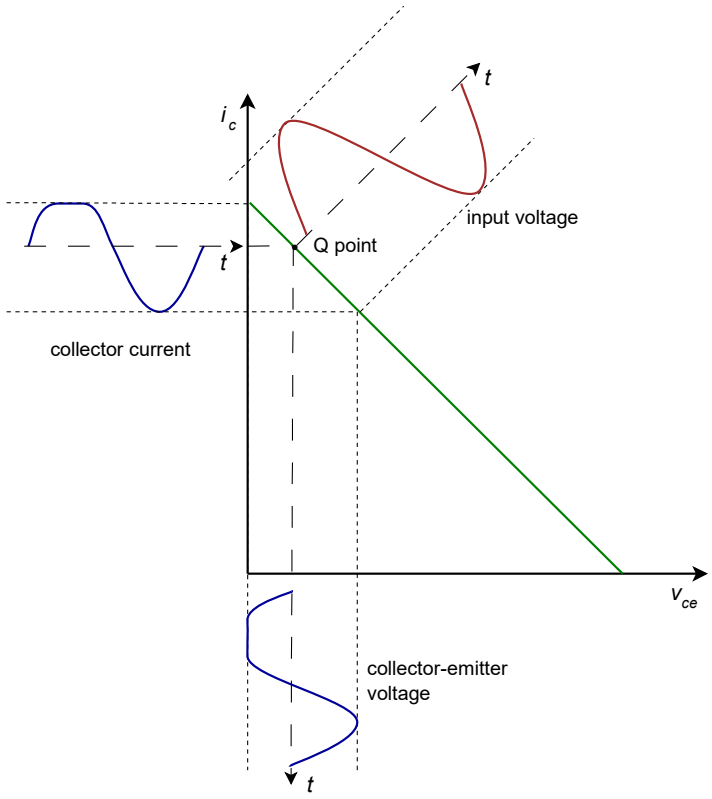


Figure 8.3
AC load line, Q point closer to saturation.

If we shift the Q point toward cutoff, we solve the saturation clipping problem but now we have a new problem, as illustrated in Figure 8.4. It should come as no surprise that we now have cutoff clipping.

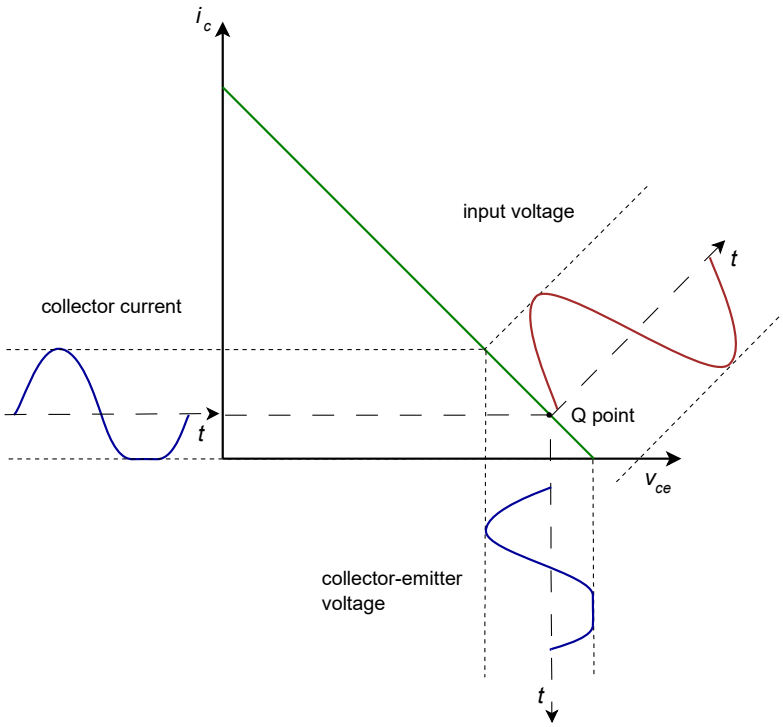


Figure 8.4
AC load line, Q point closer to cutoff.

In this version the largest unclipped peak voltage swing is $v_{CE(cutoff)} - V_{CEQ}$ (or alternately, $I_{CQ}(r_E+r_C)$) and the largest peak current swing is I_{CQ}. What's important here is that the waveform has been clipped. It doesn't really matter which side has been clipped, either way it's gross distortion. Eventually, every amplifier will have a limit but we will be able to produce the largest unclipped voltage swing if the Q point is centered on the AC load line. This is shown in Figure 8.5.

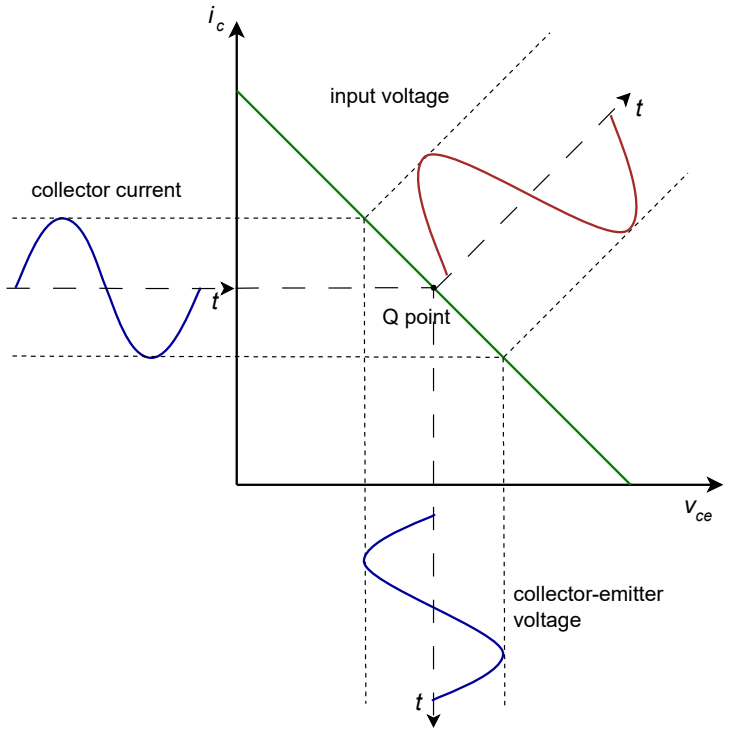


Figure 8.5
AC load line, centered Q point.

With a centered Q point, the largest unclipped peak voltage swing is V_{CEQ} and the largest unclipped peak current swing is I_{CQ}. By examining equations 8.1 and 8.2 it is apparent that in order to achieve a centered Q point on the AC load line, the following must be true:

$$\frac{V_{CEQ}}{I_{CQ}} = r_E+r_C \qquad (8.3)$$

Of course, while it is useful to determine the maximum voltage across the transistor, it is more important to determine the maximum voltage across the load. Looking back at the circuit of Figure 8.1, most times the maximum load voltage (i.e., the compliance) will equal the maximum transistor voltage. This will be the case in voltage followers and unswamped amplifiers. The only time there will be a noticeable reduction is with very heavily swamped amplifiers. In this case the compliance will be reduced by the voltage divider between the load and swamping resistors. For example, a swamped amplifier with a voltage gain of 4 would lose about 20% of the maximum swing. Swamping has to be *very* heavy resulting in *very* low gains before appreciable signal is lost.

Thus we arrive at the following general rule:

Peak compliance is the smaller of V_{CEQ} *or* $I_{CQ}(r_E+r_C)$ (8.4)

Knowing the compliance, the maximum load power may be determined using power law. Power is determined using RMS values, so the peak compliance will need to be divided by √2 (or multiplied by 0.707) before continuing.

$$P_{load(max)} = \frac{Compliance_{RMS}^{\ 2}}{R_L} \qquad (8.5)$$

There is something important to note about this equation. It uses the load resistance value, not the total AC effective value (i.e., **not** r_L which is R_L in parallel with a biasing resistor). If r_L was used, we'd be calculating the power in the load *plus* the power in the biasing resistor.

We would also like to determine the maximum power dissipated by the transistor. Because the transistor's current and voltage are fluctuating with the input signal, we need to determine the magnitude of the load voltage that produces maximum power in the transistor. Intuitively, we might guess that this occurs at maximum load power but it turns out that this guess is incorrect. Under no-signal conditions the transistor is operating statically at the Q point. Therefore, quiescent power dissipation is

$$P_{DQ} = V_{CEQ} I_{CQ} \qquad (8.6)$$

In contrast, at full load for a centered Q point, we have

$$\begin{aligned} v_{CE} &= V_{CEQ}(1-\sin 2\pi ft) \\ i_C &= I_{CQ}(1+\sin 2\pi ft) \end{aligned}$$

$$\begin{aligned} P_D &= v_{CE} i_C \\ P_D &= V_{CEQ}(1-\sin 2\pi ft) \times I_{CQ}(1+\sin 2\pi ft) \\ P_D &= V_{CEQ} I_{CQ}(1-\sin^2 2\pi ft) \\ P_D &= V_{CEQ} I_{CQ}(.5+.5\cos 4\pi ft) \\ P_D &= \frac{P_{DQ}}{2} + \frac{P_{DQ}}{2}\cos 4\pi ft \end{aligned} \qquad (8.7)$$

The first term of Equation 8.7 is a fixed offset while the second term is a sinusoid at twice the signal frequency. Because the peak amplitude of this sinusoid is the same as the fixed offset, the average over time is simply the offset value. These waveforms are illustrated in Figure 8.6.

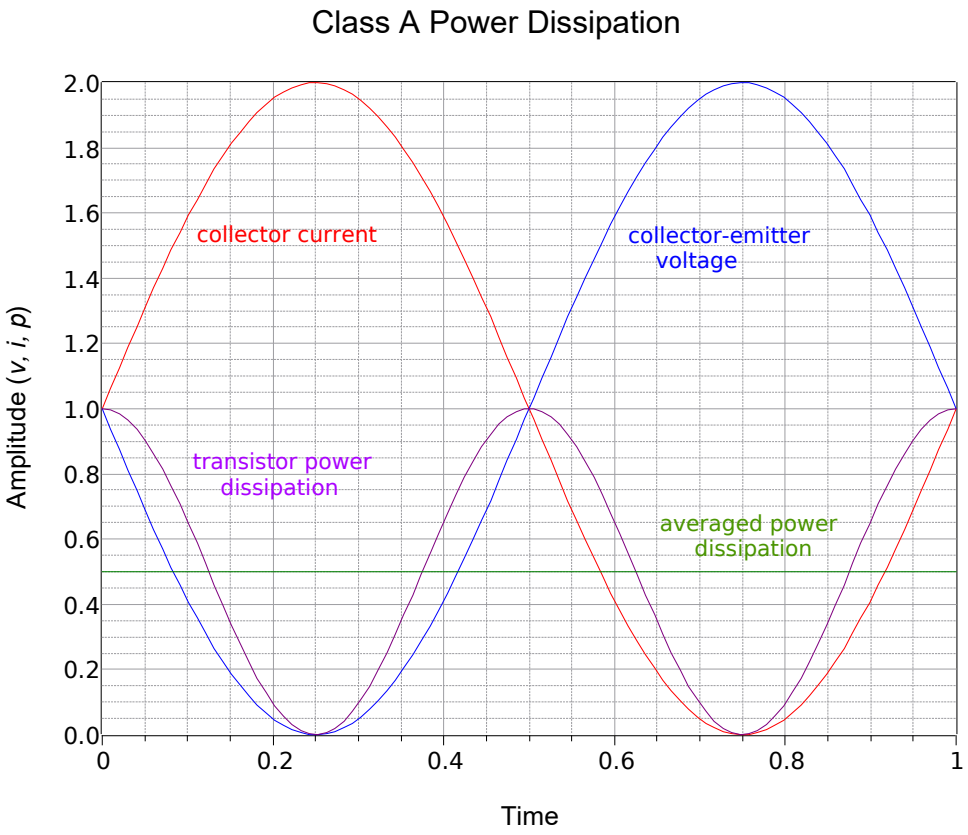


Figure 8.6
Transistor power dissipation at full load power.

The result is that the transistor only dissipates half the power at full load that it dissipates under idle conditions. This makes perfect sense if you stop to consider that the class A amplifier always draws the same power from the DC supplies, regardless of the size of the load signal. Without clipping, the average current will be I_{CQ}. That current times the supply voltage yields the supplied power. What's happening is that as the signal increases in amplitude, more and more of the power dissipated by the transistor is shifted to the load. At maximum load swing, both the transistor and the load will be dissipating $P_{DQ}/2$. As strange as it might seem, if you want to keep the output transistor of a class A amplifier cool, don't turn the volume down, turn it up.

The foregoing implies that class A designs are not power efficient. This is indeed the situation. As we have just seen, the best case maximum load power will be one half of P_{DQ}, assuming a centered Q point (non-centered will be worse). To achieve this swing, the power supply will have to be at least twice as large as V_{CEQ} because it has to cover the peak-to-peak swing, while V_{CEQ} represents the peak swing for a centered Q point.[30] In any event, the best case efficiency turns out to be dismal, as follows.

30 This is the case if the AC and DC load lines are identical. This is atypical. Consequently, the power supply will tend to be larger than twice V_{CEQ} which makes the situation even worse.

$$\eta = \frac{P_{out}}{P_{in}} = \frac{P_{load}}{P_{DC}}$$
$$\eta = \frac{P_{DQ}/2}{2V_{CEQ}I_{CQ}}$$
$$\eta = \frac{P_{DQ}/2}{2P_{DQ}}$$
$$\eta = 25\%$$

This represents the maximum or best case efficiency for an *RC* coupled class A amplifier. It may be considerably less depending on precisely how it is biased. This, truly, is the Achilles heel of the class A topology: it is wasteful. It draws full power from the supply regardless if signal is present and, at best, will translate only one quarter of that power into useful load power. At the same time, the power dissipation of the transistor will need to be at least twice that of the delivered load power, and might need to be much greater. Why use it then? To its advantage, it is a relatively simple design so if large output powers are not needed, it can prove useful. This is most definitely the case for the early stages of a multi-stage amplifier where the amount of load power is very small (basically the power delivered to the following stage). In that instance, the increased complexity of more power efficient designs is not warranted or cost effective.

Example 8.1

For the amplifier shown in Figure 8.7, determine the compliance, maximum load power, worst case transistor dissipation and efficiency.

$$I_{CQ} = \frac{|V_{EE}| - V_{BE}}{R_E}$$
$$I_{CQ} = \frac{15\,\text{V} - 0.7\,\text{V}}{120\,\Omega}$$
$$I_{CQ} = 119\,\text{mA}$$

By inspection, V_{CEQ} = 5.7 V. The AC cutoff voltage is

$$v_{CE(cutoff)} = V_{CEQ} + I_{CQ}(r_C + r_E)$$
$$v_{CE(cutoff)} = 5.7\,\text{V} + 119\,\text{mA}(0 + 120\,\Omega \parallel 32\,\Omega)$$
$$v_{CE(cutoff)} = 5.7\,\text{V} + 119\,\text{mA}(25.3\,\Omega)$$
$$v_{CE(cutoff)} = 5.7\,\text{V} + 3\,\text{V}$$
$$v_{CE(cutoff)} = 8.7\,\text{V}$$

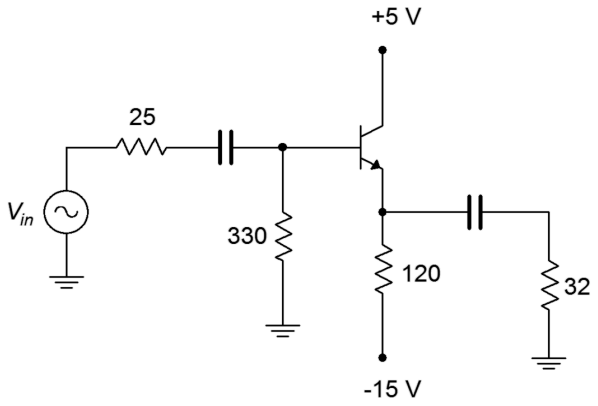


Figure 8.7
Schematic for Example 8.1.

The smaller of V_{CEQ} and $I_{CQ}(r_C + r_E)$ is the peak compliance, so

compliance = 3 V peak

Given the compliance, we can use power law to find the load power

$$P_{load(max)} = \frac{Compliance_{RMS}^{2}}{R_L}$$
$$P_{load(max)} = \frac{(.707 \times 3\,\mathrm{V})^2}{32\,\Omega}$$
$$P_{load(max)} = 141\,\mathrm{mW}$$

This is not a lot of power for something like a loudspeaker but is a fair amount to drive something like a pair of headphones.

The transistor's worst case power dissipation is

$$P_{D(max)} = P_{DQ} = I_{CQ} V_{CEQ}$$
$$P_{D(max)} = 119\,\mathrm{mA} \times 5.7\,\mathrm{V}$$
$$P_{D(max)} = 678\,\mathrm{mW}$$

The supplied circuit power is the average current draw times the total supplied voltage differential

$$P_{DC} = I_{CQ}(V_{CC} - V_{EE})$$
$$P_{DC} = 119\,\mathrm{mA} \times 20\,\mathrm{V}$$
$$P_{DC} = 2.38\,\mathrm{W}$$

The efficiency is the ratio of maximum load power to supplied DC power

$$\eta = \frac{P_{load(max)}}{P_{DC}}$$
$$\eta = \frac{141\,\mathrm{mW}}{2.38\,\mathrm{W}}$$
$$\eta = 5.9\,\%$$

This is much worse than the theoretical best case. This is due, at least in part, to the fact that the Q point is not centered on the AC load line.

To complete the analysis, note that the transistor's breakdown rating (BV_{CEO}) should be at least as large as $v_{CE(cutoff)}$ (8.7 volts), and the maximum current rating should be at least as large as $i_{C(sat)}$ (119 mA+5.7 V/25.3 Ω = 344 mA).

Computer Simulation

A computer simulation of a class A emitter follower using a Darlington pair is examined next. Of primary interest here is the verification of the output compliance so a transient analysis will be used. The simulator schematic is shown in Figure 8.8.

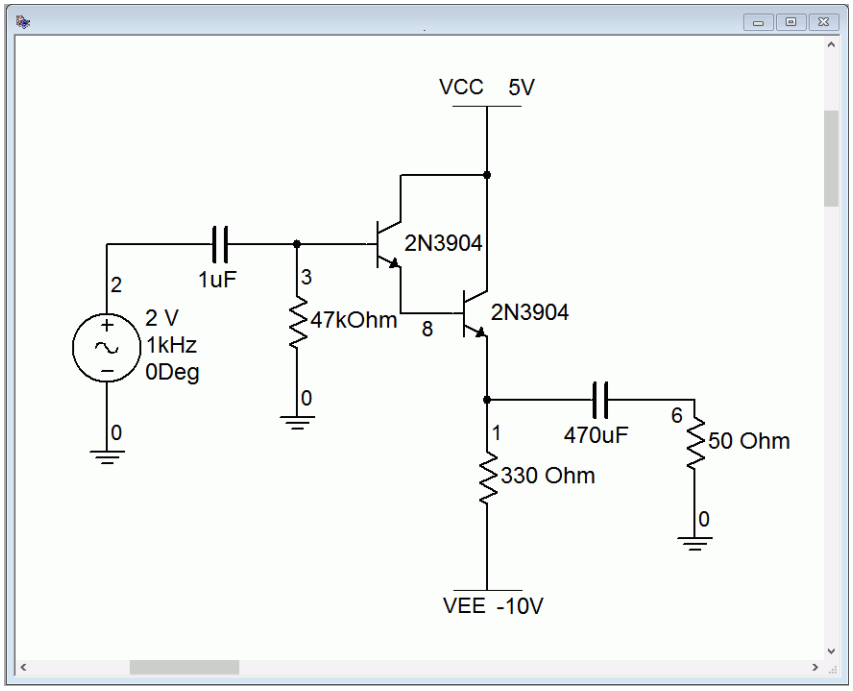


Figure 8.8
Class A follower in simulator.

We can make a few quick computations to determine the compliance. First, we find the collector Q point current.

$$I_{CQ} = \frac{|V_{EE}| - V_{BE}}{R_E}$$
$$I_{CQ} = \frac{10\,\text{V} - 1.4\,\text{V}}{330\,\Omega}$$
$$I_{CQ} = 26\,\text{mA}$$

By inspection, the emitter is two base-emitter junction potentials below ground, or −1.4 V. As the collectors are tied to V_{CC}, this means that $V_{CEQ} = 6.4$ V. The other half of the swing, from V_{CEQ} to $v_{CE(cutoff)}$ is

$$v_{CE(cutoff)} - V_{CEQ} = I_{CQ}(r_C + r_E)$$
$$v_{CE(cutoff)} - V_{CEQ} = 26\,\text{mA}(0 + 330\,\Omega \parallel 50\,\Omega)$$
$$v_{CE(cutoff)} - V_{CEQ} = 26\,\text{mA}(43.4\,\Omega)$$
$$v_{CE(cutoff)} - V_{CEQ} = 1.13\,\text{V}$$

The Q point is not centered and is closer to cutoff. This means that the amplifier will produce cutoff clipping around 1.1 volts and saturation clipping around 6 volts. In

other words, there is more room for the current to swing up to saturation than to swing down to zero. As this is the current flowing through the load and we have a non-inverting follower, we expect to see the load voltage echo this. That is, the negative portion of the load voltage should clip before the positive portion.

The transient analysis results are shown in Figure 8.9. A two volt peak input signal is applied (blue trace). The negative portion of the load voltage clips at approximately 1.1 volts as expected (red trace). The input signal is not large enough to cause saturation clipping. This was done on purpose to verify the voltage gain of the follower. It should be very close to unity. In fact, the trace shows that the gain is around 0.95 or so.

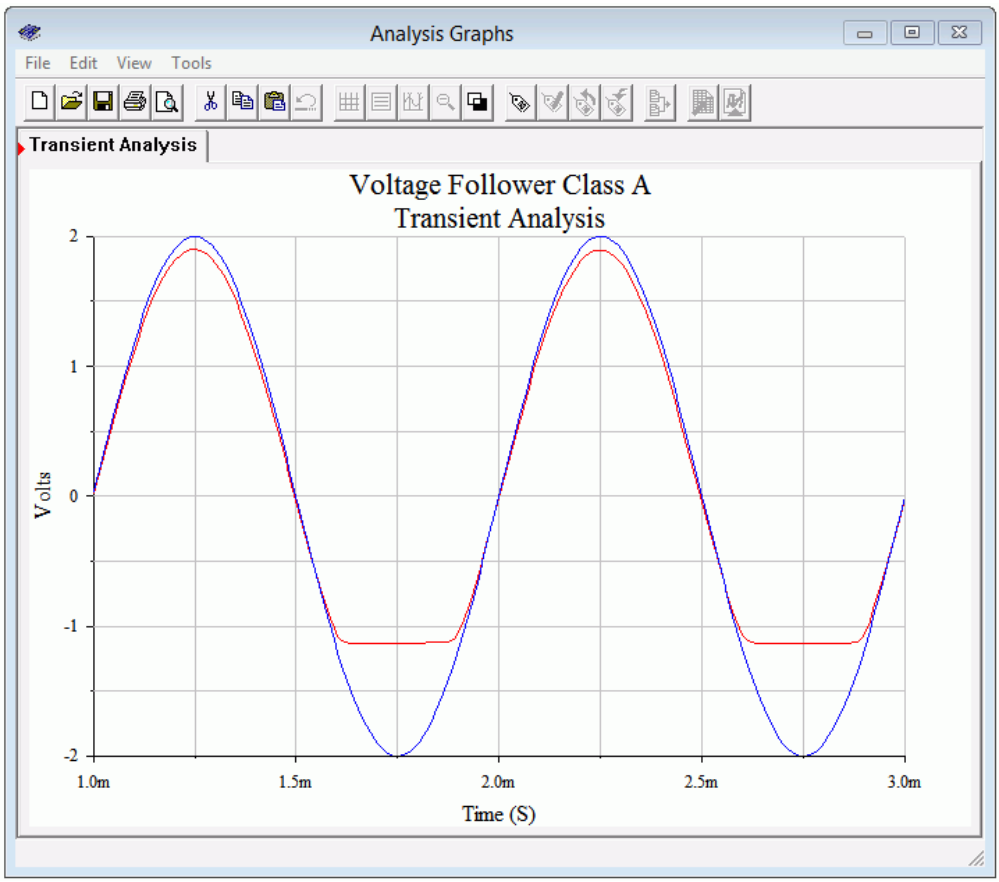


Figure 8.9
Class A follower transient analysis.

If this had been a voltage amplifier instead of a follower, these waveforms would appear flipped vertically. To verify this, the circuit is modified to produce a voltage amplifier with a gain of approximately one. This is achieved by moving the load to the collector and adding a 330 Ω biasing resistor. This will result in the same AC load impedance. To maintain a similar V_{CEQ}, V_{CC} is raised by 10 volts. Finally, the original 330 Ω emitter biasing resistor is split in two: 287 Ω and 43 Ω. This will yield the same I_{CQ} and achieve a voltage gain of unity. As a result, we expect to see clipping at approximately 1.1 volts on the positive portion. The modified circuit is shown in Figure 8.10 and the resulting transient simulation in Figure 8.11.

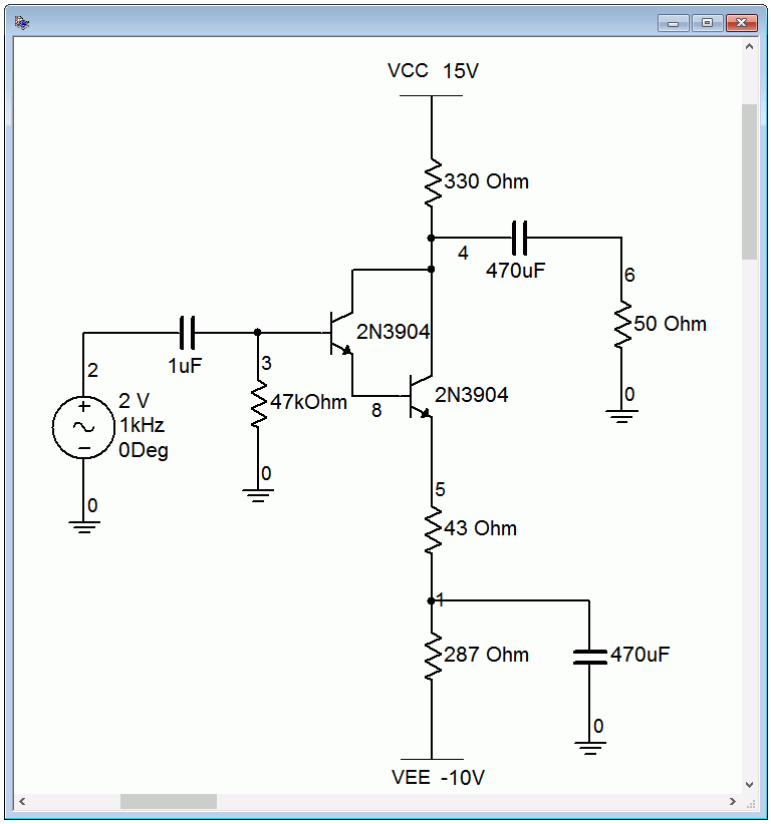


Figure 8.10
Class A amplifier in simulator.

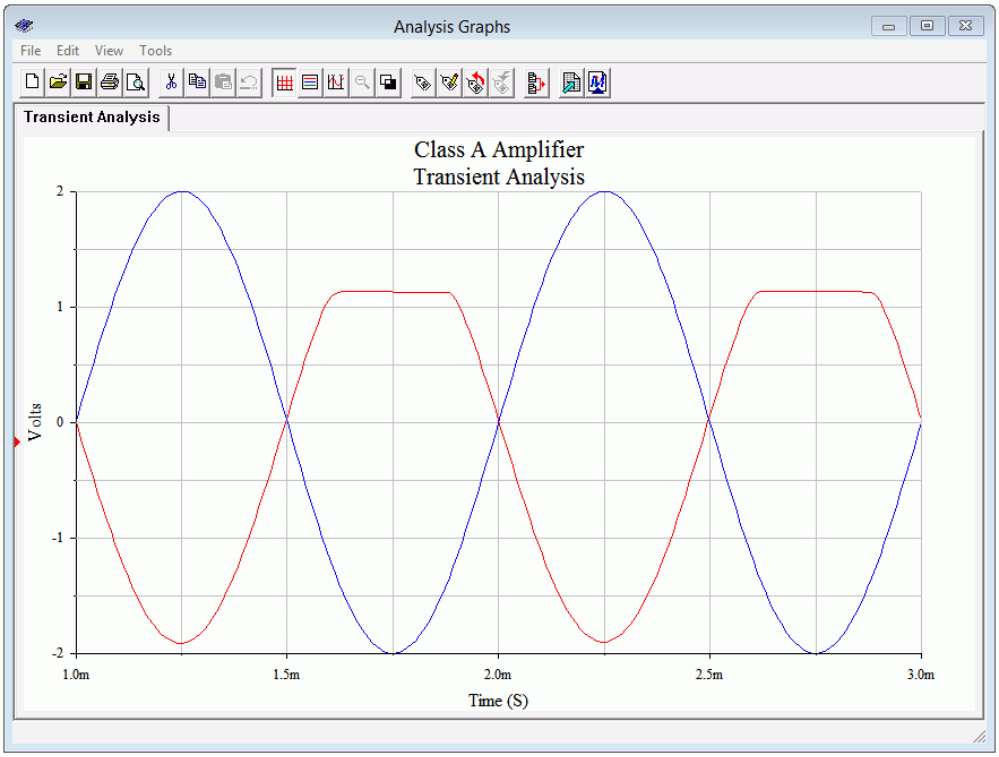


Figure 8.11
Class A amplifier transient analysis.

One final item of interest regarding the simulations: If the input level is increased in an attempt to see clipping on the other half of the waveform, something strange happens. At first it will appear as though it never clips. A careful examination reveals something different, though. Given the values in these circuits, they will exhibit a certain amount of clamping action (clamping was presented in Chapter 3). This will cause the waveform to shift. If you inspect the peak-to-peak value, it will be close to the value of $v_{CE(cutoff)}$. It will be a little less due to the fact that, particularly for a Darlington pair, $V_{CE(sat)}$ is not 0 V.

8.4 Loudspeakers

One of the more common loads for amplifiers is a loudspeaker. It makes sense then to look at how they are constructed and note anything interesting or peculiar as far as their electrical characteristics are concerned. The most common form of loudspeaker is the *dynamic loudspeaker*[31]. All dynamic loudspeakers share certain common elements regardless of size or acoustic output capability. A cutaway view of a low frequency driver is shown in Figure 8.12.

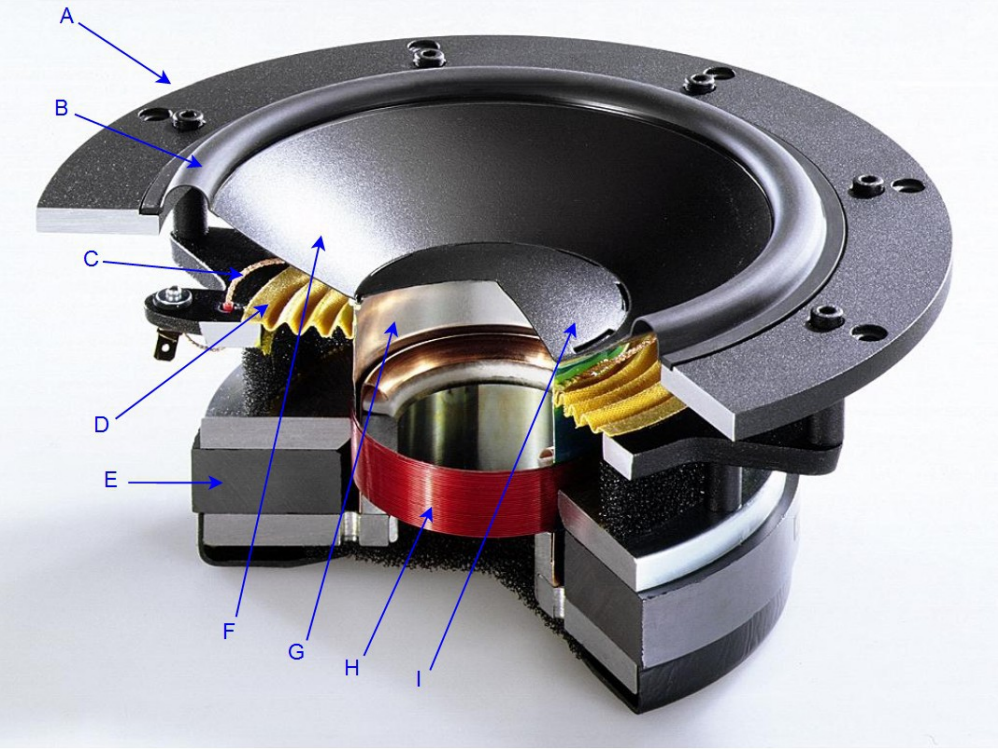


Figure 8.12
Dynamic loudspeaker.
A. Frame
B. Suspension
C. Lead wire
D. Spider
E. Magnet
F. Diaphragm
G. Voice coil former
H. Voice coil
I. Dust cap
Image courtesy of Audio Technology.

The idea behind its operation is magnetic repulsion and attraction. The heart of the unit is the *voice coil* (H). This is a coil of magnet wire wound around a former (G) that typically is made of aluminum or some other high temperature material. The

31 It's a rather odd name given that all loudspeakers are dynamic in some respect. If they weren't they wouldn't produce sound.

voice coil might be a single layer of edge-wound ribbon wire or perhaps several layers of ordinary round wire. Depending on the design, the voice coil might be anywhere from a fraction of an inch to several inches in diameter. The coil ends are connected to flexible lead wires (C) that terminate on the loudspeaker frame (A). Ultimately, that's what the amplifier will connect to.

The voice coil is fixed to a diaphragm (F) and is freely suspended by an outer edge suspension (B) and an inner element known as a *spider* (D). The voice coil sits in a strong magnetic field that is created by a powerful permanent magnet (E) that commonly uses ceramic, alnico or rare earth construction. When current from the amplifier flows through the coil, it will create it's own magnetic field that will either aid or oppose the fixed field created by the permanent magnet, depending on the direction of the current. This results in a force that causes the coil to move within the fixed field. As the coil moves, the diaphragm moves with it, pushing on the surrounding air and creating sound. The larger the current, the stronger the newly created field and the greater the resulting aid or opposition, which results in greater movement of the diaphragm and a larger sound pressure. This fundamental design has changed little since its invention in the 1920s. Modern magnets, suspension and diaphragm materials have improved considerably in the intervening years but the operational principle is pretty much the same.

It is very difficult to create a driver that can cover the full audio spectrum of 20 Hz to 20 kHz while achieving sufficient listening volume at low distortion. Consequently, drivers are often designed to cover a limited portion of the audio spectrum. Low frequency drivers are commonly referred to as *woofers* while high frequency drivers are called *tweeters*. Drivers that cover the middle range of frequencies are given the highly inventive name *midranges* (although once upon a time they were called *squawkers*). A combination of these devices will be wired together with other components to create a complete home or auto loudspeaker system. Although very high quality systems can be produced, virtually all direct radiating dynamic loudspeaker systems suffer from low conversion efficiency. For a typical consumer system, only about 1% to 2% of the applied electrical power is turned into useful acoustic output power. The vast majority of the applied power simply makes the voice coil hot.

Loudspeaker Impedance

Loudspeakers are given a nominal impedance value. The most common impedance for home use is 8 Ω while 4 Ω is common in automotive systems. It is important to remember that this is a nominal value and the true value varies with frequency. While it is common to test power amplifiers with large power resistors, they are only a coarse approximation of a real loudspeaker.

The electrical and mechanical characteristics of a loudspeaker combine to create an equivalent circuit with resistive, inductive and capacitive elements. A typical

electrical circuit model[32] of a single loudspeaker driver is shown in Figure 8.13. R_{VC} and L_{VC} are the resistance and inductance of the voice coil, respectively. The other components are electrical equivalents of mechanical properties such as suspension losses.

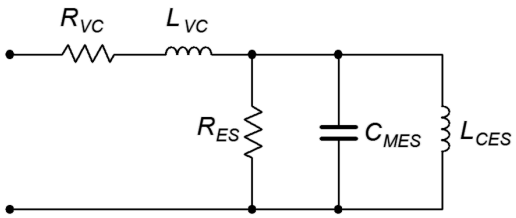


Figure 8.13
Dynamic loudspeaker electrical model.

Clearly, this is not a simple 8 Ω resistor. In fact, we see a very complex impedance: The three parallel elements will create a resonant peak and the series inductance will cause the impedance to rise with frequency. Typically, the resonant peak occurs at the lower end of the spectrum and the associated resonant frequency is denoted on a spec sheet as f_S, the free-air resonance. For a nominal 8 Ω woofer, the peak impedance can be over 30 Ω. An example of a loudspeaker impedance plot is shown in Figure 8.14.

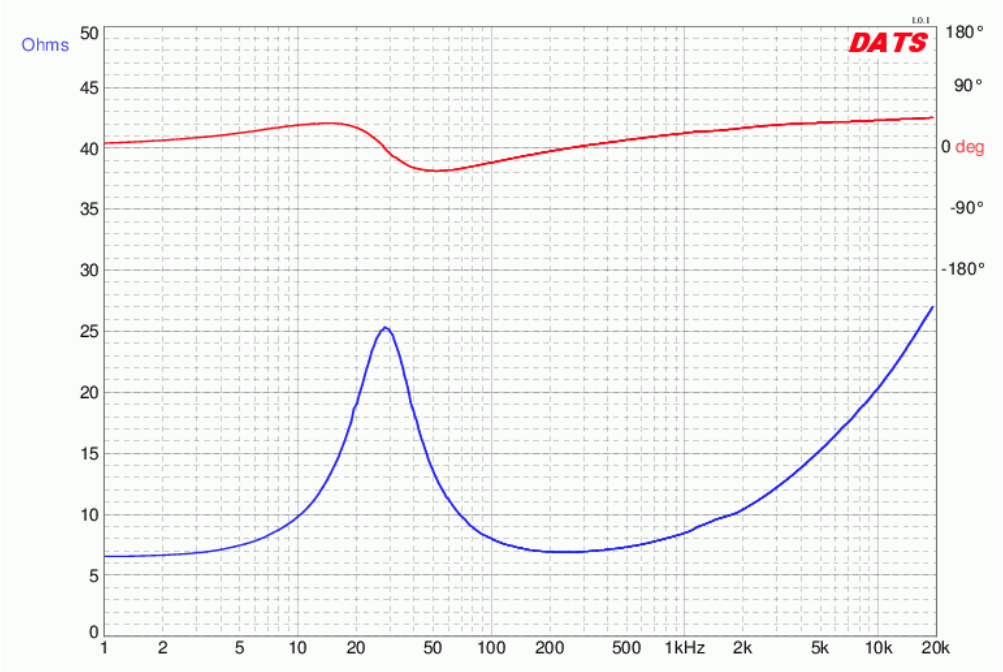


Figure 8.14
Dynamic loudspeaker impedance plot.
Courtesy of Dayton Audio.

This loudspeaker is a nominal 8 Ω unit yet the impedance can be many times this value, and at some frequencies, less than 7 Ω. This plot also includes the phase angle of the loudspeaker and we can see that it can be upwards of 40° capacitive or

32 Adapted from R. H. Small, "Direct Radiator Loudspeaker System Analysis", *Journal of the Audio Engineering Society*, June, 1972.

inductive, depending on the frequency. What makes this more interesting is that a consumer loudspeaker system is a combination of multiple drivers plus other electrical components, and this can result in an even more complex impedance plot.

The obvious question is, “Does this have any effect on the analysis of the power amplifier?” The simple answer is, “Yes”. Areas of the spectrum where the impedance magnitude drops below the nominal value will require more current for any given load voltage. Further, the phase shift caused by a partly reactive load will impact the power dissipation of the transistor.

Consider the power graph shown in Figure 8.6 for a purely resistive load. If we repeat the plot but add a noticeable phase shift to simulate a partly reactive load, something interesting happens, as shown in Figure 8.15.

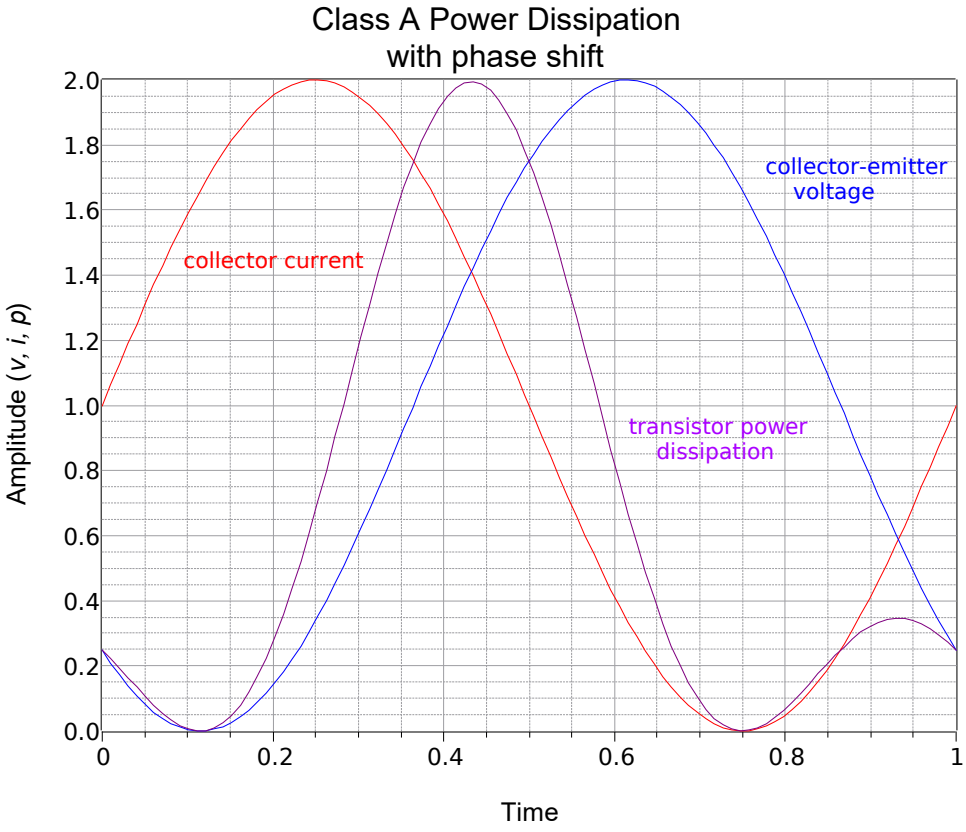


Figure 8.15
Power dissipation with reactive load.

The purple trace represents the power dissipation of the transistor. The peak current-voltage product is twice the value seen with a purely resistive load. This combination might lay outside the safe operating area of the transistor. Ultimately, reactive loads are somewhat more “challenging” than simple resistive loads. Therefore, the transistors may need to be rated higher than the values computed for an idealized resistive load.

Another way of looking at the issue of phase shift induced by loudspeakers and other complex loads is to examine the AC load line. Our previous work with load lines always assumed that the load was purely resistive. What happens in the complex impedance case?

If we examine a generic complex load at a single frequency, our former straight line load line turns into an ellipse, as shown in Figure 8.16.

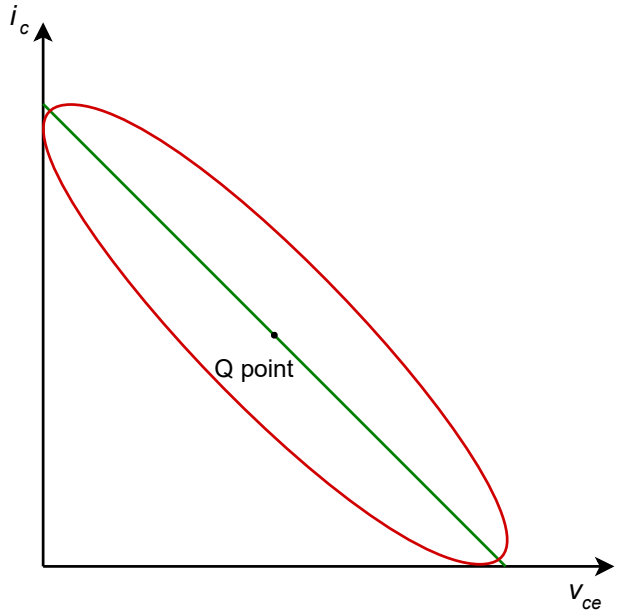


Figure 8.16
AC load line with complex load.

This plot assumes that the circuit has a centered Q point. The normal resistive load line is shown in green. We can visualize the signal starting at zero amplitude, meaning all we see is the Q point. As the signal gets larger and larger, it swings along the green line until, eventually, it maxes out at the two axes. In the complex impedance case, we also start at the Q point. As the signal increases, it traces out an ellipse around the Q point. Further increases would create a larger ellipse, and then a larger ellipse, and so on. Eventually, we would see the maximum swing just touching the axes. This is what is plotted above in red, the maximum case (i.e., at full compliance).

If the impedance angle changes, the aspect ratio of the ellipse changes in reaction to it. The larger the angle, the more open the ellipse becomes. The extremes are 0°, the purely resistive case that yields a collapsed ellipse or straight line; and 90°, the purely reactive case that yields a fully open ellipse, or circle. We have already seen that the phase angle of a loudspeaker changes with frequency, therefore, the load line also changes with frequency. As we sweep the input frequency from low to high, we can imagine the load line wavering back and forth between straight lines and various elliptical shapes. The important thing, though, is that some of these new operating regions (the areas where the red curve is above and to the right of the green line) may go outside the safe operating area of the transistor.

8.5 Power Transistor Data Sheet Interpretation

The data sheet for a popular NPN power transistor, the 2N3055, is shown in Figure 8.17. This model is available from several different manufacturers. Due to the high power dissipation, the TO-92 plastic case that is used for small signal devices is not appropriate. Instead, this device uses the all-metal TO-3 case. Under the maximum ratings we find the device has a maximum power dissipation of 115 W at a case temperature of 25° C, a maximum collector current of 15 A and a maximum collector-emitter voltage of 60 V. Obviously, the device cannot withstand maximum current and voltage simultaneously.

2N3055(NPN), MJ2955(PNP)

Preferred Device

Complementary Silicon Power Transistors

Complementary silicon power transistors are designed for general−purpose switching and amplifier applications.

Features

- DC Current Gain − h_{FE} = 20−70 @ I_C = 4 Adc
- Collector−Emitter Saturation Voltage − $V_{CE(sat)}$ = 1.1 Vdc (Max) @ I_C = 4 Adc
- Excellent Safe Operating Area
- Pb−Free Packages are Available*

MAXIMUM RATINGS

Rating	Symbol	Value	Unit
Collector−Emitter Voltage	V_{CEO}	60	Vdc
Collector−Emitter Voltage	V_{CER}	70	Vdc
Collector−Base Voltage	V_{CB}	100	Vdc
Emitter−Base Voltage	V_{EB}	7	Vdc
Collector Current − Continuous	I_C	15	Adc
Base Current	I_B	7	Adc
Total Power Dissipation @ T_C = 25°C Derate Above 25°C	P_D	115 0.657	W W/°C
Operating and Storage Junction Temperature Range	T_J, T_{stg}	−65 to +200	°C

Maximum ratings are those values beyond which device damage can occur. Maximum ratings applied to the device are individual stress limit values (not normal operating conditions) and are not valid simultaneously. If these limits are exceeded, device functional operation is not implied, damage may occur and reliability may be affected.

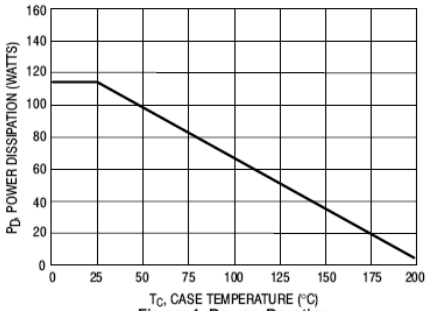


Figure 1. Power Derating

*For additional information on our Pb−Free strategy and soldering details, please download the ON Semiconductor Soldering and Mounting Techniques Reference Manual, SOLDERRM/D.

ON Semiconductor®

http://onsemi.com

15 AMPERE POWER TRANSISTORS COMPLEMENTARY SILICON 60 VOLTS, 115 WATTS

TO−204AA (TO−3)
CASE 1−07
STYLE 1

MARKING DIAGRAM

xxxx55 = Device Code
xxxx = 2N30 or MJ20
G = Pb−Free Package
A = Location Code
YY = Year
WW = Work Week
MEX = Country of Orgin

ORDERING INFORMATION

Device	Package	Shipping
2N3055	TO−204AA	100 Units / Tray
2N3055G	TO−204AA (Pb−Free)	100 Units / Tray
MJ2955	TO−204AA	100 Units / Tray
MJ2955G	TO−204AA (Pb−Free)	100 Units / Tray

Preferred devices are recommended choices for future use and best overall value.

Figure 8.17a
2N3055 data sheet.
Used with permission from SCILLC dba ON Semiconductor.

In the drawing of the TO-3 case, only two leads are shown. These are for the emitter and base. The entire body of the device is the collector. This is because the device will most likely be attached to a metal *heat sink* (see next section) to help dissipate the heat generated. The greater the contact area, the more effective the heat flow will be. The curves presented in Figure 8.17b indicate that β is considerably lower than what we saw for small signal devices. Further, $I_{C(sat)}$ tends to be larger for higher power transistors. For very high currents, β might fall to less than 20 while $I_{C(sat)}$ can be upwards of half of a volt.

Figure 8.17b
2N3055 data sheet (cont).

2N3055(NPN), MJ2955(PNP)

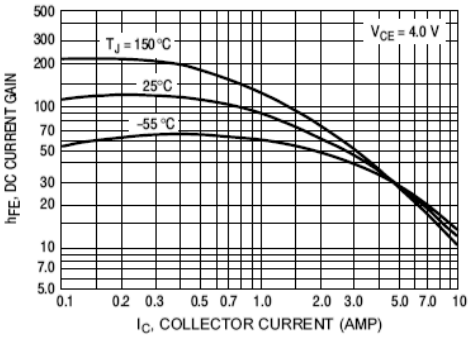


Figure 3. DC Current Gain, 2N3055 (NPN)

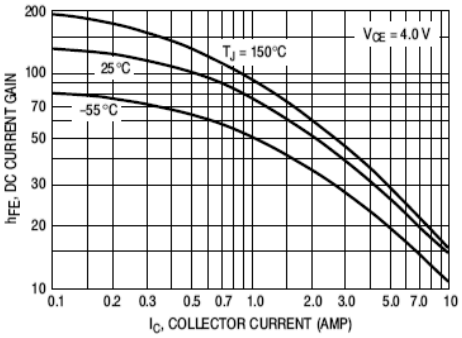


Figure 4. DC Current Gain, MJ2955 (PNP)

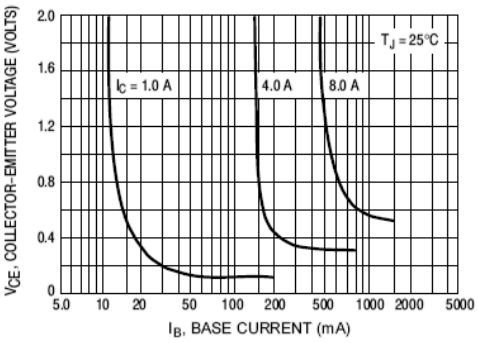


Figure 5. Collector Saturation Region, 2N3055 (NPN)

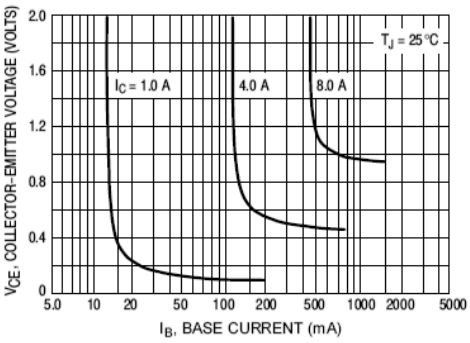


Figure 6. Collector Saturation Region, MJ2955 (PNP)

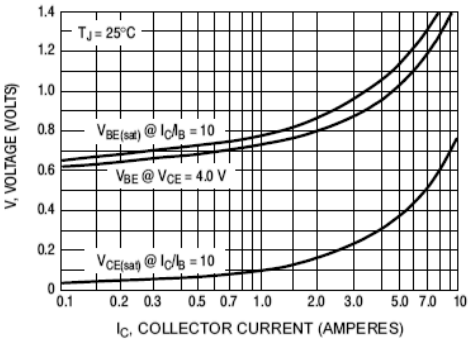


Figure 7. "On" Voltages, 2N3055 (NPN)

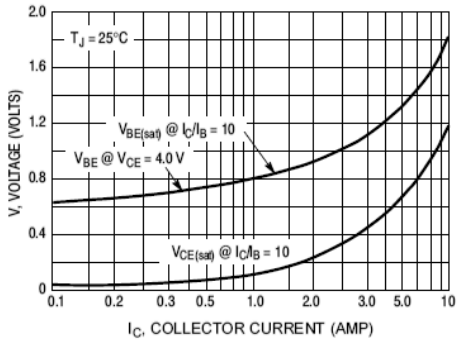


Figure 8. "On" Voltages, MJ2955 (PNP)

Figure 8.17c
2N3055 data sheet (cont).

THERMAL CHARACTERISTICS

Characteristic	Symbol	Max	Unit
Thermal Resistance, Junction−to−Case	$R_{\theta JC}$	1.52	°C/W

ELECTRICAL CHARACTERISTICS (T_C = 25°C unless otherwise noted)

Characteristic	Symbol	Min	Max	Unit
OFF CHARACTERISTICS*				
Collector−Emitter Sustaining Voltage (Note 1) (I_C = 200 mAdc, I_B = 0)	$V_{CEO(sus)}$	60	–	Vdc
Collector−Emitter Sustaining Voltage (Note 1) (I_C = 200 mAdc, R_{BE} = 100 Ω)	$V_{CER(sus)}$	70	–	Vdc
Collector Cutoff Current (V_{CE} = 30 Vdc, I_B = 0)	I_{CEO}	–	0.7	mAdc
Collector Cutoff Current (V_{CE} = 100 Vdc, $V_{BE(off)}$ = 1.5 Vdc) (V_{CE} = 100 Vdc, $V_{BE(off)}$ = 1.5 Vdc, T_C = 150°C)	I_{CEX}	 – –	 1.0 5.0	mAdc
Emitter Cutoff Current (V_{BE} = 7.0 Vdc, I_C = 0)	I_{EBO}	–	5.0	mAdc
ON CHARACTERISTICS* (Note 1)				
DC Current Gain (I_C = 4.0 Adc, V_{CE} = 4.0 Vdc) (I_C = 10 Adc, V_{CE} = 4.0 Vdc)	h_{FE}	 20 5.0	 70 –	–
Collector−Emitter Saturation Voltage (I_C = 4.0 Adc, I_B = 400 mAdc) (I_C = 10 Adc, I_B = 3.3 Adc)	$V_{CE(sat)}$	 –	 1.1 3.0	Vdc
Base−Emitter On Voltage (I_C = 4.0 Adc, V_{CE} = 4.0 Vdc)	$V_{BE(on)}$	–	1.5	Vdc
SECOND BREAKDOWN				
Second Breakdown Collector Current with Base Forward Biased (V_{CE} = 40 Vdc, t = 1.0 s, Nonrepetitive)	$I_{s/b}$	2.87	–	Adc
DYNAMIC CHARACTERISTICS				
Current Gain − Bandwidth Product (I_C = 0.5 Adc, V_{CE} = 10 Vdc, f = 1.0 MHz)	f_T	2.5	–	MHz
*Small−Signal Current Gain (I_C = 1.0 Adc, V_{CE} = 4.0 Vdc, f = 1.0 kHz)	h_{fe}	15	120	–
*Small−Signal Current Gain Cutoff Frequency (V_{CE} = 4.0 Vdc, I_C = 1.0 Adc, f = 1.0 kHz)	f_{hfe}	10	–	kHz

*Indicates Within JEDEC Registration. (2N3055)
1. Pulse Test: Pulse Width ≤ 300 μs, Duty Cycle ≤ 2.0%.

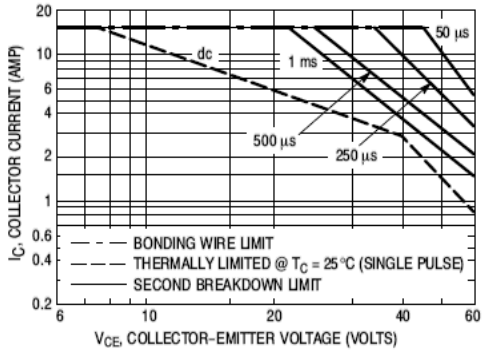


Figure 2. Active Region Safe Operating Area

There are two limitations on the power handling ability of a transistor: average junction temperature and second breakdown. Safe operating area curves indicate $I_C - V_{CE}$ limits of the transistor that must be observed for reliable operation; i.e., the transistor must not be subjected to greater dissipation than the curves indicate.

The data of Figure 2 is based on T_C = 25°C; $T_{J(pk)}$ is variable depending on power level. Second breakdown pulse limits are valid for duty cycles to 10% but must be derated for temperature according to Figure 1.

One item of note in Figure 8.17c is the small graphic at the bottom of the sheet. This is a plot of safe operating area. Basically, the combination of V_{CE} and I_C must fall within the lower-left zone. What is of particular interest is that the safe zone extends out further if the current/voltage combination is the result of a short pulse rather than a continuous condition.

Power Derating

One final item of concern is the graph found in Figure 8.17a, and magnified in Figure 8.18. This is a power derating curve.

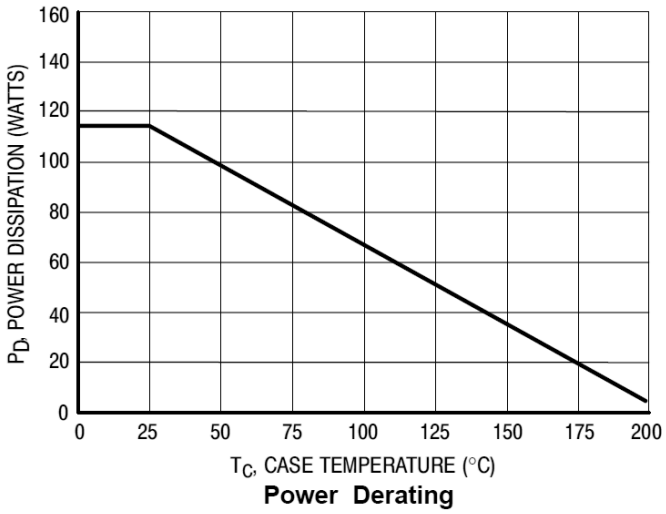


Figure 8.18
Power derating curve for 2N3055.
Used with permission from SCILLC dba ON Semiconductor.

Although the device is rated for 115 watts, that is only true at case temperatures of 25°C or lower. At higher temperatures, the power dissipation capability decreases. For example, at 100°C this device can only dissipate about 65 watts. A precise value can be computed via the following formula:

$$P_D = P_{25} - D(T_{case} - 25°C) \qquad (8.8)$$

Where
P_D is the power dissipation at the new case temperature,
P_{25} is the power dissipation at 25°C,
D is the derating factor (units of W/C°),
T_{case} is the new case temperature.

Example 8.2

Determine the power dissipation of a 2N3055 at 75°C.

Using the graph, this is a little over 80 watts.

For a more accurate result, we'll use Equation 8.8. From the data sheet the dissipation at 25°C, P_{25} , is 115 watts. The derating factor, D, is 0.657 W/C° (the derating factor is found directly above the graph of power derating in Figure 8.17a).

$$P_D = P_{25} - D\left(T_{ambient} - 25\,°\text{C}\right)$$
$$P_D = 115\,\text{W} - 0.657\,\text{W/C°}\left(75\,°C - 25\,°\text{C}\right)$$
$$P_D = 82.1\,\text{W}$$

8.6 Heat Sinks

The issue with power transistors is always heat. As noted in Example 8.2, as the transistor heats up due to internal power dissipation, its ability to dissipate heat is compromised. The trick, then, is to efficiently move the heat from the transistor to someplace else. This is normally achieved through the use of a *heat sink*.

A heat sink is a metal device that is attached to the power transistor. Typically, they are made of aluminum and feature an array of fins. By increasing the surface area, heat can be moved away from the transistor more efficiently than by the transistor alone.

Heat sinks are designed to mount specific device case styles. The most common case styles include the TO-3 "can" along with the various "power tab" styles such as the TO-220 and TO-202. Special mounting hardware and insulation spacers are also required in order to maintain electrical isolation between the transistor and the heat sink as we do not want the heat sink to be electrically live. This usually takes the form of a mica sheet, and plastic washers and bushings for the mounting machine screws (for small heat sinks, sometimes nylon machine screws are used).

There are a few general rules that should be followed when using heat sinks:

- Always use some form of heat sink grease or thermally conductive pad between the heat sink and the device. This will increase the thermal transfer between the two parts, however, *excessive* quantities of heat sink grease will *decrease* performance.
- Mount fins in the vertical plane for optimum natural convective cooling.
- Do not overcrowd or obstruct devices that use heat sinks.
- Do not block air flow around heat sinks – particularly directly above and below items that rely on natural convection.
- If thermal demands are particularly high, consider using forced convection (i.e., a small fan directed at the heat sink).

Some typical heat sinks are shown below. Figure 8.19 shows a heat sink and thermal data plot for use with a single TO-3 case device.

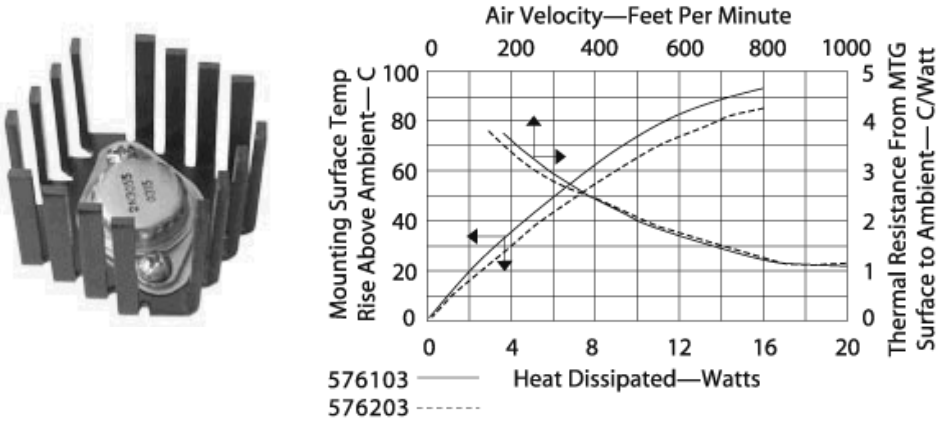


Figure 8.19
Heat sink for TO-3.
Reprinted courtesy of Aavid Thermalloy, Inc.

Figure 8.20 shows a heat sink designed for a pair of transistors using TO-220 cases. In this photo, the white insulating pads can be seen between the transistors and heat sink.

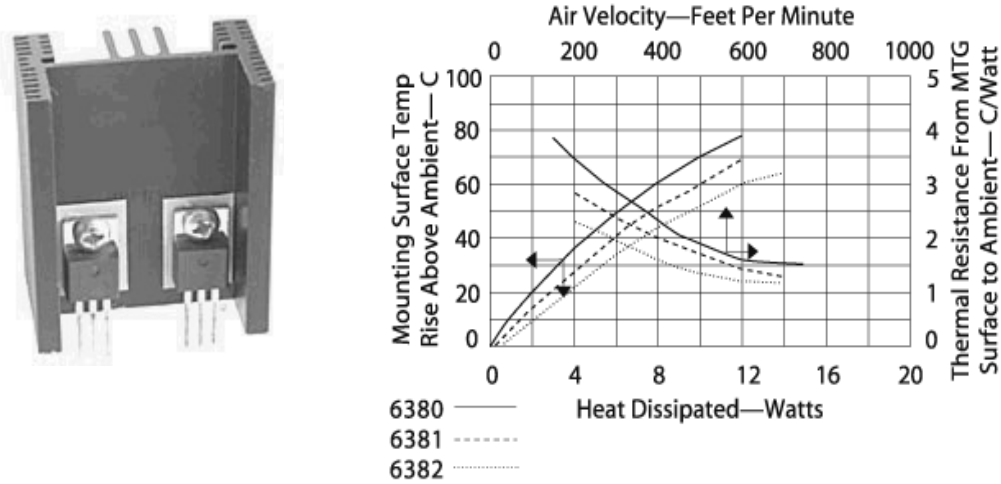


Figure 8.20
Heat sink for dual TO-220.
Reprinted courtesy of Aavid Thermalloy, Inc.

Thermal Resistance

In order to specify a particular heat sink for a given application, a more technical explanation is in order. What we are going to do is create a *thermal circuit equivalent*. In this model, the concept of *thermal resistance* is used. Thermal resistance denotes how easy it is to transfer heat energy from one mechanical part to another. The symbol for thermal resistance is θ, and the units are Celsius degrees per

watt. In this model, temperature is analogous to voltage, and thermal power dissipation is analogous to current. A useful equation is,

$$P_D = \frac{\Delta T}{\theta_{total}} \tag{8.9}$$

Where P_D is the power dissipated by the semiconductor device in watts, ΔT is the temperature differential, and θ_{total} is the sum of the thermal resistances. Basically, this is a thermal version of Ohm's law.

In order to construct our model, let's take a closer look at the power-device/heat-sink combination. This is shown in Figure 8.21. The subscript *j* stands for junction, *c* is for case (of the transistor), *s* is for heat sink and *a* is for the ambient air. T_j is the semiconductor junction temperature and is created by the product of the transistor's current and voltage. This thermal source heats the device case to T_c. The thermal resistance between the two entities is θ_{jc}. The case, in turn, heats the heat sink via the interconnection. This thermal resistance is θ_{cs}, and the resulting temperature is T_s. Finally, the heat sink passes the thermal energy to the surrounding air which is sitting at T_a. The thermal resistance of the heat sink to the air is θ_{sa}. The equivalent thermal model is shown in Figure 8.22. Although this "thermal circuit" does not have perfect correspondence with normal circuit analysis, it does illustrate the main points.

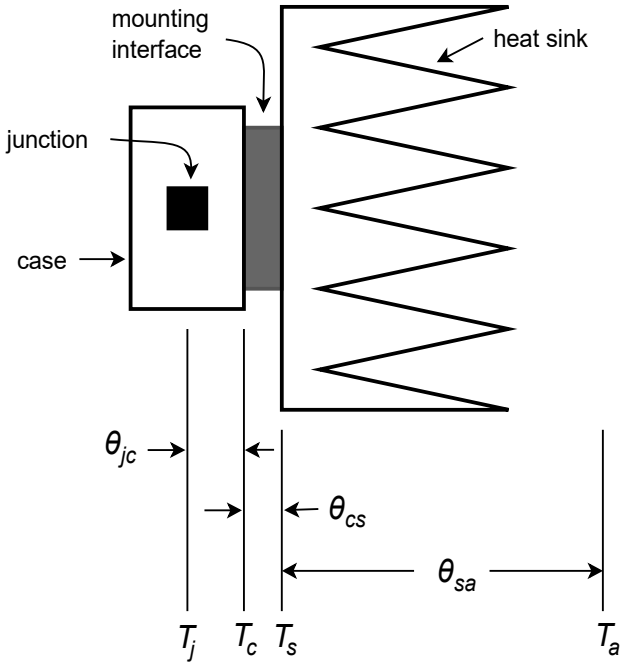


Figure 8.21
Device and heat sink.

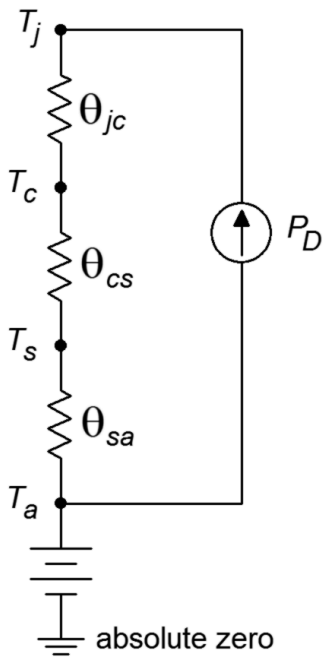


Figure 8.22
Equivalent thermal model of Figure 8.21.

In this model, ground represents a temperature of absolute zero. The circuit is sitting at an ambient temperature T_a, thus a voltage source of T_a is connected to ground and the heat sink. The three thermal resistances are in series and are driven by a current source that is set by the present power dissipation of the device. Note that if the power dissipation is high, the resulting "voltage drops" across the thermal resistances are high. Voltage is analogous to temperature in this model, so this indicates that a high temperature is created. Because there is a maximum limit to T_j, a higher power dissipation requires lower thermal resistances. As θ_{jc} is set by the device manufacturer, we have no control over that element. However, θ_{cs} is a function of the case style and the insulation material used, so we do have some control (but not a lot) over that. On the other hand, as the person who specifies the heat sink, we have a great deal of control over θ_{sa}. Values for θ_{sa} are given by heat sink manufacturers. A useful variation of Equation 8.9 is

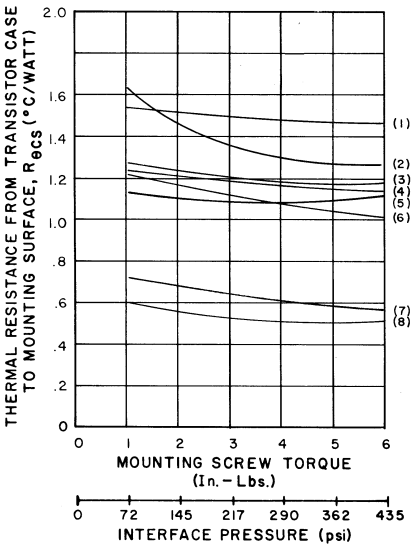
$$P_D = \frac{T_j - T_a}{\theta_{jc} + \theta_{cs} + \theta_{sa}} \tag{8.10}$$

Normally, power dissipation, junction and ambient temperatures, θ_{jc} and θ_{cs} are known. The idea is to determine an appropriate heat sink. Both T_j and θ_{jc} are given by the semiconductor device manufacturer. The ambient temperature, T_a, may be determined experimentally. Due to localized warming, it tends to be higher than the actual "room temperature". Standard graphs, such as those found in Figure 8.23, may be used to determine θ_{cs}. Note the generally lower values of θ_{cs} for the TO-3 case relative to the TO-220. This is one reason why TO-3 cases are used for higher power devices. This case also makes it easier for the manufacturer to reduce θ_{jc}.

Figure 8.23
θ_{CS} for TO-3 and TO-220
Reprinted courtesy of Thermalloy, Inc.

WITHOUT THERMAL GREASE

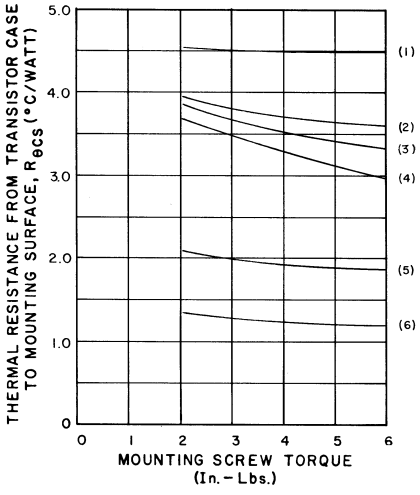
JEDEC TO-3 Interface Thermal Resistance versus Mounting Screw Torque for a TO-3 Semiconductor Device using Various Insulating Materials. No Thermal Joint Compound Used in the Interface Area.



LEGEND:

(1) THERMALFILM, .002(.05) THK.
(2) MICA, .003(.08) THK.
(3) MICA, .002(.05) THK.
(4) HARD ANODIZED, .020(.51) THK.
(5) THERMALSIL, .008(.20) THK.
(6) ALUMINUM OXIDE, .062(1.57) THK.
(7) BERYLLIUM OXIDE, .062(1.57) THK.
(8) BARE JOINT - NO FINISH

JEDEC TO-220 Interface Thermal Resistance versus Mounting Screw Torque for a TO-220 Semiconductor Device using Various Insulating Materials. No Thermal Joint Compound Used in the Interface Area.

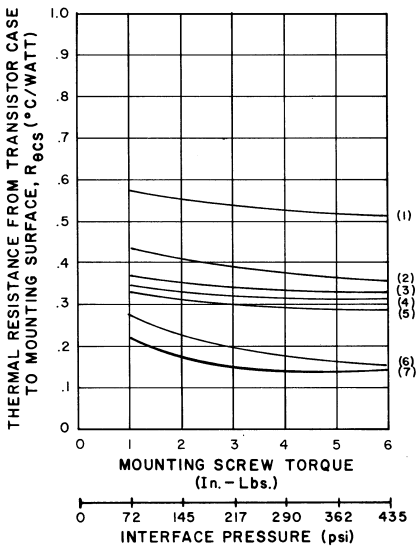


LEGEND:

(1) THERMALFILM, .002(.05) THK.
(2) MICA, .003(.08) THK.
(3) MICA, .002(.05) THK.
(4) HARD ANODIZED, .020(.51) THK.
(5) THERMALSIL, .008(.20) THK.
(6) BARE JOINT - NO FINISH

WITH THERMAL GREASE

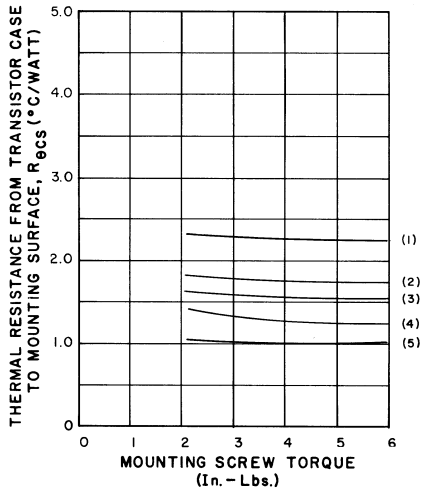
JEDEC TO-3 Interface Thermal Resistance versus Mounting Screw Torque for a TO-3 Semiconductor Device Using Various Insulating Materials. Thermalcote Thermal Joint Compound Used in the Interface Area.



LEGEND:

(1) THERMALFILM, .002(.05) THK.
(2) MICA, .003(.08) THK.
(3) MICA, .002(.05) THK.
(4) ALUMINUM OXIDE, .062(1.57) THK.
(5) HARD ANODIZED, .020(.51) THK.
(6) BERYLLIUM OXIDE, .062(1.57) THK.
(7) BARE JOINT - NO FINISH

JEDEC TO-220 Interface Thermal Resistance versus Mounting Screw Torque for a TO-220 Semiconductor Device using Various Insulating Materials. Thermalcote Thermal Joint Compound Used in the Interface Area.



LEGEND:

(1) THERMALFILM, .002(.05) THK.
(2) MICA, .003(.08) THK.
(3) MICA, .002(.05) THK.
(4) HARD ANODIZED, .020(.51) THK.
(5) BARE JOINT - NO FINISH

Example 8.3

Determine the appropriate heat sink rating for a power device rated as follows: $T_{j(max)}$ = 175°C, TO-3 case style, θ_{jc} = 1.5 C°/W. The device will be dissipating a maximum of 15 W in an ambient temperature of 40°C. Assume that the heat sink will be mounted with heat sink grease and a 0.002 mica insulator.

First, find θ_{cs} from the TO-3 "With Thermal Grease" graph in Figure 8.23. Curve 3 is used. The approximate value is 0.35 C°/W.

$$P_D = \frac{T_j - T_a}{\theta_{jc} + \theta_{cs} + \theta_{sa}}$$

$$\theta_{sa} = \frac{T_j - T_a}{P_D} - \theta_{jc} - \theta_{cs}$$

$$\theta_{sa} = \frac{175\,°\text{C} - 40\,°\text{C}}{15\,\text{W}} - 1.5\,\text{C°/W} - 0.35\,\text{C°/W}$$

$$\theta_{sa} = 7.15\,\text{C°/W}$$

This is the maximum acceptable value for the heat sink's thermal resistance. Note that the use of heat sink grease gives us an extra 0.8 C°/W or so. For this application, the heat sink pictured in Figure 8.19 will most likely be sufficient without added forced air cooling (the graph stops at less than 4 C°/W with an air flow of under 200 feet/minute).

If we repeat this problem with a much higher P_D, things work out a little differently. Let's use 40 W this time.

$$P_D = \frac{T_j - T_a}{\theta_{jc} + \theta_{cs} + \theta_{sa}}$$

$$\theta_{sa} = \frac{T_j - T_a}{P_D} - \theta_{jc} - \theta_{cs}$$

$$\theta_{sa} = \frac{175\,°\text{C} - 40\,°\text{C}}{40\,\text{W}} - 1.5\,\text{C°/W} - 0.35\,\text{C°/W}$$

$$\theta_{sa} = 1.53\,\text{C°/W}$$

If we hope to use that same heat sink, we will have to add forced air cooling of at least 700 feet/minute. The other option would be to find a more thermally efficient (and probably much larger) heat sink if we hope to use natural convection alone.

Summary

Class A operation is defined as having collector current flow for 360° of the cycle. This means that a single output device can be used to amplify the entire input waveform. To determine the maximum signal swing, or compliance, an AC load line is used. This is similar to a DC load line and plots all possible transistor voltage and current coordinate pairs. The efficiency of the class A amplifier tends to be low. The maximum theoretical efficiency is only 25%. Further, the amplifier draws full current from the power supply regardless of whether or not a signal is present. As a consequence, the transistor runs hottest when there is no signal. With an applied signal, some of the power formerly dissipated within the transistor is shifted to the load.

Loudspeakers offer a complex impedance as a load. As such, they are more challenging than simple resistive loads and may require the output transistor to be rated for a higher-than-normal power dissipation.

Heat sinks are used to efficiently move heat from the transistor's internal structure to the surrounding air. The thermal effectiveness of a heat sink is measured by thermal resistance, θ. The lower the value of θ, the more effective the heat sink is at transferring heat from the transistor to the surrounding air. For high power applications where a good deal of heat is generated, heat sinks are augmented with forced air cooling.

Review Questions

1. Define class A operation.
2. Why are voltage followers generally preferred over voltage amplifiers for power output applications?
3. How does an AC load line differ from a DC load line?
4. What is the advantage of having a centered Q point on the AC load line?
5. What effect does a reactive load have on an AC load line?
6. Describe the operation of a dynamic loudspeaker.
7. What are heat sinks? What are they used?
8. What is thermal resistance?

Problems

Analysis Problems

1. Draw the AC load line for the circuit of Figure 8.24. Also determine the compliance, maximum load power, maximum transistor dissipation and efficiency. V_{CC} = 6 V, V_{EE} = −12 V, R_{gen} = 50 Ω, R_B = 2.2 kΩ, R_E = 470 Ω , R_L = 75 Ω.

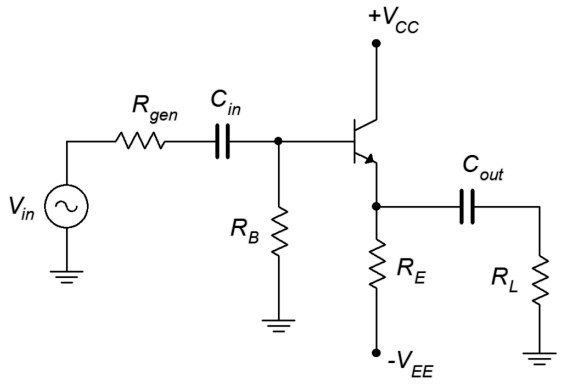


Figure 8.24

2. Recalculate Problem 1 if the load is halved.

3. Determine if the circuit of Figure 8.25 has a centered Q point on its AC load line. $V_{CC} = -10$ V, $V_{EE} = 15$ V, $R_B = 1$ kΩ, $R_E = 330$ Ω, $R_L = 50$ Ω.

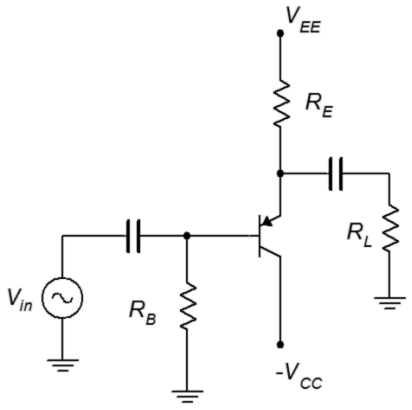


Figure 8.25

4. Draw the AC load line for the circuit of Figure 8.25. Also determine the compliance, maximum load power, maximum transistor dissipation and efficiency. $V_{CC} = -8$ V, $V_{EE} = 12$ V, $R_B = 1$ kΩ, $R_E = 330$ Ω, $R_L = 32$ Ω.

5. Draw the AC load line for the circuit of Figure 8.26. Also determine the compliance, maximum load power, maximum transistor dissipation and efficiency. $V_{CC} = 15$ V, $V_{EE} = -20$ V, $R_B = 10$ kΩ, $R_E = 100$ Ω, $R_L = 16$ Ω.

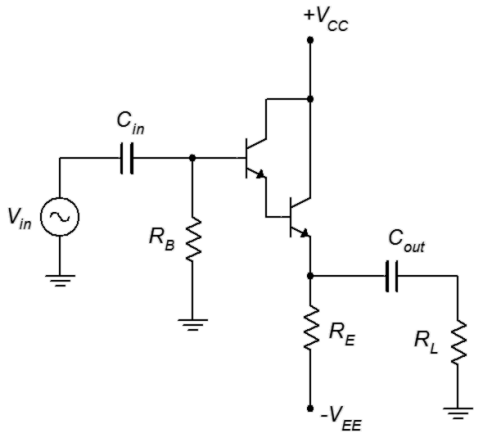


Figure 8.26

6. Determine if the circuit of Figure 8.27 has a centered Q point on its AC load line. V_{CC} = 30 V, R_1 = 3.9 kΩ, R_2 = 3.3 kΩ, R_E = 560 Ω, R_L = 50 Ω.

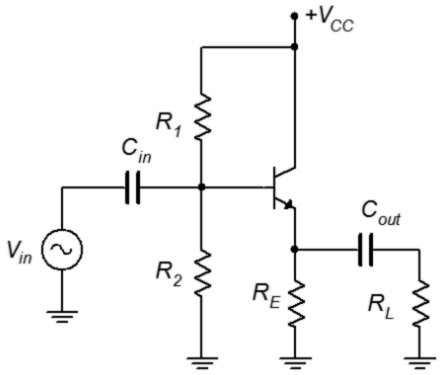
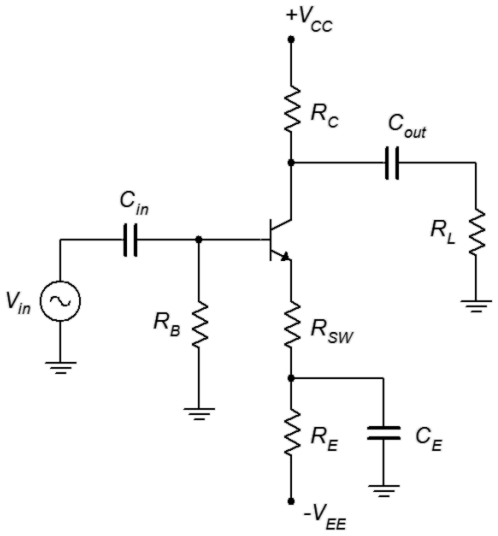


Figure 8.27

7. Draw the AC load line for the circuit of Figure 8.27. Also determine the compliance, maximum load power, maximum transistor dissipation and efficiency. V_{CC} = 30 V, R_1 = 2.2 kΩ, R_2 = 2.2 kΩ, R_E = 470 Ω, R_L = 32 Ω.

8. Determine if the circuit of Figure 8.28 has a centered Q point on its AC load line. V_{CC} = 15 V, V_{EE} = −15 V, R_B = 1 kΩ, R_E = 510 Ω, R_{SW} = 10 Ω, R_C = 270 Ω, R_L = 50 Ω.

Figure 8.28



9. Draw the AC load line for the circuit of Figure 8.28. Also determine the compliance, maximum load power, maximum transistor dissipation and efficiency. V_{CC} = 25 V, V_{EE} = −15 V, R_B = 1 kΩ, R_E = 270 Ω, R_{SW} = 6.8 Ω, R_C = 330 Ω, R_L = 16 Ω.

10. A power transistor has a $P_{D(max)}$ of 50 watts at 25°C. It has a derating factor of 0.4 W/C°. Will this transistor be sufficient for a circuit that needs to dissipate 40 watts at 85°C?

11. A power transistor has a $P_{D(max)}$ of 100 watts at 25°C. It has a derating factor of 0.6 W/C°. Will this transistor be sufficient for a circuit that needs to dissipate 65 watts at 75°C?

12. Determine the appropriate heat sink rating for a power device rated as follows: $T_{j(max)}$ = 175°C, TO-3 case style, θ_{jc} = 1.5 C°/W. The device will be dissipating a maximum of 25 W in an ambient temperature of 35°C. Assume that the heat sink will be mounted with heat sink grease and a 0.003 mica insulator.

13. Determine the appropriate heat sink rating for a power device rated as follows: $T_{j(max)}$ = 165°C, TO-220 case style, θ_{jc} = 3 C°/W. The device will be dissipating a maximum of 15 W in an ambient temperature of 35°C. Assume that the heat sink will be mounted with heat sink grease and a 0.002 mica insulator.

Design Problems

14. Alter the emitter power supply in the circuit described in Problem 1 to achieve a centered Q point.

15. Alter the emitter power supply in the circuit described in Problem 4 to achieve a centered Q point.

Challenge Problems

16. Find a heat sink (make and model number) that will meet the thermal resistance requirement for Problem 12 with no more than 400 feet/minute of forced air.

17. Alter the voltage divider in the circuit described in Problem 6 to achieve a centered Q point.

Computer Simulation Problems

18. Perform a transient analysis for the circuit described in Problem 1 to verify the compliance.

19. Perform a transient analysis for the circuit described in Problem 4 to verify the compliance.

20. Perform a transient analysis for the circuit described in Problem 9 to verify the compliance.

9 BJT Class B Power Amplifiers

9.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Define class B operation.
- Determine AC load lines for class B amplifier stages.
- Determine the compliance, maximum load power, efficiency and required device ratings for class B circuits.
- Discuss the advantages and disadvantages of class B operation versus class A operation.
- Discuss the origin of *notch distortion* and methods used to mitigate it.
- Explain the operation of a *current mirror*.
- Explain the operation of a *Sziklai pair*.
- Explain the operation and use of a V_{BE} *multiplier*.
- Outline the operation of *fully complimentary* and *quasi complimentary* output stages utilizing direct coupled driver stages.
- Discuss methods to protect the output devices from overload.

9.1 Introduction

Class B amplifiers have long been a mainstay of linear amplifier design. Compared to class A operation, class B amplifiers offer much greater power efficiency. That is, a much larger percentage of the applied DC power can be turned into useful AC output to the load. This also means that the power dissipation requirements for the transistors are lowered. Further, unlike class A operation, class B designs do not continuously draw full power from their DC supplies. Instead, they draw current as it is needed, and therefore run relatively cool at idle and at low output power.

The downside for this improved efficiency is added complexity. For starters, two transistors are required for linear class B operation. Also, the biasing can be a little trickier which requires modifications to the relatively straightforward class A biasing circuits we have already examined. Also, class B amplifiers suffer from a unique form of distortion that class A amplifiers do not.

In this chapter we shall also examine the use auxiliary device configurations and sub-circuits to improve performance. These include the *current mirror*, *Sziklai pair*, V_{BE} *multiplier*, and overload protection and prevention circuitry. As with the class A amplifier, the common collector or voltage follower configuration tends to be the most widely used configuration for class B circuits, hence, we shall focus on followers and not on voltage amplifiers.

9.2 The Class B Configuration

Class B operation is defined as having AC collector current flow 180° out of the cycle. Consequently, in order to amplify the entire signal, two devices will be needed. Further, we will need to pay attention to how the two waveform halves are "stitched together" as this could be a problem area. The obvious question at this point is, why do we bother separating the positive and negative half-waves if it leads to circuit complexity and possible waveform issues? The answer is improved efficiency.

In the previous chapter we discovered that class A amplifiers are not efficient. In fact, at best they only transform 25% of the DC input power into useful load power. Why does this occur and how does the class B topology address this situation?

With a class A design, we determined that the optimal location for the Q point was midway along the AC load line. In this way, the signal could swing up toward saturation and down toward cutoff equally. Unfortunately, this also means that if no signal is present, then the transistor is still pulling half of the maximum current. It would make more sense if the transistor idled at zero current (or some small nominal value) than idling at half of the maximum.[33]

The basic idea of class B is to push the Q point down so that it is sitting right at cutoff on the AC load line. This means that I_{CQ} is 0 A and virtually no power is drawn from the supply at idle. Locating the Q point at cutoff also means that the transistor will immediately clip the negative portion of the wave. Consequently, we will need a mirror image circuit to produce that portion (and which will clip the positive portion).

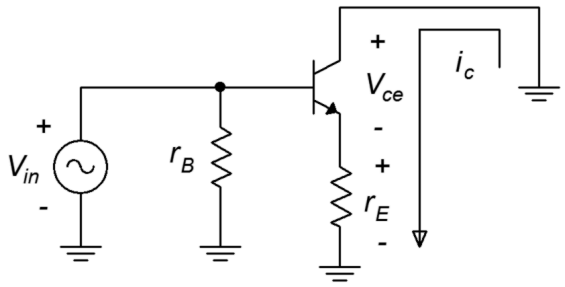


Figure 9.1
Voltage follower simplified AC circuit.

To gain a better understanding of how class B operation works, consider the simplified AC circuit of a voltage follower shown in Figure 9.1. If we situate the Q point directly at $v_{CE(cutoff)}$ then the associated I_{CQ} is 0 A. As the input signal swings positive, the collector current increases. As it does so, the voltage across the load (r_E) begins to increase and the voltage across the transistor's collector-emitter begins to

33 Think of a car engine: How much sense would it make to have an engine with a 6000 RPM maximum run at 3000 RPM when you're sitting motionless at a red light?

decrease (due to KVL). When the input signal swings negative, the transistor is turned off. As a result, no collector current is created, no voltage is developed across the load and v_{CE} stays at cutoff. It is as if the input waveform has been half-wave rectified. This action is shown in Figure 9.2. As the input signal swings positive, the operating point slides up the load line, moving toward saturation, and this increased current creates a load voltage that follows the input signal. In contrast, when the input tries to swing negative, there is no place else to go on the load line and the negative portion of the wave is simply clipped.

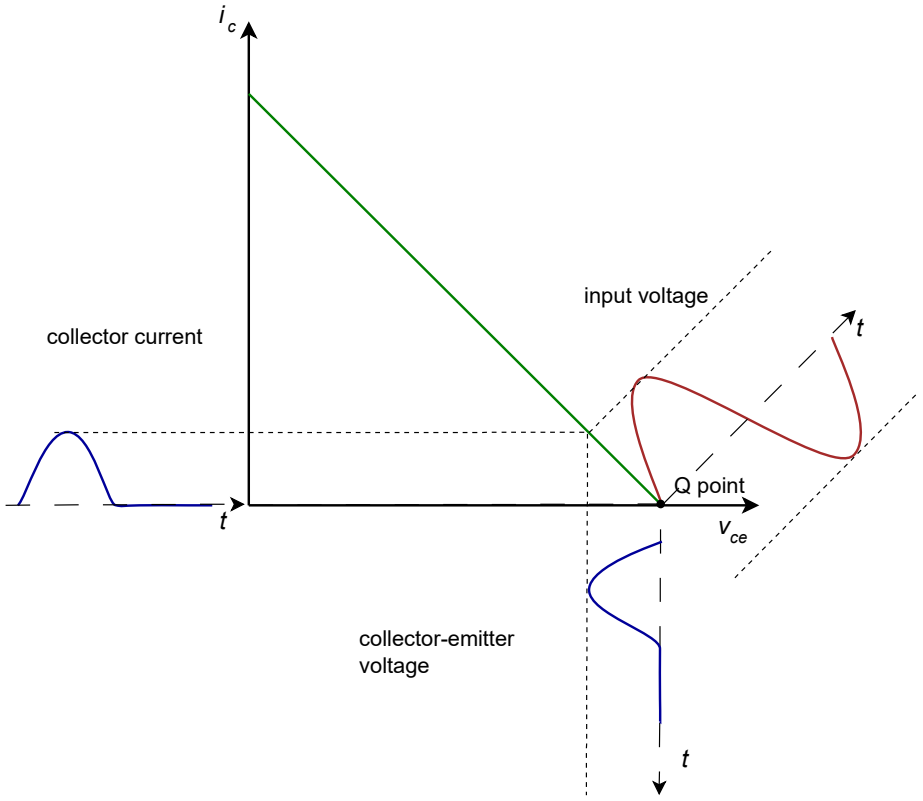


Figure 9.2
AC load line for class B operation.

If we made a PNP version of the circuit depicted in Figure 9.1, the exact opposite would happen: the amplifier would reproduce the negative portion of the wave and clip the positive portion. The next question is, how do we bias the transistor at cutoff and splice together the NPN and PNP versions into a workable whole?

Let's begin by making two bare-bones emitter followers, one NPN and the other PNP. We'll connect their emitters together and tie that to the load. We'll connect the NPN's collector directly to a DC supply and the PNP's collector to ground. Remember, collector resistors will not be needed because these are followers. We will include no biasing components on the base because we want to set I_{CQ} to 0 A. We will also have to add input and output capacitors to prevent the source and load from inadvertently shorting out or shunting portions of the DC circuit. The result is seen in Figure 9.3.

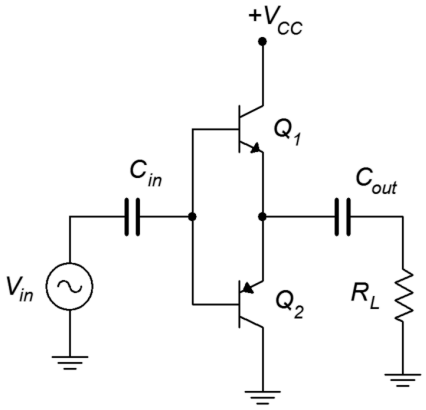


Figure 9.3
Prototype Class B circuit.

With no signal supplied, both transistors must be off. This is because their bases are tied together, and without some other applied potential, both base-emitter voltages must be zero. Assuming Q_1 and Q_2 are matched, the power supply voltage should split evenly between them, leaving half of V_{CC} at the emitters. This potential also appears across C_{out}, preventing the DC voltage from reaching R_L.

When the input signal goes positive, it raises V_{B1} and V_{B2} above $0.5V_{CC}$. This keeps Q_2 off but turns on Q_1. Current is now free to flow down through Q_1 and into the load. When the input signal swings negative, the inverse happens: Q_1 is turned off and Q_2 is turned on. This allows current to flow up from the load and down through Q_2 (if this is confusing, remember that a DC voltage had already been established across C_{out} that is equal to $0.5V_{CC}$ and this is what allows the current to flow from ground up through R_L and then down through Q_2 as Q_2 begins to conduct). You can think of Q_1 pushing current into the load (sourcing) and Q_2 pulling current from the load (sinking). Consequently, class B amplifiers are sometimes called a *push-pull* amplifiers.

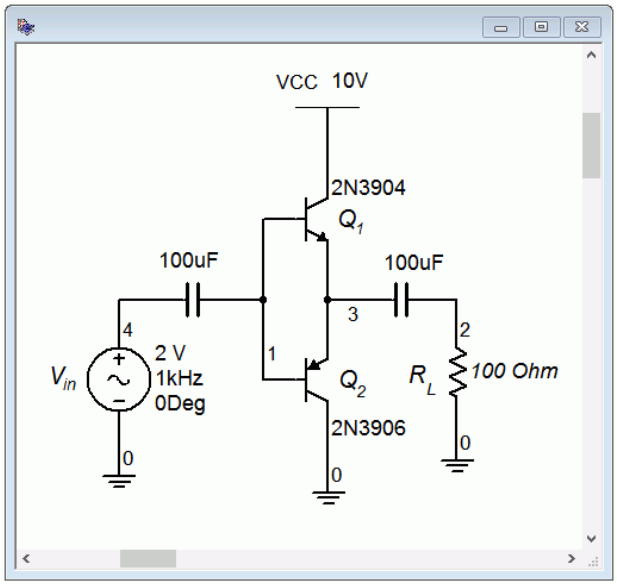


Figure 9.4
Prototype Class B circuit in simulator.

To see how well this prototype works, we'll enter a version into a simulator, as shown in Figure 9.4. A transient analysis is performed with the results shown in Figure 9.5.

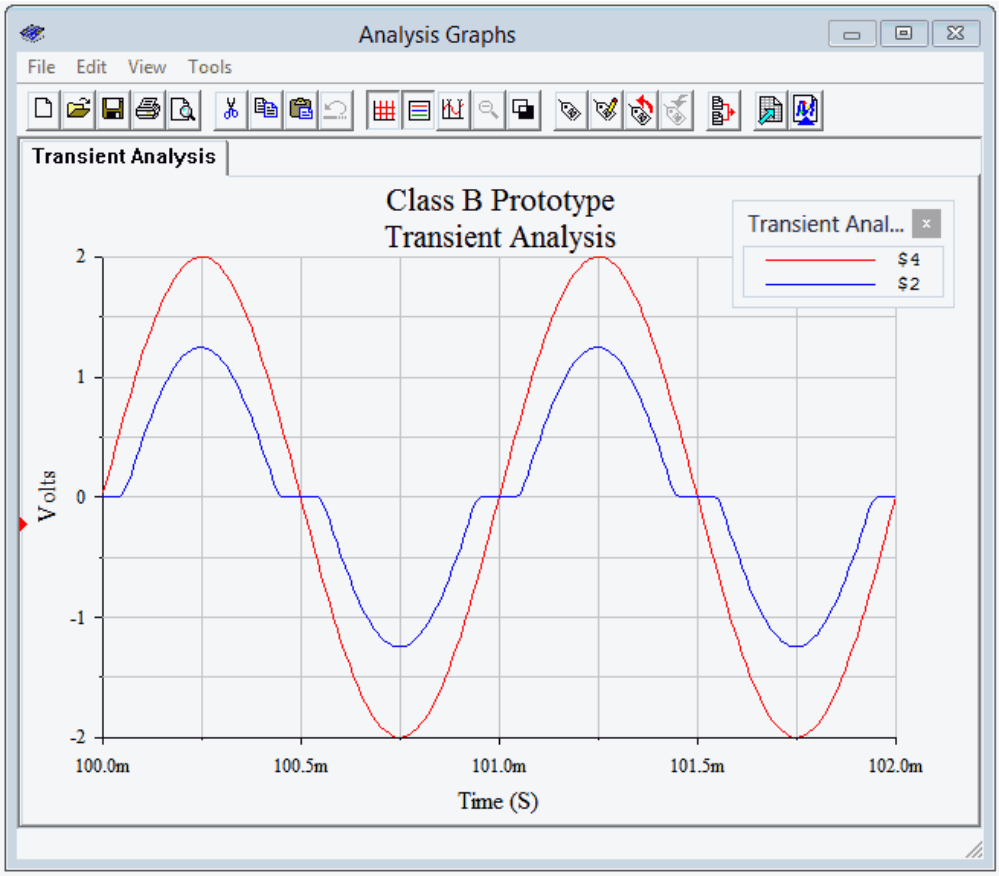


Figure 9.5
Transient analysis of prototype Class B circuit.

Clearly, there are issues with the output waveform (node 2, in blue). First off, the signal amplitude is noticeably smaller than the 2 volt peak input signal. The second issue is the bizarre "flat spotting" of the output waveform near the zero-crossing points. It turns out that these two problems are manifestations of the same root cause. If we look carefully at the peaks, we can get a clue as to what is going on. The peak output voltage is about 0.75 volts below the input, or just about one PN junction forward potential. The problem is that the input signal will not truly turn on the NPN transistor until the signal exceeds approximately 0.7 volts or drops below −0.7 volts for the PNP side. That region between −0.7 volts and +0.7 volts is a dead zone that the amplifier will not respond to. Essentially, the amplifier "rips out" anything between ±0.7 volts. This is a gross form of distortion and goes by many names including *notch distortion* and *cross-over distortion*. The particularly nasty part about this form of distortion is that it hits small signals worse than large signals. Most other forms of nonlinearities tend to get worse as the signal level increases.

Class AB Operation

The basic solution to this problem is to provide a small idle current so that the transistors are almost on. In this way, only a very small input signal will be needed to turn on the devices. As this would slightly increase the conduction angle, this form of operation is referred to as *class AB operation*. One potential solution is to add a voltage divider as depicted in Figure 9.6.

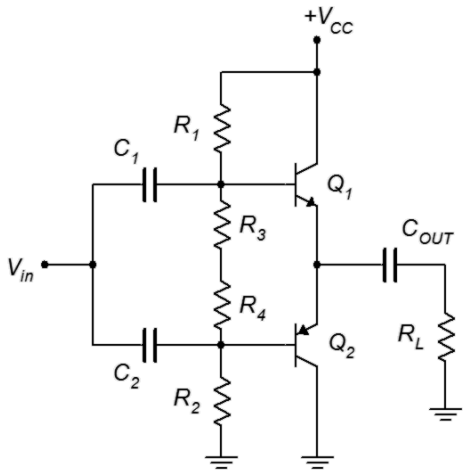


Figure 9.6
Prototype Class AB circuit.

This circuit uses a symmetrical layout: everything above a horizontal line drawn through the middle is echoed by a component below the line. In other words, $R_1 = R_2$, $R_3 = R_4$, $C_1 = C_2$, and Q_1 and Q_2 are a complimentary pair. The divider is configured so that the voltage drops across R_3 and R_4 are about 0.7 volts each.

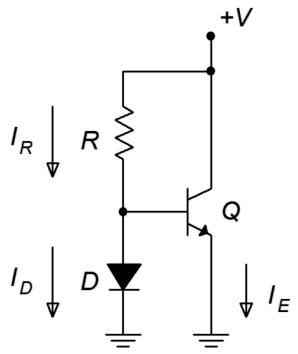
Properly designed, the circuit of Figure 9.6 will reduce notch distortion. Unfortunately, it has other problems. The first issue involves the three capacitors. C_{OUT} in particular might be quite large. These can be removed if we went to a symmetrical bipolar power supply. Instead of running the PNP's collector to ground, we'll tie it to a negative DC supply. To keep the same total voltage, we'll set Q_1's collector to half of the original V_{CC} and Q_2's collector to half of the original V_{CC} but negative. On the input side, we could run the input signal to the junction of R_3 and R_4.

The second issue plaguing the circuit of Figure 9.6 is the stability of the bias. The voltage divider resistors have to be very accurate in order to set the transistors right where we want them and stability is a concern. The problem is that we are trying to use a device with a linear current-voltage characteristic (a resistor) to match the exponential current-voltage characteristic of a PN junction. This problem is exacerbated by the fact that these devices will drift with temperature, and drift in different ways. The solution to this problem is to use a device with better matching characteristics. What better device to match a PN junction than another PN junction?

The Current Mirror

Consider the circuit shown in Figure 9.7. This is called a *current mirror*. Here is how it works: First, look at the divider between R and D. The voltage across R must equal the supply voltage minus the diode drop, or approximately $V - 0.7$. This sets up a current, I_R. If the base current is small enough to ignore, this same current flows down through the diode as I_D. This diode current sets up a specific voltage across the diode (somewhere in the vicinity of 0.7 volts although the exact voltage is not important). Because the diode is in parallel with the base-emitter junction, then $V_{BE} = V_D$. If the transconductance curve (I-V curve) of the transistor is identical to that of the diode, then the emitter current must be the same as the diode current. Any change in the diode current would cause a slight change in diode voltage, and since diode voltage and base-emitter voltage are the same, then the emitter current must change in response. In other words, the emitter current mirrors the diode current. We can program the diode current (and hence, the emitter current) by setting an appropriate value for R. Whatever the current through R is, that's also the collector current.

Figure 9.7
A simple current mirror.



The idea of the current mirror is used with great effect in integrated circuits where it is easy to match device characteristics. When it comes to discrete components, it is not nearly so easy to match a diode to a transistor. Fortunately, we don't have to have a perfect match. Simply using any signal diode will provide a much better match for V_{BE} than using a resistor. Although the diode current won't match the collector current precisely, the bias will be more stable and the components will track much better than when using a resistor.

Combining these ideas leads us to the circuit of Figure 9.8. This is our first practical class B amplifier with the potential for decent values of distortion and stability.

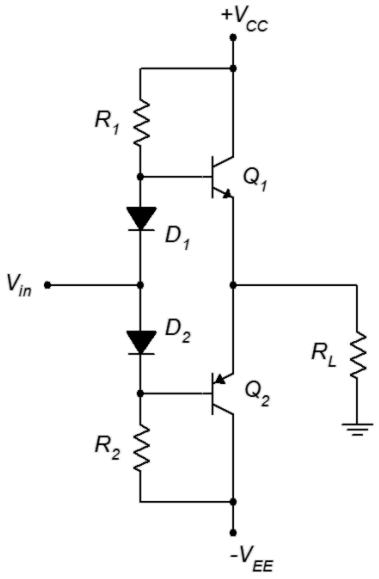


Figure 9.8
Class B amplifier with diode bias and bipolar supplies.

One thing that sometimes bothers people when they first see a diode biased amplifier is how the AC signal will pass through the diodes to the bases. At first glance, it appears that the positive portion of the input signal would be "going the wrong way" against diode D_1. What we need to remember is that D_1 is *already* forward-biased due to the DC supply and surrounding resistors. The signal won't see an open, it will see the dynamic resistance of the diode.[34]

Class B Circuit Maximums

Now that we have a workable circuit, we need to derive formulas for the endpoints of the AC load line, meaning $v_{CE(cutoff)}$ and $i_{C(sat)}$, and determine the compliance. The first item of note is cutoff. Because the two transistors will split the available supply, V_{CEQ} will always equal half of the total supply. Further, we have biased these devices at cutoff, and therefore

$$v_{CE(cutoff)} = V_{CEQ} = 0.5 \cdot Total\ DC\ Supply \qquad (9.1)$$

In the case of a bipolar supply, that's the same as one of the two sides. Because the class B uses two transistors, peak compliance will be the same as cutoff.

$$Compliance_{peak} = V_{CEQ} = 0.5 \cdot Total\ DC\ Supply \qquad (9.2)$$

The maximum voltage rating of the transistors will occur when they are off. In that instance, if the opposite transistor is fully conducting it will have a negligible voltage across its collector-emitter. Consequently, the off-state transistor can see the entire power supply.

$$BV_{CEO} = Total\ DC\ Supply \qquad (9.3)$$

The saturation current is dictated by the compliance and the load. The only thing that limits AC current is the load. Therefore

$$i_{C(sat)} = \frac{Compliance_{peak}}{r_L} \qquad (9.4)$$

Based on that, we can say

$$P_{load(max)} = \frac{Compliance_{RMS}^{2}}{r_L} \qquad (9.5)$$

34 Of course, the dynamic resistance is a function of the current flowing through the diode so it will fluctuate as the signal changes. This is best thought of as a distortion generating mechanism because it will slightly alter the signal that reaches the base. This distortion is most likely orders of magnitude smaller than the notch distortion the diode mitigates, so it's a good trade.

Before we go any further, there is an important item to note about the circuit of Figure 9.8 (and the variants we shall discuss). If you take another look at the circuit you will see that there is *nothing* in the collector-emitter line to limit DC current. In fact, if we were to plot the DC load alongside the AC load, we'd get something like Figure 9.9.

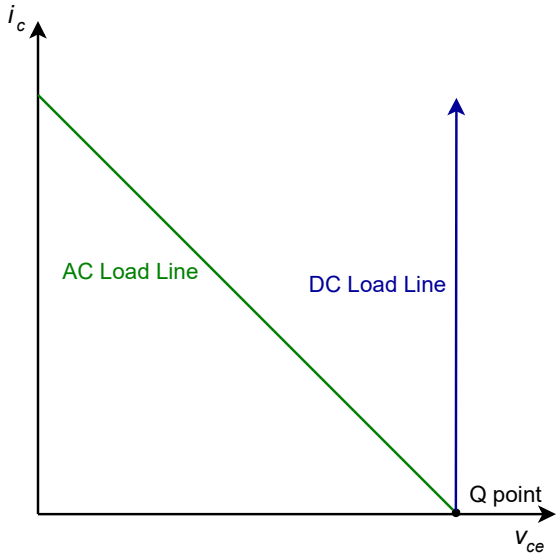


Figure 9.9
Comparison of AC and DC load lines for class B operation.

The DC load line goes straight up. There is no saturation current value short of infinity. What this really means is that if we're not careful with the bias, it is possible to destroy the transistors. The same is true if we accidentally short the load. A quick examination of Equations 9.4 and 9.5 shows that a shorted load condition would lead to huge currents and powers, and destroy the transistors in the process. With an audio amplifier, this could happen if one of the strands of loudspeaker wire unraveled and touched the adjacent lead. Obviously, not a happy situation unless you prefer the smell of burnt silicon over the sound of music. We will examine means of protecting the transistors from accidental overloads later in the chapter.

Class B Power Dissipation

Transistor power dissipation for the class B configuration is a bit trickier than it was for the class A. The idle (Q point) power dissipation is very low. It is found by multiplying I_{CQ} by V_{CEQ}. I_{CQ} is generally set at a few percent of $i_{C(sat)}$ so it's obviously way below the maximum load power. Unlike the class A design, P_{DQ} does **not** represent the worst case for class B.

To determine the worst case transistor power dissipation, we begin by describing the current and voltage waveforms during the conduction phase of the NPN. The collector current will appear as the positive half of a sine wave. The maximum case has the current starting at zero and peaking at $i_{C(sat)}$ (or alternately, V_{CEQ}/r_L). For v_{CE}, it starts at V_{CEQ} and then swings down to zero as a negative half sine.

Nothing says that the maximum load power case must cause the maximum transistor dissipation. In fact, we saw this wasn't the case for class A. Consequently, we will introduce a coefficient, *k*, that represents the percentage of the maximum current. We now arrive at our general equations for transistor current and voltage for the first half cycle.

$$i_C = k\frac{V_{CEQ}}{r_L}\sin 2\pi ft \qquad (9.6)$$

$$v_{CE} = V_{CEQ}(1-k\sin 2\pi ft) \qquad (9.7)$$

Where $0 \le k \le 1$

For convenience, we'll set $2\pi f$ to 1. To get the power dissipation, we find the product of the transistor's current and voltage.

$$P_D = i_C v_{CE}$$
$$P_D = k\frac{V_{CEQ}}{r_L}\sin t \times V_{CEQ}(1-k\sin t)$$
$$P_D = \frac{V_{CEQ}^2}{r_L}k\sin t - \frac{V_{CEQ}^2}{r_L}k^2\sin^2 t$$
$$P_D = \frac{V_{CEQ}^2}{r_L}\left(k\sin t - k^2\sin^2 t\right)$$

Remove ugly $\sin^2$ term...

$$P_D = \frac{V_{CEQ}^2}{r_L}\left(k\sin t + \frac{k^2}{2}\cos 2t - \frac{k^2}{2}\right)$$

..and integrate to get:

$$P_D = \frac{V_{CEQ}^2}{r_L}\left(-k\cos t - \frac{k^2}{4}\sin 2t - \frac{k^2 t}{2}\right)\Bigg|_0^{\pi}$$

Note that $\frac{V_{CEQ}^2}{r_L} = 2P_{load(max)}$ This is a constant, so replace it to simplify and then evaluate the expression. Finally, divide by 2π to find the average over one full cycle.

$$P_D = \frac{2\,P_{load(max)}\left(2\mathrm{k} - \frac{k^2\pi}{2}\right)}{2\,\pi}$$

$$P_D = 2\,P_{load(max)}\left(\frac{k}{\pi} - \frac{k^2}{4}\right) \qquad (9.8)$$

Equation 9.8 is the general case. For the worst case, we need the min/max k value. We will take the derivative of Equation 9.8 and then set it to zero to find the worst case value of k.

$$P_D = 2\,P_{load(max)}\left(\frac{k}{\pi} - \frac{k^2}{4}\right)$$

$$\frac{d\,P_D}{d\,k} = 2\,P_{load(max)}\left(\frac{1}{\pi} - \frac{k}{2}\right)$$

Worst case occurs at $k = 2/\pi$. This means that only $2/\pi$, or 63.7%, of the load line is used at maximal heating of the transistor. 63.7% of the load line corresponds to a load power of about 40% of $P_{load(max)}$ (i.e., 0.637^2). We now substitute this value back into Equation 9.8 to find the worst case P_D.

$$P_D = 2\,P_{load(max)}\left(\frac{2/\pi}{\pi} - \frac{(2/\pi)^2}{4}\right)$$

$$P_D = \frac{2}{\pi^2}P_{load(max)} \approx \frac{P_{load(max)}}{5} \qquad (9.9)$$

The end result is that when the load is receiving about 40% of its maximum power, the transistors will be at their hottest and will be dissipating approximately 20% of the maximum load power (or about half of the power delivered to the load at that point). Thus, if a class B amplifier is rated to produce a maximum load power of 100 watts, the transistors will get their hottest when the load is receiving 40 watts, and each transistor will be dissipating 20 watts. The transistors will be dissipating *less* power when the load is at maximum. This is easily verified by substituting $k = 1$ into Equation 9.8. The result is a power dissipation of 13.7% of maximum load power, or 13.7 watts for the preceding example.

To help gain a deeper understanding of precisely what's happening here, the transistor waveforms are plotted in Figures 9.10 and 9.11. The maximum load power case is presented in Figure 9.10. Here, we see the full i_C and v_{CE} swings. Note that when i_C is maximum, v_{CE} is 0, hence the power is 0. In contrast, Figure 9.11 shows the worst case. Collector current peaks at just under 64% of maximum but v_{CE} drops to only about 36% rather than 0% of its maximum. This results in a power dissipation curve with considerably greater area underneath it, indicating a higher average power.

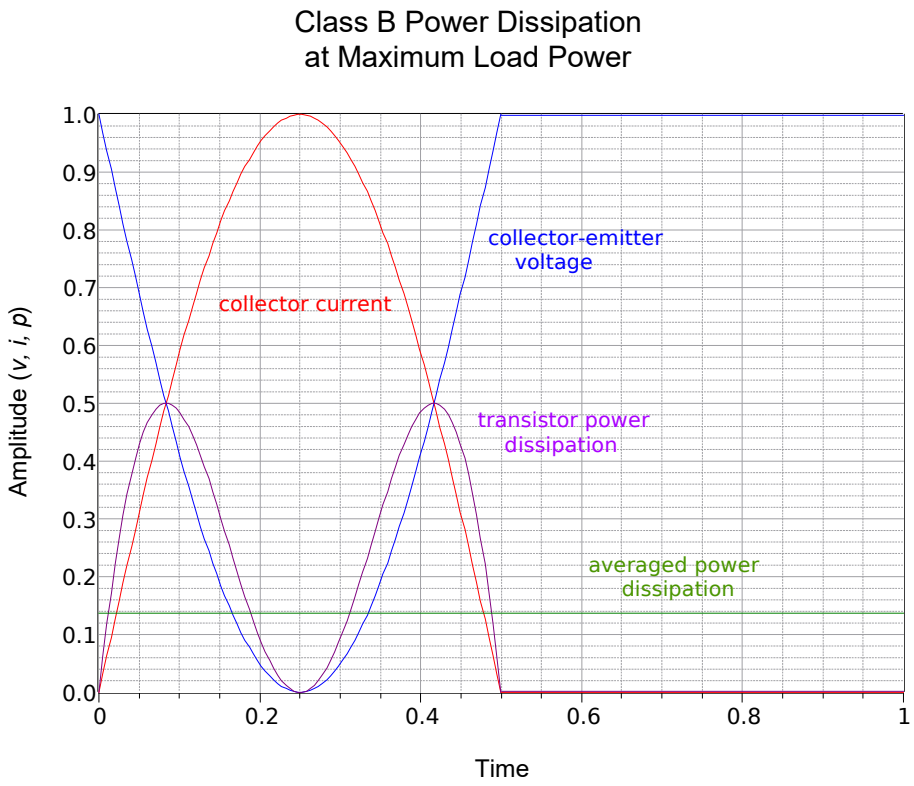


Figure 9.10
Class B transistor power dissipation at $P_{Load(max)}$.

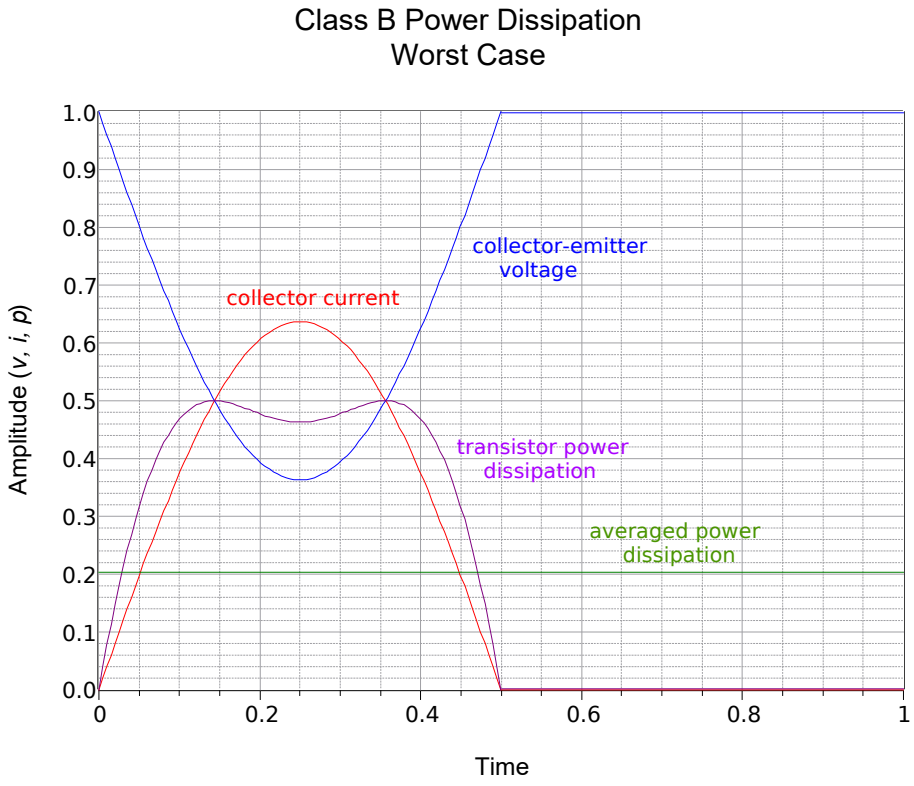


Figure 9.11
Class B transistor power dissipation, worst case.

Finally, it is worth remembering that the reactive load issues discussed for class A amplifiers still apply to class B amplifiers. Loads with a complex impedance may be harder to drive than the ideal purely resistive loads examined here. Therefore, we may have to over-rate the transistors above $P_{load(max)}/5$. As a side note, the AC load line for a class B amplifier with a reactive load will appear as an ellipse that has been cut in half (refer back to Figure 8.16 and imagine a horizontal cut line running through the Q point).

Class B Efficiency

Efficiency is defined as useful output or load power versus supplied DC power.

$$\eta = \frac{P_{out}}{P_{in}} = \frac{P_{load(max)}}{P_{DC}}$$

This is dynamic for class B amplifiers. At $P_{load(max)}$ the DC supply is delivering the full voltage of 2 V_{CEQ} and the corresponding peak value of the current is $i_{C(sat)}$.

$$i_{C(sat)} = \frac{V_{CEQ}}{r_L}$$

The average of this over a half-cycle is

$$i_{C(avg)} = \frac{1}{\pi} \times \frac{V_{CEQ}}{r_L}$$

Therefore, the supplied power must be

$$P_{DC} = 2V_{CEQ} i_{C(sat)}$$
$$P_{DC} = 2V_{CEQ} \times \frac{1}{\pi} \frac{V_{CEQ}}{r_L}$$
$$P_{DC} = \frac{2}{\pi} \times \frac{V_{CEQ}^2}{r_L}$$

As noted previously $\frac{V_{CEQ}^2}{r_L} = 2P_{load(max)}$ therefore

$$P_{DC} = \frac{4}{\pi} \times P_{load(max)}$$

Finally, substitute this expression back into the original definition for efficiency.

$$\eta_{max} = \frac{P_{load(max)}}{P_{DC}}$$
$$\eta_{max} = \frac{P_{load(max)}}{\frac{4}{\pi} P_{load(max)}}$$
$$\eta_{max} = \frac{\pi}{4}$$
$$\eta_{max} \approx 78.5\%$$

We find that the maximum theoretical efficiency of a class B amplifier is over three times that of a class A amplifier.

Time for an example.

Example 9.1

The amplifier shown in Figure 9.12 is driving a nominal 8 Ω loudspeaker. Determine the compliance, maximum load power and worst case transistor dissipation. Also estimate Z_{in} assuming β = 50 and determine the transistor ratings for maximum current and BV_{CEO}.

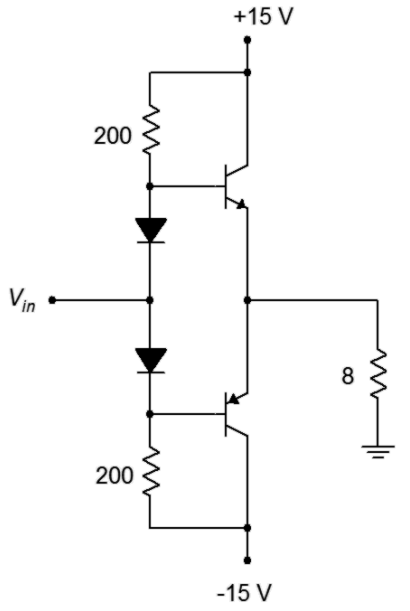


Figure 9.12
Schematic for Example 9.1.

By inspection, V_{CEQ} = 15 V. This is the peak compliance.

$$compliance = 15 \text{ V peak} = 10.6 \text{ V RMS}$$

Given the compliance, we can use power law to find the load power

$$P_{load(max)} = \frac{Compliance_{RMS}{}^2}{R_L}$$
$$P_{load(max)} = \frac{(0.707 \times 15\,\text{V})^2}{8\,\Omega}$$
$$P_{load(max)} = 14\,\text{W}$$

That's not huge but it might be enough to irritate the neighbors.

The transistors' worst case power dissipation is

$$P_D = \frac{P_{load(max)}}{5}$$
$$P_D = \frac{14\,\text{W}}{5}$$
$$P_D = 2.8\,\text{W}$$

The breakdown voltage is the entire supply so $BV_{CEO} > 30$ V. The maximum current through the transistors is the same as the maximum load current or $i_{C(sat)}$.

$$i_{C(sat)} = \frac{V_{CEQ}}{r_L}$$
$$i_{C(sat)} = \frac{15\,\text{V}}{8\,\Omega}$$
$$i_{C(sat)} = 1.88\,\text{A}$$

The input impedance is found in the usual manner but with a minor twist. $Z_{in(base)}$ is approximately equal to $\beta\, r_L$, or about 400 Ω. Only one transistor is on at any given time, though, so this is in parallel with the two 200 Ω biasing resistors but not the other $Z_{in(base)}$ (the "off" transistor has a very high input impedance because it is not conducting). This leaves a $Z_{in(base)}$ of about 80 Ω, or ten times the load impedance.

Computer Simulation

To verify the basic operation of a class B amplifier, the circuit of Figure 9.13 is entered into a simulator. The amplifier should clip just below the ±10 volt power rails so a 10 volt peak source is used to verify this. Also, A_v should be about 1.

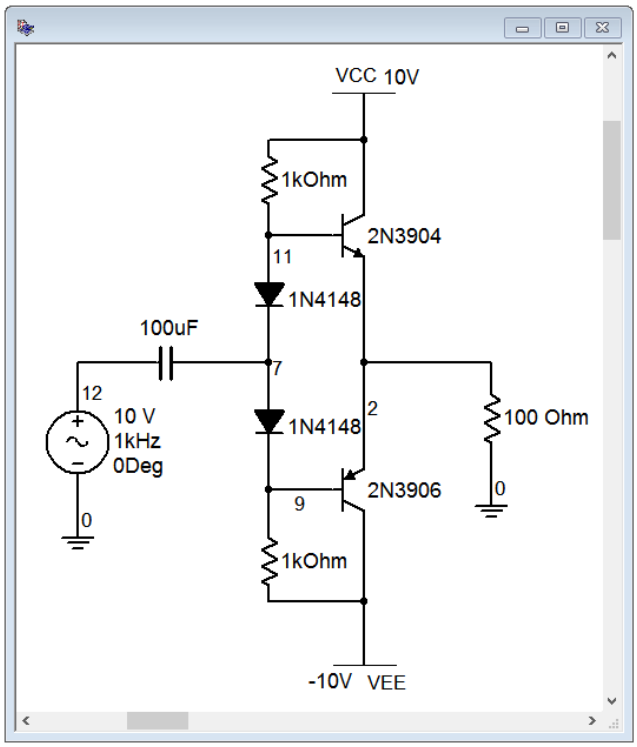


Figure 9.13
Class B amplifier in simulator.

The results of a transient analysis are shown in Figure 9.14. First, it is apparent that the voltage gain is approximately unity as the input and output waves are nearly coincident (with the exception of the clipped portion). Clipping occurs at about 8.5 volts. This is largely due to limiting from the biasing diodes. Once the input gets close in value to the power supply, the diodes become reverse-biased and the signal does not make it to the base. Thus, the output clips prematurely.

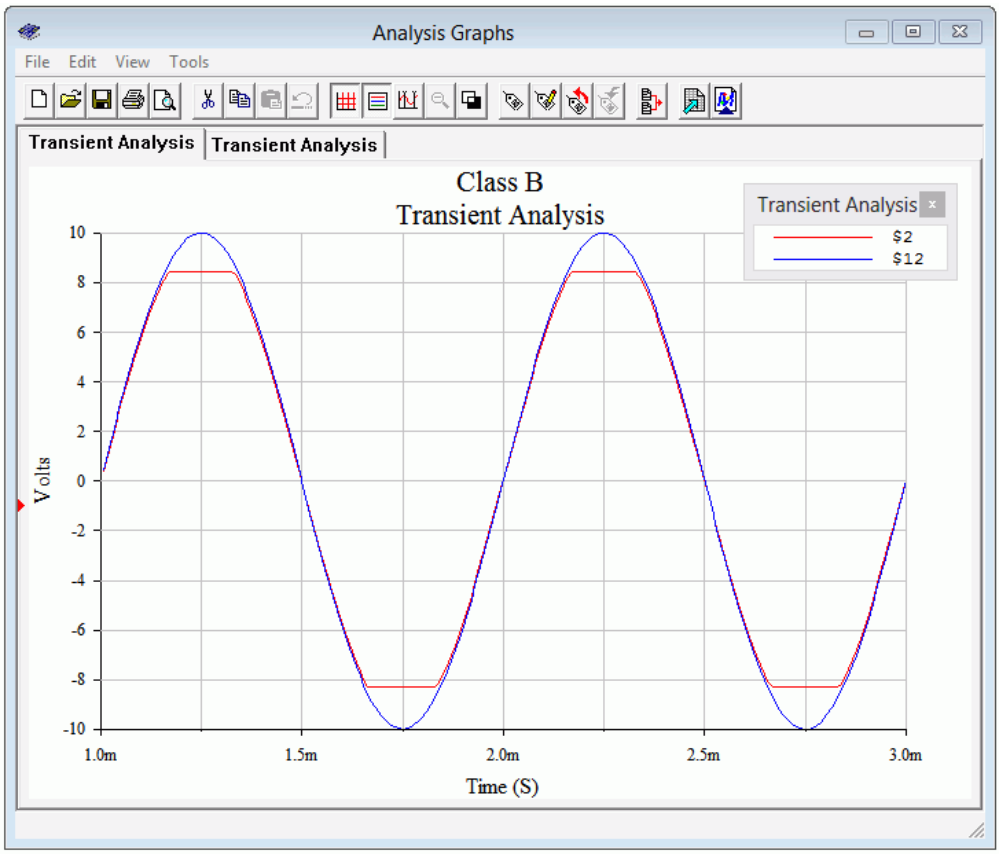


Figure 9.14
Class B amplifier transient analysis.

To verify earlier commentary regarding the biasing diodes, the circuit is modified so that the diodes are shorted out and the transient analysis is run again. The results are shown in Figure 9.15.

Two things should be apparent in the new simulation. First, the output waveform is suffering from obvious notch distortion. Therefore, we can be pleased that the biasing diodes have done their job to reduce this effect. The second item involves the compliance. The new output is not clipped. Granted, the lack of biasing diodes has reduced the signal by about 0.7 volts, but a careful examination of the output waveform shows that it has reached over 9 volts peak without clipping. In fact, if we increase the input to 12 volts peak, as in Figure 9.16, we can see that clipping occurs just under the power rails. Thus, the premature clipping is due to the diodes and their interaction with the power supplies and surrounding resistors. In short, we lose a little bit of compliance due to the diodes but that's much better than getting notch distortion.

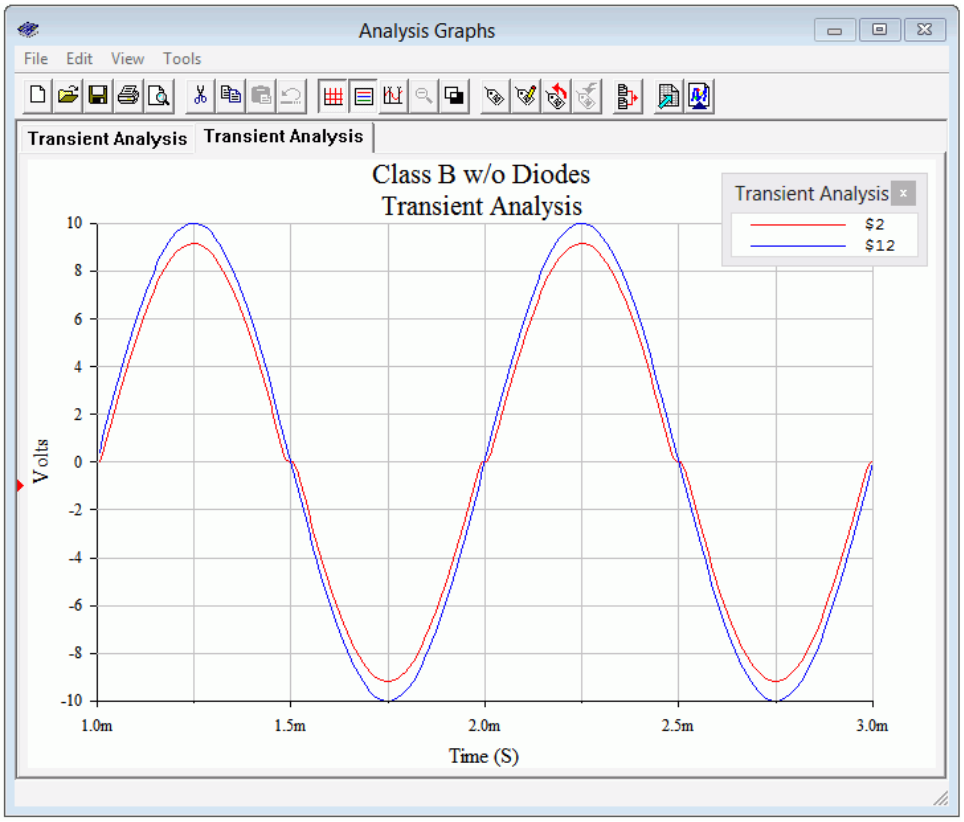


Figure 9.15
Class B amplifier transient analysis, without biasing diodes.

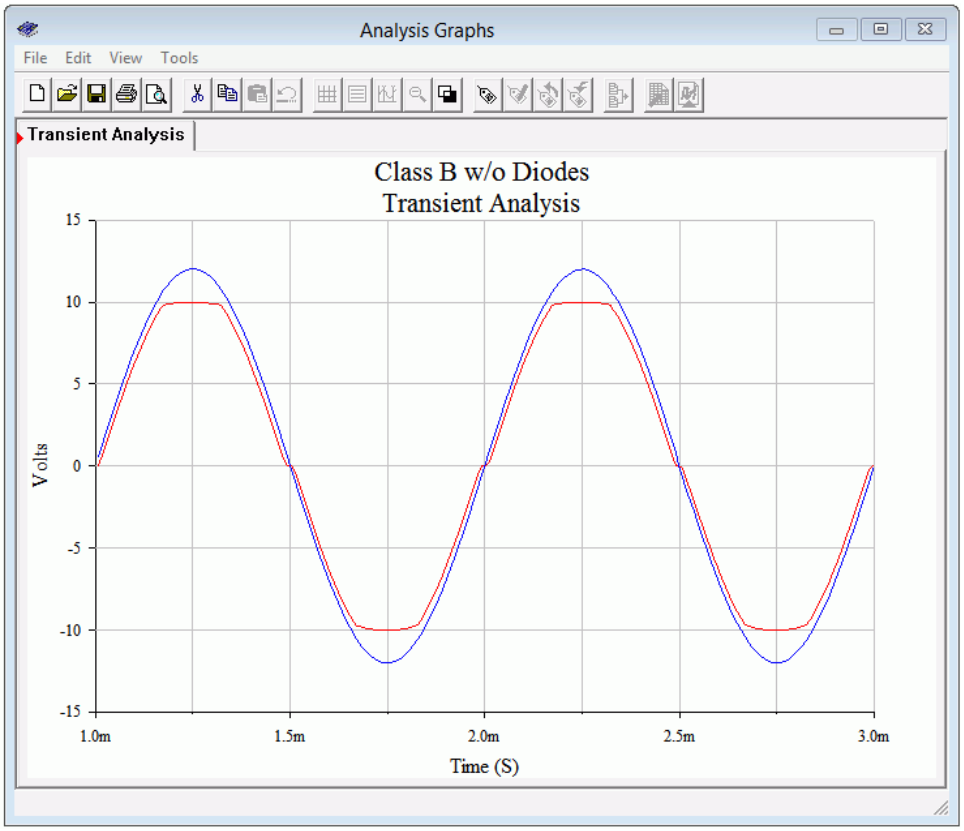


Figure 9.16
Class B amplifier transient analysis, without biasing diodes and showing clipping.

Direct Coupled Driver

The circuits we have examined offer both current gain and power gain but not voltage gain. In order to increase the signal voltage, some preceding voltage gain stages will most likely be needed. These stages can be connected to the previous class B follower circuits with coupling capacitors but this is not the most effective method. A more common technique is the use of a *direct coupled driver.*

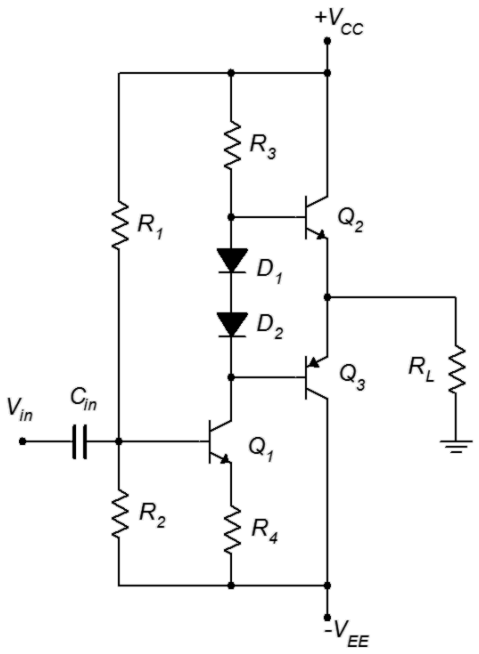


Figure 9.17
Class B amplifier with direct coupled driver.

A class B follower with a direct coupled driver stage is shown in Figure 9.17. What we've done here is combined an ordinary class A common emitter amplifier (in this case, using voltage divider bias) with a class B follower. The follower is positioned where the common emitter stage's collector resistor would be normally. This eliminates three components: the collector resistor, the interstage coupling capacitor, and the lower base biasing resistor of the class B output stage. The removal of the resistors raises the effective load resistance for the first stage, thus producing a higher voltage gain from the driver stage.

Biasing the direct coupled driver is not difficult if we remember one thing: the DC voltage across R_3 must be equal to approximately $V_{CC} - 0.7$ V. If this is not the case, the class B stage will not be symmetrical, or in other words, $V_{CEQ2} \neq V_{CEQ3}$, and the final output will not be sitting at 0 VDC as it needs to. Knowing the value of R_3 and its voltage, we can determine its current. This current flows down into Q_1 as I_{CQ1}. Knowing I_{CQ1}, we can determine the voltage across R_4 and eventually determine the appropriate divider ratio for R_1 and R_2 to achieve this value.

Example 9.2

Using the two-stage amplifier of Figure 9.17, first determine values for R_1 and R_2 to obtain proper system bias. Determine the output compliance, maximum load power and worst case transistor dissipation. Also estimate A_v. Assume β = 50 for the output devices and 100 for the first stage. V_{CC} = 20 V, V_{EE} = −20 V, R_L = 16 Ω, R_3 = 560 Ω, R_4 = 75 Ω.

For the output section (assuming it will be biased properly), by inspection, V_{CEQ} = 20 V. This is the peak compliance.

$$compliance = 20 \text{ V peak} = 14.1 \text{ V RMS}$$

Given the compliance, we can use power law to find the load power

$$P_{load(max)} = \frac{Compliance_{RMS}^{\ 2}}{R_L}$$
$$P_{load(max)} = \frac{(14.1\,\text{V})^2}{16\,\Omega}$$
$$P_{load(max)} = 12.5\,\text{W}$$

The transistors' worst case power dissipation is

$$P_D = \frac{P_{load(max)}}{5}$$
$$P_D = \frac{12.5\,\text{W}}{5}$$
$$P_D = 2.5\,\text{W}$$

To determine the biasing resistors, we start with R_3. To achieve bias symmetry, all of V_{CC} drops across R_3 with the exception of 0.7 volts for D_1. This is the same as I_{CQ1}.

$$I_{CQ_1} = \frac{V_{CC} - 0.7\,\text{V}}{R_3}$$
$$I_{CQ_1} = \frac{19.3\,\text{V}}{560\,\Omega}$$
$$I_{CQ_1} = 34.5\,\text{mA}$$

The drop across R4 is found via Ohm's law

$$V_{R_4} = I_{CQ_1} R_4$$
$$V_{R_4} = 34.5\,\text{mA} \times 75\,\Omega$$
$$V_{R_4} = 2.6\,\text{V}$$

This implies that the voltage across R_2 must be 0.7 volts more, or 3.3 volts. If we ignore the base current of Q_1, then the ratio of R_1 to R_2 must be the same as the ratio of their voltages, 36.7 to 3.3, or 11.1 to 1. In other words, R_1 must be 11.1 times larger than R_2. For good bias stability we don't wish to set R_2 too much larger than R_4. If we set it to 200 Ω, for example, then R_1 would need to be about 2.2 k Ω. Due to component tolerances, one of these resistors would need to be a potentiometer (connected as a rheostat) in order to "tweak" the final output to 0 VDC. A resistor/pot combo might be even better as it won't be so "touchy". For example, the 2.2 k Ω could be replaced with a series combination of a 1.8 k Ω and a 1 k Ω pot. This would be much easier to adjust than if the resistor was replaced with a standard 5 k Ω pot.

Now for the system voltage gain. The gain of the follower is approximately one so we need only concern ourselves with the first stage common emitter amplifier. This amplifier is swamped by R_4 and given that the collector current is over 34 mA, r'_e will be less than an ohm and can be ignored. All we need to do is find the effective load at the collector of Q_1. This is R_3 in parallel with a single $Z_{in(base)}$ (remember that only one transistor is on at any given time and the off transistor will appear as a high impedance).

$$Z_{in(base)} = \beta r_E$$
$$Z_{in(base)} = 50 \times 16\,\Omega$$
$$Z_{in(base)} = 800\,\Omega$$

$$A_v = -\frac{r_L}{r_E}$$
$$A_v = -\frac{800\,\Omega \parallel 560\,\Omega}{75\,\Omega}$$
$$A_v = -4.4$$

Before leaving this section, there are a few items to note. First, the load power calculations have assumed that the entire power supply can be used by the output devices. As we saw with the diode bias version, this is not always the case. There is another situation that can limit the output swing. The class B stage is a follower and thus has a voltage gain of one. It the driver stage can't produce the full swing then the output stage can't either. Consequently, a class A analysis (i.e., AC load line) needs to be performed on the driver in order to determine just how large the signal can be before clipping. Indeed, it's quite likely that the driver stage will clip before the output stage.

Also, the direct coupled driver does not have to be an NPN as depicted in Figure 9.17. A PNP can be used instead, it simply needs to be shifted to the upper section rather than the lower section, as shown in Figure 9.18.

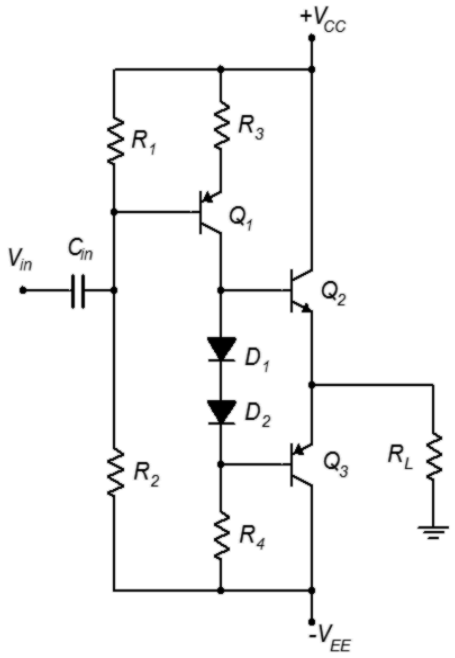


Figure 9.18
Class B amplifier with direct coupled driver, PNP version.

9.3 Extensions and Refinements

The foregoing discussion has covered the basics of class B amplifier design and operation but there are a variety of things we might add.

Current Limiting

One of the more useful additions to an amplifier is some form of protection circuit. As noted earlier, there is nothing in the basic class B amplifier to limit current, so if the load is accidentally shorted, the transistor current will spike to very high levels and possibly damage the output transistors. How might we prevent this from happening?

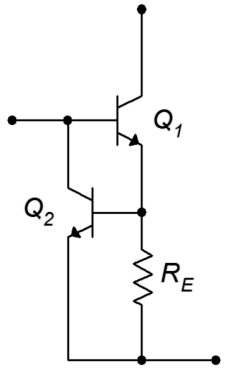
Perhaps the most basic protection technique is to place a fuse in-line with the load. This is a somewhat tricky proposition because voice and music waveforms are very dynamic. Fuses are not fast response devices and selection of the proper current rating is a trade-off. If we want to catch large, fast transients, the fuse will have to be rated on the low side. Unfortunately, the fuse might then blow on moderately loud sustained low frequency content. Relays suffer from similar problems. Also, there is the issue that if the fuses are replaceable by the consumer, they may use the wrong value or some other item[35] that will allow too much current, resulting in blown

35 Such as bits of wire, aluminum foil, small bolts or screws, etc. Just ask any seasoned repair technician.

output transistors. This can be avoided by having a second set of fuses mounted on the PC board, connected to the power rails, but this is also an imperfect solution.

One means of dealing with this situation is to employ an active current limiter, such as the circuit shown in Figure 9.19. In this circuit, Q_1 is the main NPN output transistor, the PNP side not shown. At the top is the power supply connection and the load is connected off to the right. A small resistor, R_E, is inserted into the emitter current path. The resistance is small enough that the voltage across it is normally less than 0.5 volts or so. Across the resistor is another transistor, Q_2. Under normal operation Q_2 is not on and does not affect the circuit. If the load current gets large enough (beyond the safe limit), the voltage drop across R_E will reach 0.7 V. At this point Q_2 turns on and begins to conduct current away from the base of Q_1, limiting the amount of load current to approximately 0.7 V/R_E. Unlike a fuse, once the fault is removed and the load current falls to safe levels, Q_2 disengages and normal amplifier function resumes.

Figure 9.19
Active current limiter.



Two of these circuits are required for the amplifier; one for the positive half and one for the negative half. Figure 9.20 shows a class B amplifier with the added current protection circuits (within the dashed red box).

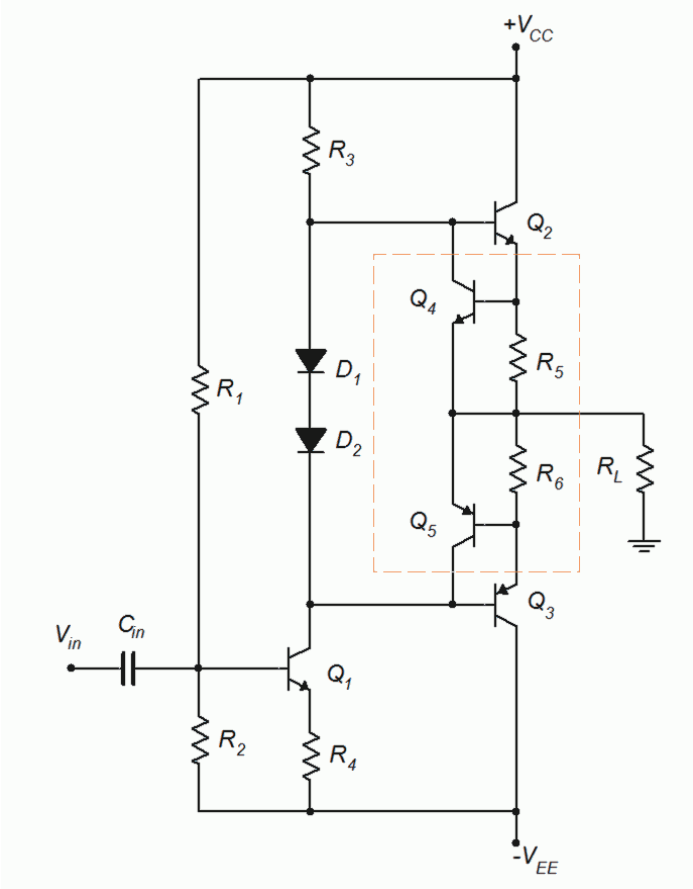


Figure 9.20
Class B amplifier with current limiters.

High Current Gain Configurations

For some applications the β of typical power transistors may prove to be insufficient. In such cases we can use high current gain configurations. There are several ways to configure the output devices. A Darlington scheme is shown in Figure 9.21. Q_3 and Q_5 are the main output devices. Q_2 and Q_4 can be thought of as drive transistors. Although their BV_{CEO} rating will need to be as high as that of Q_3/Q_5, they will be handling less current and therefore will dissipate less power. Sometimes Q_2/Q_4 are configured as seen here and sometimes emitter resistors may be added so that the output appears to be a cascade of emitter followers. Either way, there are now four base-emitter junctions to be compensated for, thus the addition of D_3 and D_4.

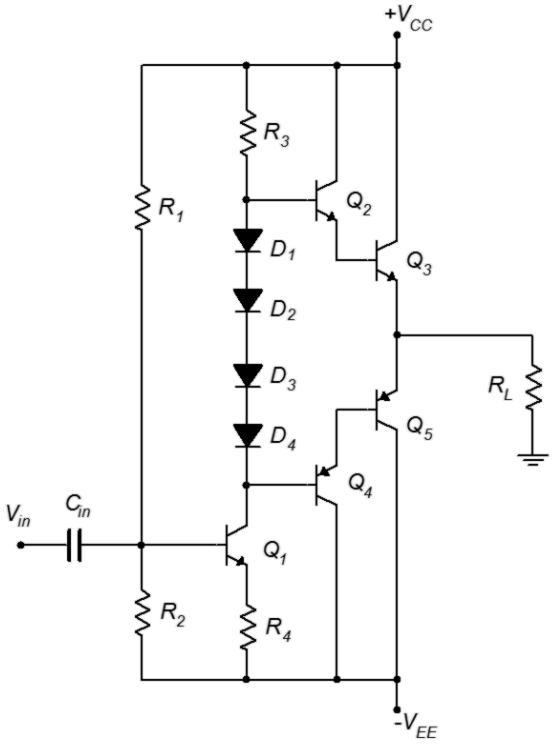
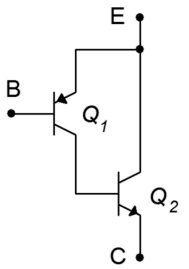


Figure 9.21
Darlington output devices for high current gain.

Some years ago, high power PNP transistors exhibiting decent audio quality were not available. Instead of using a PNP Darlington configuration, a *Sziklai pair* was used. This dual transistor configuration is named after Hungarian, and later, American, engineer George Sziklai. The Sziklai pair is also known as a *composite pair*. A composite PNP is shown in Figure 9.22. The operation is similar to that of a Darlington pair. The input transistor, Q_1, drives its collector current into the base of Q_2, the output transistor. Q_2's base current is multiplied by its β, thus its collector current is the product of I_{B1}, β_1 and β_2, just like a Darlington. The differences are that the main power device for the composite PNP is an NPN, and that there is only a single V_{BE} to compensate for. An example amplifier based on the earlier direct coupled driver circuit is shown in Figure 9.23. This configuration is known as a

Figure 9.22
Sziklai pair, AKA composite PNP.



quasi-complementary output[36]. Note the use of three compensating diodes; two for the Darlington NPN and one for the composite PNP/Sziklai pair. One advantage here is that the power devices, Q_3 and Q_5, can be identical models.

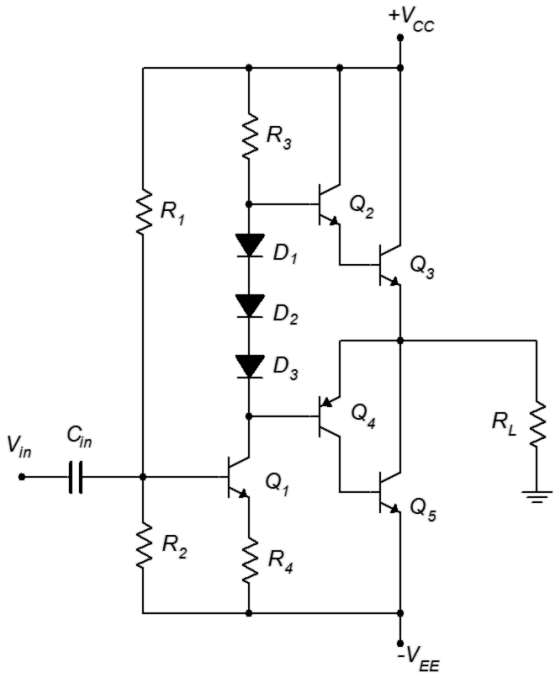


Figure 9.23
Quasi-complementary output.

The current limiting circuit of Figure 9.19 can be added to both the Darlington and quasi-complementary output stages. Current limiting can also be added to the circuits presented in the next section, but have been left off for clarity.

Current Sharing

For higher output powers, and especially for amplifiers driving very low load impedances such 1 or 2 ohms, it may not be possible to use a single NPN and PNP for the output stage. Instead, the duties of each device will be shared among two or three transistors operating in parallel. Generally, current-sharing schemes will also require a Darlington-based approach due to the very high total load current.

Initially, it may be tempting to just place two or three transistors in parallel, each base terminal being fed by a common drive transistor. There is a fatal flaw with this approach. The issue is referred to as either *thermal runaway* or *current hogging*. The basic problem is that the temperature coefficient of transconductance for a bipolar transistor is positive. In other words, r'_e gets smaller as temperature rises. This means that, as the device heats up, it tends to more easily conduct current. Of course, if the device draws more current, it will dissipate more power, which means that it

36 Not to be confused with the *quasi-complimentary output* of which we have only partially-nice things to say.

will get hotter, which means that it will conduct more current, which means that it will dissipate more power, and so on. This process spirals out of control and is particularly evil when we have multiple devices in parallel. Devices are never perfectly matched so the load current will never be perfectly shared among the devices. This means that one device will tend to get hotter than the other(s) which leads to it getting even hotter and grabbing a larger and larger share of the total current. Eventually, this device "hogs" all of the available current and winds up destroying itself. What we have is a positive thermal feedback loop. Transistor hara-kiri is a decidedly sub-optimal performance characteristic. To circumvent this problem we can add small resistors to the emitters of the devices. As all of the paralleled devices are being driven from a common drive transistor, they all see the same base voltage. If one output transistor starts to grab a larger share of the load current, the voltage drop across the emitter resistor will increase, thus forcing a reduction in that transistor's base-emitter voltage. This reduction compensates for the initial tendency of the current to increase. This technique is referred to as *local negative feedback*. We have seen this concept before, for example, a swamping resistor falls into this category as does the collector feedback bias scheme.

An example of a current-sharing output section is shown in Figure 9.24. This version uses a Darlington scheme. The newly added transistors are shown in red while the thermal/current feedback resistors are shown in blue.

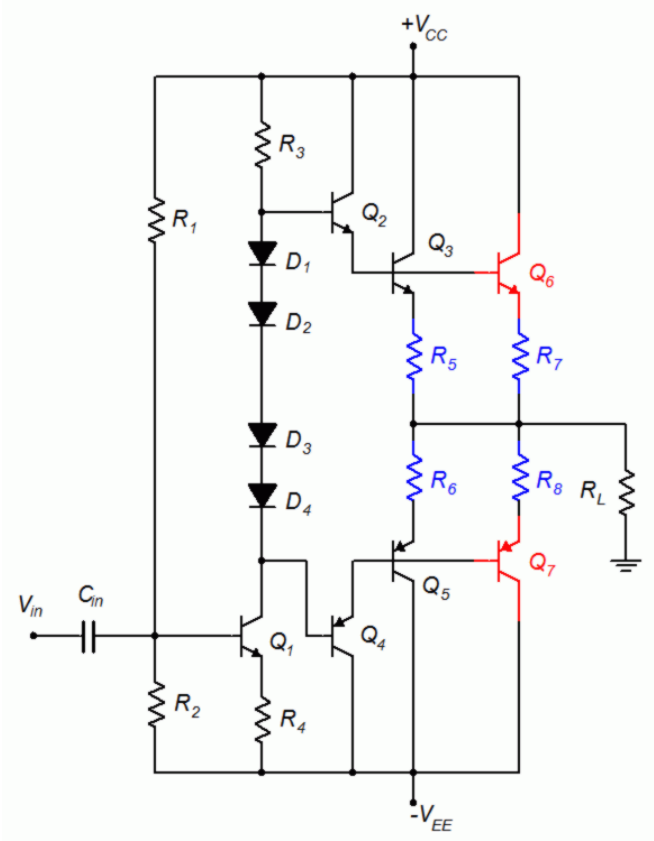
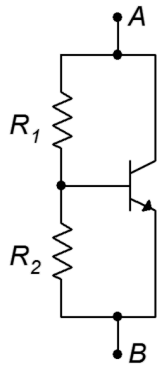


Figure 9.24
Current-sharing output section.

V_{BE} Multiplier and Miller Capacitor

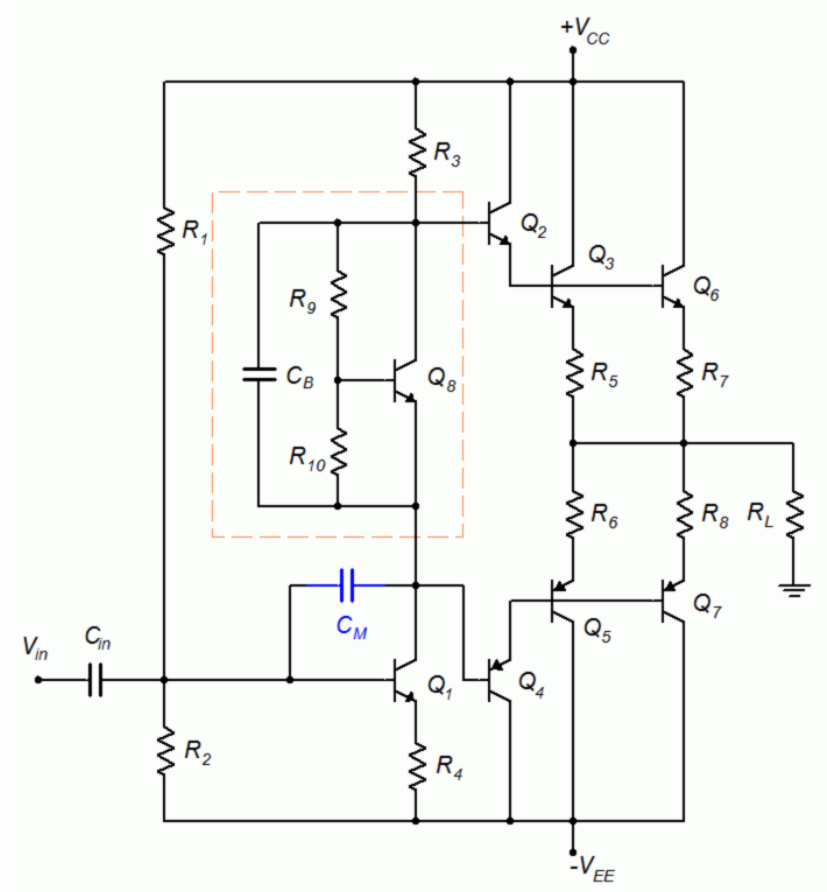
With the ever increasing complexity of the output section, the simple diode biasing scheme lacks the ability to set the optimal idle bias to achieve minimum crossover distortion. A more flexible approach is to use a V_{BE} multiplier, as illustrated in Figure 9.25.

Figure 9.25
V_{BE} multiplier.



If we ignore base current, the currents through the two resistors are identical. Therefore, their voltages have the same ratio as their resistances. R_2 is in parallel with V_{BE} so its voltage must be equal to V_{BE}. Consequently, the voltage across R_1 must be a multiple of V_{BE}. For example, if we want to generate the equivalent of four base-emitter drops from point A to point B, we make R_1 three times as large as R_2. What makes this particularly useful is that we can set any ratio we want, and by replacing either resistor with a potentiometer, we can make that voltage adjustable. An example is shown in Figure 9.26, using the current-sharing amplifier of Figure 9.24. The V_{BE} multiplier is shown in the dashed red box and replaces the four biasing diodes. A capacitor, C_B, shunts the multiplier, making sure it behaves as a short for AC signals.

Figure 9.26
Amplifier with V_{BE} multiplier and Miller compensation capacitor.



The circuit of Figure 9.26 also includes another capacitor, C_M. This is a Miller compensation capacitor. It uses the Miller Effect (see Chapter 6, Miller's Theorem) to create a much larger equivalent input capacitance. This capacitance appears in parallel with the input of Q_1 and creates a lag network. The inclusion of C_M allows the designer to tailor the high frequency response of the amplifier.

Bridging

Some amplifiers employ a *bridged* output scheme. This is particularly true where power supply voltages are limited, for example, in automobiles (nominally 12 VDC but closer to 13.8 VDC from the alternator). Without resorting to an expensive DC-to-DC converter, an amplifier designed for automotive use is in a bit of a bind. If we assume a +12 volt DC source, then a basic class B amplifier would use a capacitor coupled configuration like Figure 9.6. This would yield a V_{CEQ} of 6 volts and a compliance of about 4.2 volts RMS. If we were to use a standard home loudspeaker of 8 Ω, the maximum load power would be a mere 2.25 watts (this is one reason why car audio systems typically use 4 Ω loudspeakers, because it doubles the output power, in this case to 4.5 watts). A bridged output can double this again.

A block diagram of a bridged drive is shown in Figure 9.27. The left half is what we would see normally. On the right half we have a second identical amplifier but we drive it with an inverted copy of the original signal. What ends up happening is that, as the left output goes positive, the right output goes negative by the same amount. The end result is that the load sees twice the voltage it would have from the left amplifier alone. Power varies as the square of voltage so doubling the voltage quadruples the power. The 2.25 watt output of that car amplifier jumps up to 9 watts[37]. If we want to go higher than this our only options are to further decrease the loudspeaker impedance (there are practical limits that will not let us go much further) or increase the DC power supply. In the home, increasing the supply is relatively easy as we have an AC power source. In an automotive system this is a much more expensive proposition because only DC is generally available.

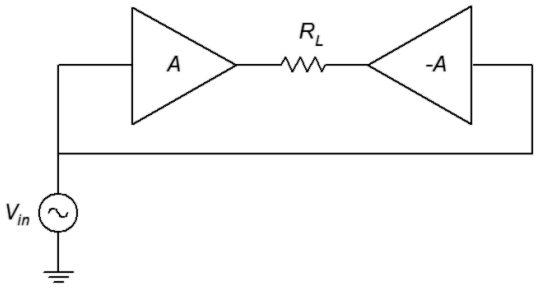


Figure 9.27
Block diagram of bridged amplifier.

37 OK, not huge, but probably still loud enough to obscure the sirens of emergency vehicles until they're right behind you.

An illustration of a bridged system at the transistor level is shown in Figure 9.28. For simplicity, this schematic uses basic class B stages. Input and right-side inversion circuitry are not shown.

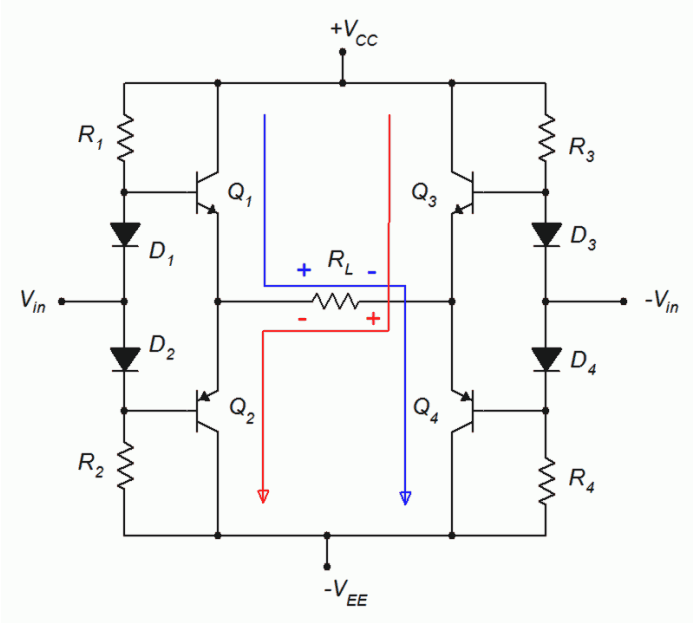


Figure 9.28
Bridged amplifier, transistor level.

The four transistors create a classic H bridge with the load in the center. For a positive V_{in}, Q_1 turns on. At the same time, because it is being fed by an inverted version of V_{in}, Q_4 turns on. Thus, current flows down from V_{CC}, through Q_1, across R_L from left-to-right, down Q_4, and finally to V_{EE} (blue trace). At full swing, the voltage across the load will be $V_{CC} - V_{EE}$ from left to right. For a negative V_{in} the opposite happens: Q_3 and Q_2 turn on, allowing current to flow through R_L from right left. The maximum voltage will be the same but with inverted polarity, hence an effective doubling of load voltage and a quadrupling of load power.

This increased output does not come without downsides. The most obvious problem is a doubling of the output circuitry. The second issue may not be immediately apparent: the load is not grounded. In simple automobile audio systems, the chassis of the vehicle is used as the system common or "ground". Therefore, it is possible to just run a single wire to a loudspeaker from an amplifier's output. The other loudspeaker terminal is then connected to the chassis at a convenient location. With a bridged output, two wires must be run to the loudspeaker. If a single-lead loudspeaker system is upgraded with a bridged output amplifier, new loudspeaker wires must be installed. Otherwise, the chassis return lead winds up shorting out one side of the bridge[38].

38 Oopsy!

Power Supply Bypassing

A final comment regarding high power amplifiers, particularly audio amplifiers, is in order. All through this chapter we have assumed that the DC power supplies will act ideally, mainly that they will present themselves as good AC grounds. This is not always easy to achieve given the practical limitations of PC board layouts, wiring constraints and so forth. Consequently, power supply bypass capacitors are often used to ensure a good AC ground. While this was mentioned in Chapter 7 regarding small signal amplifiers, it is perhaps more important for power amplifiers. The power supply bypass capacitors used for power amplifiers tend to be larger and the quality more stringent. A simple 1 μF bypass capacitor would hardly be the norm for a high power audio amplifier. The practical problem is that large, high quality capacitors are not inexpensive. For example, a 10 μF polypropylene capacitor will be at least an order of magnitude more expensive than a similar size aluminum electrolytic (dollars versus cents). Unfortunately, the electrolytic will have higher leakage and ESR, and will not behave nearly as well at high frequencies (indeed, the impedance will actually start to *increase* due to inductive effects once the frequency gets high enough). To both improve performance and save money, bypass capacitors are sometimes doubled or tripled. For example, a large aluminum electrolytic might be placed in parallel with a much smaller polyester or polypropylene capacitor. The aluminum electrolytic will give the small X_C needed at lower frequencies, and when it starts to behave less ideally at higher frequencies, the higher quality poly capacitor effectively shunts it and extends the operating range. The result is nearly as effective as the single large, high quality capacitor but much less expensive.

Summary

Class B operation is defined as having transistor collector current active for 180° out of the waveform cycle. In order to amplify the entire 360° in linear fashion, two devices are required. Each transistor of a complementary pair is biased at cutoff to achieve a 180° conduction angle. This means that no-signal collector current is zero, leading to very low power consumption at idle and, unlike class A amplifiers, dynamic power consumption. Unfortunately, pure class B operation also results in notch or crossover distortion as the switch over from one transistor to the other is not seamless. This can be mitigated by biasing the devices slightly “on”, which is known as class AB operation. While simple resistor voltage dividers may be used for this purpose, a generally superior method uses diodes as they mimic the base-emitter current-voltage characteristic.

The compliance of a class B amplifier is based on its power supplies. Ideally, the peak-to-peak compliance of the amplifier will equal the total power supply differential. Worst case power dissipation and efficiency are far superior when compared to class A topology: Device power dissipation is only one-fifth of maximum load power and the theoretical efficiency at maximum load power is

78.5%. Maximum heating of the transistors occurs at approximately 40% of maximum load power.

A direct coupled drive stage is often used as it reduces parts count and improves performance. Amplifiers can also be extended through the use of Darlington pairs and current sharing schemes. Other refinements include the use of active current limiting for device protection, and V_{BE} multipliers in place of simple biasing diodes for greater flexibility.

Bridging is a technique used to increase output power. It relies on driving a floating load from both sides. Two amplifiers are needed for this configuration but it can offer a quadrupling of load power for the same power supplies.

Review Questions

1. Define class B operation and compare it to class A.
2. How do the AC and DC load lines differ between class A and class B operation?
3. What is the purpose of the biasing diodes in a class B amplifier?
4. Explain the source of notch distortion and discuss how it can be reduced.
5. What is thermal runaway? How might it be controlled?
6. What is a Sziklai pair?
7. What is a V_{BE} multiplier?
8. Explain the operation of an active current limiter.
9. What is a bridged output configuration? What are its benefits?
10. What's the difference between a complement and a compliment?

Problems

Analysis Problems

1. For the circuit of Figure 9.29, determine compliance, $P_{load(max)}$, $P_{D(max)}$, BV_{CEO} and $I_{C(max)}$. $V_{CC} = 15$ V, $V_{EE} = -15$ V, $\beta = 75$, $R_L = 16\ \Omega$, $R_1 = 680\ \Omega$, $R_2 = 680\ \Omega$.

2. For the circuit of Figure 9.29, determine Z_{in}. $V_{CC} = 15$ V, $V_{EE} = -15$ V, $\beta = 75$, $R_L = 16\ \Omega$, $R_1 = 680\ \Omega$, $R_2 = 680\ \Omega$.

3. For the circuit of Figure 9.29, determine Z_{in}. $V_{CC} = 25$ V, $V_{EE} = -25$ V, $\beta = 70$, $R_L = 8\ \Omega$, $R_1 = 560\ \Omega$, $R_2 = 560\ \Omega$.

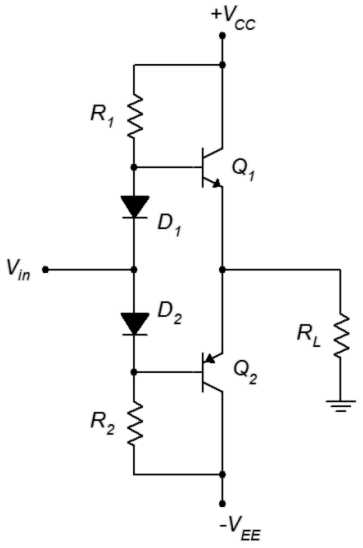


Figure 9.29

4. For the circuit of Figure 9.29, determine compliance $P_{load(max)}$, $P_{D(max)}$, BV_{CEO} and $I_{C(max)}$. V_{CC} = 25 V, V_{EE} = −25 V, β = 70, R_L = 8 Ω, R_1 = 560 Ω, R_2 = 560 Ω.

5. For the circuit of Figure 9.30, determine $P_{load(max)}$, $P_{D(max)}$, BV_{CEO} and $I_{C(max)}$. V_{CC} = 15 V, β = 75, R_L = 16 Ω, R_1 = 630 Ω, R_2 = 630 Ω.

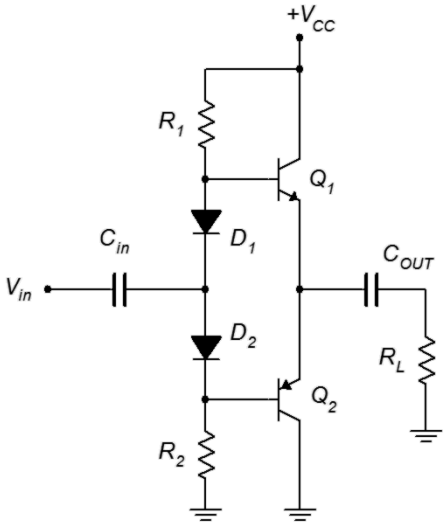


Figure 9.30

6. For the circuit of Figure 9.30, determine Z_{in}. V_{CC} = 15 V, β = 75, R_L = 16 Ω, R_1 = 630 Ω, R_2 = 630 Ω.

7. For the circuit of Figure 9.30, determine Z_{in}. V_{CC} = 25 V, β = 70, R_L = 8 Ω, R_1 = 560 Ω, R_2 = 560 Ω.

8. For the circuit of Figure 9.30, determine $P_{load(max)}$, $P_{D(max)}$, BV_{CEO} and $I_{C(max)}$. $V_{CC} = 25$ V, $\beta = 70$, $R_L = 8\ \Omega$, $R_1 = 510\ \Omega$, $R_2 = 510\ \Omega$.

9. For the circuit of Figure 9.31, determine $P_{load(max)}$, $P_{D(max)}$, BV_{CEO} and $I_{C(max)}$ for the output transistors. $V_{CC} = 24$ V, $V_{EE} = -24$ V, $\beta = 75$, $R_L = 8\ \Omega$, $R_1 = 2.5\ k\Omega$, $R_2 = 300\ \Omega$, $R_3 = 330\ \Omega$, $R_4 = 63\ \Omega$.

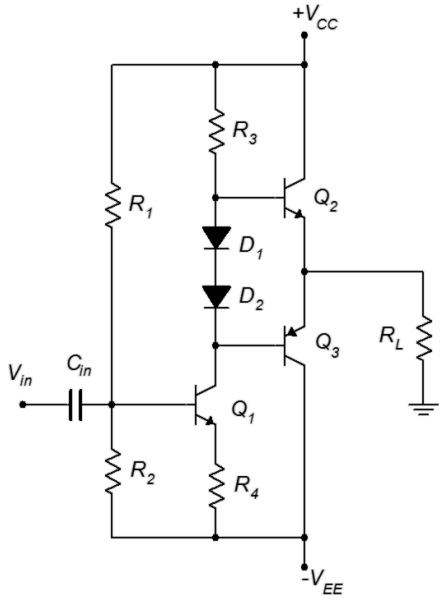


Figure 9.31

10. For the circuit of Figure 9.31, determine A_v and Z_{in}. $V_{CC} = 24$ V, $V_{EE} = -24$ V, $\beta = 75$, $R_L = 8\ \Omega$, $R_1 = 2.5\ k\Omega$, $R_2 = 300\ \Omega$, $R_3 = 330\ \Omega$, $R_4 = 63\ \Omega$.

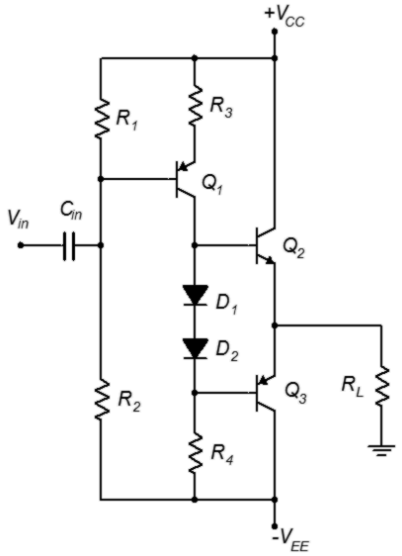


Figure 9.32

11. For the circuit of Figure 9.32, determine $P_{load(max)}$, $P_{D(max)}$, BV_{CEO} and $I_{C(max)}$ for the output transistors. $V_{CC} = 24$ V, $V_{EE} = -24$ V, $\beta = 75$, $R_L = 16\ \Omega$, $R_1 = 600\ \Omega$, $R_2 = 5\ \text{k}\Omega$, $R_3 = 63\ \Omega$, $R_4 = 330\ \Omega$.

12. For the circuit of Figure 9.32, determine A_v and Z_{in}. $V_{CC} = 24$ V, $V_{EE} = -24$ V, $\beta = 75$, $R_L = 16\ \Omega$, $R_1 = 600\ \Omega$, $R_2 = 5\ \text{k}\Omega$, $R_3 = 63\ \Omega$, $R_4 = 330\ \Omega$.

13. Determine the limit current for the circuit of Figure 9.33 if $R_E = 0.2\ \Omega$.

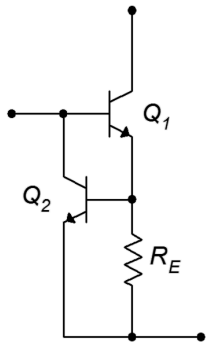


Figure 9.33

14. Determine $P_{load(max)}$, and $P_{D(max)}$, BV_{CEO} and $I_{C(max)}$ for the output and driver transistors of Figure 9.34. $V_{CC} = 50$ V, $V_{EE} = -50$ V, $\beta = 75$, $R_L = 8\ \Omega$, R_5 through $R_8 = 0.05\ \Omega$. Assume all other components produce proper bias.

Design Problems

15. For the circuit of Figure 9.31, determine values for R_1 and R_2 for proper bias. $V_{CC} = 32$ V, $V_{EE} = -32$ V, $\beta = 75$, $R_L = 8\ \Omega$, $R_3 = 330\ \Omega$, $R_4 = 63\ \Omega$.

16. Determine a value for R_E to set the limit current for the circuit of Figure 9.33 to 2 A.

Challenge Problems

17. For the circuit of Figure 9.34, determine values for R_1 and R_2 for proper bias. $V_{CC} = 50$ V, $V_{EE} = -50$ V, $\beta = 85$, $R_L = 8\ \Omega$, R_5 through $R_8 = 0.05\ \Omega$, $R_3 = 2.2\ \text{k}\Omega$, $R_4 = 330\ \Omega$.

18. For the circuit of Figure 9.35, determine a value for R_5 for proper bias. $V_{CC} = 30$ V, $V_{EE} = -30$ V, $\beta = 100$, $R_L = 16\ \Omega$, $R_1 = 2.2\ \text{k}\Omega$, $R_2 = 8.2\ \text{k}\Omega$, $R_3 = 1.2\ \text{k}\Omega$, $R_4 = 47\ \Omega$, $R_5 = 330\ \Omega$, $R_6 = 470\ \Omega$, $R_7 = 68\ \Omega$.

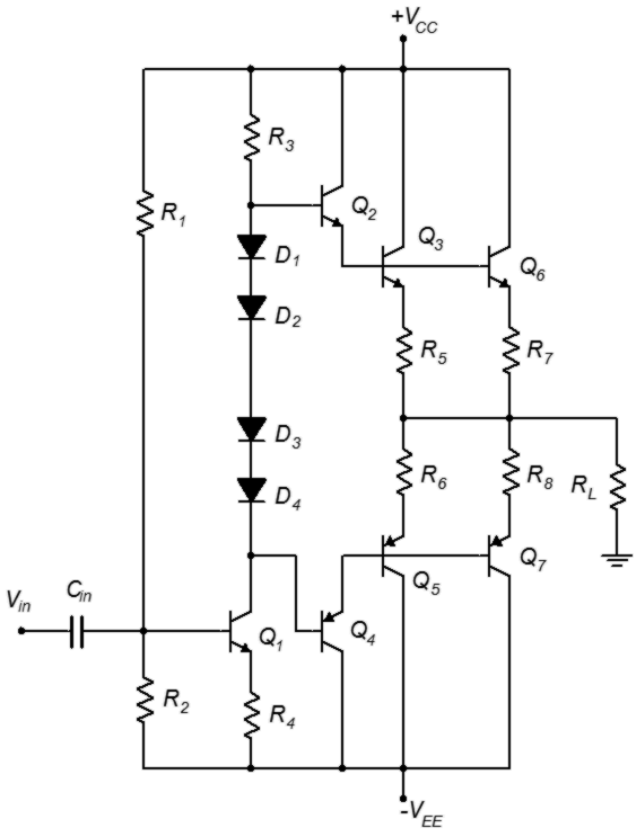


Figure 9.34

Computer Simulation Problems

19. Perform a transient analysis on the circuit of Problem 1 to verify the compliance.

20. Perform a transient analysis on the circuit of Problem 4 to verify the compliance.

21. Perform a DC analysis on the design from Problem 15 to verify the results.

22. Perform a DC analysis on the design from Problem 17 to verify the results.

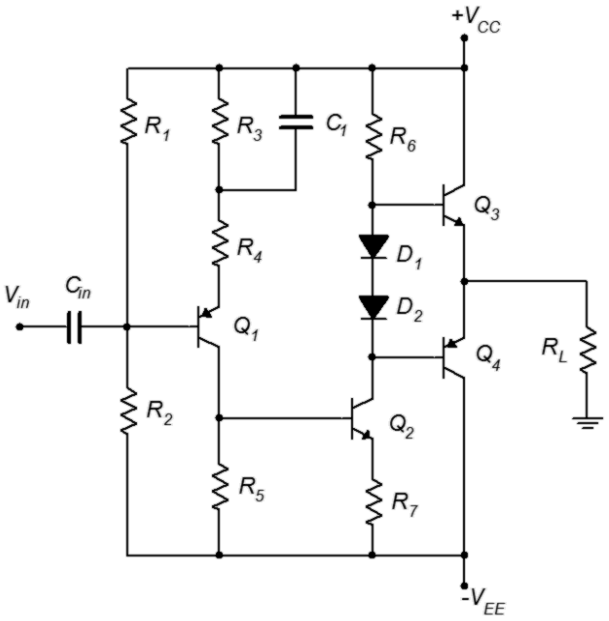


Figure 9.35

The person who invented sliced bread invented neither bread nor knife.

10 Junction Field Effect Transistors (JFETs)

10.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Explain the basic operation of the JFET, detailing the different operating regions.
- Draw and explain a basic DC bias model for a JFET.
- Explain the terms *transconductance* and *pinch-off voltage*.
- Draw a JFET characteristic curve, detailing pertinent elements and points.
- Discuss the advantages and disadvantages of various JFET biasing topologies.
- Analyze basic DC biasing circuits for JFETs.
- Compare and contrast JFET operation to that of BJTs.

10.1 Introduction

The field effect transistor, or FET, is a semiconductor device that serves as an alternative to the bipolar junction transistor. FETs are available in two broad types: the junction FET, or JFET, and the metal oxide semiconductor FET, or MOSFET.

It is best not to think of FETs as either better or worse than the BJT. They have different characteristics and lend themselves to applications where BJT performance might be wanting. The inverse is also true and for some applications the judicious use of a combination of BJTs and FETs can produce superior performance when compared to either device used alone. Like the BJT's NPN and PNP variants, FETs comes in two "flavors": the N-channel type and the P-channel type. We shall cover JFETs first and then discuss MOSFETs in subsequent chapters. In this chapter we will cover the internal structure of the JFET, its theory of operation and biasing techniques. In the next chapter we shall discuss small signal JFET amplifiers; both voltage amplifiers and voltage followers.

The JFET is fundamentally different from a bipolar junction transistor. While the JFET, like the BJT, relies on the PN junction for operation, the JFET is modeled as a voltage-controlled current source while the BJT is modeled as a current-controlled current source. Further, the BJT relies on a forward-biased base-emitter junction for proper operation while the JFET achieves current control via a reverse-biased junction. Consequently, JFET biasing circuits tend to be incompatible with BJT biasing schemes and one device cannot be swapped out for the other.

10.2 JFET Internals

A simplified internal model of a JFET is shown in Figure 10.1. The main portion of the device is called the *channel*. The diagram illustrates an N-channel device. The channel is built upon a *substrate* (i.e., base layer) of oppositely doped material. Attached to the opposing ends of the channel are two terminals; the *source* and the *drain*. Embedded within the channel is a region using the opposite material type. A lead is attached to this as well and is called the *gate*. Although there is not perfect correspondence between them, the drain, source and gate are roughly analogous to the BJT's collector, emitter and base, respectively.

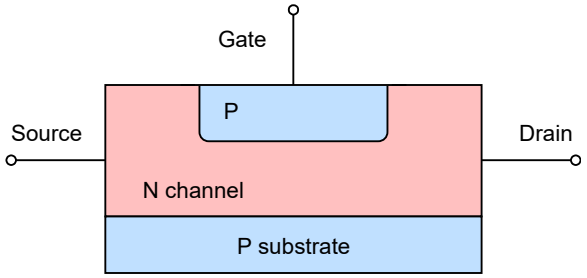


Figure 10.1
Simplified internal structure of an N-channel JFET.

This diagram is drawn symmetrically. Some devices are designed in this fashion and their drain and source terminals can be swapped with no change in operation. This is not true for all devices, though. For small values of drain-source voltage, the channel exhibits a certain amount of resistance that is dependent on the doping level and physical layout of the device. Further, under normal operation, I_D will equal I_S.

To understand how the device behaves, refer to Figure 10.2. Here we shall consider electron flow (shown as a dashed line). First, a positive voltage, V_{DD}, is attached to the drain terminal along with a current limiting resistor, R_D. A negative supply, V_{GG}, is applied to the gate terminal via resistor R_G. Let's start with the gate supply set to zero. If we start V_{DD} at zero, here is what happens as we increase its value. Initially, an increase in drain-source voltage will elicit a proportional increase in the current flowing through the channel. In other words, the channel acts like a resistor. As the voltage across the drain-source increases further, at some point the current will saturate, and no further increases in current will occur in spite of further increases in V_{DD} and V_{DS}. At this point the device is behaving as a constant current source. The drain-source voltage where this transition occurs is called the *pinch-off voltage*, V_p. If the drain-source voltage increases too much, breakdown will occur and current will begin to increase rapidly.

What's particularly interesting is what happens when the gate supply is increased in the negative direction. This reverse-biases the gate-source PN junction and results in a larger depletion region being formed. The depletion region widens into the channel, thus restricting current flow sooner and at a lower level. The more negative we make V_{GG}, the lower I_D becomes. Eventually, when V_{GG} goes negative enough, the

drain current will turn off. This voltage is called $V_{GS(off)}$ and it has the same magnitude as V_P (i.e., $V_P = |V_{GS(off)}|$). The action can be thought of as operating like a water valve: turning the gate source voltage more negative is like turning off the spigot and decreasing the flow.

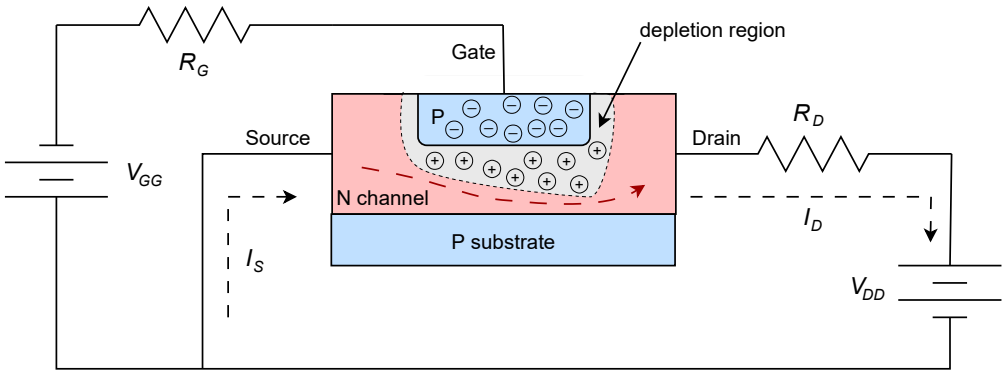


Figure 10.2
Electron flow in an N-channel JFET.

The operation of the JFET can visualized nicely by plotting a set of drain curves, as shown in Figure 10.3.

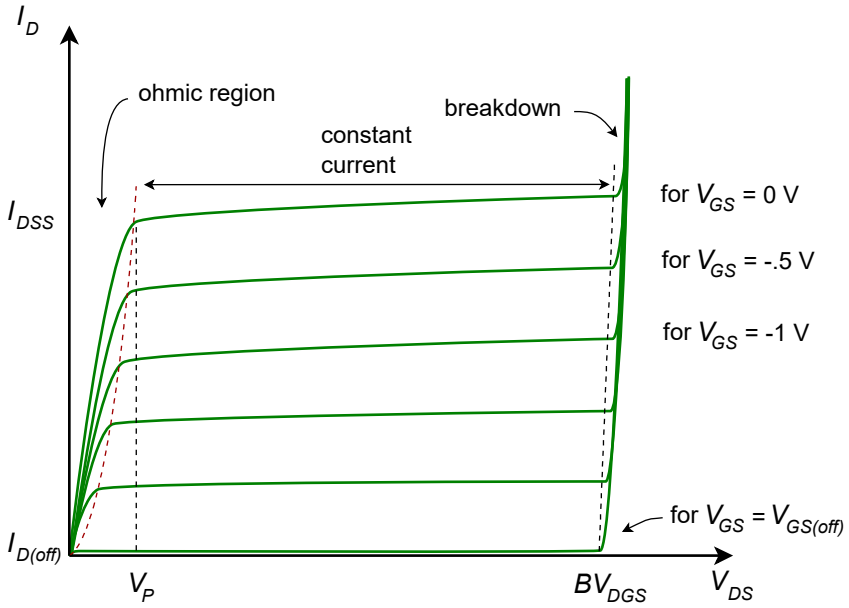


Figure 10.3
JFET drain curves.

The drain curve family plots drain current, I_D, versus drain-source voltage, V_{DS}. We begin with the top-most curve. This is generated by setting the gate-source voltage, V_{GS}, to zero. We then cycle V_{DS} from zero to some higher value. Initially, we see a proportional rise in I_D as V_{DS} increases. This is called the *ohmic* or *triode region*. Eventually, the channel saturates and the current levels out. This is the *constant current* or *saturation region* and it occurs for $V_{DS} > V_P$. The breakdown voltage is called BV_{DGS}, or alternately, $V_{(BR)DS}$. Above this voltage the current increases rapidly. As usual, we do not wish to operate the device in this *breakdown region*.

If we now repeat the process but this time use a small negative value for V_{GS}, we will trace out a curve of very similar shape. The transition to constant current mode will happen at a slightly lower voltage and the current value will be somewhat lower as well. This process continues in like fashion as we make V_{GS} more and more negative. Eventually, when $V_{GS} = V_{GS(off)}$, the drain current drops to virtually zero (in fact, a small leakage current flows called $I_{D(off)}$). In contrast, if V_{GS} was allowed to go positive, operation would be lost because the PN junction would become forward-biased and we would lose control of the current via the depletion region. This means that the JFET's current control is entirely in the second quadrant and the largest drain current flows when $V_{GS} = 0$ V. This current is called I_{DSS}, which stands for the drain current with a shorted gate-source (i.e, if it's shorted, then $V_{GS} = 0$ V). The JFET cannot produce a continuous current larger than I_{DSS} safely.

The characteristic equation relating drain current and gate-source voltage is shown below. This is valid for the constant current region (i.e., $V_{DS} > V_P$).

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)^2 \qquad (10.1)$$

Where
V_{GS} is the gate-source voltage ($V_{GS(off)} \leq V_{GS} \leq 0$),
I_D is the drain current,
I_{DSS} is the maximum current,
$V_{GS(off)}$ is the turn-off voltage.

From this we see that the JFET is a *square-law* device rather than like the BJT which has a logarithmic characteristic.[39] In essence, this curve is a portion of a parabola. This means that the JFET's characteristic curve is much more gradual in slope than that of a BJT. This will have important implications when it comes to voltage gain potential and distortion, as we shall see in the following chapter.

It is useful to remember that $V_{GS(off)}$ and I_{DSS} are unique to a given device, rather like β is for a BJT. There can also be a fairly large variation in these parameters. For example, a particular model of JFET might show an I_{DSS} variation between 2 mA and 20 mA, and a $V_{GS(off)}$ variation between −2 V and −8 V. Generally, the most negative $V_{GS(off)}$ values will be associated with the largest I_{DSS} values.

Equation 10.1 is plotted in Figure 10.4. Compare this curve to the curve generated by the Shockley equation for BJTs, Figure 7.1. The graph is shown in normalized form. Instead of plotting for specific values of $V_{GS(off)}$ and I_{DSS}, the axes are presented as fractional portions of the maximums (i.e., the horizontal axis is $-V_{GS}/V_{GS(off)}$ and the vertical axis is I_D/I_{DSS}).

39 As evidenced in the Shockley equation, Equation 2.1.

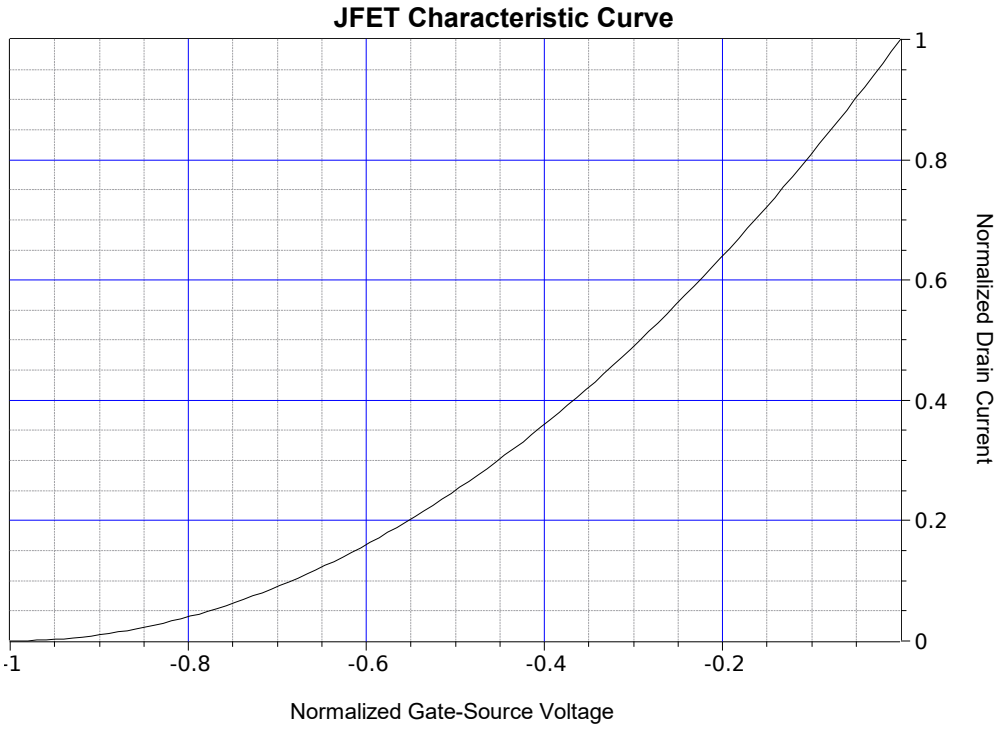


Figure 10.4
JFET normalized characteristic curve (note: this uses $-V_{GS}/V_{GS(off)}$ for the normalized voltage so that the curve does not appear reversed compared to a typical device curve).

Example 10.1

Using both Equation 10.1 and the graph of Figure 10.4, determine the drain current if the gate-source voltage is −1 V and the JFET specs are I_{DSS} = 8 mA and $V_{GS(off)}$ = −2 V.

First, using Equation 10.1

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)^2$$

$$I_D = 8\,\text{mA}\left(1 - \frac{-1\,\text{V}}{-2\,\text{V}}\right)^2$$

$$I_D = 2\,\text{mA}$$

Using the graph, $V_{GS}/V_{GS(off)}$ is 1 V/−2 V, or −0.5. Find this value on the horizontal axis, follow up to the curve and then across to the right vertical axis. The normalized drain current is 0.25, thus I_D is 0.25 I_{DSS}, or 2 mA.

As the characteristic curve plots output current versus input voltage, the slope of this represents the *transconductance*, an important characteristic for biasing and signal analysis. Device transconductance is denoted as g_m, or alternately as g_{fs}, and given

units of siemens. We can derive an equation for transconductance by taking the derivative of Equation 10.1.

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)^2$$

$$\frac{d\,I_D}{d\,V_{GS}} = -\frac{2\,I_{DSS}}{V_{GS(off)}}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)$$

The coefficient $-2\,I_{DSS}/V_{GS(off)}$ is defined as g_{m0}, the transconductance when $V_{GS} = 0$ V. This is the maximum transconductance of the device. Substituting, we arrive at

$$g_{m0} = -\frac{2\,I_{DSS}}{V_{GS(off)}} \tag{10.2}$$

$$g_m = g_{m0}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right) \tag{10.3}$$

A normalized plot of transconductance versus V_{GS} is shown in Figure 10.5. The horizontal axis is $-V_{GS}/V_{GS(off)}$ and the vertical axis is g_m/g_{m0}.

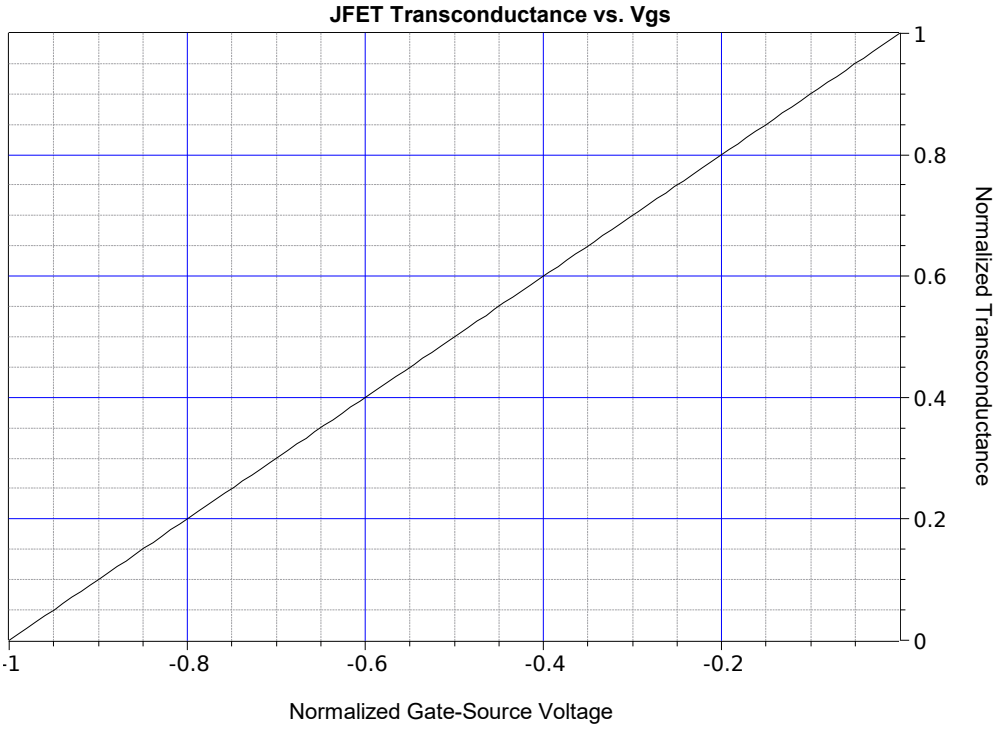


Figure 10.5
Curve of transconductance.

From this graph we see that the transconductance is a linear function.

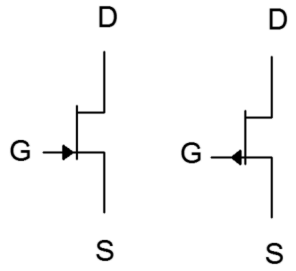
Another item of interest regarding these device equations: If we combine Equations 10.1 and 10.3, we generate two equations that will prove useful in upcoming work.

$$\frac{g_m}{g_{m0}} = \sqrt{\frac{I_D}{I_{DSS}}} \tag{10.4}$$

$$\frac{I_D}{I_{DSS}} = \left(\frac{g_m}{g_{m0}}\right)^2 \tag{10.5}$$

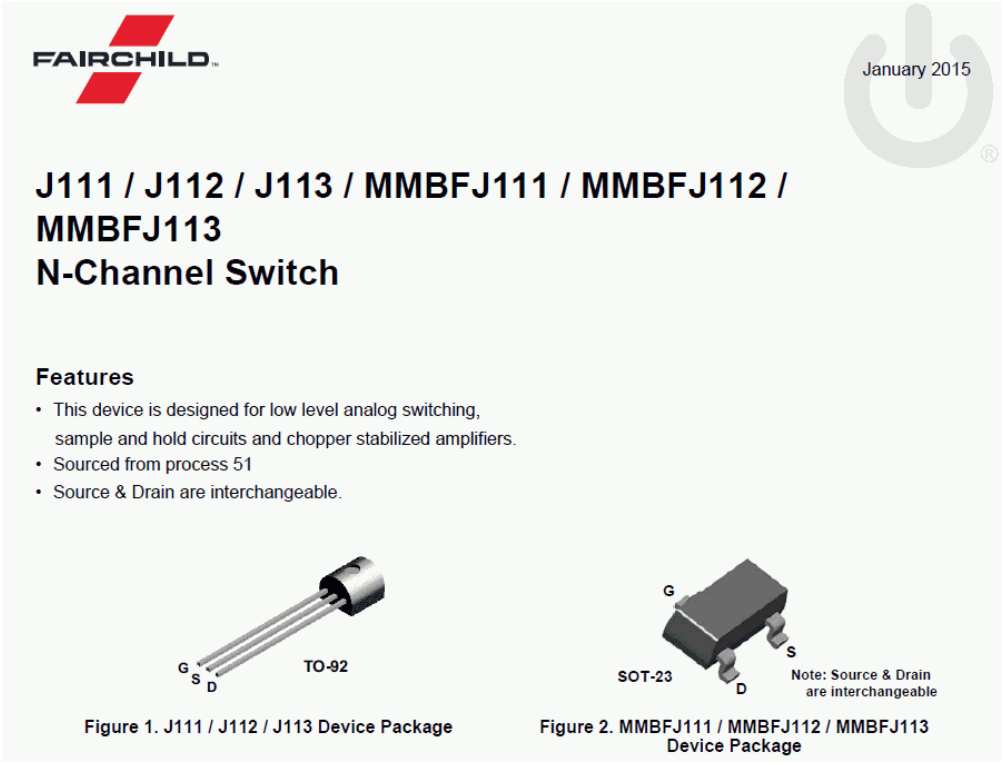
Before moving on, the schematic symbols for JFETs are shown in Figure 10.6. The middle vertical line represents the channel, and as is usually the case, the arrow points to N material. Sometimes the gate arrow is draw in the middle rather than toward the source. Also, as is the case the BJT, sometimes these symbols are drawn within a circle.

Figure 10.6
JFET schematic symbols:
N-Channel (left)
P-Channel (right)



10.3 JFET Data Sheet Interpretation

A data sheet for the J111 series N-channel JFET is shown in Figure 10.7. This is a small signal device designed for audio frequency circuits. It is available in the common TO-92 through-hole package as well as in the surface mount SOT-23 package. Note that the source and drain are interchangeable for this device.



FAIRCHILD

January 2015

J111 / J112 / J113 / MMBFJ111 / MMBFJ112 / MMBFJ113
N-Channel Switch

Features

- This device is designed for low level analog switching, sample and hold circuits and chopper stabilized amplifiers.
- Sourced from process 51
- Source & Drain are interchangeable.

Figure 1. J111 / J112 / J113 Device Package

Figure 2. MMBFJ111 / MMBFJ112 / MMBFJ113 Device Package

Figure 10.7a
J111 series N-channel JFET data sheet.
Used with permission from SCILLC dba ON Semiconductor.

Examining the absolute maximum ratings and thermal characteristics, we find values typical of small signal devices. Maximum drain-gate and gate-source voltages are 35 volts and the maximum power dissipation is 625 milliwatts.

Absolute Maximum Ratings(1), (2)

Stresses exceeding the absolute maximum ratings may damage the device. The device may not function or be operable above the recommended operating conditions and stressing the parts to these levels is not recommended. In addition, extended exposure to stresses above the recommended operating conditions may affect device reliability. The absolute maximum ratings are stress ratings only. Values are at T_A = 25°C unless otherwise noted.

Symbol	Parameter	Value	Unit
V_{DG}	Drain-Gate Voltage	35	V
V_{GS}	Gate-Source Voltage	-35	V
I_{GF}	Forward Gate Current	50	mA
T_J, T_{STG}	Operating and Storage Junction Temperature Range	-55 to 150	°C

Notes:

1. These ratings are based on a maximum junction temperature of 150°C.
2. These are steady-state limits. Fairchild Semiconductor should be consulted on applications involving pulsed or low-duty-cycle operations.

Thermal Characteristics

Values are at T_A = 25°C unless otherwise noted.

Symbol	Parameter	Max.		Unit
		J111 / J112 / J113(3)	MMBFJ111 / MMBFJ112 / MMBFJ113(4)	
P_D	Total Device Dissipation	625	350	mW
	Derate Above 25°C	5.0	2.8	mW/°C
$R_{\theta JC}$	Thermal Resistance, Junction-to-Case	125		°C/W
$R_{\theta JA}$	Thermal Resistance, Junction-to-Ambient	200	357	°C/W

Notes:

3. PCB size: FR-4, 76 mm x 114 mm x 1.57 mm (3.0 inch x 4.5 inch x 0.062 inch) with minimum land pattern size.
4. Device mounted on FR-4 PCB 36mm × 18mm × 1.5mm; mounting pad for the collector lead minimum 6cm^2.

Figure 10.7b
J111 series N-channel JFET data sheet (cont).

Electrical Characteristics

Values are at T_A = 25°C unless otherwise noted.

Symbol	Parameter	Conditions		Min.	Max.	Unit
Off Characteristics						
$V_{(BR)GSS}$	Gate-Source Breakdown Voltage	I_G = -1.0 µA, V_{DS} = 0		-35		V
I_{GSS}	Gate Reverse Current	V_{GS} = -15 V, V_{DS} = 0			-1.0	nA
V_{GS}(off)	Gate-Source Cut-Off Voltage	V_{DS} = 15 V, I_D = 1.0 µA	**111**	-3.0	-10.0	V
			112	-1.0	-5.0	
			113	-0.5	-3.0	
I_D(off)	Drain Cutoff Leakage Current	V_{DS} = 5.0 V, V_{GS} = -10 V			1.0	nA
On Characteristics						
I_{DSS}	Zero-Gate Voltage Drain Current(5)	V_{DS} = 15 V, V_{GS} = 0	**111**	20		mA
			112	5.0		
			113	2.0		
r_{DS}(on)	Drain-Source On Resistance	V_{DS} ≤ 0.1 V, V_{GS} = 0	**111**		30	Ω
			112		50	
			113		100	

Figure 10.7c
J111 series N-channel JFET data sheet (cont).

From the electrical characteristics, note the large variation in $V_{GS(off)}$. For the J111, this runs from a minimum of −3 V to a maximum of −10 V. The J112 and J113 exhibit even wider min/max ratios. Also, note how the larger $V_{GS(off)}$ ranges are associated with larger maximums for I_{DSS}.

Finally, let's take a look at a series of performance curves shown in Figure 10.7d.

Figure 10.7d
J111 series N-channel JFET data sheet (cont).

Typical Performance Characteristics

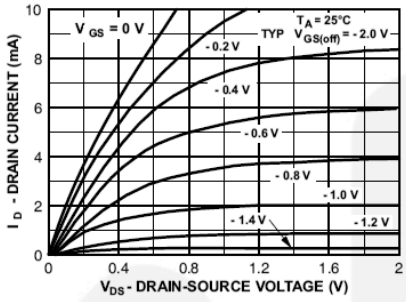


Figure 3. Common Drain-Source

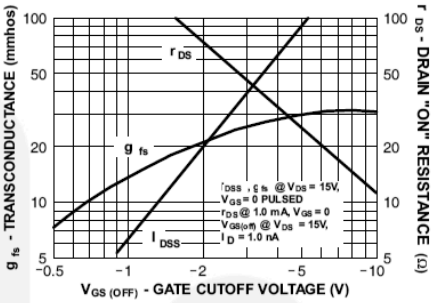


Figure 4. Parameter Interactions

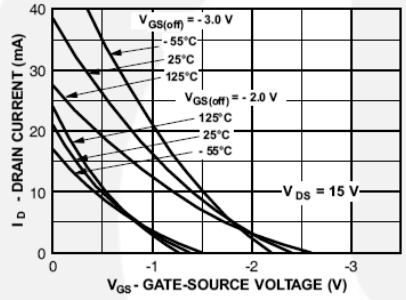


Figure 5. Transfer Characteristics

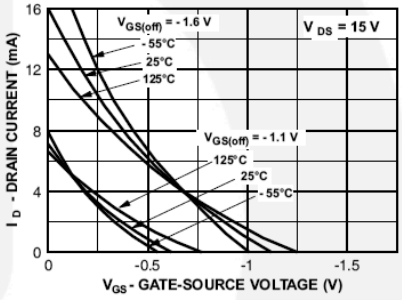


Figure 6. Transfer Characteristics

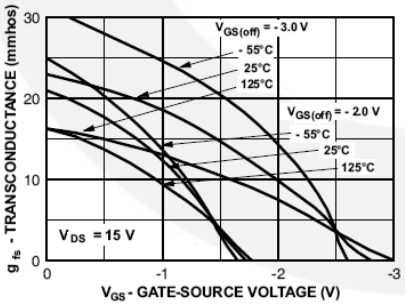


Figure 7. Transfer Characteristics

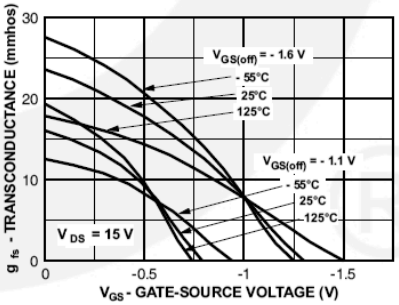


Figure 8. Transfer Characteristics

The upper-left graph is a family of drain curves and corresponds to Figure 10.3, presented earlier. Note that the horizontal scale for V_{DS} only shows the first 2 volts. The point here is to examine the ohmic region. The two middle graphs plot the

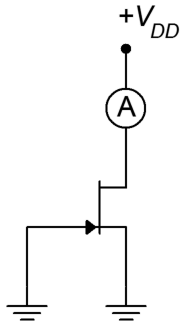
characteristic curve of the device and correspond to Figure 10.4, although these graphs are drawn rotated around the vertical axis (note that V_{GS} is still shown as a negative value). Two important things may be noted here. First, as already mentioned, large values of I_{DSS} tend to be associated with large values of $V_{GS(off)}$. This graph shows that individual plots tend to scale both horizontally and vertically away from the origin. Second, thermal variations are very much apparent: As the temperature increases, the characteristic curve tends to become less steep.

Finally, the two bottom-most graphs plot the variation of g_m with V_{GS}. These correspond to Figure 10.5, although again, the horizontal axis has been rotated around the vertical. Once again we see considerable variation due to temperature. Also, none of the plots exhibit perfect linearity. Further, at lower temperatures, the linearity of the plots decreases even more, warping a relatively straight line into a complex curve.

10.4 JFET Biasing

There are several different ways of biasing a JFET. For many configurations, I_{DSS} and $V_{GS(off)}$ will be needed. A simple way to measure these parameters in the lab is shown in Figure 10.8. To measure I_{DSS} we simply ground the gate and source terminals as this forces V_{GS} to be 0 V. We insert an ammeter between V_{DD} and the drain, and then set V_{DD} to a value higher than V_P (+15 VDC generally being sufficient). The resulting ammeter reading is I_{DSS}. Obtaining $V_{GS(off)}$ is only slightly more work. Leaving the ammeter in the drain, unhook the gate from ground and instead connect it to an adjustable negative power supply. Turn the supply more negative until the ammeter reads zero (practically speaking, < 1% of I_{DSS}). At that point the voltage source will be equal to $V_{GS(off)}$.

Figure 10.8
Measuring I_{DSS} and $V_{GS(off)}$.



DC Model

Before we begin examining the bias circuits themselves, we need a basic DC model of the JFET. A model sufficient for our analyses is shown in Figure 10.9.

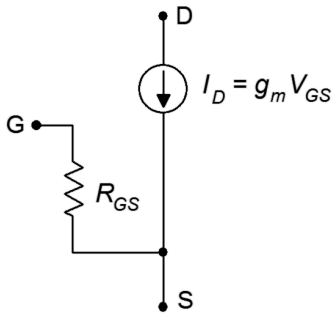


Figure 10.9
DC model of JFET.

The model consists of a voltage-controlled current source, I_D, that is equal to the product of the gate-source voltage, V_{GS}, and the transconductance, g_m. The resistance between the gate and source, R_{GS}, is that of the reverse-biased PN junction, in other words, ideally infinity for DC. As a consequence, in most practical circuits we can assume that gate current, I_G, is zero. Therefore, $I_D = I_S$.

Constant Voltage Bias

The simplest form of bias is the *constant voltage bias*. The prototype is shown in Figure 10.10 with current directions and voltage polarities shown.

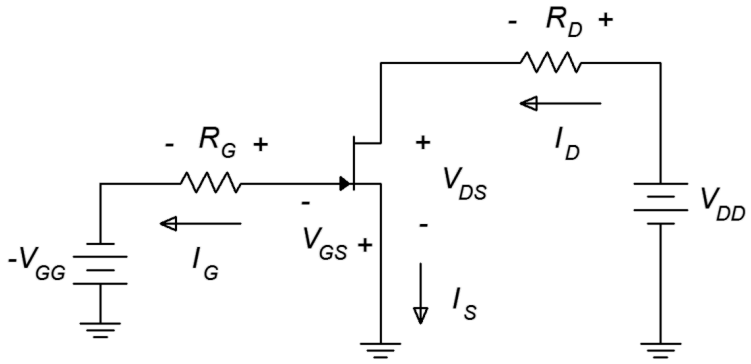


Figure 10.10
Constant voltage bias prototype.

This is a fairly straightforward design using only a couple of resistors and power sources. Figure 10.11 shows the same circuit but with the JFET model inserted, ready for analysis.

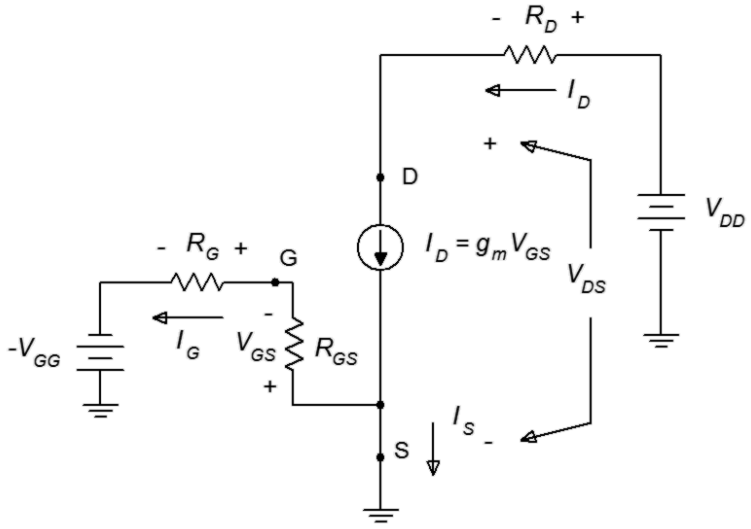


Figure 10.11
Constant voltage bias with model.

Ultimately, the goal here is to determine a means for finding the transistor's drain current and drain-source voltage, along with the potentials across any other components.

To begin, consider the gate-source loop. By KVL, the V_{GG} source must drop across R_G and the gate-source junction, V_{GS}.

$$V_{GG} = V_{R_G} + V_{GS}$$
$$V_{GG} = I_G R_G + V_{GS}$$

I_G is approximately zero so this simplifies to

$$V_{GS} = V_{GG}$$

Given the transconductance, g_m, we can find I_D. Alternately, I_D may be found using Equation 10.1 along with the device parameters I_{DSS} and $V_{GS(off)}$. For this circuit, the latter technique tends to be more practical. Once I_D is found, the voltage drop across R_D may be found, and then V_{DS} is determined from KVL.

Example 10.2

For the circuit of Figure 10.12, determine I_D and V_{DS}. Assume $I_{DSS} = 10$ mA and $V_{GS(off)} = -5$ V.

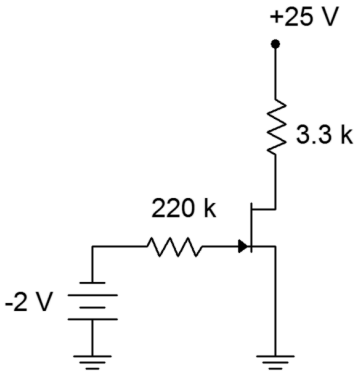


Figure 10.12
Schematic for Example 10.2.

First, because $I_G \approx 0$, the drop across R_G is ≈ 0 and $V_{GS} = V_{GG}$. Using Equation 10.1

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)^2$$
$$I_D = 10\,\text{mA}\left(1 - \frac{-2\,\text{V}}{-5\,\text{V}}\right)^2$$
$$I_D = 3.6\,\text{mA}$$

Looking at the drain-source loop, KVL shows

$$V_{DD} = I_D R_D + V_{DS}$$
$$V_{DS} = V_{DD} - I_D R_D$$
$$V_{DS} = 25\,\text{V} - 3.6\,\text{mA} \times 3.3\,\text{k}\Omega$$
$$V_{DS} = 13.1\,\text{V}$$

While the computation for the constant voltage bias is relatively simple, it does not exhibit a stable Q point. For example, if Example 10.2 is repeated with another JFET, this one with $I_{DSS} = 12$ mA and $V_{GS(off)} = -6$ V, the results are starkly different: I_D grows to 5.33 mA and V_{DS} shrinks to 7.4 V. These are considerable changes given the relatively modest shifts in the device parameters. In this regard, the constant voltage bias is reminiscent of the simple base bias configuration used with BJTs.

To get a better understanding of the Q point stability issue, refer to Figure 10.13.

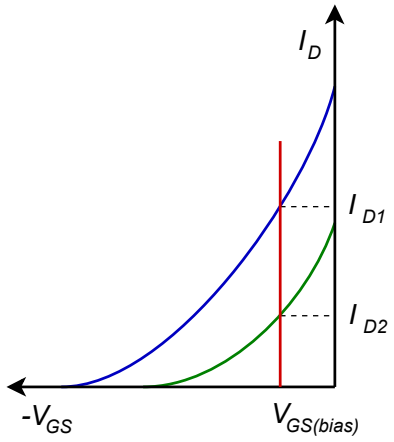


Figure 10.13
Variation for constant voltage bias.

Characteristic curves are plotted here for two different devices, one in green and one in blue. These represent the sort of device parameter variations we might expect to see across a product model. The fixed value of gate bias voltage is shown in red. From this graph it should be obvious that this form of bias will produce a wide variation in drain current, and thus, is not a good choice for applications that require a stable Q point. If the application does not have this requirement, constant voltage bias offers the advantage of requiring a minimum of components.

Self Bias

Self bias uses a small number of components and only a single power supply, yet it offers better stability than constant voltage bias. The name comes from the fact that the drain current will be used to create a voltage drop that sets up the gate-source, hence the circuit "biases itself". It is also referred to as *automatic bias*. The self bias prototype is shown in Figure 10.14.

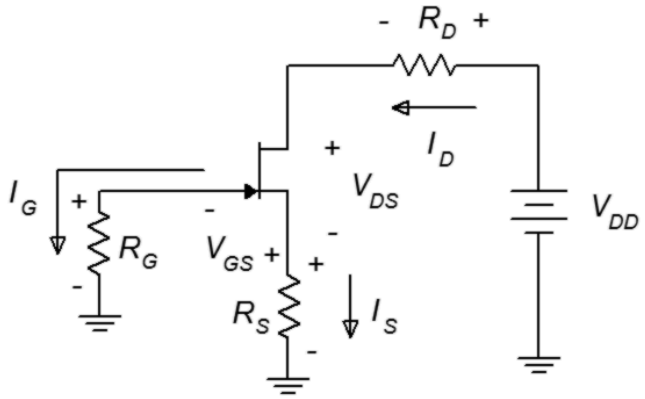


Figure 10.14
Self bias prototype.

Once again, we may assume that I_G is 0. As R_G is connected directly to ground, this means that $V_G \approx 0$ V. This being true, inspection of the schematic reveals that the magnitude of V_{GS} must be the same as the voltage across R_S. Because $I_D = I_S$ then

$$V_{GS} = -I_D R_S \tag{10.6}$$

This value of V_{GS} is what generates the drain current. The definition is self-referential. This being the case, how do we analyze the circuit? A proper derivation of the equation for drain current is not trivial. We start with the characteristic equation (Equation 10.1) and expand it.

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)^2$$

$$I_D = I_{DSS}\left(1 - \frac{2\,V_{GS}}{V_{GS(off)}} + \frac{{V_{GS}}^2}{{V_{GS(off)}}^2}\right)$$

$$I_D = I_{DSS} - \frac{2\,I_{DSS} V_{GS}}{V_{GS(off)}} + \frac{I_{DSS}\,{V_{GS}}^2}{{V_{GS(off)}}^2}$$

Substitute using Equation 10.2

$$g_{m0} = -\frac{2\,I_{DSS}}{V_{GS(off)}}$$

$$I_D = I_{DSS} + g_{m0} V_{GS} + \frac{I_{DSS}\,{V_{GS}}^2}{{V_{GS(off)}}^2}$$

Using Equation 10.6 this can be expanded to

$$I_D = I_{DSS} - g_{m0} I_D R_S + \frac{I_{DSS}\,{I_D}^2 {R_S}^2}{{V_{GS(off)}}^2}$$

Rearranging yields

$$0 = \frac{I_{DSS}\,{R_S}^2}{{V_{GS(off)}}^2}\,{I_D}^2 - (1 + g_{m0} R_S)\,I_D + I_{DSS}$$

This is a quadratic equation in the form ax^2+bx+c and can be solved using the quadratic formula:

$$y = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

The positive option in the numerator may be ignored as this occurs for V_{GS} beyond $V_{GS(off)}$. The result is

$$I_D = 2\,I_{DSS}\left(\frac{1+g_{m0}\,R_S-\sqrt{1+2\,g_{m0}\,R_S}}{(g_{m0}\,R_S)^2}\right) \qquad (10.7)$$

Although this is an accurate analytical solution, it's certainly not the sort of equation most people want to memorize or derive as needed. As the $g_{m0}\,R_S$ term is repeated in this equation multiple times, it is useful to plot this equation in terms of normalized I_D versus $g_{m0}\,R_S$. This curve is plotted in Figure 10.15.

To use this curve, the first step is to find $g_{m0}\,R_S$. The value of R_S is determined by inspection and g_{m0} may be determined by Equation 10.2, repeated below for convenience.

$$g_{m0} = -\frac{2\,I_{DSS}}{V_{GS(off)}}$$

The value of $g_{m0}\,R_S$ is found on the horizontal axis, traced up to the curve and then over to the normalized I_D ratio. This number is multiplied by I_{DSS} to determine the value of I_D.

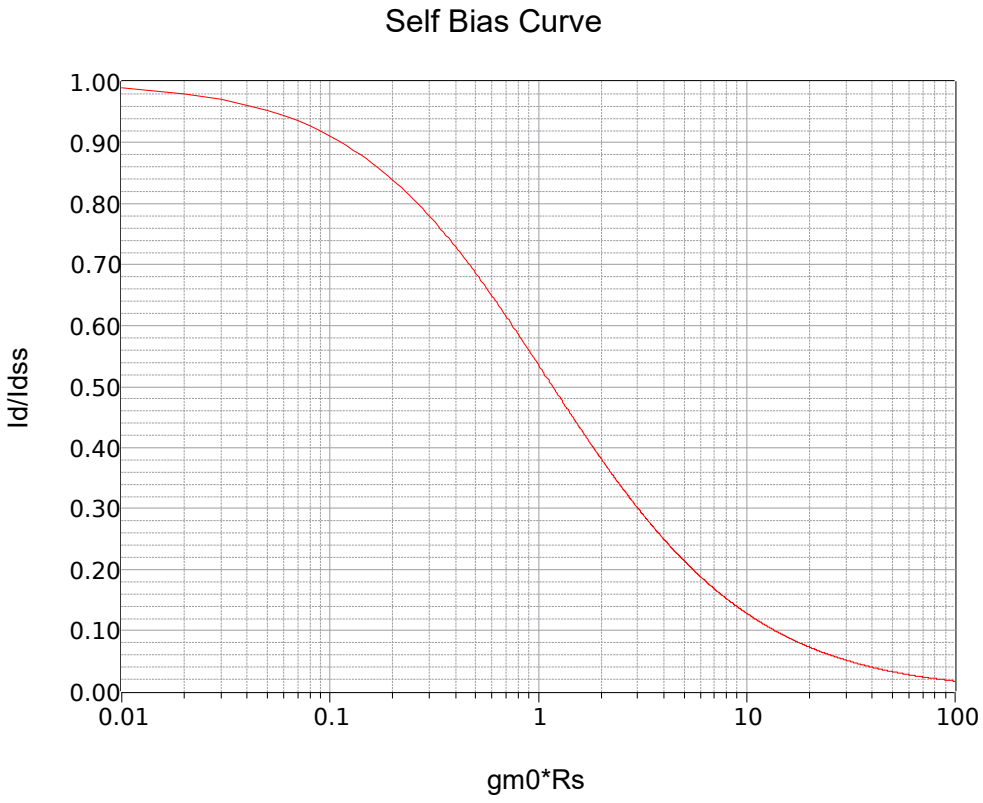
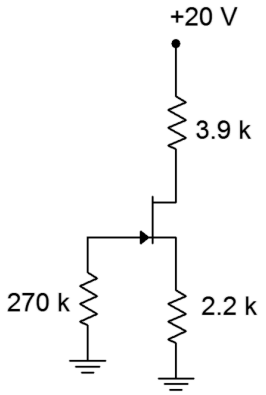


Figure 10.15
Self bias curve.

Example 10.3

Determine I_D and V_{DS} for the circuit shown in Figure 10.16. Assume $I_{DSS} = 10$ mA and $V_{GS(off)} = -4$ V.

Figure 10.16
Schematic for Example 10.3.



Using the graphical method, first determine $g_{m0} R_S$.

$$g_{m0} = -\frac{2 I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{2 \times 10\,\text{mA}}{-4\,\text{V}}$$
$$g_{m0} = 5\,\text{mS}$$

Therefore $g_{m0} R_S = 5$ mS · 2.2 k Ω = 11. The self bias graph yields approximately 0.12 for the normalized current ratio. Therefore

$$I_D = 0.12\, I_{DSS}$$
$$I_D = 0.12 \times 10\,\text{mA}$$
$$I_D = 1.2\,\text{mA}$$

Using Ohm's law and KVL

$$V_D = V_{DD} - I_D R_D$$
$$V_D = 20\,\text{V} - 1.2\,\text{mA} \times 3.9\,\text{k}\Omega$$
$$V_D = 15.32\,\text{V}$$

$$V_S = I_D R_S$$
$$V_S = 1.2\,\text{mA} \times 2.2\,\text{k}\Omega$$
$$V_S = 2.64\,\text{V}$$

$$V_{DS} = V_D - V_S$$
$$V_{DS} = 15.32\,\text{V} - 2.64\,\text{V}$$
$$V_{DS} = 12.68\,\text{V}$$

An alternate technique is to make an initial guess for V_{GS}, typically one half of $V_{GS(off)}$. The value of I_D is then computed from the characteristic equation (Equation 10.1) and compared with the Ohm's law relation, Equation 10.6, rewritten as $I_D = -V_{GS}/R_S$. Chances are, the two results will not agree so adjust the V_{GS} estimate and repeat the process. If done properly, the currents should be closer. Iterate this process until you converge on the answer.

To use this technique for the preceding problem we'd start by assuming $V_{GS} = -2$ V (half of $V_{GS(off)}$). Using this in Equation 10.1 yields $I_D = 2.5$ mA, while using Equation 10.6 produces $I_D = 910$ µA. Obviously the initial estimate was not correct. The second estimate for V_{GS} needs to increase negatively as this will decrease the result from Equation 10.1 and increase the result from Equation 10.6, hopefully meeting in the middle. We might try −2.5 volts. This will yield 1.4 mA from Equation 10.1 and 1.14 mA from Equation 10.6. As the gap has narrowed, the adjustment for the third estimate will be smaller, so we could try −2.6 volts. This would be relatively close to the value as computed in Example 10.3 ($V_{GS} = -V_S$).

This approximation technique also offers a clue as to how self bias gains stability over constant voltage bias. If for some reason I_D was to increase, this would create a larger voltage drop across R_S. Because this voltage is the same magnitude as V_{GS}, this means that V_{GS} grows negatively. A more negative V_{GS} reduces I_D, thus opposing the initial change in drain current. This feedback mechanism is similar in function to the BJT collector feedback bias. The stability issue is visualized in Figure 10.17.

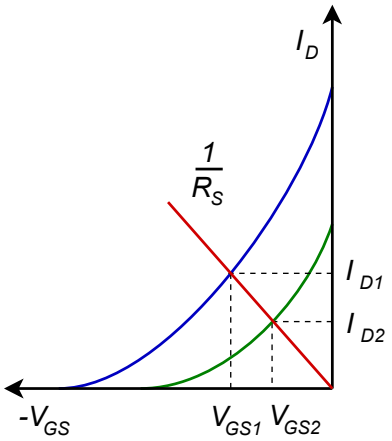


Figure 10.17
Variation for self bias.

Two device curves are plotted to represent parameter variation (green and blue). Equation 10.6 shows the relationship between I_D and V_{GS}. If we put this in the form $y = mx + b$, we find that the line goes through the origin and has a slope of $1/R_S$. This line is plotted in red. Where the line intersects the device curve yields the drain current and gate-source voltage for that particular device. Unlike constant voltage bias, self bias shifts some variation over to V_{GS}, making I_D more stable. In fact, if there is a particular design target for I_D or V_{GS}, a rearrangement of Equation 10.6 can be used to find the needed value of R_S along with the characteristic curve or equation.

$$R_S = -\frac{V_{GS}}{I_D}$$

For example, if a certain I_D is desired, this value could be used with Equation 10.1 to determine the corresponding V_{GS}. These values are then used to find the required R_S. Alternately, the normalized values could be obtained via Figure 10.4.

Example 10.4

Determine a value for R_S to set $V_{GS} = -2$ V for the circuit shown in Figure 10.18. Assume $I_{DSS} = 20$ mA and $V_{GS(off)} = -4$ V.

We can determine the drain current using Equation 10.1.

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)^2$$

$$I_D = 20\,\text{mA}\left(1 - \frac{-2\,\text{V}}{-4\,\text{V}}\right)^2$$

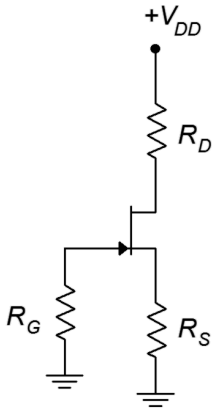
$$I_D = 5\,\text{mA}$$

$$R_S = -\frac{V_{GS}}{I_D}$$

$$R_S = -\frac{-2\,\text{V}}{5\,\text{mA}}$$

$$R_S = 400\,\Omega$$

Figure 10.18
Schematic for Example 10.4.



In sum, self bias is a minimal parts count circuit that offers modest stability. The stability can be improved with the addition of other components, as we shall see with the next bias configuration.

Combination Bias

The *combination bias* configuration (AKA *source bias*) is based on self bias but adds a negative power supply connected to R_S, hence its name. This will enhance the stability of I_D, V_{DS} and g_m. The combination bias prototype is shown in Figure 10.19.

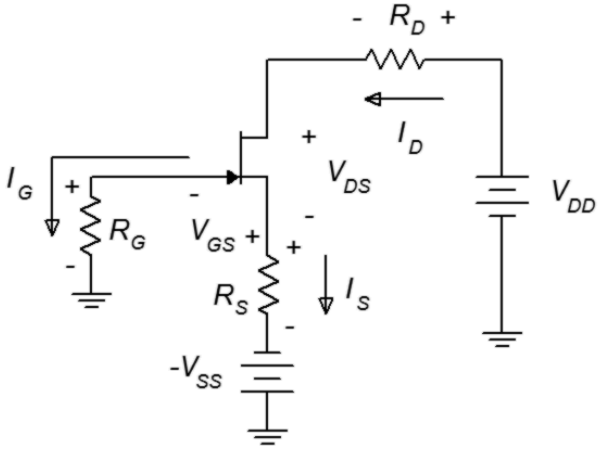


Figure 10.19
Combination bias prototype.

The analysis is similar to that of self bias but with one major twist: the source power supply increases the voltage drop across R_S. This stabilizes the voltage (and hence, the current) because it is no longer equal to $-V_{GS}$, but rather

$$V_{R_S} = I_D R_S = |V_{GS}| + |V_{SS}| \tag{10.8}$$

If $V_{SS} >> V_{GS}$, then we can approximate I_D as V_{SS}/R_S. As with self bias, an analytical solution for I_D is possible. In order to do so, we would begin with the characteristic equation and Equation 10.8. The derivation is left as an exercise.

$$I_D = 2 I_{DSS} \left(\frac{1 + g_{m0} R_S (1+k) - \sqrt{1 + 2 g_{m0} R_S (1+k)}}{(g_{m0} R_S)^2} \right) \tag{10.9}$$

The formula is very similar to the self bias formula but with the addition of a factor, k. k is a "swamping factor" and is defined as the ratio of V_{SS} to $V_{GS(off)}$. If $k = 0$, there is no source power supply and the formula reverts back to the simpler self bias formula. On the other hand, if k is very large, $I_D \approx V_{SS}/R_S$.

As was the case with self bias, we can plot Equation 10.9 using the $g_{m0}R_S$ factor. A series of three plots for $k = 2$, 3 and 4 are rendered in Figure 10.20.[40]

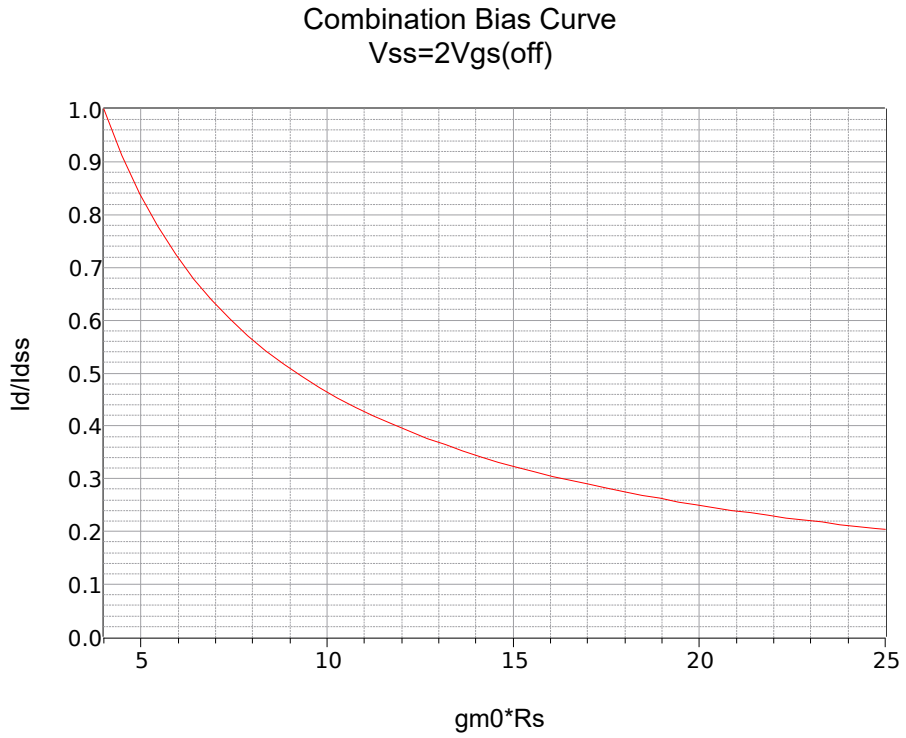


Figure 10.20a
Combination bias curve, k = 2.

40 We could add a third axis for k and plot a surface, and while it might be pretty, a 3D plot like this rendered onto a 2D surface, such as a page in a textbook, is of marginal utility.

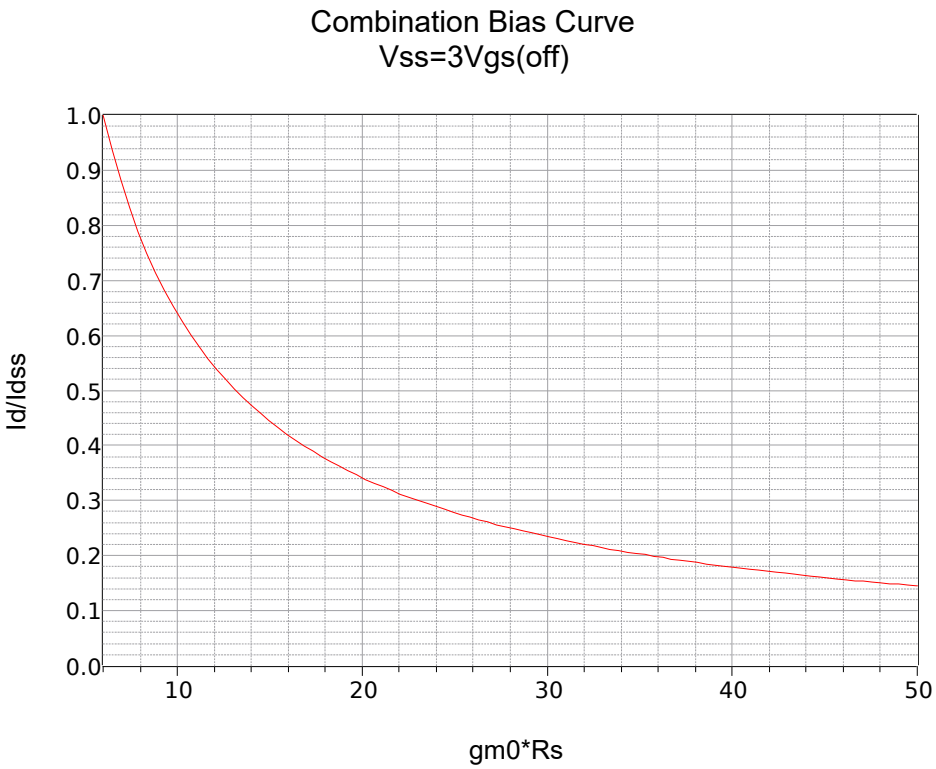


Figure 10.20b
Combination bias curve, k = 3.

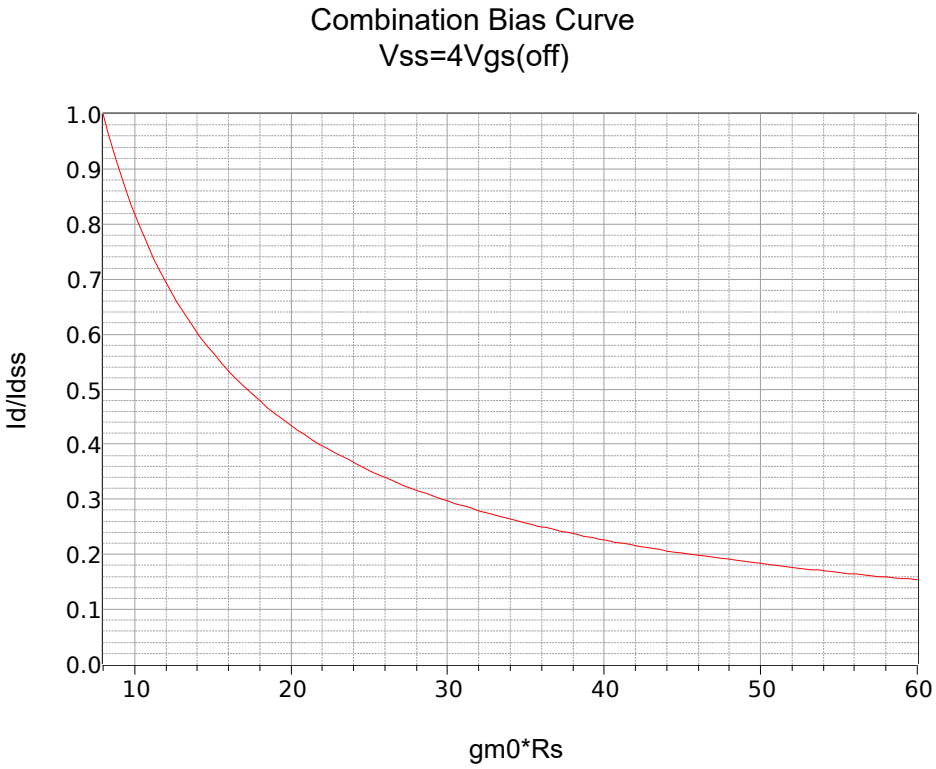


Figure 10.20c
Combination bias curve, k = 4.

Example 10.5

Determine I_D and V_{DS} for the circuit shown in Figure 10.21. Assume I_{DSS} = 12 mA and $V_{GS(off)}$ = −4 V.

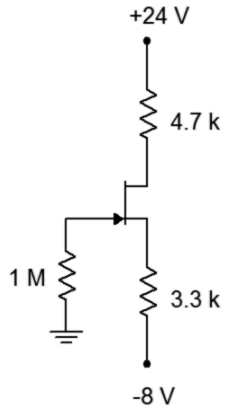


Figure 10.21
Schematic for Example 10.5.

Using the graphical method, first determine $g_{m0}\,R_S$.

$$g_{m0} = -\frac{2\,I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{2\times 12\,\text{mA}}{-4\,\text{V}}$$
$$g_{m0} = 6\,\text{mS}$$

Therefore $g_{m0}\,R_S$ = 6 mS · 3.3 k Ω = 19.8. The swamping ratio, k, is $V_{SS}/V_{GS(off)}$ = −8/−4 = 2. This requires the graph in Figure 10.20a. This graph yields approximately 0.25 for the normalized current ratio. Therefore

$$I_D = 0.25\,I_{DSS}$$
$$I_D = 0.25\times 12\,\text{mA}$$
$$I_D = 3\,\text{mA}$$

Using Ohm's law and KVL

$$V_D = V_{DD} - I_D R_D$$
$$V_D = 24\,\text{V} - 3\,\text{mA}\times 4.7\,\text{k}\Omega$$
$$V_D = 9.9\,\text{V}$$

$$V_S = V_{SS} + I_D R_S$$
$$V_S = -8\,\text{V} + 3\,\text{mA}\times 3.3\,\text{k}\Omega$$
$$V_S = 1.9\,\text{V}$$

$$V_{DS} = V_D - V_S$$
$$V_{DS} = 9.9\,\text{V} - 1.9\,\text{V}$$
$$V_{DS} = 8\,\text{V}$$

As a crosscheck, using Equation 10.9 yields 3.028 mA for I_D. The deviation is no doubt due to inaccuracy in reading the graph. In any case, using this value of drain current we find V_S to be 1.992 volts, a little higher than calculated above. This indicates that V_{GS} is −1.992 volts (because $V_G \approx 0$ V). If we plug this value of V_{GS} into Equation 10.1, I_D = 3.024 mA; an excellent match with the deviation being due to accumulated rounding errors.

In order to show the increased Q point stability of the combination bias, we'll repeat the preceding problem using a JFET with a significantly lower I_{DSS}.

Example 10.6

Determine I_D for the circuit shown in Figure 10.21. Assume $I_{DSS} = 8$ mA and $V_{GS(off)} = -4$ V.

For this version we'll use Equation 10.9. First determine $g_{m0} R_S$.

$$g_{m0} = -\frac{2 I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{2 \times 8\,\text{mA}}{-4\,\text{V}}$$
$$g_{m0} = 4\,\text{mS}$$

Therefore $g_{m0} R_S = 4$ mS · 3.3 k Ω = 13.2. The swamping ratio, k, is $V_{SS}/V_{GS(off)} = -8/-4 = 2$.

$$I_D = 2 I_{DSS}\left(\frac{1 + g_{m0} R_S (1+k) - \sqrt{1 + 2 g_{m0} R_S (1+k)}}{(g_{m0} R_S)^2}\right)$$
$$I_D = 2 \times 8\,\text{mA}\left(\frac{1 + 13.2(1+2) - \sqrt{1 + 2 \times 13.2(1+2)}}{(13.2)^2}\right)$$
$$I_D = 2.906\,\text{mA}$$

For the graphical method, a reasonable estimate for the normalized I_D would be around 0.36, yielding a drain current of 2.88 mA. Stability is apparent because the drain current has dropped only a few percent in spite of the fact that I_{DSS} decreased by 33%.

The graph of Figure 10.22 illustrates nicely the increased stability of the Q point. Once again, we plot two representative device curves in green and blue. As was the case with self bias, a plot line can be drawn, the slope of which is equal to the reciprocal of R_S. This plot line does not go though the origin, though. Instead, the x axis intercept is the voltage $|V_{SS}|$. Thus, the red plot line is shifted along the V_{GS} axis.

As can be seen in the graph, the variation in I_D is reduced (although at the expense of variation in V_{GS}). For large values of V_{SS} with correspondingly large values of R_S, the bias plot line becomes nearly horizontal, indicating a very stable Q point. With two variables in play, this bias proves to be very flexible. It can also be realized by using a positive voltage divider at the gate and removing V_{SS} (returning R_S to ground).

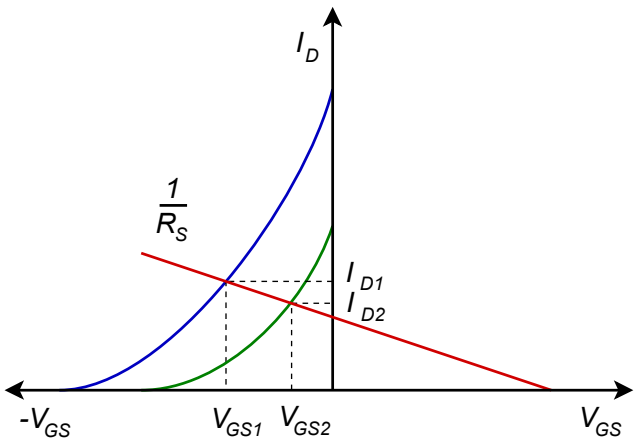
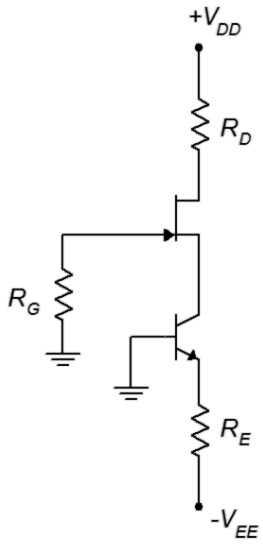


Figure 10.22
Variation for combination bias.

Constant Current Bias

The most stable bias for JFETs relies, oddly enough, on a current source made with a BJT. It is called constant *current bias*, yet another imaginative tag. Interestingly, although this will keep the Q point very stable, a fixed I_D does *not* guarantee the most stable value of voltage gain. In fact, it might be easier to achieve that goal using combination bias. The prototype constant current bias circuit is shown in Figure 10.23. An NPN BJT is used for an N-channel JFET and a PNP would be used with a P-channel JFET, typically driven from above (i.e., circuit flipped top to bottom).

Figure 10.23
Constant current bias prototype.



Ignoring the JFET for a moment, the BJT is configured as in two-supply emitter bias. In this case the base is tied directly to ground, leaving the emitter at about −0.7 VDC. The remainder of the V_{EE} supply drops across R_E, establishing the emitter current. As the collector is connected directly to the JFET's source terminal, this means that $I_S \approx I_E$. The source current winds up being just as stable as the emitter current, which we have already seen is very stable. The only requirement is that I_E should not be programmed to be larger than I_{DSS}. This being true, I_D will set up a corresponding V_{GS}. This also establishes V_S because $V_G \approx 0$ V. Therefore, the source terminal will be a small positive voltage and this is precisely what the BJT needs in order to guarantee that its collector-base junction is reverse-biased.

Computation of circuit currents and voltages is straightforward and does not involve the use of graphical aides. The first step is to examine the BJT's emitter loop and determine I_E. Once this is found, I_S and I_D are known, and all remaining component potentials may be found using Ohm's law and KVL.

This technique does not involve the calculation of V_{GS}. In fact, because I_D is very stable, V_{GS} will show the widest variation of all biasing circuits when the JFET is changed. If V_{GS} is needed, it can be determined via a little algebraic manipulation on Equation 10.1.

Example 10.7

Detemine I_D, V_{DS} and V_{GS} in the circuit of Figure 10.24. I_{DSS} = 15 mA and $V_{GS(off)} = -3$ V.

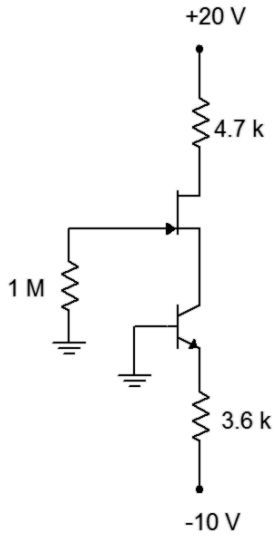


Figure 10.24
Schematic for Example 10.7.

We begin by finding I_E.

$$I_E = \frac{|V_{EE}| - 0.7\,\text{V}}{R_E}$$
$$I_E = \frac{10\,\text{V} - 0.7\,\text{V}}{3.6\,\text{k}\,\Omega}$$
$$I_E = 2.58\,\text{mA}$$

I_E is the same as I_S and I_D, therefore

$$V_D = V_{DD} - I_D R_D$$
$$V_D = 20\,\text{V} - 2.58\,\text{mA} \times 4.7\,\text{k}\,\Omega$$
$$V_D = 7.87\,\text{V}$$

To find V_S we note that $V_S = -V_{GS}$ and rearrange Equation 10.1.

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)^2$$

$$V_{GS} = V_{GS(off)}\left(1 - \sqrt{\frac{I_D}{I_{DSS}}}\right)$$
$$V_{GS} = -3\,\text{V}\left(1 - \sqrt{\frac{2.58\,\text{mA}}{15\,\text{mA}}}\right)$$
$$V_{GS} = -1.24\,\text{V}$$

Therefore V_S = 1.24 V and

$$V_{DS} = V_D - V_S$$
$$V_{DS} = 7.87\,\text{V} - 1.24\,\text{V}$$
$$V_{DS} = 6.63\,\text{V}$$

We turn next to a computer simulation of a similar circuit to validate our methodology.

Computer Simulation

A constant current bias circuit is entered into a simulator as shown in Figure 10.25.

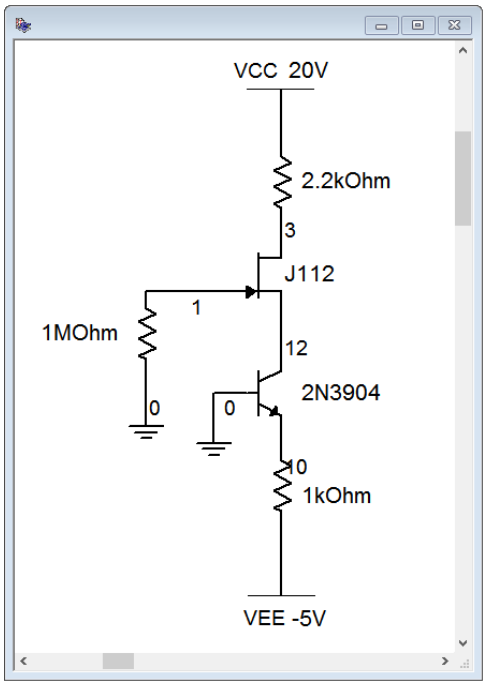


Figure 10.25
Constant current bias circuit in simulator.

A cursory estimate shows that I_E and I_D should be around 4.3 mA. Also, V_D should be approximately 20 V − 4.3 mA · 2.2 kΩ, or about 10.54 volts. The results of a DC operating point analysis are shown in Figure 10.26.

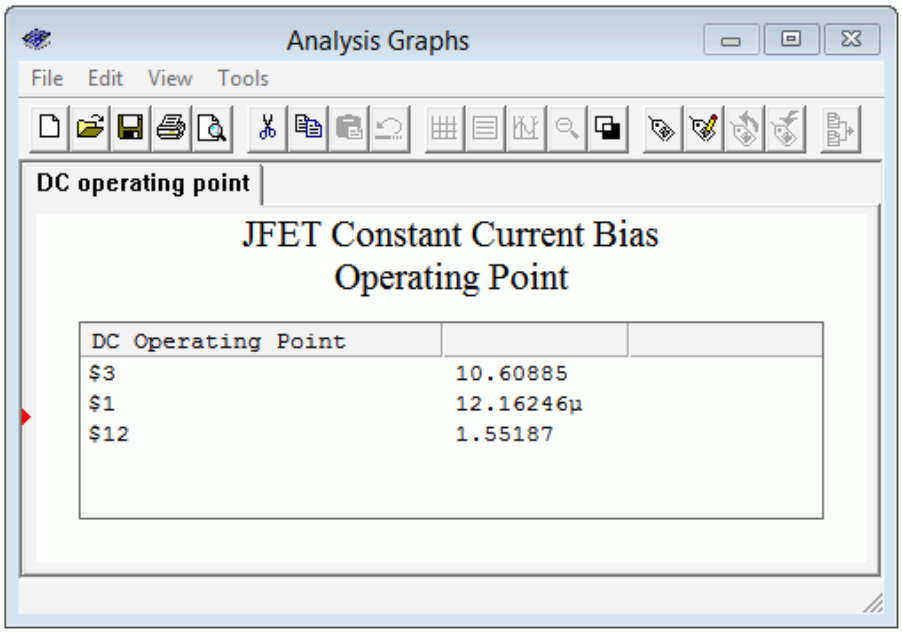


Figure 10.26
Constant current bias DC operating point simulation results.

The drain voltage (node 3) is just over 10.6 volts, agreeing with our estimate. Also, note the minuscule gate voltage (node 1) of 12 µV which verifies our continuing assumption in these circuits that $V_G \approx 0$ VDC. Finally, we see a modest potential of about 1.5 volts at the source terminal (node 12). This shows the proper reverse-biasing of both the gate-source and collector-base junctions.

Finally, we can examine the Q point variation using Figure 10.27. Here, the plot line is perfectly horizontal and all device variation is manifest in V_{GS}.

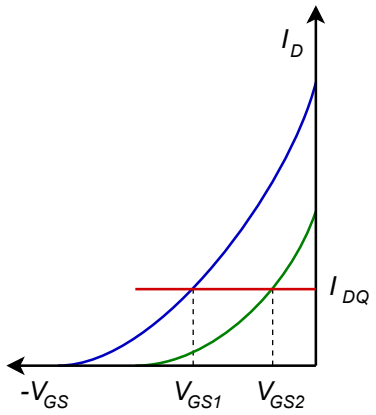


Figure 10.27
Variation for constant current bias.

Summary

The junction field effect transistor is an altogether different device from the bipolar junction transistor. Instead of relying on a forward-biased PN junction to control current, the JFET utilizes a reverse-biased PN junction. Furthermore, the JFET uses voltage control rather than the BJT's current control. In spite of this, a family of JFET drain curves offers similarity to the BJT's collector curves, exhibiting three regions: ohmic, constant current and breakdown.

The DC model of a JFET includes a voltage-controlled current source in the drain and a very, very large resistance, R_{GS}, from gate to source. This resistance models that of a reverse-biased PN junction. The characteristic equation of the JFET is square-law and is consequently much more gentle in slope than the corresponding equation for a BJT. The maximum current produced by a JFET is I_{DSS} and occurs when $V_{GS} = 0$ V. V_{GS} must always be negative to ensure proper operation and all negative values will lead to a drain current less than I_{DSS}. Once the gate-source becomes negative enough (at $V_{GS(off)}$), drain current goes to zero.

There are several methods to bias JFETs. Perhaps the most simple method is to apply a fixed potential to the gate while grounding the source. This is called constant voltage bias and is the least stable bias in terms of Q point. Self bias uses a minimum of components and offers modest stability. It is a decent general-purpose bias. The addition of a negative power supply to the source resistor leads to the combination

bias topology. This circuit offers improvements in stability over self bias. The most stable bias is the constant current bias. This form relies on a BJT to establish a very stable current.

Review Questions

1. Compare the operation of the JFET to the BJT.
2. Compare the regions of JFET drain curves to those of BJT collector curves.
3. Why is the JFET referred to as a square-law device?
4. Rank the biasing schemes presented in this chapter in terms of Q point stability.
5. What is pinch-off voltage?
6. How does the JFET DC biasing model differ from the BJT DC model?

Problems

Analysis Problems

1. For the circuit of Figure 10.28, determine I_D and V_{DS}. $I_{DSS} = 40$ mA, $V_{GS(off)} = -4$ V, $V_{DD} = 26$ V, $V_{GG} = -2$ V, $R_G = 220$ kΩ, $R_D = 1.2$ kΩ.

2. For the circuit of Figure 10.28, determine I_D and V_{DS}. $I_{DSS} = 20$ mA, $V_{GS(off)} = -3$ V, $V_{DD} = 22$ V, $V_{GG} = -1$ V, $R_G = 390$ kΩ, $R_D = 1$ kΩ.

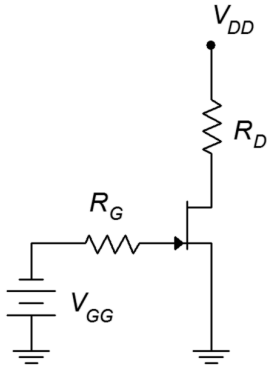


Figure 10.28

3. For the circuit of Figure 10.29, determine I_D, V_G and V_D. $I_{DSS} = 24$ mA, $V_{GS(off)} = -6$ V, $V_{DD} = 36$ V, $R_G = 220$ kΩ, $R_S = 2$ kΩ, $R_D = 1.8$ kΩ.

4. For the circuit of Figure 10.29, determine I_D, V_S and V_{DS}. $I_{DSS} = 18$ mA, $V_{GS(off)} = -3$ V, $V_{DD} = 30$ V, $R_G = 270$ kΩ, $R_S = 2.7$ kΩ, $R_D = 3.3$ kΩ.

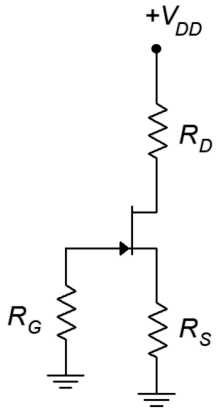


Figure 10.29

5. For Figure 10.30, determine I_D, V_G and V_D. I_{DSS} = 16 mA, V_{DD} = 25 V, $V_{GS(off)}$ = −3 V, V_{SS} = −6 V, R_G = 560 kΩ, R_S = 2 kΩ, R_D = 3.6 kΩ.

6. For Figure 10.30, determine I_D, and V_{DS}. I_{DSS} = 16 mA, V_{DD} = 25 V, $V_{GS(off)}$ = −3 V, V_{SS} = −9 V, R_G = 680 kΩ, R_S = 2 kΩ, R_D = 2.7 kΩ.

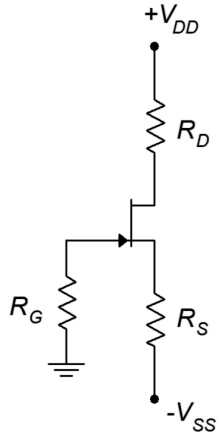


Figure 10.30

7. For Figure 10.31, determine I_D, V_G and V_D. I_{DSS} = 16 mA, V_{DD} = 25 V, $V_{GS(off)}$ = −3 V, V_{EE} = −9 V, R_G = 810 kΩ, R_E = 2 kΩ, R_D = 2.7 kΩ.

8. For the circuit of Figure 10.31, determine I_D and V_{DS}. I_{DSS} = 40 mA, $V_{GS(off)}$ = −4 V, V_{DD} = 30 V, V_{EE} = −6 V, R_G = 750 kΩ, R_E = 500 Ω, R_D = 1.8 kΩ.

Design Problems

9. Using the circuit of Figure 10.29, determine a value for R_S to set I_D to 4 mA. I_{DSS} = 10 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 20 V, R_G = 430 kΩ, R_D = 1.8 kΩ.

10. Using the circuit of Figure 10.28, determine a value for V_{GG} to set I_D to 2 mA. I_{DSS} = 10 mA, $V_{GS(off)}$ = −4 V, V_{DD} = 28 V, R_G = 470 kΩ, R_D = 4.7 kΩ.

11. Using the circuit of Figure 10.31, determine a value for R_E to set I_D to 4 mA. I_{DSS} = 18 mA, $V_{GS(off)}$ = −3 V, V_{DD} = 25 V, V_{EE} = −12 V, R_G = 330 kΩ, R_D = 2.2 kΩ.

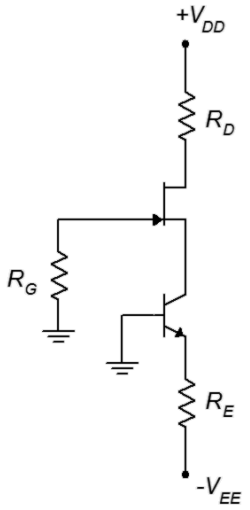


Figure 10.31

12. Using the circuit of Figure 10.31, determine values for R_E and R_D to set I_D to 5 mA and V_D to 6 V. I_{DSS} = 20 mA, $V_{GS(off)}$ = −4 V, V_{DD} = 32 V, V_{EE} = −10 V, R_G = 390 kΩ.

Challenge Problems

13. Following the derivation of Equation 10.7, derive Equation 10.9.

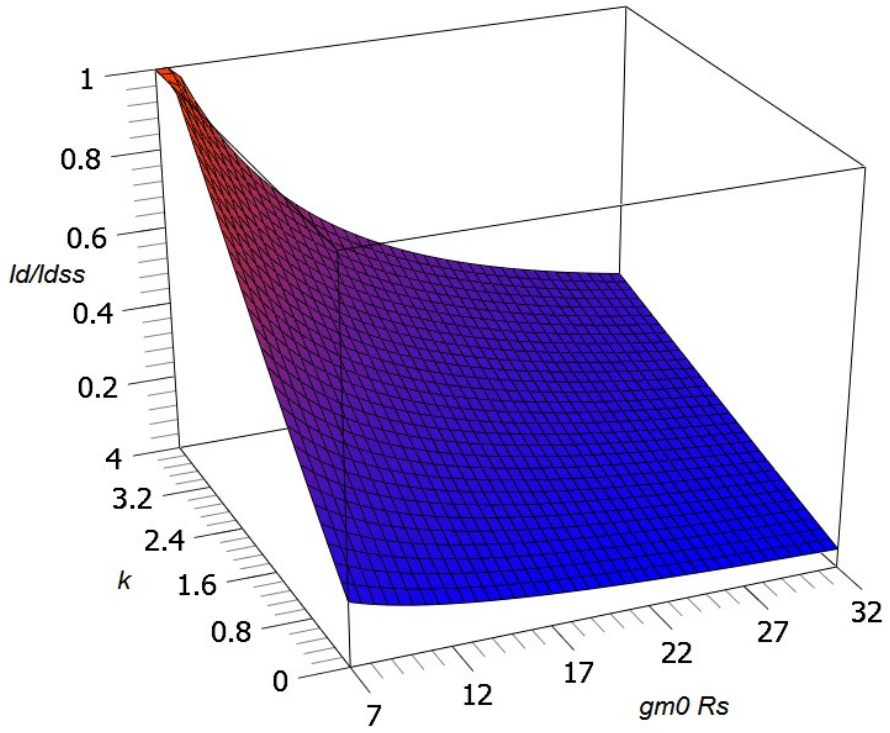
14. Using the circuit of Figure 10.30, determine values for R_S and V_{SS} to set I_D to 4 mA. I_{DSS} = 16 mA, $V_{GS(off)}$ = −4 V, V_{DD} = 30 V, R_G = 680 kΩ, R_D = 2 kΩ.

Computer Simulation Problems

15. Perform a DC operating point simulation on the circuit of Problem 7 to verify the results. The J111 will be sufficient.

16. Perform a DC operating point simulation on the circuit of Problem 10 to verify the results. The J111 will be sufficient.

Combination Bias Surface Plot



The graphs of Figure 10.20 represent three slices from this surface.

Looks cool, but...

11 JFET Small Signal Amplifiers

11.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Determine the voltage gain, input impedance and output impedance of basic JFET amplifiers.
- Draw and explain a basic AC model of a JFET.
- Compare and analyze JFET voltage amplifiers and voltage followers.
- Discuss the advantages and disadvantages of JFET circuits with those of comparable BJT circuits.
- Analyze small signal combination BJT/JFET amplifier circuits.
- Discuss applications that make use of the JFET's ohmic region.

11.1 Introduction

The JFET can be used to create both voltage amplifiers and voltage followers. In comparison with the BJT, the JFET tends to have less voltage gain potential. On the other hand, JFET circuits offer the possibility of a much higher input impedance, lower noise and better high frequency performance. There are many other similarities with BJT amplifiers. For example, the possibility of swamping still exists as a means of lowering distortion at the expense of voltage gain. Also, the JFET voltage amplifier inverts the signal, just like the BJT version. When it comes to AC analysis, a key element for the BJT is r'_e. For the JFET, the comparable parameter is transconductance, g_m.

JFET amplifiers and followers can be used with their BJT cousins. Indeed, the combination of the two, each playing to their strengths, has the potential to outperform a design using only one type of device.

Alongside their use in amplifiers and followers, JFETs can also be used in their ohmic region. This includes applications as voltage-controlled resistors and analog switches. In this mode, the device no longer behaves as a constant current source. Instead, the channel resistance becomes a function of the gate-source voltage and can be used as a control element within a voltage divider. As such, it has the capability of changing resistance value much faster than a mechanical potentiometer.

11.2 Simplified AC Model of the JFET

An AC model of the JFET is shown in Figure 11.1. This is essentially the same model as was used for DC analysis. Once again, we have a voltage-controlled current source situated in the drain. The reverse-biased junction shows up as a very large resistance, r_{GS}.

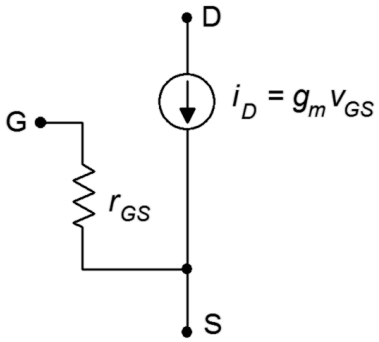


Figure 11.1
AC model of JFET.

It is worth mentioning that this model is suitable only for low frequencies. At higher frequencies, device capacitances can play a major role in the response of the amplifier. There are three device capacitances not shown in the Figure that shunt each pair of terminals: C_{GS}, C_{DG} and C_{DS}. On a data sheet, the "lumped" capacitances are often given. These are C_{iss}, the capacitance looking into the gate with the source and drain shorted to ground: $C_{iss} = C_{GS} + C_{DG}$; and C_{rss}, the capacitance seen from the drain with the gate and source shorted to ground: $C_{rss} = C_{DS} + C_{DG}$. As we shall see, these capacitances can have a sizable impact on amplifier characteristics such as Z_{in}.

The value of transconductance, g_m, will prove to be of particular interest. It is roughly of equal importance to r'_e in a BJT.[41]

11.3 Common Source Amplifier

The *common source amplifier* is analogous to the common emitter amplifier. The prototype amplifier circuit with device model is shown in Figure 11.2.

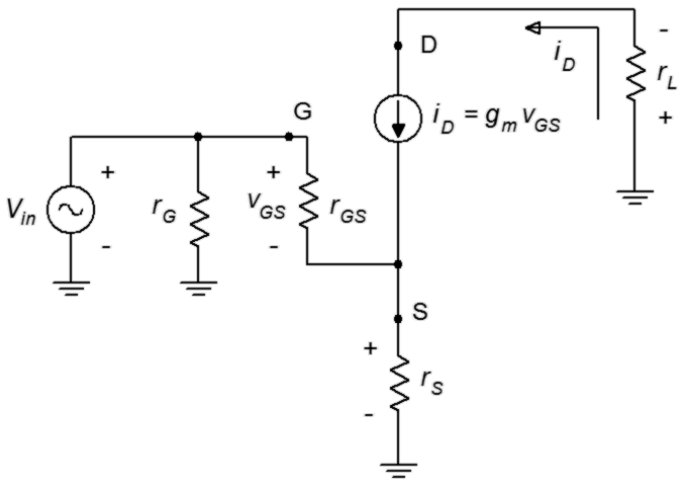


Figure 11.2
Common source amplifier with model.

41 In fact, we can say that $1/r'_e$ is g_m for a BJT.

This circuit includes a swamping resistor, r_S. The input signal is presented to the gate terminal while the output is taken from the drain.

Voltage Gain

An equation for the voltage gain, A_v, is developed as follows. First, we start with the fundamental definition, namely that voltage gain is the ratio of v_{out} to v_{in}, and proceed by expressing these voltages in terms of their Ohm's law equivalents.

$$A_v = \frac{v_{out}}{v_{in}} = \frac{v_D}{v_G}$$
$$A_v = \frac{-i_D r_L}{i_D r_S + v_{GS}}$$
$$A_v = \frac{-g_m v_{GS} r_L}{g_m v_{GS} r_S + v_{GS}}$$
$$A_v = -\frac{g_m r_L}{g_m r_S + 1} \qquad (11.1)$$

If there is no swamping resistor, the first portion of the denominator drops out and the gain simplifies to $-g_m \cdot r_L$. The swamping resistor in the source, r_S, plays the same role here as it did in the BJT: it helps to stabilize the gain and reduce distortion. It does so at the expense of voltage gain.

Input Impedance

Referring back to Figure 11.2, the input impedance of the amplifier will be r_G in parallel with the impedance looking into the gate terminal, $Z_{in(gate)}$. For the non-swamped case, this will be r_{GS}. At low frequencies r_{GS} is very large, well into the megohms. In most practical circuits, r_G will be much lower, hence

$$Z_{in} = r_G \parallel r_{GS} \approx r_G \qquad (11.2)$$

Theoretically, for swamped amplifiers $Z_{in(gate)}$ will be higher than r_{GS} but this is a moot point. In either case, it is relatively easy to obtain a high input impedance, certainly much easier than it is for typical single-device BJT amplifiers.

It might be easy to become complacent and simply assume that r_G sets the input impedance and that's the end of it. This would be a mistake. As mentioned earlier, with impedances this high, we cannot ignore items such as junction capacitance. For example, for a general purpose device a typical value for C_{iss}, the total input capacitance, may be in the vicinity of 5 to 10 pF. This capacitance appears in parallel with r_G. If this amplifier is used for ultrasonic signals, the capacitive reactance, X_C, would be as low as 160 kΩ at 100 kHz. Although this is high compared to typical BJT circuits, it's less than the R_G values commonly used for biasing. At higher

frequencies, the situation is even worse as X_C decreases with frequency. Also, we are ignoring the Miller effect here which makes the situation even worse than even worse, so perhaps we can say that it's even worser, which is a claim we could also make regarding the grammar of this sentence.

Output Impedance

To investigate the output impedance, we'll refer to Figure 11.3. This circuit is very similar to that of Figure 11.2. The major difference is that the AC load equivalent has been split into its two components, the load itself, R_L, and the drain biasing resistor, R_D.

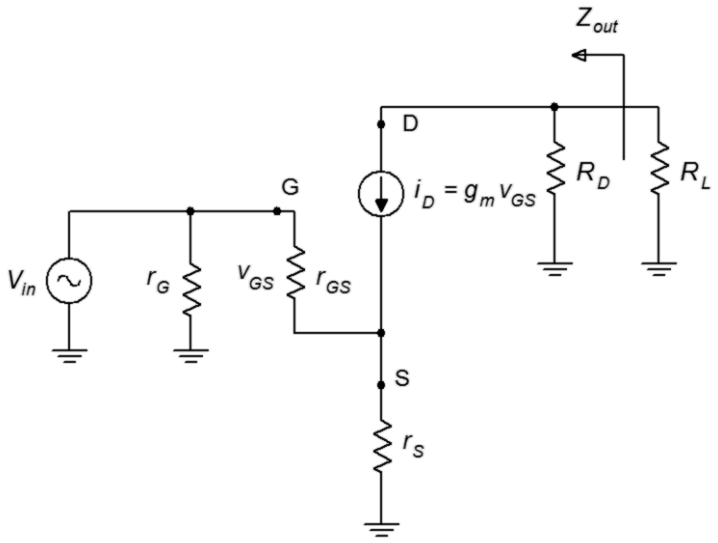


Figure 11.3
Output impedance of common source amplifier.

From the vantage point of R_L, peering back into the amplifier we see R_D in parallel with the impedance at the drain. At the drain we find the current source, i_D. The internal impedance of this equivalent current source is very high compared to typical values for R_D (hundreds of kΩ), therefore we can approximate the output impedance as

$$Z_{out} \approx R_D \tag{11.3}$$

It should be noted that all forms of DC bias discussed in the previous chapter are game here. There are a few limitations to be aware of, though. For example, when using constant voltage bias, swamping is not possible as that bias form does not use a source resistor. In contrast, self bias and combination bias include a source resistor so swamping is a possibility, however, R_S may need to be split and partially bypassed to achieve the desired results. Finally, constant current bias is not well-positioned to use swamping as that would require some additional work to fit in a new R_S along with the current source. More typically, the current source will just be bypassed with a capacitor to produce a non-swamped amplifier.

Determine the voltage gain and input impedance for the circuit shown in Figure 11.4. Assume I_{DSS} = 15 mA and $V_{GS(off)}$ = −3 V.

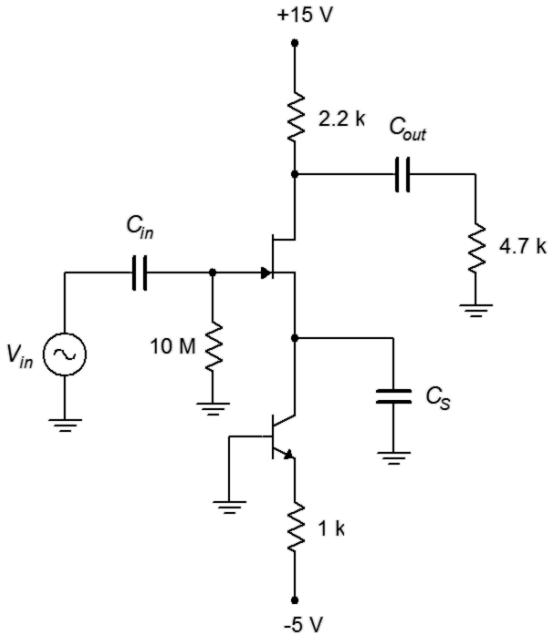


Figure 11.4
Circuit for Example 11.1.

This is an unswamped common source amplifier with constant current bias. We can determine Z_{in} via inspection.

$$Z_{in} = Z_{in(gate)} \parallel R_G$$
$$Z_{in} \approx 10\,\text{M}\Omega$$

To find the voltage gain, we'll first need to find I_D and g_{m0} in order to find g_m.

$$I_D = \frac{|V_{EE}| - 0.7\,\text{V}}{R_E}$$
$$I_D = \frac{5\,\text{V} - 0.7\,\text{V}}{1\,\text{k}\Omega}$$
$$I_D = 4.3\,\text{mA}$$

$$g_{m0} = -\frac{2I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{30\,\text{mA}}{-3\,\text{V}}$$
$$g_{m0} = 10\,\text{mS}$$

Knowing the current and maximum transconductance, we can find g_m through the use of Equation 10.4.

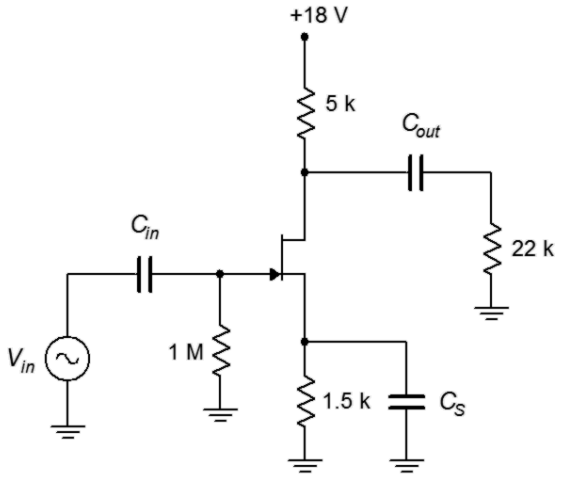
$$\frac{g_m}{g_{m0}} = \sqrt{\frac{I_D}{I_{DSS}}}$$
$$g_m = g_{m0}\sqrt{\frac{I_D}{I_{DSS}}}$$
$$g_m = 10\,\text{mS}\sqrt{\frac{4.3\,\text{mA}}{15\,\text{mA}}}$$
$$g_m = 5.35\,\text{mS}$$

$$A_v = -\frac{g_m r_L}{g_m r_S + 1}$$
$$A_v = -\frac{5.35\,\text{mS}(2.2\,\text{k}\Omega \parallel 4.7\,\text{k}\Omega)}{5.35\,\text{mS} \times 0\,\Omega + 1}$$
$$A_v = -8.02$$

Example 11.2

Determine the voltage gain and input impedance for the circuit shown in Figure 11.5. Assume I_{DSS} = 24 mA and $V_{GS(off)}$ = −4 V.

Figure 11.5
Circuit for Example 11.2



This is an unswamped common source amplifier with self bias. Once again, we can determine Z_{in} via inspection.

$$Z_{in} = Z_{in(gate)} \parallel R_G$$
$$Z_{in} \approx 1\,\text{M}\Omega$$

To find the voltage gain, we'll first need to find g_{m0}. Also, we can take a shortcut and find the normalized drain current from the self bias graph instead of finding I_D itself.

$$g_{m0} = -\frac{2\,I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{48\,\text{mA}}{-4\,\text{V}}$$
$$g_{m0} = 12\,\text{mS}$$

R_S is 1.5 kΩ, therefore $g_{m0}R_S$ = 18. From the self bias graph this produces a normalized drain current of 0.08.

$$g_m = g_{m0}\sqrt{\frac{I_D}{I_{DSS}}}$$
$$g_m = 12\,\text{mS}\sqrt{0.08}$$
$$g_m = 3.4\,\text{mS}$$

Again, there is no swamping so r_S = 0. The gain formula reduces to

$$A_v = -g_m r_L$$
$$A_v = -3.4\,\text{mS}\left(22\,\text{k}\,\Omega \parallel 5\,\text{k}\,\Omega\right)$$
$$A_v = -13.9$$

We will now turn our attention to the effect of swamping. As in the BJT case, we expect to sacrifice gain and in return, see an improvement in distortion. We shall examine this through the use of a simulation.

Computer Simulation

A common source amplifier using self bias is entered into the simulator as shown in Figure 11.6.

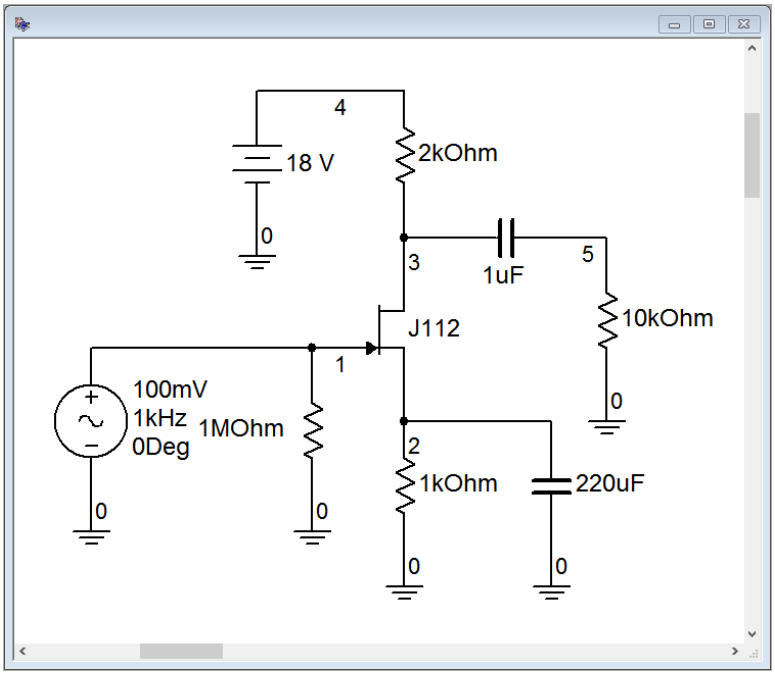


Figure 11.6
Common source amplifier in simulator.

Reasonable device values for this model are $I_{DSS} = 40$ mA and $V_{GS(off)} = -2.3$ V. Based on these we find

$$g_{m0} = -\frac{2\,I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{80\,\text{mA}}{-2.3\,\text{V}}$$
$$g_{m0} = 34.8\,\text{mS}$$

Given $R_S = 1$ kΩ, the self-bias equation yields

$$I_D = 2\,I_{DSS}\left(\frac{1+g_{m0}R_S-\sqrt{1+2\,g_{m0}R_S}}{(g_{m0}R_S)^2}\right)$$
$$I_D = 2\,I_{DSS}\left(\frac{1+34.8-\sqrt{1+2\times 34.8}}{(34.8)^2}\right)$$
$$I_D = 1.81\,\text{mA}$$

$$g_m = g_{m0}\sqrt{\frac{I_D}{I_{DSS}}}$$
$$g_m = 34.8\,\text{mS}\sqrt{\frac{1.81\,\text{mA}}{40\,\text{mA}}}$$
$$g_m = 7.4\,\text{mS}$$

$$A_v = -g_m r_L$$
$$A_v = -7.4\,\text{mS}\left(2\,\text{k}\,\Omega \parallel 10\,\text{k}\,\Omega\right)$$
$$A_v = -12.3$$

The results of a transient analysis are shown in Figure 11.7 for a 100 mV peak input signal.

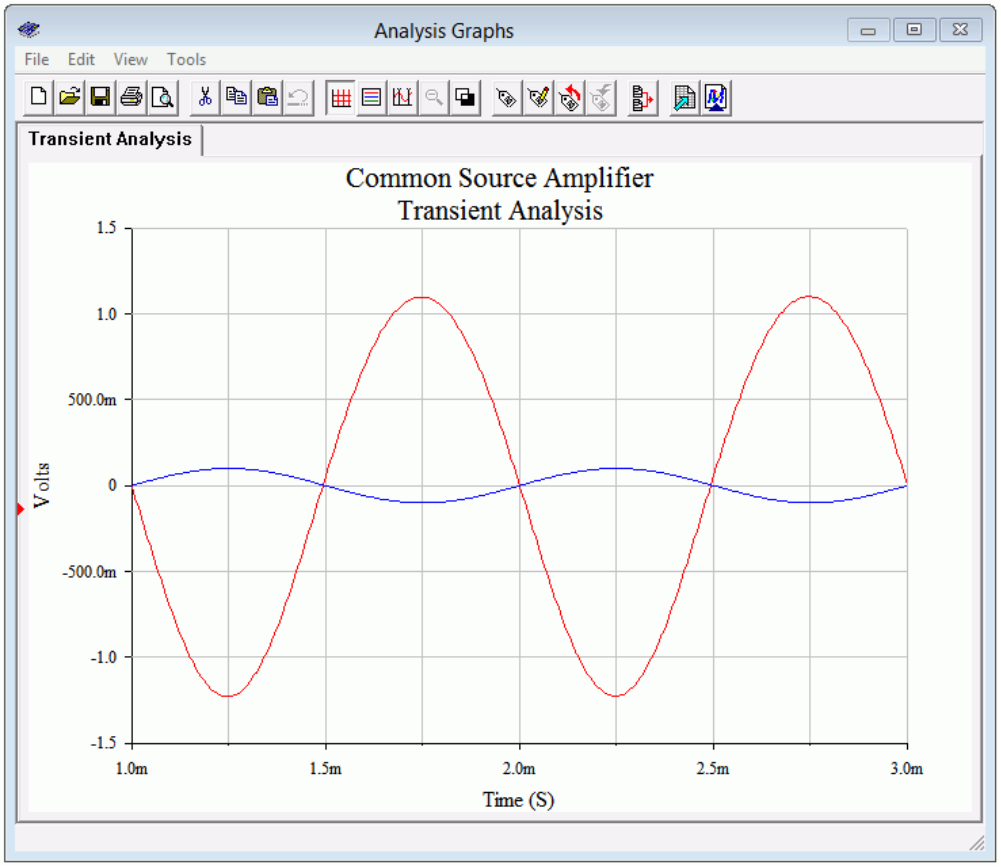


Figure 11.7
Transient analysis of the common source amplifier.

First, the inversion between the input (blue trace) and output (red trace) is obvious. Also, some distortion is evident in the output waveform. Inspecting the positive and negative peaks shows that the positive peaks are are a little more broad than the negative ones and don't quite reach the same magnitude. The averaged gain is about −11.75, around 5% low from the estimate and not too bad considering the distortion.

The simulation is run a second time, but this time around the 1 kΩ source resistor is split into a 200 Ω/800 Ω pair with only the 800 Ω being bypassed. The DC bias does

not change leaving g_m untouched, but the new 200 Ω swamping resistor drops the gain to

$$A_v = -\frac{g_m r_L}{g_m r_S + 1}$$
$$A_v = -\frac{7.4\,\text{mS}(2\,\text{k}\Omega \parallel 10\,\text{k}\Omega)}{7.4\,\text{mS} \times 200\,\Omega + 1}$$
$$A_v = -4.96$$

The output signal from the simulator is shown in Figure 11.7.

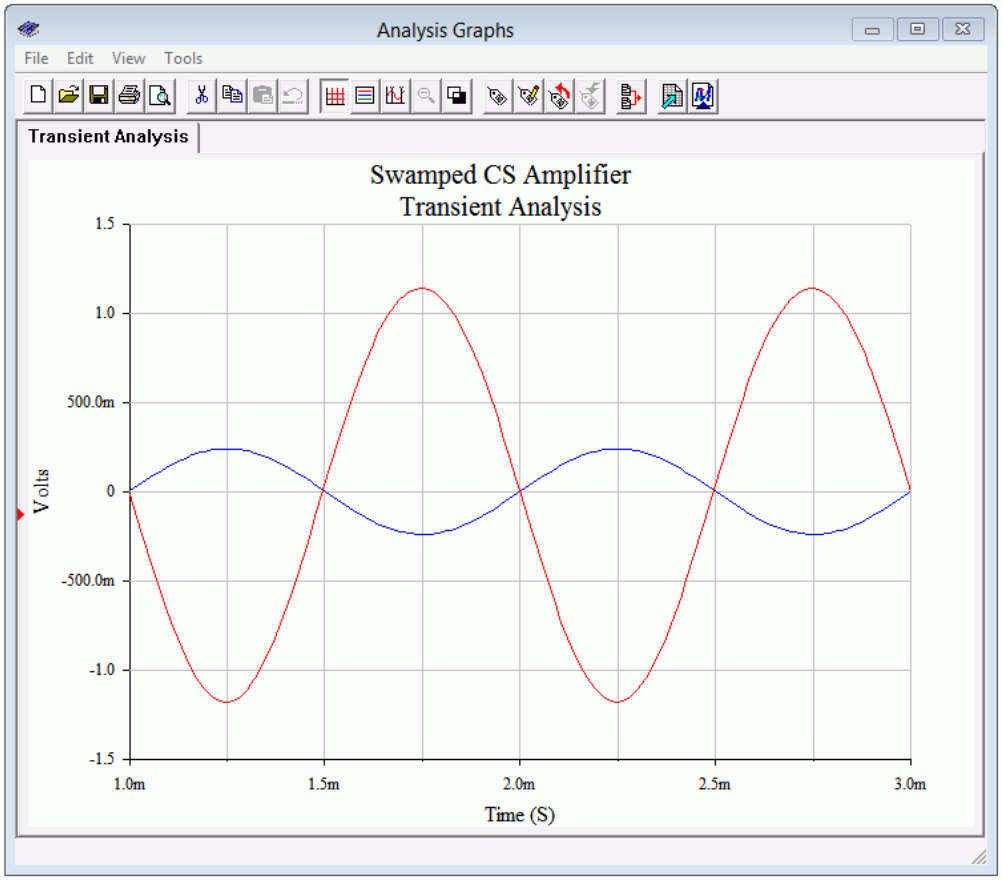


Figure 11.8
Transient analysis of the swamped common source amplifier.

The input signal was raised to 240 mV peak in order to keep the output signals of the two versions at the same amplitude. The symmetry appears to be better here and the gain works out to −4.85, just a few percent low.

Total harmonic distortion (THD) analysis is performed next. The results are shown in Figures 11.9 and 11.10. To keep the comparison fair, the input levels are adjusted to maintain similar output voltages. The non-swamped results are seen in Figure 11.9, and as expected based on the waveform asymmetry, the THD is relatively high at roughly 4%. The swamped version scores better at just over 1.6%, although this is still not stellar performance.

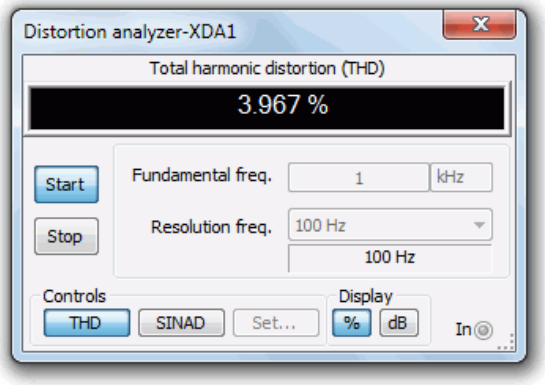


Figure 11.9
THD of common source amplifier.

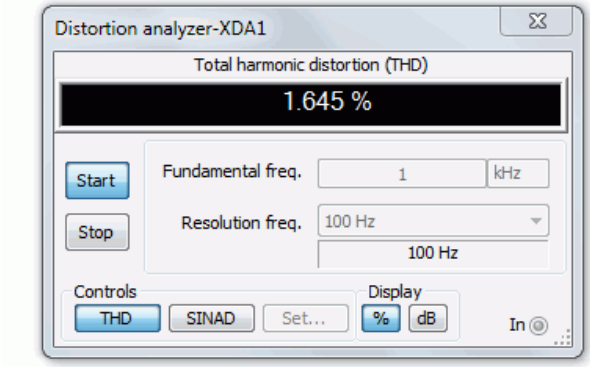


Figure 11.10
THD of swamped common source amplifier.

It is interesting to note that while the voltage gain of the swamped amplifier has dropped to 41% of the non-swamped gain, the THD has dropped to 41% of the non-swamped THD. In fact, given the square-law nature of the characteristic curve, we would expect the distortion to be lower if we used smaller signals. To verify this, the THD simulation is run again for the swamped amplifier, but now using an input signal ten times smaller at only 24 mV peak. The result is shown in Figure 11.11.

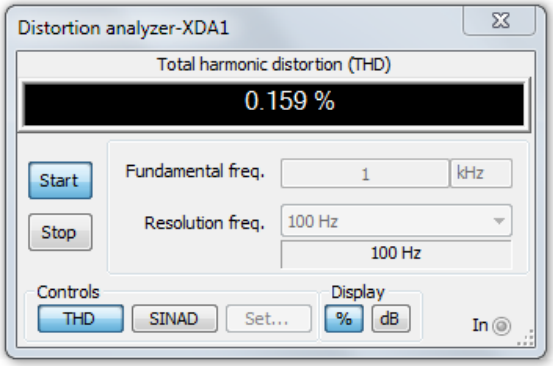


Figure 11.11
THD of swamped common source amplifier with reduced signal level.

The resulting THD is markedly lower for an order of magnitude improvement. We're now at least approaching “hi-fi” territory.

11.4 Common Drain Amplifier

The *common drain amplifier* is analogous to the common collector emitter follower. The JFET version is also known as a *source follower*. The prototype amplifier circuit with device model is shown in Figure 11.12. As with all voltage followers, we expect a non-inverting voltage gain close to unity, a high Z_{in} and low Z_{out}.

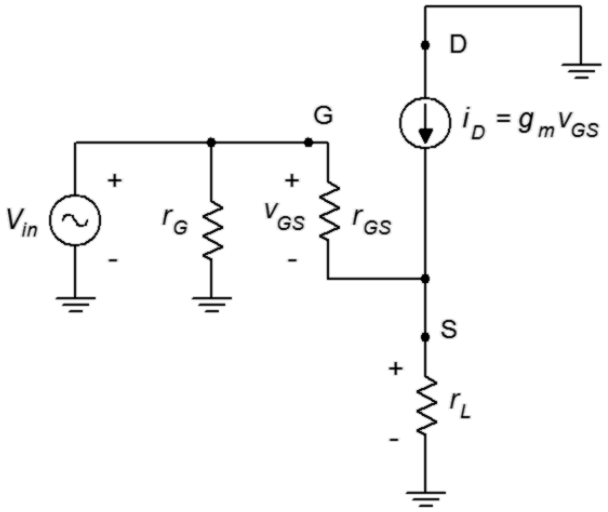


Figure 11.12
Common drain (source follower) prototype.

The input signal is presented to the gate terminal while the output is taken from the source. Many bias circuits may be used here as long as they do not have a grounded source terminal such as constant voltage bias.

Voltage Gain

In order to develop an equation for the voltage gain, A_v, we follow the same path we took with the common source amplifier earlier in this chapter. First, we start with the fundamental definition, namely that voltage gain is the ratio of v_{out} to v_{in}, and proceed by expressing these voltages in terms of their Ohm's law equivalents.

$$A_v = \frac{v_{out}}{v_{in}} = \frac{v_S}{v_G}$$

$$A_v = \frac{i_D r_L}{i_D r_L + v_{GS}}$$

$$A_v = \frac{g_m v_{GS} r_L}{g_m v_{GS} r_L + v_{GS}}$$

$$A_v = \frac{g_m r_L}{g_m r_L + 1} \qquad (11.4)$$

Equation 11.4 is very similar to the gain equation derived for the swamped common source amplifier; the notable changes being the lack of the minus sign indicating that

this circuit does not invert the signal, and r_L replacing r_S in the denominator. It is worth remembering that r_L here is the AC source resistance while in the common source amplifier r_L is the AC drain resistance. To avoid potential confusion, this equation could also be written as

$$A_v = \frac{g_m r_S}{g_m r_S + 1} \tag{11.4b}$$

In any event, the goal is to make sure that $g_m r_S >> 1$. By doing so, the voltage gain will be very close to unity.

Input Impedance

The analysis for common drain input impedance is virtually identical to that for the swamped common source amplifier. The result is replicated here for convenience.

$$Z_{in} = r_G \parallel r_{GS} \approx r_G \tag{11.5}$$

Output Impedance

In order to investigate the output impedance, we'll separate the load resistance from the source bias resistor, as shown in Figure 11.13.

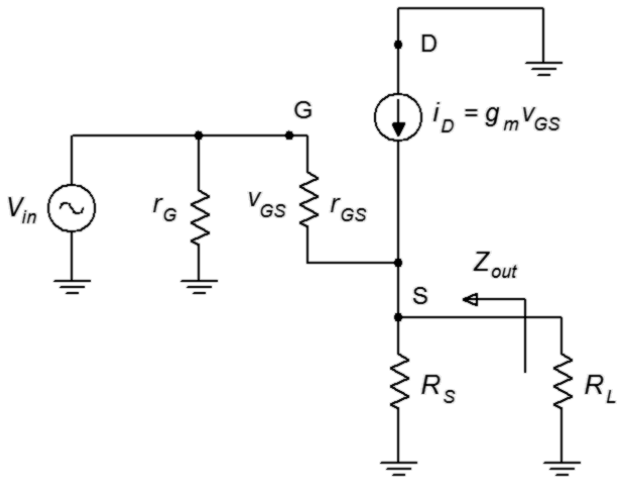


Figure 11.13
Common drain output impedance analysis.

From the position of R_L, looking back toward the source we find R_S in parallel with the impedance looking back into the source terminal. The voltage at this node is v_{GS} and the current entering this node is i_D. The ratio of the two must yield the impedance looking into the source.

$$Z_{source} = \frac{v_{GS}}{i_D}$$
$$Z_{source} = \frac{v_{GS}}{g_m v_{GS}}$$
$$Z_{source} = \frac{1}{g_m} \tag{11.6}$$

Therefore, the output impedance is

$$Z_{out} = R_S \parallel \frac{1}{g_m} \tag{11.7}$$

We can expect this value to be much smaller than the output impedance of typical common source amplifiers.

Example 11.3

For the follower shown in Figure 11.14, determine the input impedance and output voltage. Assume V_{in} = 100 mV, I_{DSS} = 30 mA, $V_{GS(off)}$ = −2 V.

This is a follower using self bias. We'll find g_m via the self bias graph.

$$g_{m0} = -\frac{2 I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{60\,\text{mA}}{-2\,\text{V}}$$
$$g_{m0} = 30\,\text{mS}$$

R_S is 1 kΩ, yielding 30 for $g_{m0} R_S$. The normalized drain current from the self bias graph is approximately 0.05.

$$g_m = g_{m0}\sqrt{\frac{I_D}{I_{DSS}}}$$
$$g_m = 30\,\text{mS}\sqrt{0.05}$$
$$g_m = 6.71\,\text{mS}$$

$$A_v = \frac{g_m r_S}{g_m r_S + 1}$$
$$A_v = \frac{6.71\,\text{mS}(1\,\text{k}\Omega \parallel 600\,\Omega)}{6.71\,\text{mS}(1\,\text{k}\Omega \parallel 600\,\Omega) + 1}$$
$$A_v = 0.716$$

Thus V_{out} is 71.6 mV. By inspection, Z_{in} may be approximated as 2.2 MΩ.

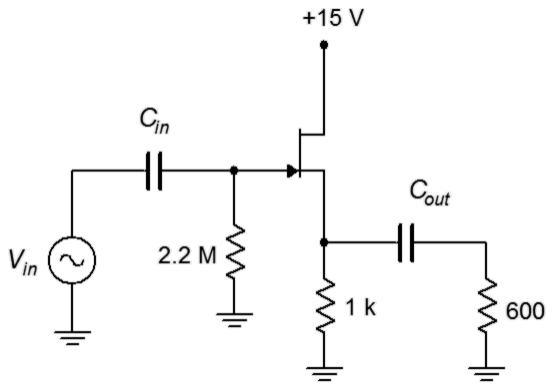
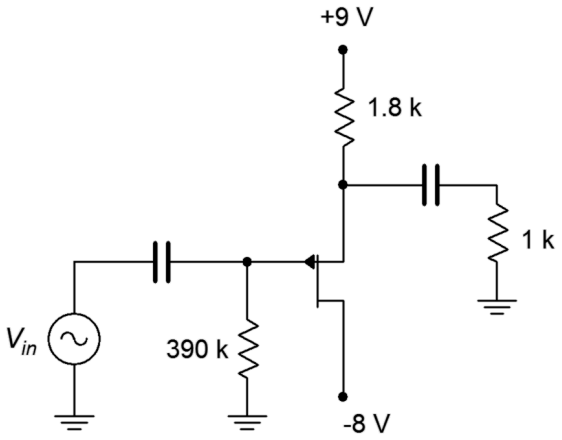


Figure 11.14
Circuit for Example 11.3.

For the circuit shown in Figure 11.15, determine the input impedance and output voltage. Assume V_{in} = 100 mV, I_{DSS} = 36 mA, $V_{GS(off)}$ = 3 V.

This follower uses combination bias with a P-channel JFET. Note that the source is at the top. We'll find g_m via the combination bias graph for k = 3 ($k = V_{SS} / V_{GS(off)}$).

Figure 11.15
Circuit for Example 11.4.



$$g_{m0} = \frac{2 I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = \frac{72\,\text{mA}}{3\,\text{V}}$$
$$g_{m0} = 24\,\text{mS}$$

R_S is 1.8 kΩ, yielding 43.2 for $g_{m0} R_S$. The normalized drain current from the k = 3 combination bias graph is approximately 0.17.

$$g_m = g_{m0}\sqrt{\frac{I_D}{I_{DSS}}}$$
$$g_m = 24\,\text{mS}\sqrt{0.17}$$
$$g_m = 9.9\,\text{mS}$$

$$A_v = \frac{g_m r_S}{g_m r_S + 1}$$
$$A_v = \frac{9.9\,\text{mS}(1\,\text{k}\Omega \parallel 1.8\,\text{k}\Omega)}{9.9\,\text{mS}(1\,\text{k}\Omega \parallel 1.8\,\text{k}\Omega)+1}$$
$$A_v = 0.864$$

Thus V_{out} is 86.4 mV. By inspection, Z_{in} may be approximated as 390 kΩ.

11.5 Common Gate Amplifier

The *common gate amplifier* is analogous to the common base BJT amplifier. It displays similar performance characteristics. An example of this amplifier is shown in Figure 11.16. As with BJT version, we expect a non-inverting voltage gain, a low Z_{in} and a high Z_{out}. In some respects, it can be considered as the opposite of the common drain amplifier.

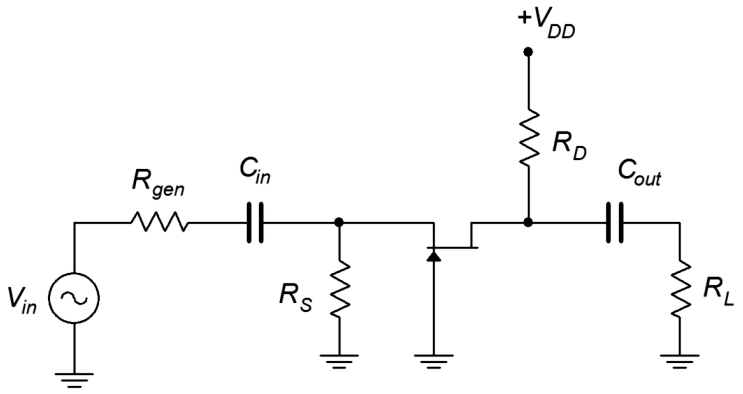


Figure 11.16
A common gate amplifier.

The input signal is presented to the source terminal while the output is taken from the drain. This version uses self bias, but other biasing schemes are available and may be used here as long as they do not have a grounded source terminal, such as constant voltage bias.

Voltage Gain

In order to develop an equation for the voltage gain, A_v, consider the simplified AC prototype shown in Figure 11.17. There is a minor change of nomenclature in this figure, namely that the voltage used for the controlled current source is defined as v_{sg} rather than as v_{gs}. This is done to simplify the analysis because the input is applied from source to gate. The practical effect of this change is that the direction of the current source is now reversed, with current flowing out of the drain.

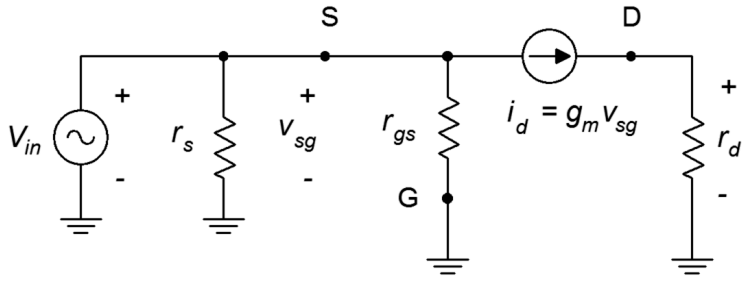


Figure 11.17
Common gate prototype.

We shall begin the analysis with the fundamental definition, namely that voltage gain is the ratio of v_{out} to v_{in}, and proceed by expressing these voltages in terms of their Ohm's law equivalents. Note that r_{gs} may be approximated as infinite, while r_s and r_d represent the simplified AC impedances at the source and drain, respectively.

$$A_v = \frac{v_{out}}{v_{in}} = \frac{v_D}{v_S}$$
$$A_v = \frac{i_D r_L}{v_{SG}}$$
$$A_v = \frac{g_m v_{SG} r_L}{v_{SG}}$$
$$A_v = g_m r_L \qquad (11.8)$$

Equation 11.8 is essentially the same as the gain equation derived for the non-swamped common source amplifier; the notable change being the lack of the minus sign indicating that this circuit does not invert the signal.

Input Impedance

Conveniently, the input impedance of the common gate amplifier is found in the same manner as the output impedance of the common drain follower, as both involve determining the impedance "looking into" the source terminal.

Referring to Figure 11.17, from the perspective of the signal generator, we find r_S in parallel with the impedance looking into the source terminal. The voltage at this node is v_{SG} and the current entering this node is i_D. The ratio of the two must yield the impedance looking into the source.

$$Z_{source} = \frac{v_{SG}}{i_D}$$
$$Z_{source} = \frac{v_{SG}}{g_m v_{SG}}$$
$$Z_{source} = \frac{1}{g_m} \qquad (11.9)$$

Therefore, the input impedance is

$$Z_{in} = r_S \parallel \frac{1}{g_m} \qquad (11.10)$$

Typically, this value will be far smaller than the input impedance of typical common source amplifiers and common drain followers.

Output Impedance

The analysis of common gate input impedance is virtually identical to that of the common drain amplifier. The result is replicated here for convenience.

$$Z_{out} \approx R_D \qquad (11.11)$$

For the amplifier shown in Figure 11.18, determine the input impedance and output voltage. Assume V_{in} is 50 mV and g_m = 5 mS.

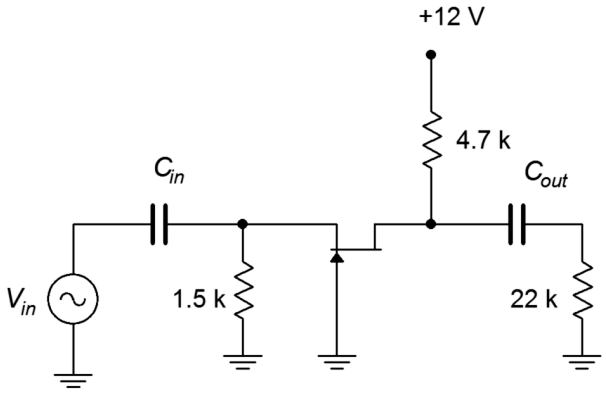


Figure 11.18
Circuit for Example 11.5.

Note that this circuit uses self bias. Fortunately, we already have the transconductance value stated in the problem, and need not go through a complete bias analysis. We'll begin with the voltage gain and proceed to the output voltage.

$$A_v = g_m r_d$$
$$A_v = 5\,\text{mS}(4.7\,\text{k}\Omega \parallel 22\,\text{k}\Omega)$$
$$A_v = 19.4$$

Thus V_{out} is 19.4 times 50 mV, or roughly 968 mV. Now for the input impedance.

$$Z_{in} = r_S \parallel \frac{1}{g_m}$$
$$Z_{in} = 1.5\,\text{k}\Omega \parallel \frac{1}{5\,\text{mS}}$$
$$Z_{in} = 176\,\Omega$$

11.6 Multi-stage and Combination Circuits

The rules for multi-stage circuits utilizing JFETs are the same as those discussed for BJTs: Steps must be taken to ensure that the bias of one stage does not adversely affect the bias of surrounding stages (typically by using coupling capacitors or going to a DC coupled system), the load for a given stage will be the input impedance of the following stage, the input impedance of the system will be the input impedance of the first stage, and the system gain will be the product of the individual stage gains.

Keeping those items in mind, there are no limits concerning mixing BJTs with JFETs, or mixing N-channel with P-channel devices. There are certain practical issues, however, that might dictate where certain devices are used. JFETs, due to their high input impedance and modest gain potential, tend to be used at the front end of amplifying systems. Their comparatively low self-noise is also a bonus at this location. BJTs, on the other hand, have high gain potential and tend to be used in the remaining stages. Their high distortion can be tamed through swamping.

To examine the possibilities, let's walk through the mixed, multi-stage amplifier presented in Figure 11.19.

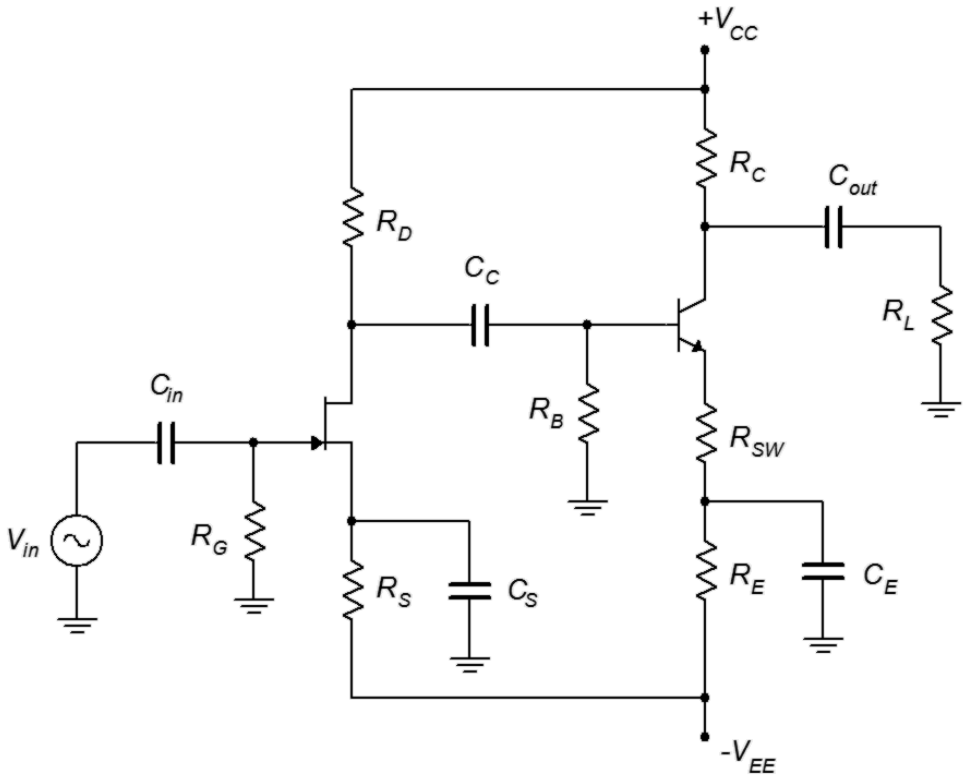


Figure 11.19
Two-stage JFET/BJT amplifier.

This amplifier uses a bipolar power supply which gives the designer a lot of flexibility. The first stage consists of a JFET common source amplifier. It utilizes combination bias (notice that R_S is connected to the shared negative supply, V_{EE}, that also serves as V_{SS}). C_S bypasses the source resistor so this stage does not use swamping. Distortion should not be an issue unless the input signal is fairly large. The load for this stage is R_D in parallel with the input impedance to the second stage (coupling capacitor C_C will appear ideally as a short for signal frequencies).

The second stage utilizes an NPN BJT configured as a swamped common emitter amplifier. It utilizes two-supply emitter bias. Its input impedance is the parallel combination of R_B and $Z_{in(base)}$. The base input impedance, in turn, is a function of β and R_{SW} (r'_e will have minimal impact due to the swamping resistor). The load for this stage will be R_L in parallel with R_C. That value divided by R_{SW} will yield the approximate stage gain (once again, ignoring r'_e). Although the second stage will be dealing with a larger signal, distortion will be mitigated by the swamping resistor.

The system gain will be the product of the two stage gains. As they both invert the signal, the inversion of the inversion will lead to an output signal that is in phase with the input signal. The system input impedance will depend on the JFET first stage and can be approximated to be equal to R_G, at least at low frequencies.

11.7 Ohmic Region Operation

As noted in the previous chapter, the JFET's operational curves span three regions. Two have been discussed: the constant current region is where the normal amplifiers and followers are biased, and breakdown is a region to be avoided due to potential damage. The third region is known as the *ohmic region*, or *triode region*. It occurs in the area where V_{DS} is less than the pinch-off voltage, V_P. In this area, the device behaves more like a resistor than like a current source. If we were to examine a family of drain curves, like those of Figure 10.3, and magnify the area near the origin, we would see something like the plot in Figure 11.20.

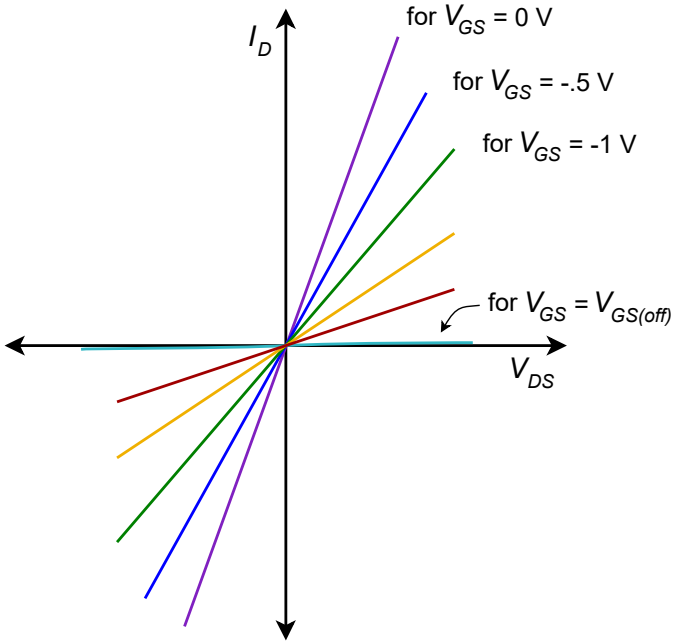


Figure 11.20
FET ohmic region.

If V_{DS} is a small value, typically less than 100 mV or so, each of the curves appears as a straight line. Further, the slope of that line is a function of the gate-source voltage, V_{GS}. The closer V_{GS} is to 0 V, the steeper the slope (violet line, labeled "for V_{GS} = 0 V") and the closer V_{GS} is to $V_{GS(off)}$, the more shallow the slope (shorter dark red line toward bottom). Finally, if $V_{GS} = V_{GS(off)}$, the slope is nearly zero (aqua line, nearly horizontal and straddling the V_{DS} axis). Because this is a plot of drain current versus drain-source voltage, the slope indicates the conductance of the channel. In somewhat more useful terms, we can say that the reciprocal of the slope indicates the resistance of the channel. Therefore, if V_{GS} = 0 V, the channel resistance will be at its minimum, and when $V_{GS} = V_{GS(off)}$, the channel resistance will be at its maximum. The maximum channel resistance can be quite high, well into the hundreds of kilo-ohms. The minimum channel resistance varies considerably from device to device. It is found on a data sheet as $r_{DS(on)}$. $r_{DS(on)}$ can be as small as a few ohms for specialized JFETs and as large as hundreds of ohms for general purpose

devices.[42] For example, the data sheet for the J111 series JFETs found in Figure 10.7 shows maximum values of 30 Ω, 50 Ω and 100 Ω for the J111, J112 and J113, respectively. The channel resistance does *not* follow a linear relation with V_{GS}.

To be more specific, in this region the drain current no longer follows the characteristic equation we used for biasing (Equation 10.1). The drain current equation in the ohmic region is:

$$I_D = \frac{V_{DS}}{V_P} 2 I_{DSS} \left(\left(1 - \frac{V_{GS}}{V_P} \right) - \frac{V_{DS}}{V_P} \right) \qquad (11.12)$$

Where $V_P = | V_{GS(off)} |$ and V_{GS} is to be taken as an absolute value and lies between 0 and V_P.

Recalling that, in general, $r_{DS} = V_{DS} / I_D$, we can substitute Equation 11.12 for I_D and, after including the definition of g_{m0}, arrive at an expression for r_{DS}:

$$r_{DS} = \frac{V_P}{g_{m0} \left(V_P - V_{GS} - \frac{V_{DS}}{2} \right)}$$

For small values of V_{DS}, this reduces to a simple equation:

$$r_{DS} = \frac{V_P}{g_{m0} \left(V_P - V_{GS} \right)} \qquad (11.13)$$

What we have created here is *voltage-controlled resistor*. Equation 11.13 shows that the resistance of the channel is a function of the gate-source voltage: the channel resistance will be at its minimum ($r_{DS(on)}$) when $V_{GS} = 0$ V, and it approaches infinity when V_{GS} equals $V_{GS(off)}$. Generally, there are two applications that make use of the ohmic region: an electronic rheostat/potentiometer and an analog switch. A simple circuit that can be used for either application is shown in Figure 11.21.

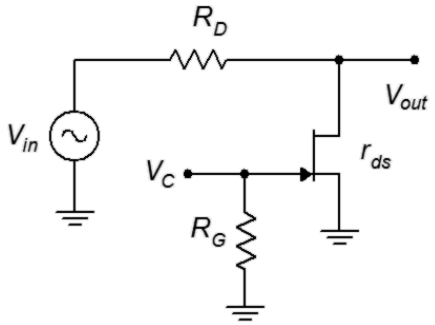


Figure 11.21
Using the JFET as a voltage-controlled resistor or switch.

42 $r_{DS(on)}$ can be as little as a few milliohms for specialized high power MOSFETs (Chapters 12 and 13).

Note that no external bias is applied to the circuit. Instead, a control voltage, V_C, is applied to the gate and the input signal is applied to a resistor attached to the drain terminal. The output is taken across the JFET's drain-source.

The idea behind this circuit is the basic resistive voltage divider. The JFET's channel resistance, r_{DS}, forms a voltage divider along with R_D.

$$v_{out} = V_{in} \frac{r_{DS}}{r_{DS} + R_D}$$

If $r_{DS} >> R_D$, v_{out} approaches v_{in}. Conversely, if $r_{DS} << R_D$, v_{out} approaches zero. Normally, R_D is set somewhere between the maximum and minimum channel resistances in order to obtain the widest range of operation.

As the control voltage V_C is V_{GS}, then V_C controls the size of v_{out}. If we set V_C to 0 V, r_{DS} is very small and thus $v_{out} \approx 0$. On the other hand, if V_C is set to a large negative potential (beyond $V_{GS(off)}$), then $v_{out} \approx V_{in}$. If V_C is set between these extremes then v_{out} will be somewhere in the middle range. If V_C is continuously variable, then the circuit behaves like a solid-state potentiometer. If, in contrast, V_C is only set at the limits, then the circuit behaves like a switch, either allowing or preventing the signal from transferring through. This is known as an *analog switch*.

A single JFET/resistor combination as shown in Figure 11.21 will have limited isolation as an analog switch and only a modest range of adjustment when used as a voltage-controlled potentiometer. To improve performance, multiple circuits can be cascaded or other JFETs can be added to create a “pi” attenuator network.

This voltage-controlled resistor has a huge advantage over traditional electro-mechanical potentiometers and switches: speed. In this circuit, the resistance can be changed at very high rates, essentially, as fast as V_C can change. Consequently, it would be no big deal to switch the input signal on and off at rates well over 100,000 times per second. No mechanical switch or potentiometer can hope to perform anywhere near that speed, and any attempt to do so would lead to the devices burning up from the friction. In general, *flaming potentiometers* are frowned upon during the design and development process, although it would make a decent name for an indie rock band. Another advantage is that a switch can be thrown “remotely”, that is, we only need to route the control voltage to the switch operator, not the signal itself. This can reduce system noise. It's also easier to implement if the switch is being “thrown” programmatically, such as via a microcontroller.

Example 11.6

For the circuit shown in Figure 11.22, if the input signal is 50 mV, determine the output voltage for $V_C = 0$ VDC and −6 VDC. Assume $V_{GS(off)} = -5$ V, $r_{DS(on)} = 30\ \Omega$ and $r_{DS(off)} = 800$ kΩ.

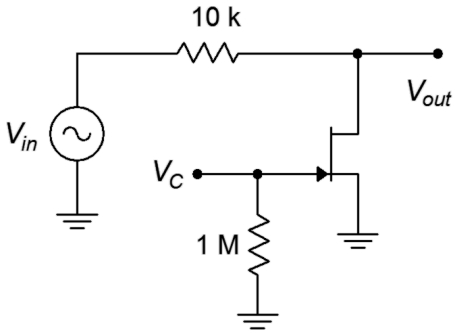


Figure 11.22
Circuit for Example 11.6.

For $V_C = 0$ VDC, the channel resistance will be at its minimum of $r_{DS(on)}$.

$$V_{out} = V_{in}\frac{r_{DS(on)}}{R_D + r_{DS(on)}}$$
$$V_{out} = 50\,\text{mV}\frac{30\,\Omega}{10\,\text{k}\,\Omega + 30\,\Omega}$$
$$V_{out} = 0.15\,\text{mV}$$

The signal has been reduced by a factor of over 330. That's not as good as a mechanical switch but if we cascaded two of these the overall reduction would be more than 100,000:1.

For $V_C = -6$ VDC, the channel resistance will be at its maximum of $r_{DS(off)}$.

$$V_{out} = V_{in}\frac{r_{DS(off)}}{R_D + r_{DS(off)}}$$
$$V_{out} = 50\,\text{mV}\frac{800\,\text{k}\,\Omega}{10\,\text{k}\,\Omega + 800\,\text{k}\,\Omega}$$
$$V_{out} = 49.4\,\text{mV}$$

This represents nearly 99% of the input signal, so the signal is passed through cleanly.

Computer Simulation

To verify the results of the preceding example, the circuit is entered into a simulator as shown in Figure 11.23. A J111 JFET model is used which has parameters similar to those used in the example.

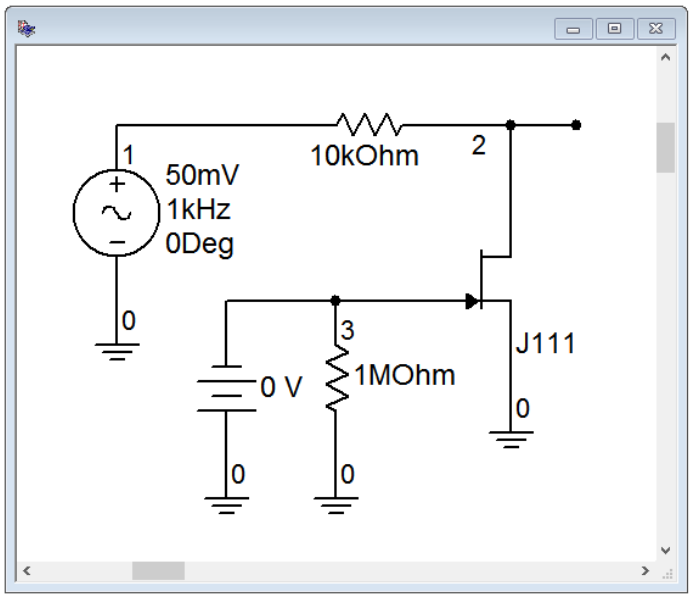


Figure 11.23
Analog switch in simulator.

A transient analysis is run twice; the first time is with a control voltage of 0 V and the second with a control voltage of −6 V. In the first case, the output should be just a small residual and in the second, we should see the full input signal. The results of the first trial are shown in Figure 11.24 while the second is shown in Figure 11.25.

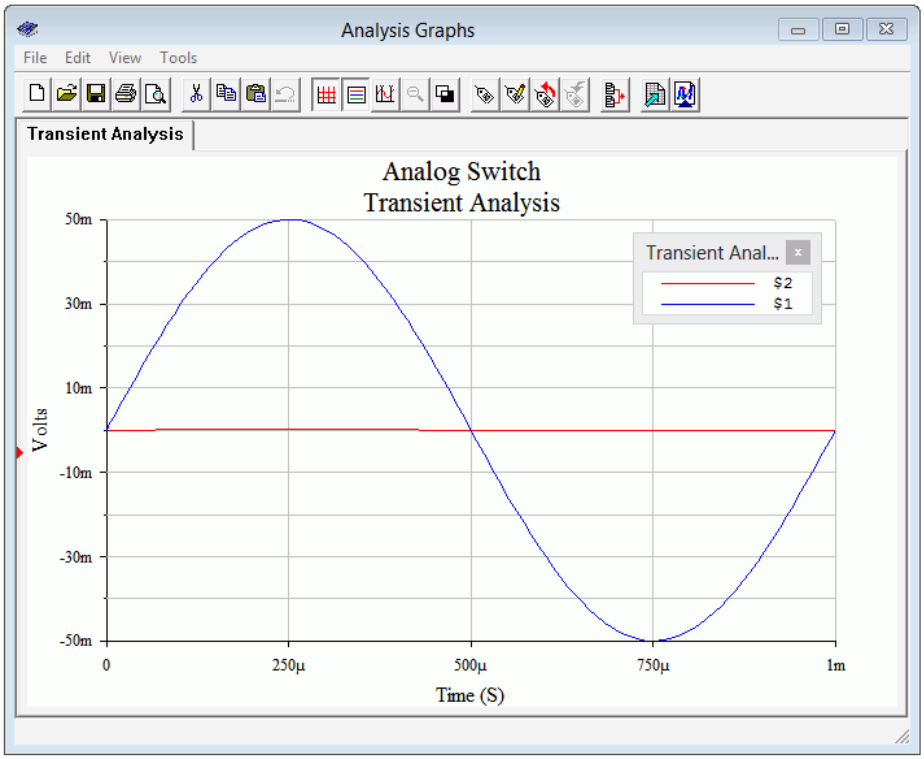


Figure 11.24
Transient analysis using $V_C = 0$ *V.*

With V_C at 0 V, the output trace (red, at node 2) is nearly flat. The precise value of its peak is 0.167 mV, not far from the value calculated in Example 11.5.

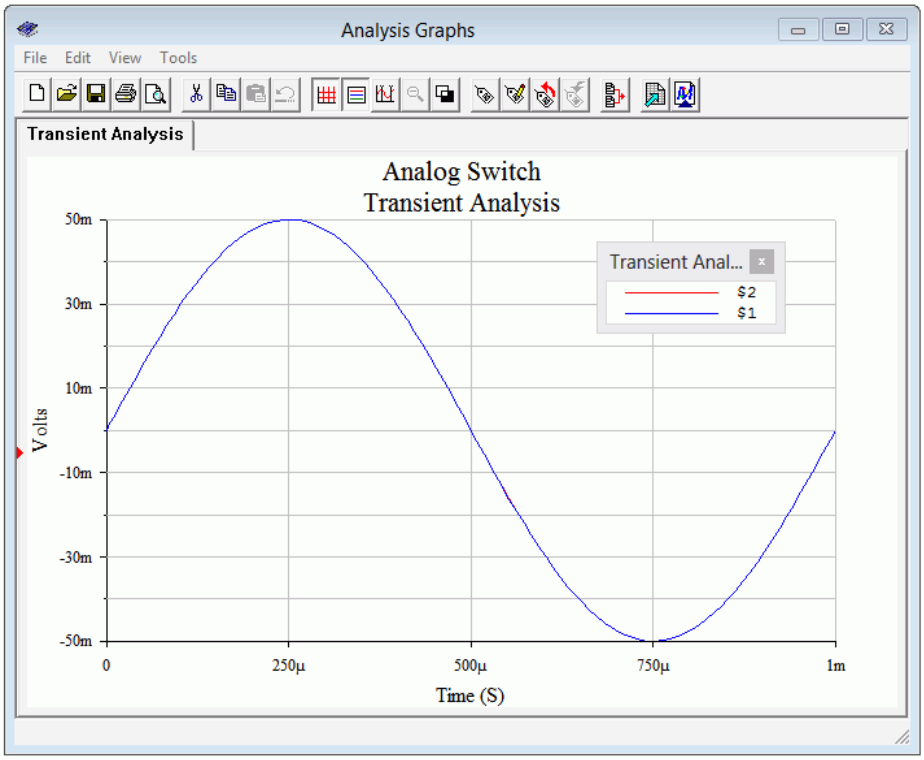


Figure 11.25
Transient analysis using $V_C = -6$ V.

In contrast, when $V_C = -6$ V the JFET is off, offering a high impedance and no loss of signal. At first glance, it may appears as though the output trace is missing but what has happened is that it is hidden behind the input trace (blue, node 1). The amplitudes are virtually identical so the blue trace completely obscures the red trace.

Summary

JFETs can be used to create both voltage amplifiers and voltage followers. The common source configuration is similar to the BJT's common emitter configuration. It offers voltage gain with signal inversion. The amplifier can be built upon any of the bias schemes presented in the preceding chapter. Biasing circuits that made use of a source resistor, such as self bias and combination bias, may also use swamping. Swamping will decrease available voltage gain but reduce distortion.

The JFET voltage follower, or source follower, is similar to the BJT's emitter follower. It offers a voltage of gain of nearly unity without inversion, a high input impedance and a low output impedance.

In general, JFETs do not offer as high of a gain as BJTs. The parameter comparable to the BJT's r'_e is the transconductance, g_m. Further, they tend to offer very high input impedance values compared to BJTs. This is due to using a reverse biased junction instead of a forward biased junction.

JFETs can also be used in their ohmic region to create voltage-controlled resistances and analog switches. A key parameter in these applications is the minimum channel resistance, $r_{DS(on)}$.

Review Questions

1. What are the functional differences between common source and common drain amplifiers?
2. Compare and contrast common source amplifiers to common emitter amplifiers.
3. Compare and contrast common drain followers to common collector followers.
4. Compare and contrast common gate amplifiers to common base amplifiers.
5. What is the ohmic region?
6. What is an analog switch and how does it function?

Problems

Analysis Problems

1. For the amplifier of Figure 11.26, determine Z_{in} and A_v. $I_{DSS} = 12$ mA, $V_{GS(off)} = -2$ V, $V_{DD} = 15$ V, $R_G = 220$ kΩ, $R_D = 2$ kΩ, $R_L = 3.3$ kΩ, $R_S = 330$ Ω.

2. For the amplifier of Figure 11.26, determine Z_{in} and V_{out}. $V_{in} = 50$ mV, $I_{DSS} = 15$ mA, $V_{GS(off)} = -3$ V, $V_{DD} = 20$ V, $R_G = 270$ kΩ, $R_D = 2$ kΩ, $R_L = 6.8$ kΩ, $R_S = 270$ Ω.

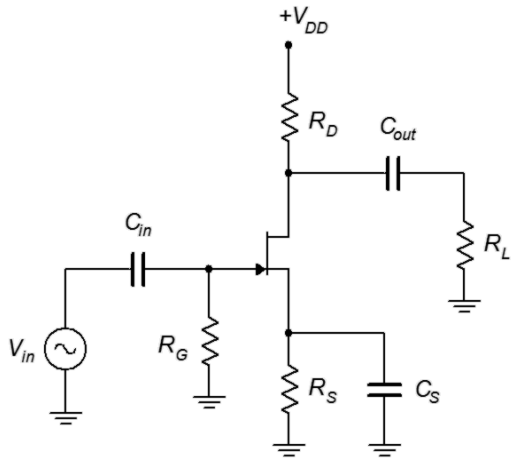


Figure 11.26

3. For the amplifier of Figure 11.27, determine Z_{in} and V_{out}. V_{in} = 60 mV, I_{DSS} = 10 mA, $V_{GS(off)}$ = −3 V, V_{DD} = 20 V, V_{SS} = −6 V, R_G = 270 kΩ, R_D = 2 kΩ, R_L = 4 kΩ, R_S = 1.8 kΩ, R_{SW} = 200 Ω.

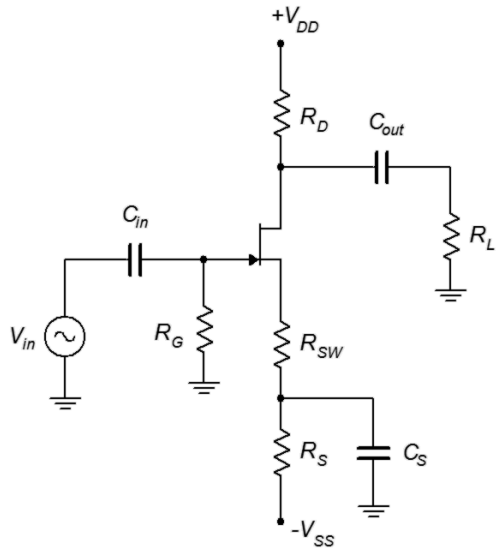


Figure 11.27

4. For the amplifier of Figure 11.27, determine Z_{in} and A_v. I_{DSS} = 12 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 18 V, V_{SS} = −4 V, R_G = 330 kΩ, R_D = 2.2 kΩ, R_L = 10 kΩ, R_S = 3 kΩ, R_{SW} = 100 Ω.

5. For the amplifier of Figure 11.28, determine Z_{in} and A_v. I_{DSS} = 12 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 18 V, V_{EE} = −4 V, R_G = 390 kΩ, R_D = 2.2 kΩ, R_E = 1 kΩ, R_L = 20 kΩ.

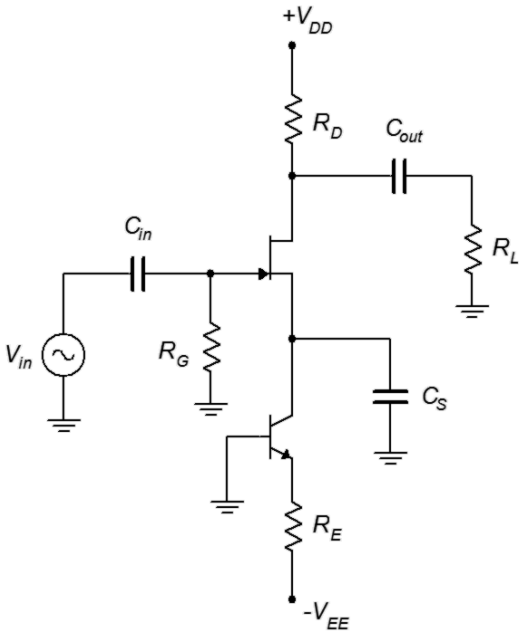
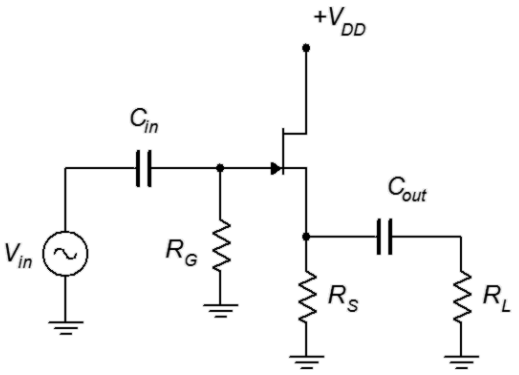


Figure 11.28

6. For the amplifier of Figure 11.28, determine Z_{in} and V_{out}. V_{in} = 70 mV, I_{DSS} = 12 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 18 V, V_{EE} = −4 V, R_G = 390 kΩ, R_D = 2.2 kΩ, R_L = 20 kΩ.

7. For the circuit of Figure 11.29, determine Z_{in} and A_v. I_{DSS} = 12 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 10 V, R_G = 220 kΩ, R_L = 3.3 kΩ, R_S = 330 Ω.

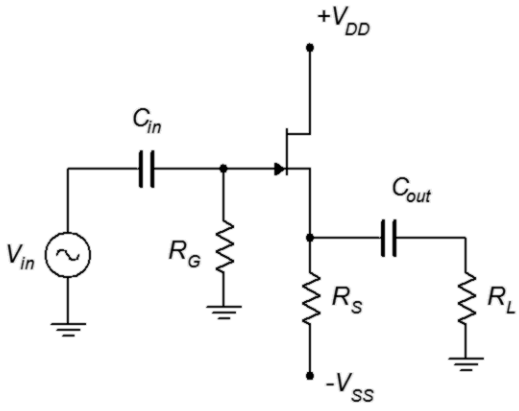
Figure 11.29



8. For the circuit of Figure 11.29, determine Z_{in} and V_{out}. V_{in} = 200 mV, I_{DSS} = 15 mA, $V_{GS(off)}$ = −3 V, V_{DD} = 12 V, R_G = 270 kΩ, R_L = 1.8 kΩ, R_S = 270 Ω.

9. For the circuit of Figure 11.30, determine Z_{in} and V_{out}. V_{in} = 100 mV, I_{DSS} = 10 mA, $V_{GS(off)}$ = −3 V, V_{DD} = 15 V, V_{SS} = −6 V, R_G = 470 kΩ, R_L = 4 kΩ, R_S = 1.8 kΩ.

Figure 11.30



10. For the circuit of Figure 11.30, determine Z_{in} and A_v. I_{DSS} = 18 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 14 V, V_{SS} = −6 V, R_G = 360 kΩ, R_L = 10 kΩ, R_S = 1 kΩ.

11. For the circuit of Figure 11.31, determine V_{out}. V_{in} = 100 mV, $r_{DS(on)}$ = 50 Ω, $r_{DS(off)}$ = 1 MΩ, $V_{GS(off)}$ = −3 V, V_C = −6 V, R_G = 270 kΩ, R_D = 6.8 kΩ.

12. For the circuit of Figure 11.31, determine V_{out}. V_{in} = 100 mV, $r_{DS(on)}$ = 75 Ω, $r_{DS(off)}$ = 750 kΩ, $V_{GS(off)}$ = −3 V, V_C = 0 V, R_G = 180 kΩ, R_D = 5.1 kΩ.

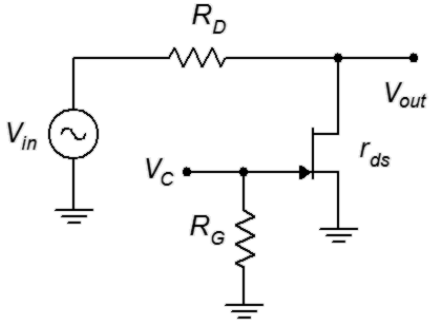


Figure 11.31

13. For the circuit of Figure 11.32, determine the voltage gain and V_{out}. Assume V_{in} = 10 mV and g_m = 2 mS.

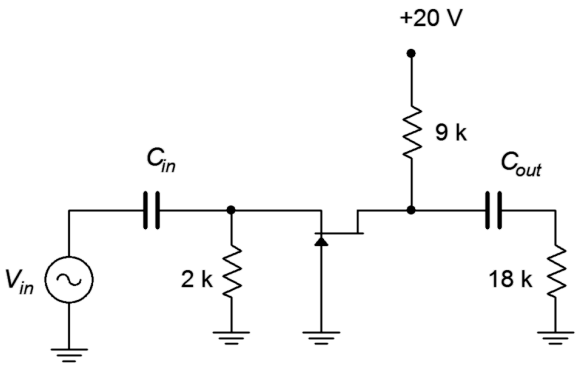


Figure 11.32

14. For the circuit of Figure 11.32, determine the input and output impedances. Assume g_m = 2 mS.

15. For the circuit of Figure 11.33, determine the input and output impedances. Assume g_m = 4 mS.

16. For the circuit of Figure 11.33, determine the voltage gain and V_{out}. Assume V_{in} = 100 mV and g_m = 4 mS.

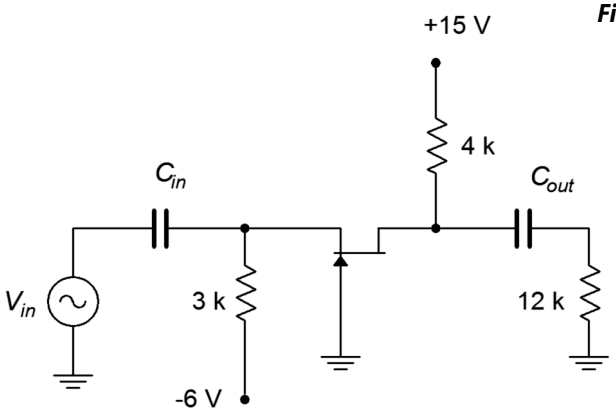


Figure 11.33

Design Challenge Problems

17. Following the circuit of Figure 11.27, design an amplifier with a gain of at least 4 and an input impedance of at least 300 kΩ. R_L = 10 kΩ. The JFET has the following parameters: $V_{GS(off)}$ = −2 V, I_{DSS} = 15 mA. Try to use standard resistor values.

18. Using the circuit of Figure 11.29, design a follower with a gain of at least 0.7 and an input impedance of at least 500 kΩ. R_L = 1 kΩ. The JFET has the following parameters: $V_{GS(off)}$ = −3 V, I_{DSS} = 20 mA. Try to use standard resistor values.

Computer Simulation Problems

19. Utilizing manufacturer's data sheets, find devices with the following specifications (typical) and verify them using the measurement techniques presented in the prior chapter. Device 1: $V_{GS(off)}$ = −2 V, I_{DSS} = 15 mA. Device 2: $V_{GS(off)}$ = −3 V, I_{DSS} = 20 mA.

20. Using the device model from the preceding problem, verify the design of Problem 17.

21. Using the device model from Problem 19, verify the design of Problem 18 .

12 Metal Oxide Semiconductor FETs (MOSFETs)

12.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Discuss the functional differences between MOSFETs and JFETs.
- Draw and explain a basic DC bias model for a MOSFET.
- Graph the transconductance curves for both DE-MOS and E-MOS transistors, and describe their functional differences.
- Perform DC bias analysis on various MOSFET circuits.
- Explain necessary ESD precautions for MOS devices.

12.1 Introduction

The MOSFET shares many similarities with the JFET including very low gate current and being modeled as a voltage-controlled current source. It is also available in N- and P-channel varieties. Unlike the JFET, it has two variations: the depletion-enhancement mode variant, or DE-MOSFET; and the enhancement-only mode variant, or E-MOSFET. All of the bias types discussed for JFETs will work for DE-MOSFETs, plus a few others. E-MOSFETs, on the other hand, require new biasing prototypes.

For AC analysis, both common source and common drain amplifier topologies may be realized with DE- and E-MOSFETs. The equations for input impedance, voltage gain and the like are generally unchanged from the JFET. E-MOSFETs are also available as power devices. They have certain advantages over power BJTs, including higher speed and a negative temperature coefficient of transconductance which means they are less likely to suffer from thermal instabilities such as current hogging.

One item of practical importance is that MOSFETs are very susceptible to ESD (electrostatic discharge) and special precautions must be taken to prevent accidental damage to the device. Unlike both the JFET and the BJT, the MOSFET does not rely on a PN junction for its operation. Instead, it uses a charge-based system not unlike a capacitor. The gate is, in fact, insulated from the channel. For this reason it is sometimes referred to as an IGFET, which stands for Insulated Gate FET. This insulation layer will lead to very, very high input resistance due to extremely low gate current but also leads to the issue of ESD susceptibility.

12.2 The DE-MOSFET

Like the JFET, the DE-MOSFET is based around the idea of modulating current flow through the drain-source channel by generating a depletion layer from a gate-source voltage. It achieves this through an entirely different process, though. To understand how the device is constructed, a simplified functional drawing of an N-channel DE-MOSFET is shown in Figure 12.1.

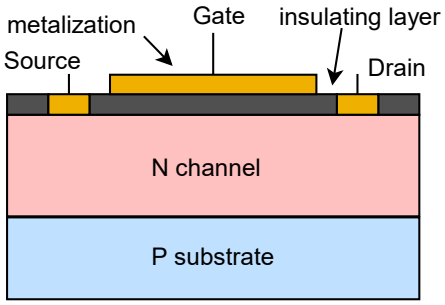


Figure 12.1
DE-MOSFET internal structure.

Here we see the N-type channel sitting on a P substrate. Drain and source leads are attached to either end. Above the channel is a very thin insulating layer (silicon dioxide is one possibility). Above this we have a metalization to which the gate terminal is attached. Note that there is no PN junction involved with the gate. To this we shall add external bias sources and limiting resistors, as shown in Figure 12.2.

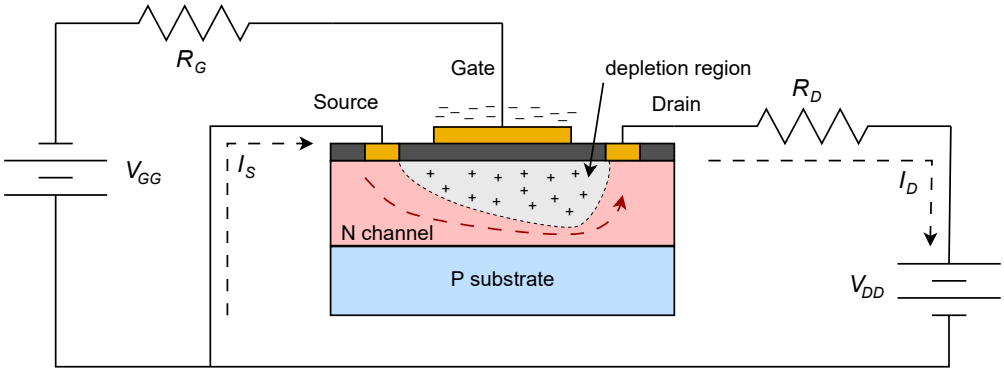


Figure 12.2
DE-MOSFET bias with electron flow.

The dashed lines represent electron current flow as in our previous device analyses. A positive supply, V_{DD}, is attached to the drain via a limiting resistor. A second supply, V_{GG}, is attached to the gate. Gate current can be approximated as zero, so $V_{GS} = V_{GG}$. If V_{GS} is zero, a certain amount of current will flow through the channel based on the channel's physical parameters and the applied drain-source potential. For relatively low values of V_{DS}, the channel will behave somewhat like a resistance. This is the same ohmic region as seen with the JFET. As V_{DS} increases, the channel will saturate and begin to behave like a constant current source. If V_{DS} is brought too high, the drain current increases sharply as the device enters the breakdown region. The general behavior mimics that of a JFET. Note that the current moves laterally, across the device, so this type of construction is referred to as a *lateral MOSFET.*

If V_{GS} is set to a modest negative voltage, a depletion region will develop inside the channel. Basically, the gate is acting like one plate of a capacitor, the channel like the other plate, and the insulating layer is the dielectric. Just like a capacitor, the negative charge on the gate "plate" leads to an equivalent positive charge on the channel "plate". As the channel is made of N-type material, this action creates a region devoid of free charges, hence a depletion region. This depletion region will lead to pinch-off sooner, and thus a lower current in the saturation region. The more negative V_{GS} is made, the greater the depletion region and the lower the corresponding drain current. Eventually, if V_{GS} is brought negative enough, the channel will be blocked and no drain current will flow. This voltage is referred to as $V_{GS(off)}$ (again). The current produced when $V_{GS} = 0$ V is likewise referred to as I_{DSS}. This mode of operation is referred to as *depletion mode* because of the depletion region that is created.

What makes the DE-MOSFET distinct from the JFET is what happens when $V_{GS} > 0$ volts. In a JFET, this would forward bias the junction and control would be lost. Here, however, a positive V_{GS} simply reverses the polarities associated with the gate and channel "plates". Thus, a positive V_{GS} enhances channel conductivity and drain current increases as V_{GS} is brought more positive. This mode of operation is called *enhancement mode*. This also means that I_{DSS} is no longer the maximum drain current of which the device is capable. A characteristic curve is shown in Figure 12.3, below.

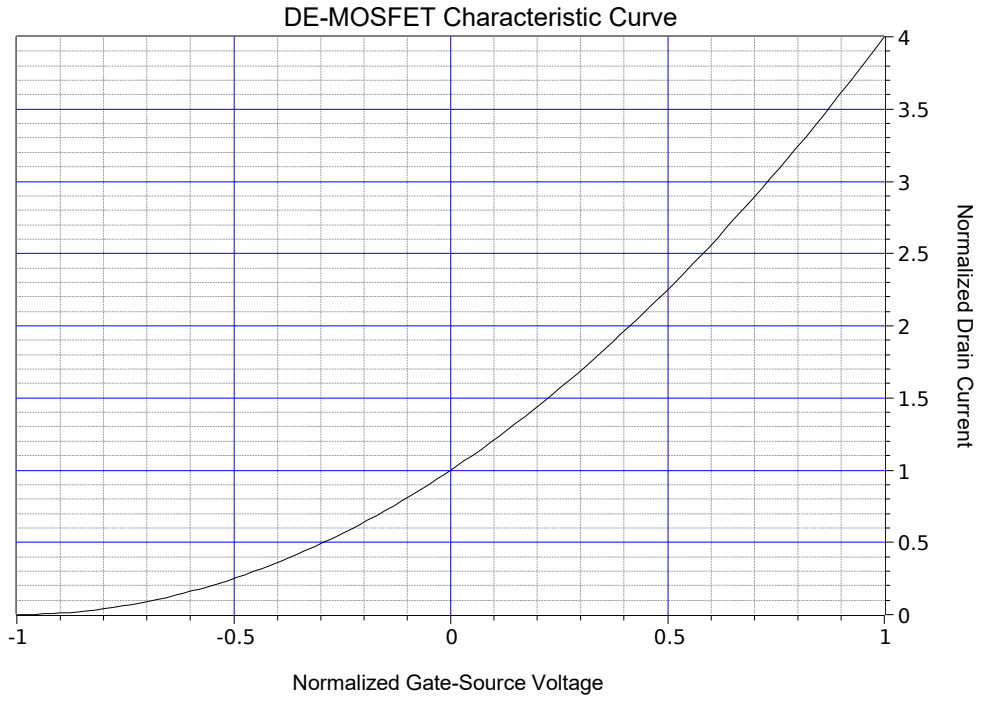


Figure 12.3
Characteristic curve for DE-MOSFET (note: this uses $-V_{GS}/V_{GS(off)}$ for the normalized voltage so that the quadrants do not appear reversed compared to a typical device curve).

This curve is essentially the same curve as presented for the JFET with the exception that it extends into the first quadrant. This makes the DE-MOSFET a unique device in that it can operate in two different quadrants.

The device equation for operation when $V_{DS} > V_P$ is also the same, but with an extended range for V_{GS}:

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)^2 \qquad (12.1)$$

Where
V_{GS} is the gate-source voltage ($V_{GS(off)} \leq V_{GS} \leq V_{GS(max)}$),
I_D is the drain current,
I_{DSS} is the maximum current,
$V_{GS(off)}$ is the turn-off voltage.

$V_{GS(max)}$ may be found on a data sheet. Above this voltage the insulating layer will be damaged and the device will cease to function properly. A typical value for this might be in the range of 20 to 30 volts. The trick is that given the very small gate current, even a simple electrostatic discharge can damage the device. It is very easy to develop hundreds of volts static on the human body. In fact, it is generally not noticeable until the potential reaches a few thousand volts (as in body hair standing up). The consequence of this is that simply picking up the device could destroy it.[43]

There are a couple different ways of dealing with this issue. The first way is to add back-to-back Zener diodes across the device during its manufacture. The problem with this is that the diode leakage current will be greater than the gate current and this degrades performance. The other technique is to prevent the charge from getting to the device in the first place. For example, the MOSFET can be shipped in conductive plastic (not to be confused with ordinary plastic or polystyrene foam). Some devices are shipped with a metal shorting that encompasses all of the leads. Also, during manufacture or prototyping, environmental controls are established to minimize the creation of static charges, optimal humidity being important as one example. Workers who handle devices may work on special conductive mats or wear wrist straps that are attached to ground. These items are only mildly conductive, that is, of high resistance, as it would not be safe to electrically ground a human working in an electrical lab. The devices are conductive enough to bleed off static charge but not so conductive as to present a shock hazard. Once installed on the circuit board, normal ESD precautions apply.

As the device's characteristic equation has not changed, many of the items derived for the JFET still apply to the DE-MOSFET. This includes the transconductance equation plot found in Chapter 10.

43 Which brings to mind the old question of what to store a universal solvent in.

As the transconductance equation is unchanged with the exception of an extended range for V_{GS}, the definition for g_{m0} is also unchanged.

$$g_{m0} = -\frac{2\,I_{DSS}}{V_{GS(off)}} \tag{12.2}$$

$$g_m = g_{m0}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right) \tag{12.3}$$

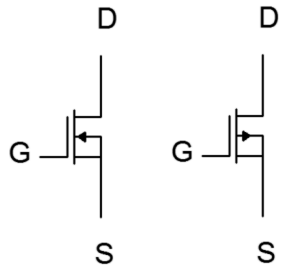
It is worth noting that g_{m0} no longer represents the maximum device transconductance because I_{DSS} no longer represents the maximum drain current as seen in Figure 12.3. To illustrate it another way,

$$g_m = g_{m0}\sqrt{\frac{I_D}{I_{DSS}}} \tag{12.4}$$

It is very important to watch the sign of V_{GS} in Equation 12.3. In enhancement mode, a positive V_{GS} will lead to a g_m greater than g_{m0} due to the double negative.

The schematic symbols for the DE-MOSFET are shown in Figure 12.4. As is the norm, the arrow points in the direction of N material, with the central vertical bar representing the channel. The arrow is attached to the substrate. In some devices this is brought out of the packaging as a fourth lead although in many it is simply tied back to the source terminal as shown here. Finally, note how the gate terminal is not drawn connected to the body of the device, emphasizing its isolated nature.

Figure 12.4
DE-MOSFET schematic symbols. N-channel (left) and P-channel (right).



12.3 DE-MOSFET Biasing

As the characteristic equations of the JFET and DE-MOSFET are the same, the DC biasing model is the same. Consequently, the DE-MOSFET can be biased using any of the techniques used with the JFET including self bias, combination bias and current source bias as these are all second quadrant biasing schemes (i.e., have a negative V_{GS}). The self bias and combination bias equations and plots from Chapter 10 may be used without modification. The DE-MOSFET also allows first quadrant operation so a couple of new biasing forms become available: zero bias and voltage divider bias. In reality, both are variations on constant voltage bias but which utilize the first quadrant.

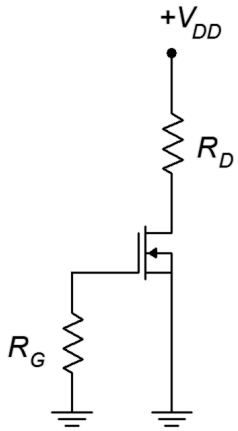
Zero Bias

Zero bias is unique. In some ways it can be thought of as a cross between self bias and constant voltage bias. Like self bias, it does not require a second DC source for the gate or source terminal. Like constant voltage bias, there is no need for a source

resistor, R_S. A prototype of zero bias is shown in Figure 12.5. There is no question that it is a minimal parts-count circuit.

Zero bias is so named because it operates at $V_{GS} = 0$ V. Recall that the gate current is ideally zero, thus there is no drop across R_G and $V_G = 0$ V as a consequence. The source is tied directly to ground, therefore V_{GS} must equal 0 V. As V_{GS} doesn't change, this can be thought of as a form of constant voltage bias. The interesting bit is that when an AC signal is applied to the gate, its negative portion will pull the MOSFET down into depletion mode and the positive portion will push the operation into enhancement mode. Because the device can operate in this fashion, conducting current while straddling zero, so to speak, DE-MOSFETs are sometimes referred to as *normally on* devices.

Figure 12.5
Zero bias prototype.

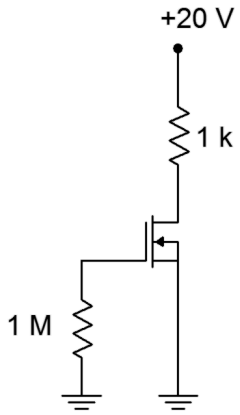


Determining the operating point for zero bias is startlingly easy. Because $V_{GS} = 0$ V, I_D must equal I_{DSS} and g_m must equal g_{m0}. Like all constant voltage biasing schemes, though, Q point stability is not very good. Another point to notice is that, as there is no source resistor, this bias is only applicable to non-swamped common source amplifiers. It cannot be used with a source follower or swamped amplifier (if a small swamping resistor is inserted into the source, technically the circuit can be classified as self bias, although the AC signal may still push operation into enhancement mode).

Example 12.1

Determine I_D, V_D and g_{m0} for the circuit shown in Figure 12.6. Assume $I_{DSS} = 12$ mA and $V_{GS(off)} = -3$ V.

Figure 12.6
Circuit for Example 12.1.



By inspection, as this is zero bias $I_D = I_{DSS}$, and therefore $I_D = 12$ mA. Using KVL and Ohm's law we can find V_D.

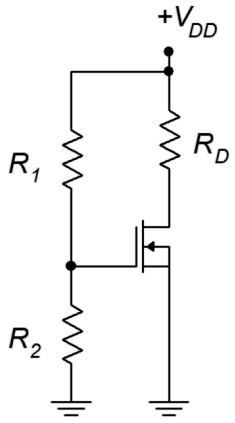
$$V_D = V_{DD} - I_D R_D$$
$$V_D = 20\,\text{V} - 12\,\text{mA} \times 1\,\text{k}\Omega$$
$$V_D = 8\,\text{V}$$

$$g_{m0} = -\frac{2 I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{2 \times 12\,\text{mA}}{-3\,\text{V}}$$
$$g_{m0} = 8\,\text{mS}$$

Voltage Divider Bias

Voltage divider bias is a form of constant voltage bias that operates in enhancement mode. A prototype circuit is shown in Figure 12.7. Note that the source terminal is connected directly to ground. This is important. If this was not the case, this would be a form of combination bias (basically shifting the V_{SS} supply up to ground and then shifting the gate voltage from ground up to a positive V_{SS} to maintain the same differential voltage). As such, it would be operating in depletion mode.

Figure 12.7
Voltage divider bias prototype.



The voltage divider comprised of R_1 and R_2 will establish a DC bias potential on the gate. As the source is at ground, $V_{GS} = V_G = V_{R2}$. Given that V_{DD} must be positive, then V_{GS} must be positive, and enhancement mode operation is a given.

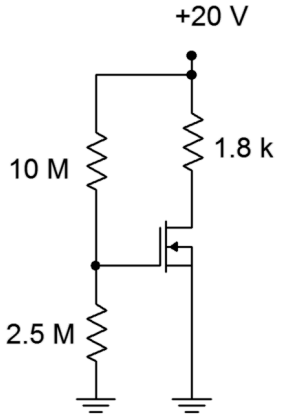
The most direct way to handle this is to determine the voltage divider potential and use either the characteristic equation (Equation 10.1) or associated graph to determine the drain current. Once I_D is found, the drain-source voltage may be found via the standard Ohm's law/KVL route.

Before continuing, note that the values of the divider resistors can be very high without creating biasing problems (unlike the BJT version of voltage divider bias). This is because the gate current is so small that even when using megohm values for the divider, the loading caused by the gate will not be noticeable.

Example 12.2

For the circuit of Figure 12.8, determine I_D and V_D. Assume $I_{DSS} = 2$ mA and $V_{GS(off)} = -6$ V.

Figure 12.8
Circuit for Example 12.2.



The voltage divider will yield V_{GS}.

$$V_{GS} = V_{DD}\left(\frac{R_2}{R_1 + R_2}\right)$$
$$V_{GS} = 20\,\text{V}\left(\frac{2.5\,\text{M}\Omega}{10\,\text{M}\Omega + 2.5\,\text{M}\Omega}\right)$$
$$V_{GS} = 4\,\text{V}$$

Use Equation 10.1 to find I_D.

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_{GS(off)}}\right)^2$$
$$I_D = 2\,\text{mA}\left(1 - \frac{4\,\text{V}}{-6\,\text{V}}\right)^2$$
$$I_D = 5.56\,\text{mA}$$

Use KVL and Ohm's law to find V_D.

$$V_D = V_{DD} - I_D R_D$$
$$V_D = 20\,\text{V} - 5.56\,\text{mA} \times 1.8\,\text{k}\Omega$$
$$V_D = 9.99\,\text{V}$$

Alternately, using the curve of Figure 12.3, we would first find the normalized gate-source voltage which is 4 V/6 V or 0.667 (note that the curve plots $-V_{GS}/V_{GS(off)}$ so that the quadrants do not appear reversed). From this the normalized drain current, I_D/I_{DSS}, may be determined to be approximately 2.8, yielding a drain current of 5.6 mA.

12.4 The E-MOSFET

The E-MOSFET is available in both low power and high power versions. It operates in enhancement mode (first quadrant) only. The construction of the low power version is similar to that of the DE-MOSFET but with one important distinction. A simplified cross-section of an N-channel E-MOSFET is shown in Figure 12.9.

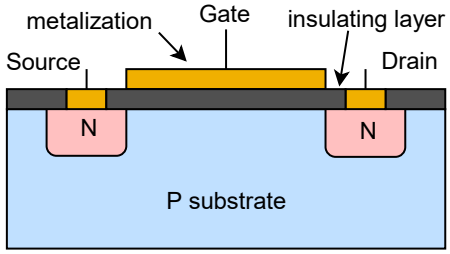


Figure 12.9
E-MOSFET internal structure.

In the E-MOSFET, the P material extends up through the channel and to the gate insulating layer. This has the effect of preventing current flow with negative gate-source voltages applied. Consequently, E-MOSFETs are sometimes referred to as *normally off* devices. In fact, the E-MOSFET will not conduct if V_{GS} is zero, or even for small positive values of V_{GS}. The P and N materials functionally create energy hills or barriers that prevent current flow through the channel. This can be compared to an NPN BJT that has an open base terminal: no collector current would flow (unless the collector-emitter voltage exceeded the breakdown limit).

To understand how the E-MOSFET functions, refer to Figure 12.10. This diagram shows the device with positive drain and gate supplies attached to it through limiting resistors. The dashed lines indicate electron current flow. As with the DE-MOSFET, the gate can be seen as one plate of a capacitor while the P material serves as the other plate. A positive voltage on the gate will lead to a negative charge on the P material side. If the charge is large enough, all of the holes in the P material can be filled leaving the portion of the material situated near the isolation layer neutral

(neither P nor N). Any further increase in gate voltage injects more negative charge into this region, this making it behave like N material. This is called an *N-type inversion layer* and it allows a path for current to flow. The more positive we make the gate voltage, the greater the effect, and the greater the current.

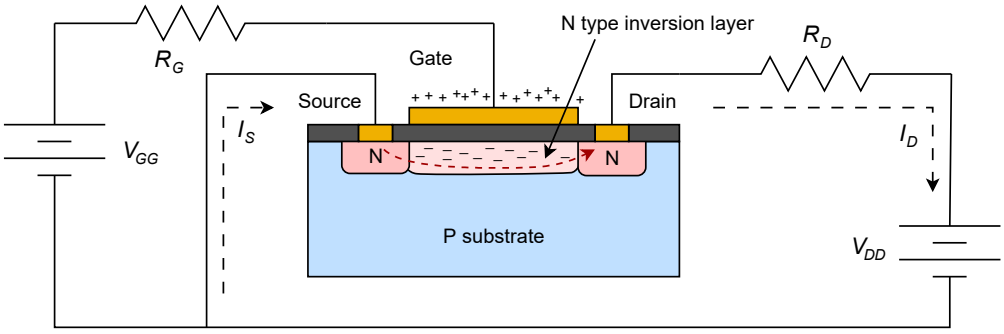


Figure 12.10
E-MOSFET bias with electron flow.

The voltage where the current begins to flow is called the threshold voltage and is usually denoted as $V_{GS(th)}$, although it is sometimes shortened to V_{th} or just V_t. Like both the JFET and DE-MOSFET, the E-MOSFET drain curve family exhibits three characteristic regions: the *ohmic* or *triode* region, the *constant current* or *saturation* region, and the *breakdown* region.

The characteristic equation for the E-MOSFET operating in its constant current region is given below. Like the other FETs examined, this is a square-law device.

$$I_D = k\left(V_{GS} - V_{GS(th)}\right)^2 \qquad (12.5)$$

Where
I_D is the drain current,
V_{GS} is the gate-source voltage ($V_{GS(th)} \leq V_{GS} \leq BV_{GS}$),
$V_{GS(th)}$ is the the threshold voltage,
k is a device parameter (a constant, units of amps/volt^2 or siemens/volt).

The derivative of Equation 12.5 yields the transconductance.

$$g_m = \frac{dI_D}{dV_{GS}} = 2k\left(V_{GS} - V_{GS(th)}\right) \qquad (12.6)$$

Equation 12.5 is plotted in Figure 12.11. The normalized gate-source voltage is $V_{GS}/V_{GS(th)}$ and the normalized drain current is the ratio of I_D to the current generated when V_{GS} is twice $V_{GS(th)}$. This curve is reminiscent of the characteristic curve of a BJT. First, they are both in the first quadrant. Second, both curves exhibit an increasing positive slope. Finally, the curves don't begin to "take off" until some specific turn-on voltage is reached. In the case of the BJT, that voltage is approximately 0.7 V for a silicon device. For the E-MOSFET, that voltage is $V_{GS(th)}$. Obviously though, the MOSFET curve does not increase as rapidly as the BJT curve.

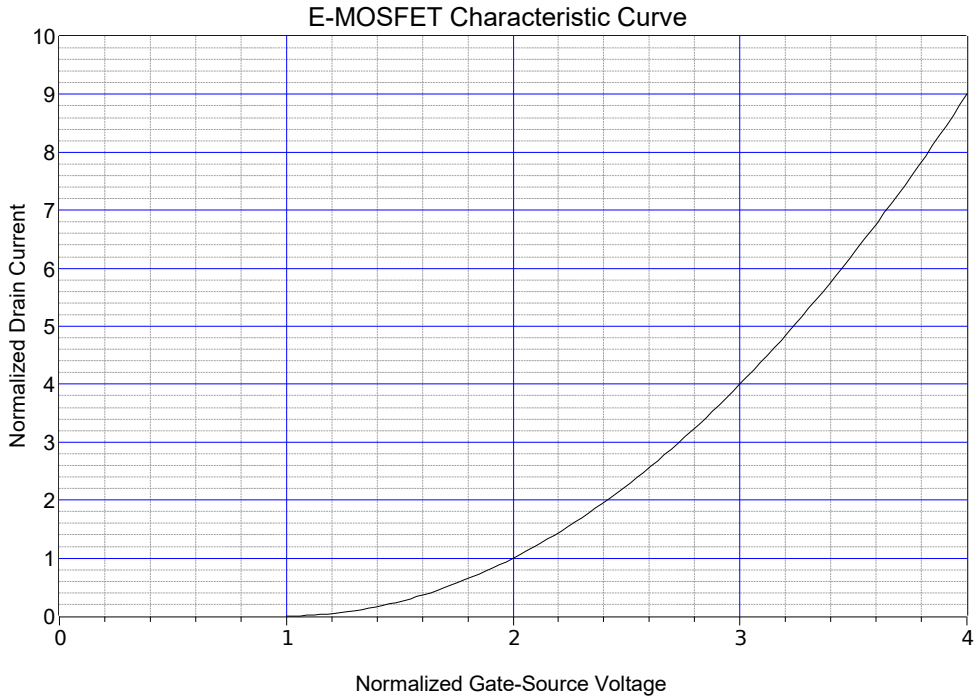


Figure 12.11
Characteristic curve for E-MOSFET.

The power E-MOSFET utilizes a different construction from low power MOSFETs and offers certain advantages over power BJTs including very fast switching speed and lower drive current demands. Consequently, they tend to be favored over BJTs in high power, high speed switching applications such as switching power supply regulators, DC-to-DC converters and class D amplifiers (Chapter 14). These devices also exhibit extremely low $r_{DS(on)}$ values, in some cases just a few milliohms. There are different methods of construction, the most recent being the trench style. A cutaway view is shown in Figure 12.12.

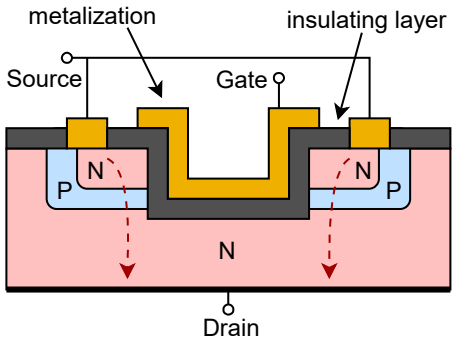


Figure 12.12
Construction of power Trench E-MOSFET.

A power E-MOSFET is made of a large number of cells, each featuring the U-shaped gate "trench" (an earlier style used a V-shaped trench). Note the location of the drain, now opposite of the gate and source. The advantage here is that the current flows *vertically* rather than laterally. This results in a much lower $r_{DS(on)}$ and considerably greater current capacity. The characteristic curve still echoes that of Figure 12.11 although it tends to be steeper when compared to low power devices.

12.5 E-MOSFET Data Sheet Interpretation

A data sheet for an E-MOSFET, the FDMS86180, is shown in Figure 12.13. This is an N-channel, high power device using trench construction.

FAIRCHILD

August 2016

FDMS86180

N-Channel Shielded Gate PowerTrench® MOSFET
100 V, 151 A, 3.2 mΩ

Features

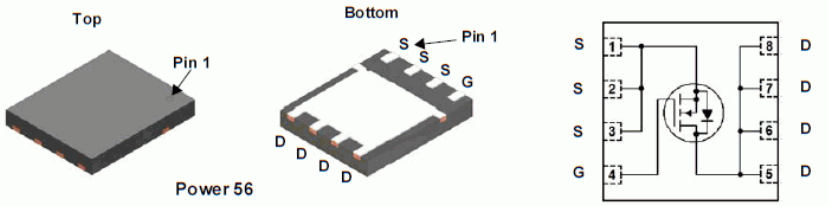
- Shielded Gate MOSFET Technology
- Max $r_{DS(on)}$ = 3.2 mΩ at V_{GS} = 10 V, I_D = 67 A
- Max $r_{DS(on)}$ = 7.9 mΩ at V_{GS} = 6 V, I_D = 33 A
- 50% Lower Qrr than Other MOSFET Suppliers
- Lowers Switching Noise/EMI
- MSL1 Robust Package Design
- 100% UIL Tested
- RoHS Compliant

General Description

This N-Channel MV MOSFET is produced using Fairchild Semiconductor's advanced PowerTrench® process that incorporates Shielded Gate technology. This process has been optimized to minimize on-state resistance and yet maintain superior switching performance with best in class soft body diode.

Applications

- Primary DC-DC MOSFET
- Synchronous Rectifier in DC-DC and AC-DC
- Motor Drive
- Solar



MOSFET Maximum Ratings T_A = 25 °C unless otherwise noted

Symbol	Parameter			Ratings	Units
V_{DS}	Drain to Source Voltage			100	V
V_{GS}	Gate to Source Voltage			±20	V
I_D	Drain Current -Continuous	T_C = 25 °C	(Note 5)	151	A
	-Continuous	T_C = 100 °C	(Note 5)	95	
	-Continuous	T_A = 25 °C	(Note 1a)	21	
	-Pulsed		(Note 4)	775	
E_{AS}	Single Pulse Avalanche Energy		(Note 3)	486	mJ
P_D	Power Dissipation	T_C = 25 °C		138	W
	Power Dissipation	T_A = 25 °C	(Note 1a)	2.7	
T_J, T_{STG}	Operating and Storage Junction Temperature Range			-55 to +150	°C

Thermal Characteristics

$R_{\theta JC}$	Thermal Resistance, Junction to Case	0.9	°C/W
$R_{\theta JA}$	Thermal Resistance, Junction to Ambient (Note 1a)	45	

Package Marking and Ordering Information

Device Marking	Device	Package	Reel Size	Tape Width	Quantity
FDMS86180	FDMS86180	Power 56	13 "	12 mm	3000 units

Figure 12.13a
FDMS86180 data sheet.
Used with permission from SCILLC dba ON Semiconductor.

One of the first things that might jump out is the "100% RoHS Compliant" green leaf logo in the upper center, meaning that the device meets the Restriction of Hazardous Substances directive. The device comes in the flat, multi-pin Power 56 package and features an $r_{DS(on)}$ of just a few milliohms. Continuous current

capability at room temperature is 151 amps with a pulsed current maximum of 775 amps. In Figure 12.13b we find a breakdown voltage of 100 volts and an I_{DSS} of only 1 µA. Recall that this is a normally off device, and thus I_{DSS} represents a leakage current. Continuing, $V_{GS(th)}$ varies between 2.0 and 4.0 volts, with 3.2 volts being typical. The forward transconductance, g_m (here referred to as g_{FS}) is 144 siemens at a drain current of 67 amps. This is orders of magnitude greater than what we might see with small signal devices. Turn-on and turn-off times are measured in the tens of nanoseconds, verifying the high speed switching ability of the device.

Figure 12.13b
FDMS86180 data sheet (cont).

Electrical Characteristics T_J = 25 °C unless otherwise noted.

Symbol	Parameter	Test Conditions	Min.	Typ.	Max.	Units
Off Characteristics						
BV_{DSS}	Drain to Source Breakdown Voltage	I_D = 250 µA, V_{GS} = 0 V	100			V
$\frac{\Delta BV_{DSS}}{\Delta T_J}$	Breakdown Voltage Temperature Coefficient	I_D = 250 µA, referenced to 25 °C		73		mV/°C
I_{DSS}	Zero Gate Voltage Drain Current	V_{DS} = 80 V, V_{GS} = 0 V			1	µA
I_{GSS}	Gate to Source Leakage Current	V_{GS} = ±20 V, V_{DS} = 0 V			100	nA
On Characteristics						
$V_{GS(th)}$	Gate to Source Threshold Voltage	V_{GS} = V_{DS}, I_D = 370 µA	2.0	3.2	4.0	V
$\frac{\Delta V_{GS(th)}}{\Delta T_J}$	Gate to Source Threshold Voltage Temperature Coefficient	I_D = 370 µA, referenced to 25 °C		-8		mV/°C
$r_{DS(on)}$	Static Drain to Source On Resistance	V_{GS} = 10 V, I_D = 67 A		2.4	3.2	mΩ
		V_{GS} = 6 V, I_D = 33 A		3.8	7.9	
		V_{GS} = 10 V, I_D = 67 A, T_J = 125 °C		4.0	5.4	
g_{FS}	Forward Transconductance	V_{DS} = 5 V, I_D = 67 A		144		S
Dynamic Characteristics						
C_{iss}	Input Capacitance	V_{DS} = 50 V, V_{GS} = 0 V, f = 1 MHz		4439	6215	pF
C_{oss}	Output Capacitance			2663	3730	pF
C_{rss}	Reverse Transfer Capacitance			24	55	pF
R_g	Gate Resistance		0.1	0.8	1.6	Ω
Switching Characteristics						
$t_{d(on)}$	Turn-On Delay Time	V_{DD} = 50 V, I_D = 67 A, V_{GS} = 10 V, R_{GEN} = 6 Ω		24	39	ns
t_r	Rise Time			12	22	ns
$t_{d(off)}$	Turn-Off Delay Time			30	48	ns
t_f	Fall Time			7	14	ns
Q_g	Total Gate Charge	V_{GS} = 0 V to 10 V; V_{DD} = 50 V, I_D = 67 A		60	84	nC
Q_g	Total Gate Charge	V_{GS} = 0 V to 6 V; V_{DD} = 50 V, I_D = 67 A		38	54	nC
Q_{gs}	Gate to Source Charge	V_{DD} = 50 V, I_D = 67 A		20		nC
Q_{gd}	Gate to Drain "Miller" Charge	V_{DD} = 50 V, I_D = 67 A		12		nC
Q_{oss}	Output Charge	V_{DD} = 50 V, V_{GS} = 0 V		175		nC
Drain-Source Diode Characteristics						
V_{SD}	Source to Drain Diode Forward Voltage	V_{GS} = 0 V, I_S = 2.1 A (Note 2)		0.7	1.2	V
		V_{GS} = 0 V, I_S = 67 A (Note 2)		0.8	1.3	
t_{rr}	Reverse Recovery Time	I_F = 33 A, di/dt = 300 A/µs		44	71	ns
Q_{rr}	Reverse Recovery Charge			109	207	nC
t_{rr}	Reverse Recovery Time	I_F = 33 A, di/dt = 1000 A/µs		33	53	ns
Q_{rr}	Reverse Recovery Charge			235	376	nC

A series of performance graphs are found in Figure 12.13c. In the upper left is a section of drain curves showing the ohmic region through V_{DS} = 5 V. The plot directly below this shows the increase in $r_{DS(on)}$ as temperature rises. There is about a three-fold variation across the temperature range. At lower left is the characteristic curve variation. Note that the curves are less steep as temperature increases, showing a decrease in g_m and thus, verifying a negative temperature coefficient of transconductance.

Figure 12.13c
FDMS86180 data sheet (cont).

Typical Characteristics T_J = 25 °C unless otherwise noted.

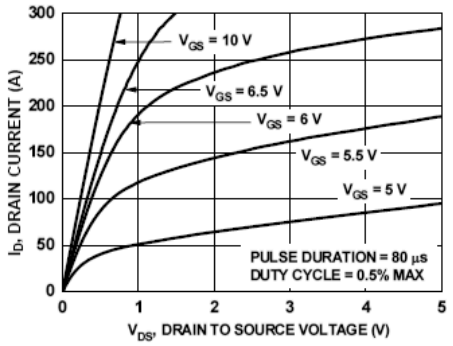


Figure 1. On-Region Characteristics

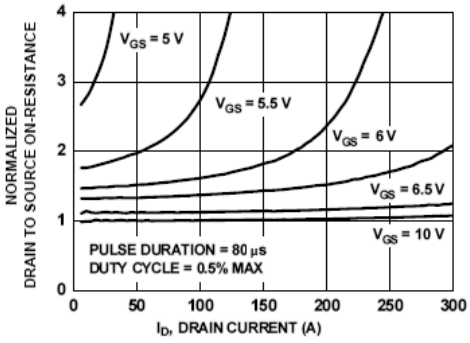


Figure2. Normalized On-Resistance vs. Drain Current and Gate Voltage

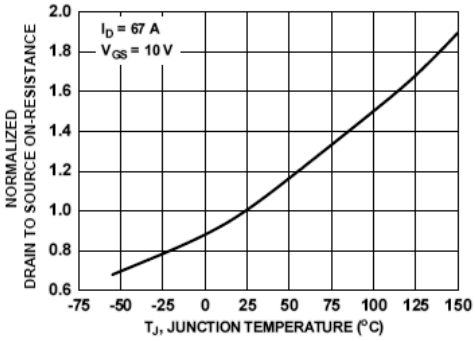


Figure3. Normalized On-Resistance vs. Junction Temperature

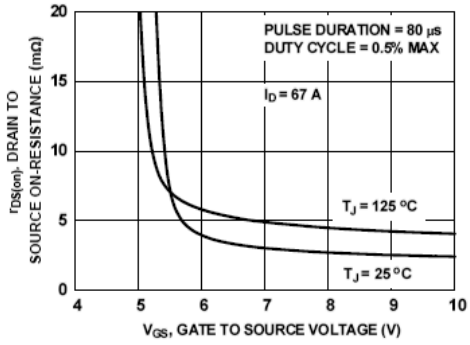


Figure4. On-Resistance vs. Gate to Source Voltage

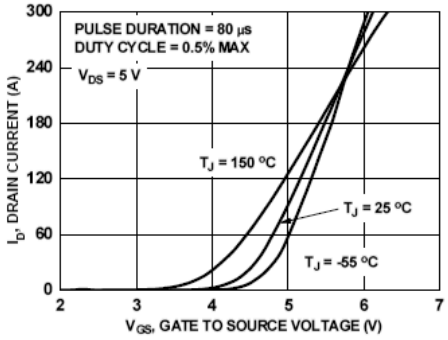


Figure 5. Transfer Characteristics

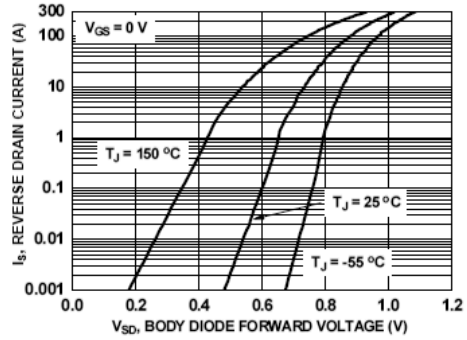


Figure6. Source to Drain Diode Forward Voltage vs. Source Current

12.6 E-MOSFET Biasing

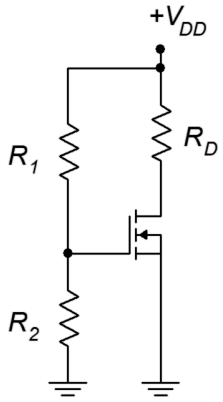
As the E-MOSFET operates only in the first quadrant, none of the biasing schemes used with JFETs will work with it. First, it should be noted that for large signal switching applications biasing is not much of an issue as we simply need to confirm that there is sufficient drive signal to turn the device on. For linear amplifiers we can use variations on constant voltage bias such as voltage divider bias, or on drain feedback bias.

Voltage Divider Bias

Voltage divider bias is reminiscent of the divider circuit used with BJTs. Indeed, the N-channel E-MOSFET requires that its gate be higher than its source, just as the NPN BJT requires a base voltage higher than its emitter. The major differences between the two are that the E-MOSFET's input gate current is negligible compared to base current and that the gate-source voltage will be most likely higher than the 0.7 volt drop seen across the base-emitter junction. Also, the gate-source voltage will not be locked to a specific voltage but will vary depending on the remainder of the circuit.

The prototype for the voltage divider bias is shown in Figure 12.14. In general, the layout it is the same as the voltage divider bias used with the DE-MOSFET. The resistors R_1 and R_2 set up the divider to establish the gate voltage. As the source terminal is tied directly to ground, this means that $V_{GS} = V_G$. The potential across R_2 needs to be set above $V_{GS(th)}$ for proper operation in accordance with Equation 12.5. Knowing the value of V_G, either the characteristic equation or the corresponding normalized drain current plot can be used to determine the drain current. The only factor missing is the device constant, k. This can be computed for any particular device based on the $I_{D(on)}$, $V_{GS(on)}$ coordinate pair specified in the data sheet (or measured in lab). An example is shown in Figure 12.15.

Figure 12.14
Voltage divider bias for E-MOSFET.



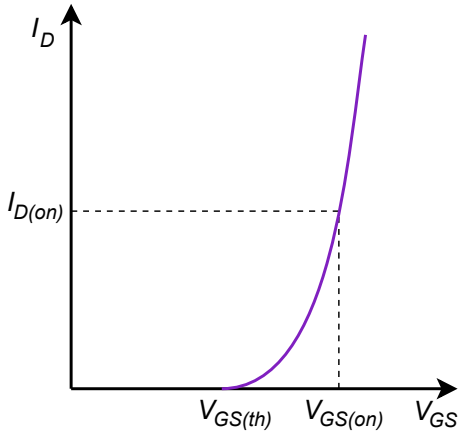


Figure 12.15
Coordinate pair on E-MOSFET curve.

The constant k is found via a rearrangement of Equation 12.5:

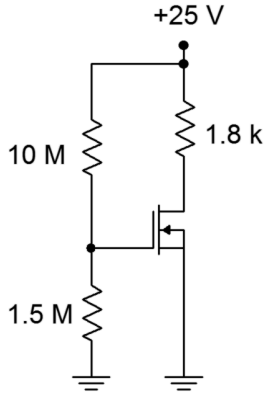
$$k = \frac{I_{D(on)}}{\left(V_{GS(on)} - V_{GS(th)}\right)^2}$$

This value can then be used for other biasing points.

Example 12.3

For the circuit and matching device curve of Figure 12.16, find I_D and V_{DS}.

Figure 12.16a
Circuit for Example 12.3.



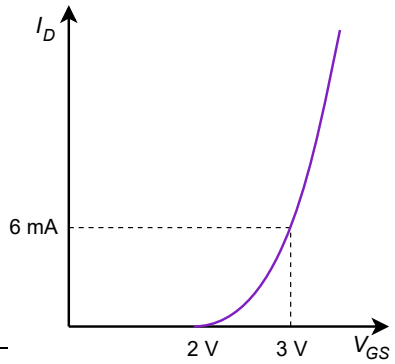
First find the value of k:

$$k = \frac{I_{D(on)}}{\left(V_{GS(on)} - V_{GS(th)}\right)^2}$$
$$k = \frac{6\,\text{mA}}{(3\,\text{V} - 2\,\text{V})^2}$$
$$k = 6\,\text{mA/V}^2$$

Now determine the gate voltage:

$$V_G = V_{DD}\frac{R_2}{R_1 + R_2}$$
$$V_G = 25\,\text{V}\frac{1.5\,\text{M}\Omega}{10\,\text{M}\Omega + 1.5\,\text{M}\Omega}$$
$$V_G = 3.26\,\text{V}$$

Figure 12.16b
Device curve for Example 12.3.



The source is grounded so $V_{GS} = V_G$.

$$I_D = k\left(V_{GS} - V_{GS(th)}\right)^2$$
$$I_D = 6\,\text{mA/V}^2(3.26\,\text{V} - 2\,\text{V})^2$$
$$I_D = 9.54\,\text{mA}$$

$$V_{DS} = V_{DD} - I_D R_D$$
$$V_{DS} = 25\,\text{V} - 9.54\,\text{mA} \times 1.8\,\text{k}\Omega$$
$$V_{DS} = 7.83\,\text{V}$$

In closing, note that it is possible to decouple the voltage divider using the same method employed with BJTs in Figure 7.13. Very large value resistors are available in only a limited variety of sizes so this technique has an added benefit. The divider resistors can use more convenient sizes because R_1 and R_2 will not set the input impedance; it will be set by the decoupling resistor.

Drain Feedback Bias

Drain feedback bias utilizes the aforementioned "on" operating point from the characteristic curve. The idea is to establish a drain current via an appropriate selection of the drain resistor and power supply. The prototype of the drain feedback circuit is shown in Figure 12.17.

This is relatively simple layout using few components. The key to understanding its operation is the KVL summation:

$$V_{DD} = V_{R_D} + V_{R_G} + V_{GS}$$
$$V_{DD} = I_D R_D + I_G R_G + V_{GS}$$

Gate current is negligible which means that

$$V_{DD} = I_D R_D + V_{GS}$$

and also

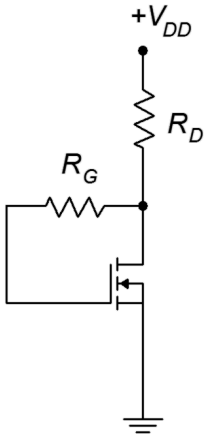
$$V_{DS} = V_{GS}$$

Therefore,

$$V_{GS} = V_{DS} = V_{DD} - I_D R_D \qquad (12.7)$$

Equation 12.7 can be used as the basis for the design of the bias circuit.

Figure 12.17
Drain feedback bias prototype.



Example 12.4

Utilizing the prototype of Figure 12.17, determine values for R_D and R_G such that the drain current is 8 mA. Assume V_{DD} = 20 V, $I_{D(on)}$ = 5 mA at $V_{GS(on)}$ = 4 V, and $V_{GS(th)}$ = 2.5 V.

First find the value of k:

$$k = \frac{I_{D(on)}}{\left(V_{GS(on)} - V_{GS(th)}\right)^2}$$
$$k = \frac{5\,\text{mA}}{\left(4\,\text{V} - 2.5\,\text{V}\right)^2}$$
$$k = 2.22\,\text{mA/V}^2$$

Now determine the required V_{GS} to obtain 8 mA of drain current by rearranging Equation 12.5.

$$I_D = k(V_{GS} - V_{GS(th)})^2$$
$$V_{GS} = V_{GS(th)} + \sqrt{\frac{I_D}{k}}$$
$$V_{GS} = 2.5\,\text{V} + \sqrt{\frac{8\,\text{mA}}{2.22\,\text{mA/V}^2}}$$
$$V_{GS} = 4.4\,\text{V}$$

And finally,

$$V_{GS} = V_{DS} = V_{DD} - I_D R_D$$
$$R_D = \frac{V_{DD} - V_{GS}}{I_D}$$
$$R_D = \frac{20\,\text{V} - 4.4\,\text{V}}{8\,\text{mA}}$$
$$R_D = 1.95\,\text{k}\Omega$$

Summary

There are two types of MOSFETs: the depletion-enhancement or DE-MOSFET and the enhancement-only or E-MOSFET. Both devices are constructed using an insulated gate instead of a PN junction and both devices exhibit a square-law characteristic curve. Like the JFET, MOSFETs are modeled as voltage-controlled current sources. Both devices show very, very small gate currents due to the insulated gate. They are static sensitive and precautions must be taken when handling them to avoid damage from ESD.

The DE-MOSFET exhibits the same characteristic curve as the JFET, however, the curve extends into the first quadrant (enhancement mode). Consequently, I_{DSS} is no longer the largest drain current possible, but rather, represents a middle ground. The DE-MOSFET can utilize all of the bias prototypes that are used with JFETs, including self bias, constant current bias and combination bias. Due to its dual quadrant capability, other biasing types are also possible including zero bias and voltage divider, both of which are variations on constant voltage bias.

The E-MOSFET operates in the first quadrant only (enhancement mode). Compared to the DE-MOSFET, its characteristic curve is shifted positive such that $V_{GS(off)}$ is now $V_{GS(th)}$, and I_{DSS} signifies the off-state leakage current. E-MOSFETs are available in both low power and high power variants. The high power versions utilize an alternate internal structure that allows drain current to flow vertically rather than horizontally. This results in very high current carrying ability and very low values for $r_{DS(on)}$.

Review Questions

1. What are the differences between JFETs and MOSFETs?
2. What are the differences between DE-MOSFETs and E-MOSFETs?
3. Why are MOSFETs sometimes referred to as "Insulated Gate" or IGFETs?
4. How does the DC bias model of the MOSFET compare to that of the BJT?
5. What biasing circuits are available for use with the DE-MOSFET?
6. What biasing circuits are available for use with the E-MOSFET?
7. What is "trench" construction and where is it used?
8. Explain typical precautions taken when handling MOSFETs and why they are necessary.

Problems

Analysis Problems

1. For the circuit of Figure 12.18, determine I_D, V_G and V_D. $I_{DSS} = 20$ mA, $V_{GS(off)} = -6$ V, $V_{DD} = 15$ V, $R_G = 470$ kΩ, $R_S = 1.2$ kΩ, $R_D = 1.8$ kΩ.

2. For the circuit of Figure 12.18, determine I_D, V_{DS} and V_D. $I_{DSS} = 20$ mA, $V_{GS(off)} = -5$ V, $V_{DD} = 30$ V, $R_G = 560$ kΩ, $R_S = 420$ Ω, $R_D = 1.5$ kΩ.

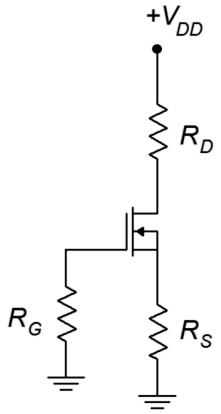


Figure 12.18

3. For Figure 12.19, determine I_D, V_G and V_D. $I_{DSS} = 15$ mA, $V_{DD} = 25$ V, $V_{GS(off)} = -3$ V, $V_{SS} = -6$ V, $R_G = 820$ kΩ, $R_S = 2$ kΩ, $R_D = 3.6$ kΩ.

4. For the circuit of Figure 12.19, determine I_D, V_{DS} and V_D. $I_{DSS} = 18$ mA, $V_{GS(off)} = -3$ V, $V_{DD} = 30$ V, $V_{SS} = -9$ V, $R_G = 910$ kΩ, $R_S = 1.2$ kΩ, $R_D = 2.7$ kΩ.

5. For the circuit of Figure 12.20, determine I_D, V_G and V_D. $I_{DSS} = 12$ mA, $V_{GS(off)} = -4$ V, $V_{DD} = 35$ V, $R_G = 680$ kΩ, $R_D = 1.8$ kΩ.

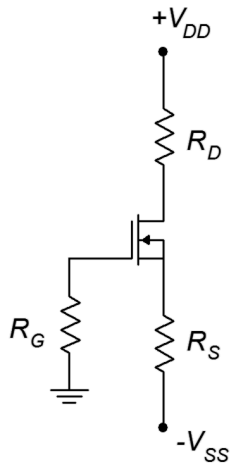


Figure 12.19

6. For the circuit of Figure 12.20, determine I_D, V_{DS} and V_D. I_{DSS} = 8 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 30 V, R_G = 750 kΩ, R_D = 2.7 kΩ.

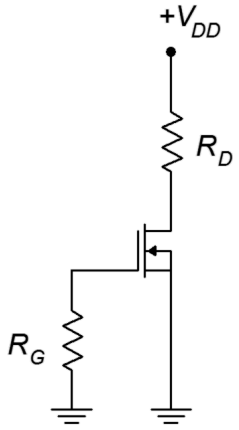


Figure 12.20

7. For the circuit of Figure 12.21, determine I_D, V_G and V_D. I_{DSS} = 8 mA, $V_{GS(off)}$ = −4 V, V_{DD} = 30 V, R_1 = 2.7 MΩ, R_2 = 110 kΩ, R_D = 470 Ω.

8. For the circuit of Figure 12.21, determine I_D, V_{DS} and V_D. I_{DSS} = 12 mA, $V_{GS(off)}$ = −6 V, V_{DD} = 20 V, R_1 = 2 MΩ, R_2 = 100 kΩ, R_D = 680 Ω.

9. For the circuit of Figure 12.22, determine I_D, V_G and V_D. $I_{D(on)}$ = 8 mA, $V_{GS(on)}$ = 5 V, $V_{GS(th)}$ = 3 V, V_{DD} = 30 V, R_1 = 2 MΩ, R_2 = 330 kΩ, R_D = 1.2 kΩ.

10. For the circuit of Figure 12.22, determine I_D, V_{DS} and V_D. $I_{D(on)}$ = 12 mA, $V_{GS(on)}$ = 6 V, $V_{GS(th)}$ = 2.5 V, V_{DD} = 25 V, R_1 = 1.5 MΩ, R_2 = 470 kΩ, R_D = 680 Ω.

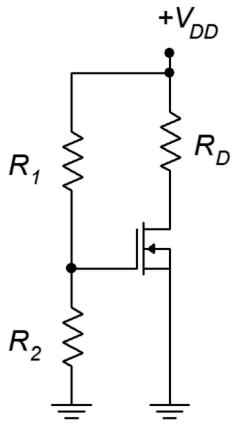


Figure 12.21

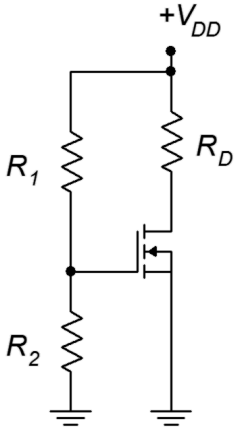


Figure 12.22

11. For the circuit of Figure 12.23, determine I_D, V_G and V_D. I_{DSS} = 12 mA, $V_{GS(off)}$ = 2 V, V_{DD} = −25 V, R_G = 470 kΩ, R_S = 800 Ω, R_D = 1.8 kΩ.

12. For the circuit of Figure 12.23, determine I_D and V_D. I_{DSS} = 10 mA, $V_{GS(off)}$ = 2 V, V_{DD} = −20 V, R_G = 560 kΩ, R_S = 680 Ω, R_D = 1.5 kΩ.

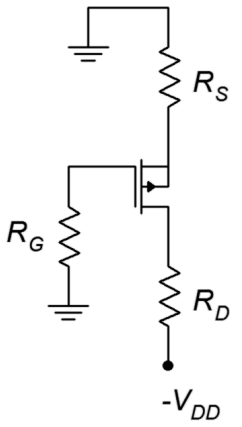


Figure 12.23

13. For the circuit of Figure 12.24, determine I_D, V_G and V_D. I_{DSS} = 14 mA, $V_{GS(off)}$ = 3 V, V_{DD} = −25 V, V_{SS} = 6 V, R_G = 780 kΩ, R_S = 2 kΩ, R_D = 3.3 kΩ.

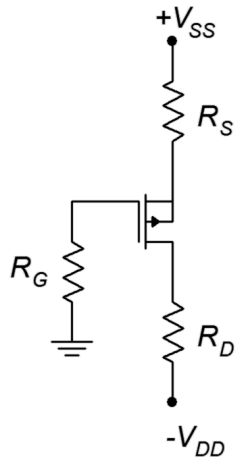


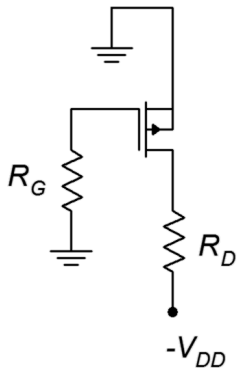
Figure 12.24

14. For the circuit of Figure 12.24, determine I_D and V_D. I_{DSS} = 16 mA, $V_{GS(off)}$ = 3.5 V, V_{DD} = −20 V, V_{SS} = 7 V, R_G = 1 MΩ, R_S = 1.5 kΩ, R_D = 2.2 kΩ.

15. For the circuit of Figure 12.25, determine I_D and V_D. I_{DSS} = 11 mA, $V_{GS(off)}$ = 2 V, V_{DD} = −24 V, R_G = 750 kΩ, R_D = 1.2 kΩ.

16. For the circuit of Figure 12.25, determine I_D and V_D. I_{DSS} = 9 mA, $V_{GS(off)}$ = 3 V, V_{DD} = −18 V, R_G = 430 kΩ, R_D = 910 Ω.

Figure 12.25



Design Problems

17. Using the circuit of Figure 12.18, determine a value for R_S to set I_D to 4 mA. I_{DSS} = 10 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 18 V, R_G = 470 kΩ, R_D = 1.5 kΩ.

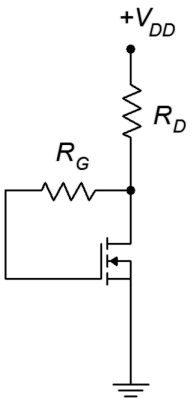


Figure 12.26

18. For the circuit of Figure 12.26, determine R_D and R_G to set I_D = 10 mA. $I_{D(on)}$ = 15 mA, $V_{GS(on)}$ = 6 V, $V_{GS(th)}$ = 2 V, V_{DD} = 20 V.

19. For the circuit of Figure 12.26, determine R_D and R_G to set I_D = 15 mA. $I_{D(on)}$ = 10 mA, $V_{GS(on)}$ = 5 V, $V_{GS(th)}$ = 2 V, V_{DD} = 25 V.

Challenge Problems

20. Using the circuit of Figure 12.19, determine values for R_D, R_S and V_{SS} to set I_D to 5 mA and V_D to 20 V. I_{DSS} = 15 mA, $V_{GS(off)}$ = −3 V, V_{DD} = 30 V, R_G = 560 kΩ.

21. Using the circuit of Figure 12.27, determine values for R_D to set V_D to 15 V. I_{DSS} = 10 mA, $V_{GS(off)}$ = 3 V, V_{SS} = 25 V, R_G = 680 kΩ.

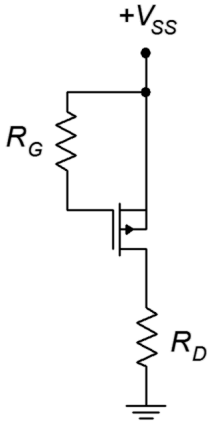
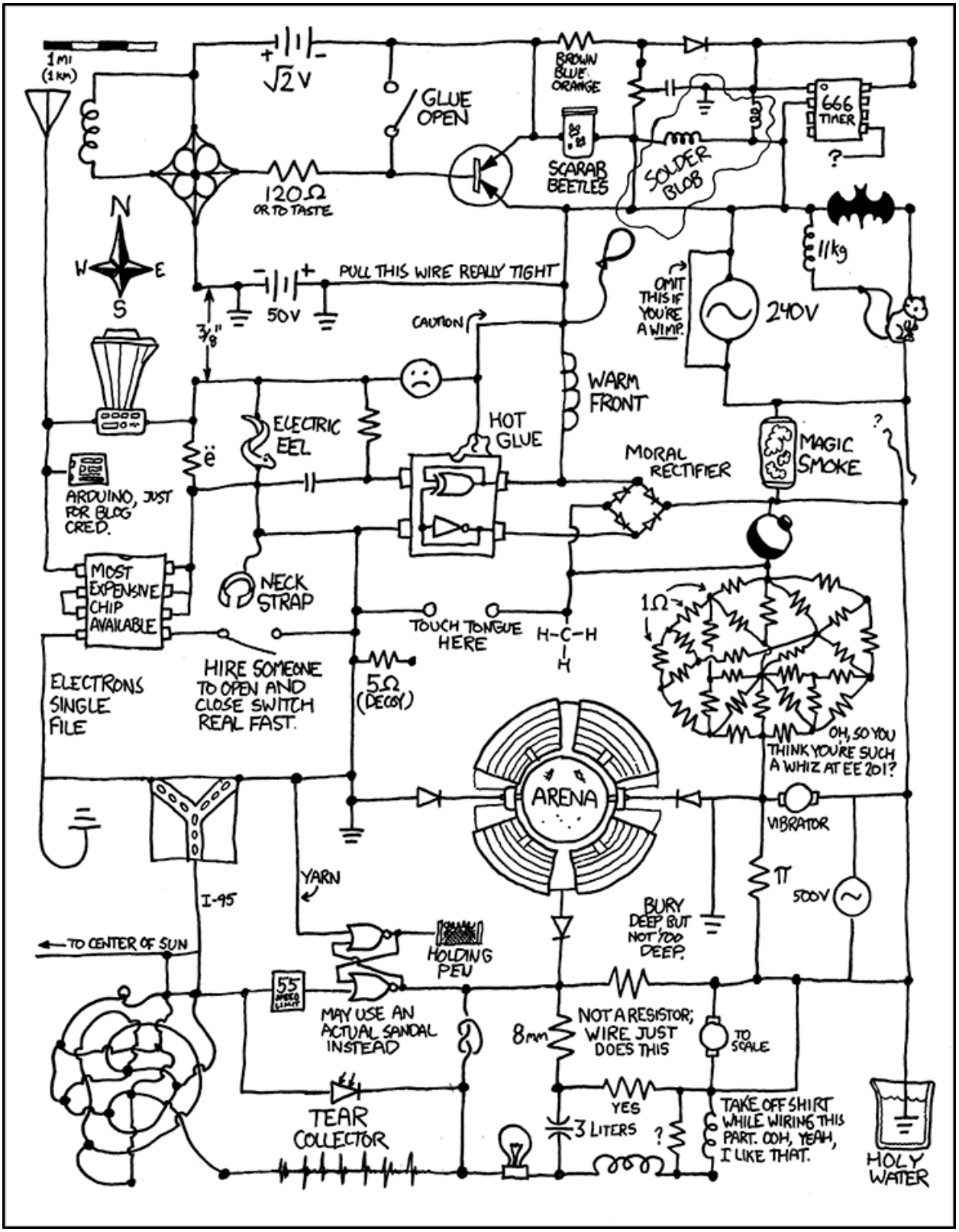


Figure 12.27



Comic courtesy of xkcd.com

13 MOSFET Small Signal Amplifiers

13.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Draw and explain a basic AC model for a MOSFET.
- Analyze basic MOSFET amplifiers for voltage gain, input impedance and output impedance.
- Analyze basic MOSFET voltage followers for voltage gain, input impedance and output impedance.

13.1 Introduction

MOSFETs can be used to create both common source voltage amplifiers and common drain voltage followers (i.e., source followers). Both circuits offer the potential for very high input impedance due to the extremely low gate current MOSFETs provide. As with JFET amplifiers, at higher frequencies input capacitance dominates and reduces the input impedance. Not all bias prototypes lend themselves to all possible AC circuits. For example, zero bias for a DE-MOSFET is not suitable for followers or swamped amplifiers as it lacks a source resistor. The same is true for voltage divider biasing used with both DE- and E-MOSFETs. These biasing schemes are suitable for non-swamped amplifiers, though.

In general, MOSFET amplifiers tend to have good high frequency performance, offer low noise and exhibit low distortion with modestly sized input signals. Compared to BJTs, their voltage gain magnitude is lower.

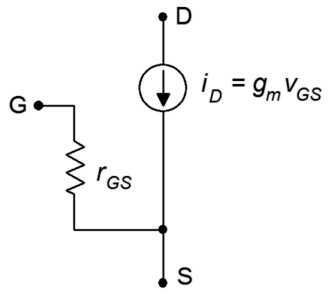
A key parameter in determining gain is the device's transconductance, g_m. Transconductance varies widely depending on the kind of MOSFET used. A small signal DE-MOSFET may exhibit a transconductance of just a few millisiemens. In contrast, a high power E-MOSFET may exhibit a transconductance of over 100 siemens.

13.2 MOSFET Common Source Amplifiers

Before we can examine the common source amplifier, an AC model is needed for both the DE- and E-MOSFET. A simplified model consists of a voltage-controlled current source and an input resistance, r_{GS}. This model is shown in Figure 13.1. The model is essentially the same as that used for the JFET. Technically, the gate-source resistance is higher in the MOSFET due to the insulated gate, and this is useful in specific applications such as in the design of electrometers, but for general purpose work it is a minor distinction. The impedance associated with the current source is not shown as it is typically large enough to ignore. Similarly, the device capacitances are not shown. It is worth noting that the capacitances associated with small signal devices might be just a few picofarads, however, a power device might exhibit values of a few nanofarads.

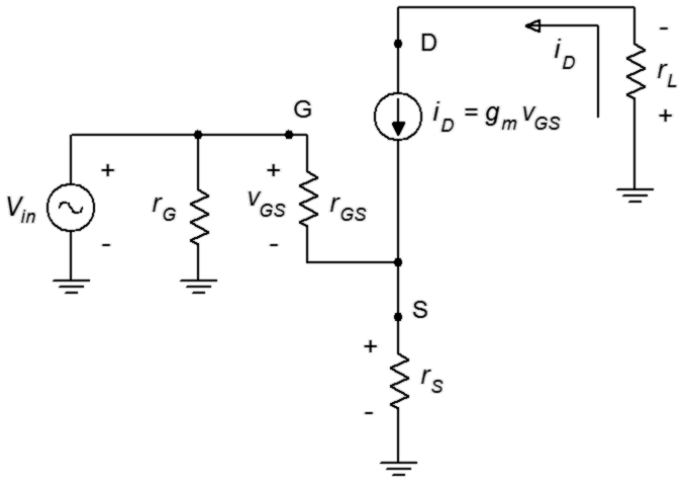
As the device model is the same for both DE- and E-MOSFETs, the analysis of voltage gain, input impedance and output impedance will apply to both devices. The only practical differences will be how the transconductance is determined, and circuit variations due to the differing biasing requirements which will effect the input impedance. In fact, there will be a great uniformity between JFET-based circuits and DE-MOSFET circuits operating in depletion mode.

Figure 13.1
AC device model for MOSFETs.



An AC equivalent of a swamped common source amplifier is shown in Figure 13.2. This is a generic prototype and is suitable for any variation on device and bias type. Ultimately, all of the amplifiers can be reduced down to this equivalent, occasionally with some resistance values left out (either opened or shorted). For example, if the amplifier is not swamped then $r_S = 0$. Similarly, r_G might correspond to a single gate biasing resistor or it might represent the equivalent of a pair of resistors that set up a gate voltage divider.

Figure 13.2
Generic common source amplifier equivalent.



Voltage Gain

In order to derive an equation for the voltage gain, we start with its definition, namely that voltage gain is the ratio of v_{out} to v_{in}. We then proceed by expressing these voltages in terms of their Ohm's law equivalents. Note that r_L can also be called r_D.

$$A_v = \frac{v_{out}}{v_{in}} = \frac{v_L}{v_G} = \frac{v_D}{v_G}$$

$$A_v = \frac{-i_D r_L}{i_D r_S + v_{GS}}$$

$$A_v = \frac{-g_m v_{GS} r_L}{g_m v_{GS} r_S + v_{GS}}$$

$$A_v = -\frac{g_m r_L}{g_m r_S + 1} \qquad (13.1)$$

or, if preferred

$$A_v = -\frac{g_m r_D}{g_m r_S + 1} \qquad (13.1b)$$

This is the general equation for voltage gain. If the amplifier is not swamped then the first portion of the denominator drops out and the gain simplifies to

$$A_v = -g_m r_L \qquad (13.2)$$

or alternately

$$A_v = -g_m r_D \qquad (13.2b)$$

The swamping resistor, r_S, plays the same role here as it did with both the BJT and JFET. Swamping helps to stabilize the gain and reduce distortion, but at the expense of voltage gain.

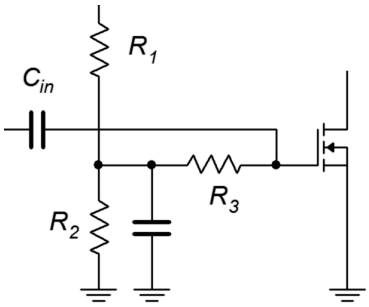
Input Impedance

Referring back to Figure 13.2, the input impedance of the amplifier will be r_G in parallel with the impedance looking into the gate terminal, $Z_{in(gate)}$. At a minimum this will be r_{GS} (it is somewhat higher when swamped but this can be ignored in most cases). At low frequencies r_{GS} is very large, perhaps as high as 10^{12} ohms. In most practical circuits, r_G will be much lower, hence

$$Z_{in} = r_G \parallel r_{GS} \approx r_G \qquad (13.3)$$

It is important to reiterate that r_G is the equivalent resistance seen prior to the gate terminal that is seen from the vantage point of V_{in}. In the case of self bias, combination bias, zero bias and constant current bias, this will be the single biasing resistor R_G. For simple voltage divider biasing, r_G will be the parallel combination of the two divider resistors (i.e., $R_1 \parallel R_2$). For decoupled voltage divider biasing, as shown in Figure 13.3, r_G will be the decoupling resistor (i.e., R_3) that is connected between the divider and the gate. This is because the divider node is bypassed to ground via a capacitor. Finally, for drain feedback biasing, r_G is the Millerized R_G that bridges the drain and gate.

Figure 13.3
Decoupled voltage divider.



Output Impedance

The derivation of output impedance is unchanged from the JFET case. From the perspective of the load, the output impedance will be the drain biasing resistor, R_D, in parallel with the internal impedance of the current source within the device model. R_D tends to be much lower than this, and thus, the output impedance can be approximated as R_D.

Therefore we may state

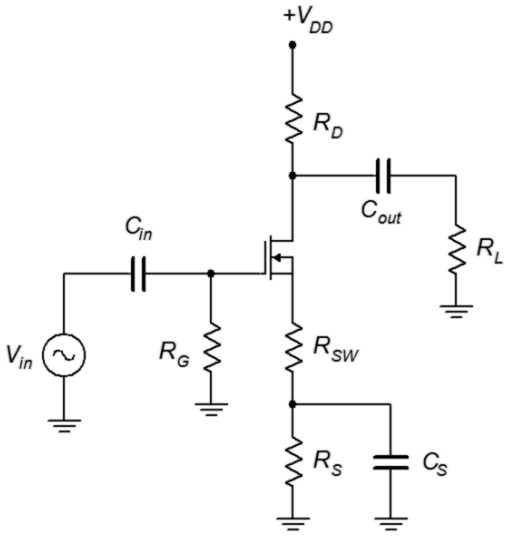
$$Z_{out} = r_{model} \,||\, R_D \approx R_D \tag{13.4}$$

At this point, a variety of examples are in order to illustrate some of the myriad combinations.

Example 13.1

For the amplifier in Figure 13.4, determine the input impedance and load voltage. $V_{in} = 20$ mV, $V_{DD} = 20$ V, $R_G = 1$ MΩ, $R_D = 1.8$ kΩ, $R_{SW} = 20$ Ω, $R_S = 400$ Ω, $R_L = 12$ kΩ, $I_{DSS} = 40$ mA, $V_{GS(off)} = -1$ V.

Figure 13.4
Circuit for Example 13.1.



This is a swamped common drain amplifier utilizing self bias. Z_{in} can be determined via inspection.

$$Z_{in} = Z_{in(gate)} \,||\, R_G$$
$$Z_{in} \approx 1\,\text{M}\Omega$$

To find the load voltage we'll need the voltage gain, and to find the gain we'll first need to find g_{m0}.

$$g_{m0} = -\frac{2\,I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{80\,\text{mA}}{-1\,\text{V}}$$
$$g_{m0} = 80\,\text{mS}$$

The combined DC value of R_S is 420 Ω, therefore $g_{m0}R_S = 33.6$. From the self bias equation or graph this produces a drain current of 1.867 mA.

$$g_m = g_{m0}\sqrt{\frac{I_D}{I_{DSS}}}$$
$$g_m = 80\,\text{mS}\sqrt{\frac{1.867\,\text{mA}}{40\,\text{mA}}}$$
$$g_m = 17.3\,\text{mS}$$

The swamping resistor, r_S, is 20 Ω. The voltage gain is

$$A_v = -\frac{g_m r_L}{g_m r_S + 1}$$
$$A_v = -\frac{17.3\,\text{mS}(1.8\,\text{k}\Omega \,||\, 12\,\text{k}\Omega)}{17.3\,\text{mS} \times 20\,\Omega + 1}$$
$$A_v = -20.1$$

And finally

$$V_{load} = A_v V_{in}$$
$$V_{load} = -20.1 \times 20\,\text{mV}$$
$$V_{load} = 402\,\text{mV}$$

Computer Simulation

The amplifier of Example 13.1 is simulated to verify the results. The circuit is entered into the simulator as shown in Figure 13.5. One issue is finding an appropriate DE-MOS device to match the parameters used in the example. The BSS229 proves to be reasonably close. This device model was tested for I_{DSS} by applying a 20 volt source to the drain and shorting the source and gate terminals to ground in the simulator. The current was just under the 40 mA target. Similarly, a negative voltage was attached to the gate and adjusted until the drain current dropped to nearly zero in order to determine $V_{GS(off)}$. The model's value was just under the desired −1 volt. Consequently, we can expect the simulation results to be close to those predicted, although not identical.

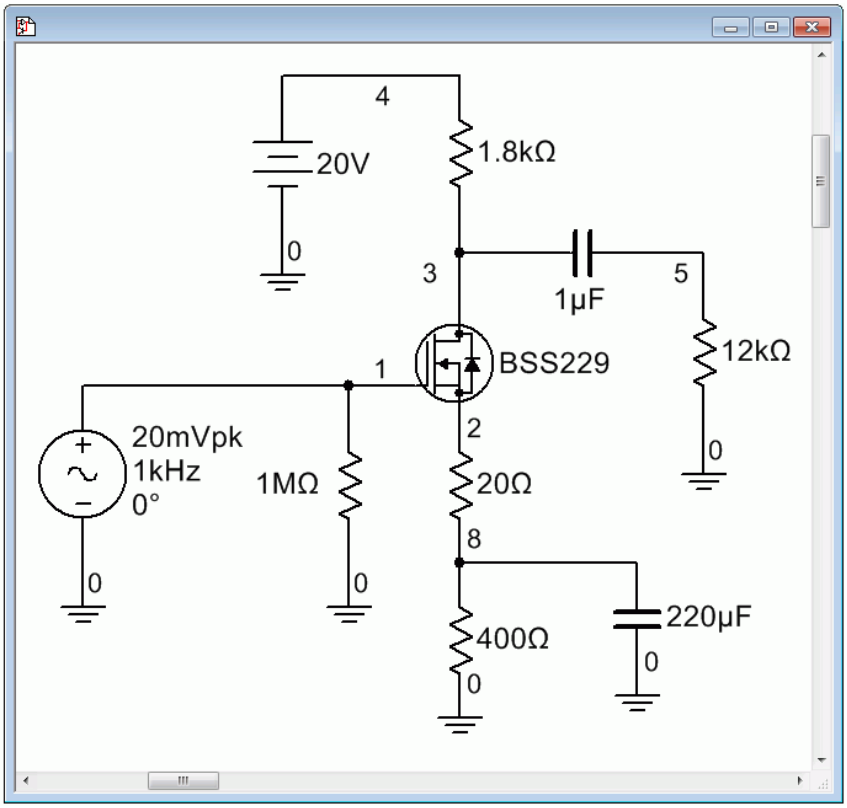


Figure 13.5
The circuit of Example 13.1 in the simulator.

The transient analysis is run next and is shown in Figure 13.6. The expected signal inversion is obvious. The peak amplitude is 417 mV, just a few percent higher than the calculated value. At least some of this deviation is due to the model's variation from the assumed device parameter values.

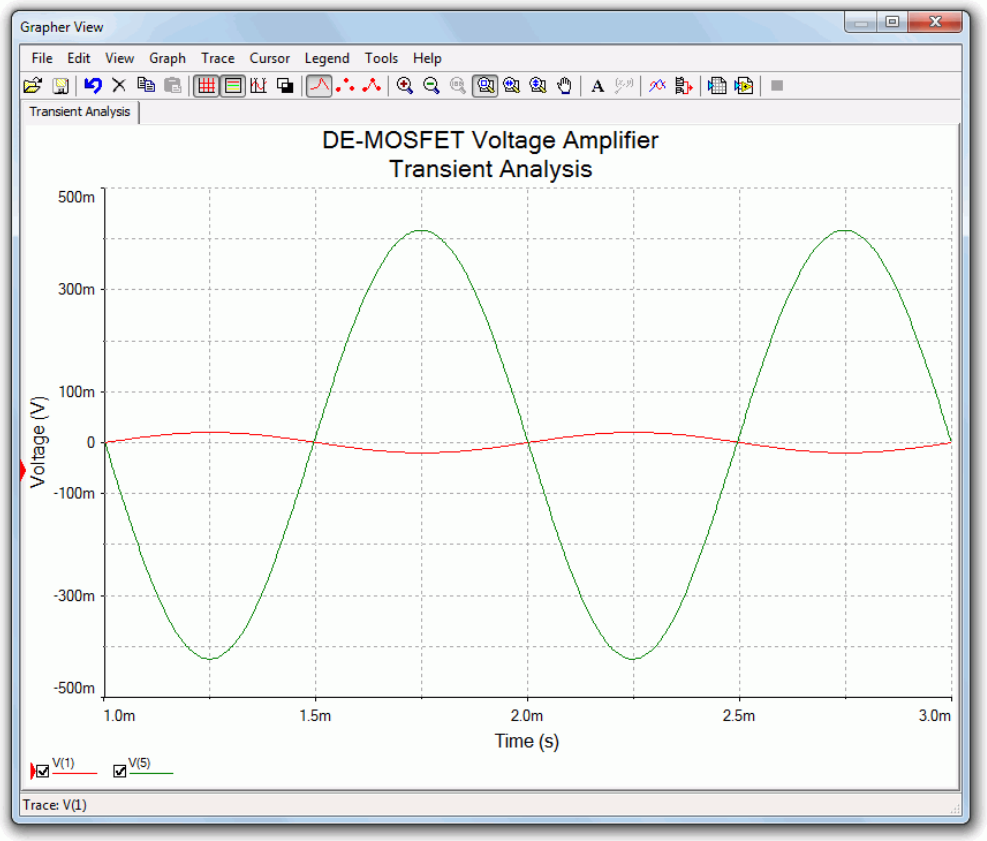


Figure 13.6
Transient analysis simulation for the circuit of Example 13.1.

A DC bias check is also performed. The drain current was calculated to be 1.867 mA. This yields an R_D voltage of a little over 3 volts, thus we expect to see a drain voltage of about 17 volts. Similarly, we would expect the source terminal to be sitting at around 700 to 800 mV and the gate at about 0 V.

The results of the DC operating point simulation are shown in Figure 13.7. The agreement with the predicted values is quite good, especially considering that the device model is not a perfect match.

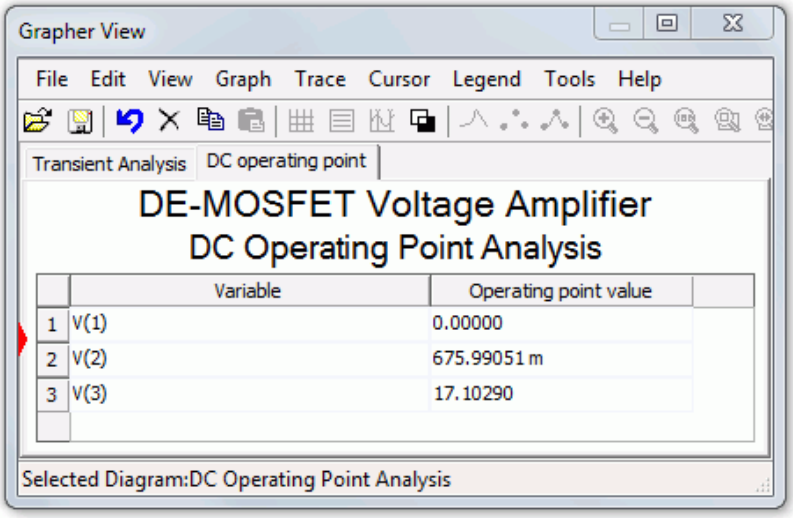


Figure 13.7
DC bias simulation for the circuit of Example 13.1.

Example 13.2

For the circuit of Figure 13.8, determine the voltage gain and input impedance. Assume $V_{GS(th)} = 2$ V, $I_{D(on)} = 50$ mA at $V_{GS(on)} = 5$ V.

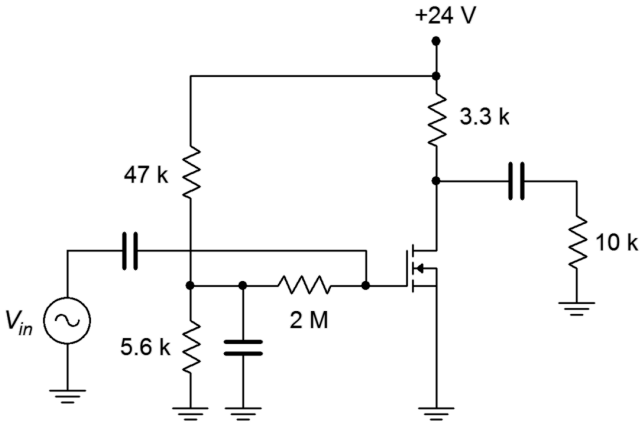
First find the value of k:

$$k = \frac{I_{D(on)}}{\left(V_{GS(on)} - V_{GS(th)}\right)^2}$$

$$k = \frac{50\,\text{mA}}{(5\,\text{V} - 2\,\text{V})^2}$$

$$k = 5.56\,\text{mA/V}^2$$

Figure 13.8
Circuit for Example 13.2.



This circuit uses power supply decoupling. The voltage drop across the 2 MΩ resistor is small enough to ignore as the current passing through it is gate current. Therefore the gate voltage is determined by the divider. Also, as the left end of the 2 MΩ resistor is tied to an AC ground due to the bypass capacitor, it represents the input impedance.

$$Z_{in} = 2\,\text{M}\Omega \,||\, Z_{in(gate)} \approx 2\,\text{M}\Omega$$

$$V_G = V_{DD}\frac{R_2}{R_1 + R_2}$$

$$V_G = 24\,\text{V}\frac{5.6\,\text{k}\Omega}{47\,\text{k}\Omega + 5.6\,\text{k}\Omega}$$

$$V_G = 2.56\,\text{V}$$

The source is grounded so $V_{GS} = V_G$.

$$I_D = k\left(V_{GS} - V_{GS(th)}\right)^2$$
$$I_D = 5.56\,\text{mA/V}^2\left(2.56\,\text{V} - 2\,\text{V}\right)^2$$
$$I_D = 1.74\,\text{mA}$$

$$g_m = 2\,k\left(V_{GS} - V_{GS(th)}\right)$$
$$g_m = 2 \times 5.56\,\text{mA/V}^2\left(2.56\,\text{V} - 2\,\text{V}\right)$$
$$g_m = 6.23\,\text{mS}$$

This amplifier is not swamped so the simplified gain equation may be used.

$$A_v = -g_m r_D$$
$$A_v = -6.23\,\text{mS}\left(3.3\,\text{k}\Omega \parallel 10\,\text{k}\Omega\right)$$
$$A_v = -15.5$$

Example 13.3

For the circuit of Figure 13.9, determine the voltage gain and input impedance. Assume $V_{GS(off)} = -0.75$ V and $I_{DSS} = 6$ mA.

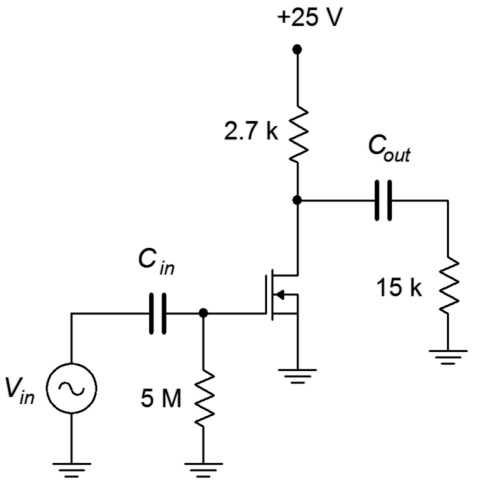


Figure 13.9
Circuit for Example 13.3.

This amplifier uses zero bias, therefore $I_D = I_{DSS}$ and $g_m = g_{m0}$.

$$g_{m0} = -\frac{2\,I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{2 \times 6\,\text{mA}}{-0.75\,\text{V}}$$
$$g_{m0} = 16\,\text{mS}$$

This amplifier is not swamped so we may use the simplified equation for voltage gain.

$$A_v = -g_m r_D$$
$$A_v = -16\,\text{mS}\left(2.7\,\text{k}\Omega \parallel 15\,\text{k}\Omega\right)$$
$$A_v = -36.6$$

Finally, for the input impedance we have

$$Z_{in} = 5\,\text{M}\Omega \parallel Z_{in(gate)} \approx 5\,\text{M}\Omega$$

13.3 MOSFET Common Drain Followers

As discussed under the section on JFETs, the *common drain amplifier* is also known as the *source follower*. The prototype amplifier circuit with device model is shown in Figure 13.10. As with all voltage followers, we expect a non-inverting voltage gain close to unity with a high Z_{in} and a low Z_{out}.

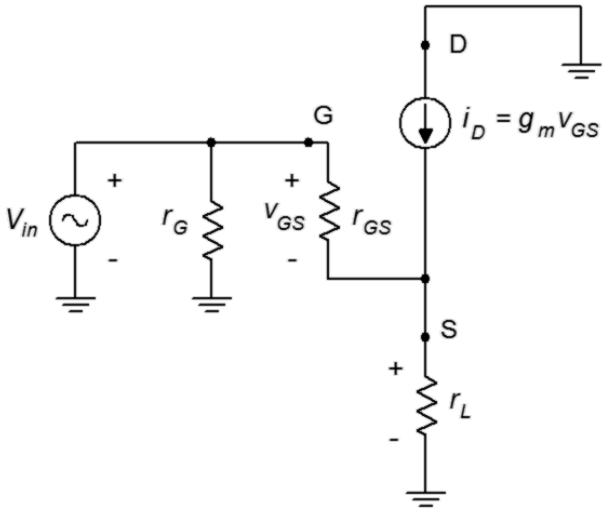


Figure 13.10
Common drain (source follower) prototype.

As is usual, the input signal is applied to the gate terminal and the output is taken from the source. Because the output is at the source, biasing schemes that have the source terminal grounded, such as zero bias and voltage divider bias, cannot be used.

Voltage Gain

The voltage gain equation for the common drain follower is developed as follows: We begin with the fundamental definition that voltage gain is the ratio of v_{out} to v_{in}, and proceed by expressing these voltages in terms of their Ohm's law equivalents. The load is now located at the MOSFET's source, and thus can be referred to as either r_L or r_S.

$$A_v = \frac{v_{out}}{v_{in}} = \frac{v_S}{v_G} = \frac{v_L}{v_G}$$

$$A_v = \frac{i_D r_L}{i_D r_L + v_{GS}}$$

$$A_v = \frac{g_m v_{GS} r_L}{g_m v_{GS} r_L + v_{GS}}$$

$$A_v = \frac{g_m r_L}{g_m r_L + 1} \qquad (13.5)$$

or, if preferred

$$A_v = \frac{g_m r_S}{g_m r_S + 1} \tag{13.5b}$$

If $g_m r_S >> 1$, the voltage gain will be very close to unity; a desired outcome.

Input Impedance

The analysis for source follower's input impedance is virtually identical to that for the common source amplifier. The same commentary applies regarding the simplification of gate biasing resistors to arrive at the value of r_G.

$$Z_{in} = r_G \parallel r_{GS} \approx r_G \tag{13.6}$$

Output Impedance

In order to determine the output impedance, we modify the circuit of Figure 13.10 by separating the load resistance from the source bias resistor. This is shown in Figure 13.11.

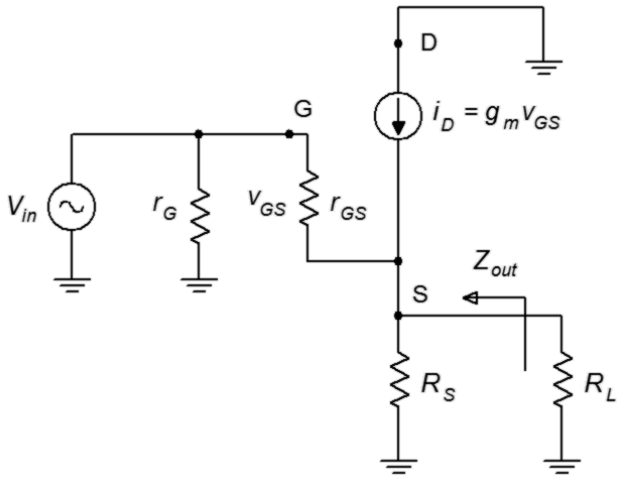


Figure 13.11
Source follower output impedance analysis.

Looking back into the source from the perspective of the load we find that the source biasing resistor, R_S, is in parallel with the impedance looking back into the source terminal.

$$Z_{out} = R_S \parallel Z_{source}$$

To find Z_{source}, note that the voltage at the source is v_{GS} and the current entering this node is i_D. The ratio of the two will yield the impedance looking back into the source.

$$Z_{source} = \frac{v_{GS}}{i_D}$$
$$Z_{source} = \frac{v_{GS}}{g_m v_{GS}}$$
$$Z_{source} = \frac{1}{g_m} \tag{13.7}$$

Therefore, the output impedance is

$$Z_{out} = R_S \parallel \frac{1}{g_m} \tag{13.8}$$

Looking at Equation 13.8 it is obvious that the higher the transconductance, the lower the output impedance. As noted earlier, a large transconductance also means that the voltage gain will be close to unity. As a general rule then, a large transconductance is desired for the source follower.

Time for a few illustrative examples.

Example 13.4

For the circuit of Figure 13.12, determine the voltage gain and input impedance. Assume $V_{GS(off)} = -0.8$ V and $I_{DSS} = 30$ mA.

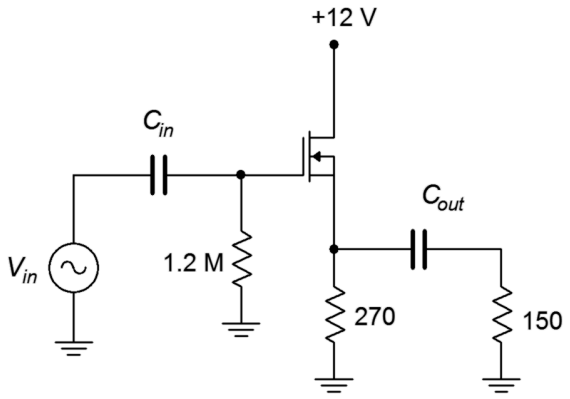


Figure 13.12
Circuit for Example 13.4.

This amplifier uses self bias so we need to determine $g_{m0}R_S$.

$$g_{m0} = -\frac{2I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{2 \times 30\,\text{mA}}{-0.8\,\text{V}}$$
$$g_{m0} = 75\,\text{mS}$$

The DC source resistance is the 270 Ω biasing resistor resulting in $g_{m0}\,R_S = 16.2$. From the self bias equation or graph this produces a drain current of 2.61 mA.

$$g_m = g_{m0}\sqrt{\frac{I_D}{I_{DSS}}}$$
$$g_m = 75\,\text{mS}\sqrt{\frac{2.61\,\text{mA}}{30\,\text{mA}}}$$
$$g_m = 22.1\,\text{mS}$$

The voltage gain is

$$A_v = \frac{g_m r_S}{g_m r_S + 1}$$
$$A_v = \frac{22.1\,\text{mS}(270\,\Omega \parallel 150\,\Omega)}{22.1\,\text{mS}\times(270\,\Omega \parallel 150\,\Omega) + 1}$$
$$A_v = 0.68$$

Finally, for the input impedance we have

$$Z_{in} = 1.2\ \text{M}\Omega \parallel Z_{in(gate)} \approx 1.2\ \text{M}\Omega$$

Example 13.5

For the circuit of Figure 13.13, determine the voltage gain and input impedance. Assume $V_{GS(off)} = -2.5$ V and $I_{DSS} = 80$ mA.

This follower uses a P-channel device with combination bias. Note that the source terminal is toward the top of the schematic. First, determine $g_{m0}R_S$ and the bias factor, k. Then the combination bias equation can be used to determine the drain current.

$$g_{m0} = -\frac{2I_{DSS}}{V_{GS(off)}}$$
$$g_{m0} = -\frac{2\times 80\,\text{mA}}{-2.5\,\text{V}}$$
$$g_{m0} = 64\,\text{mS}$$

The DC source resistance is the 1.8 kΩ biasing resistor resulting in $g_{m0}\,R_S = 115.2$. The bias factor is $V_{SS}/V_{GS(off)}$, or 4. The combination bias equation (Equation 10.9) yields $I_D = 6.67$ mA.

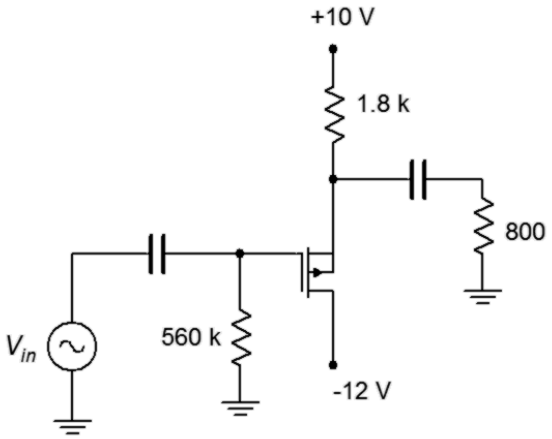


Figure 13.13
Circuit for Example 13.5.

We can now find the transconductance and voltage gain.

$$g_m = g_{m0}\sqrt{\frac{I_D}{I_{DSS}}}$$
$$g_m = 64\,\text{mS}\sqrt{\frac{6.67\,\text{mA}}{80\,\text{mA}}}$$
$$g_m = 18.5\,\text{mS}$$

The voltage gain is

$$A_v = \frac{g_m r_S}{g_m r_S + 1}$$
$$A_v = \frac{18.5\,\text{mS}\left(1.8\,\text{k}\Omega \parallel 800\,\Omega\right)}{18.5\,\text{mS}\times\left(1.8\,\text{k}\Omega \parallel 800\,\Omega\right) + 1}$$
$$A_v = 0.91$$

Lastly, the input impedance is

$$Z_{in} = 560\ \text{k}\Omega \parallel Z_{in(gate)} \approx 560\ \text{k}\Omega$$

13.4 MOSFET Common Gate Amplifiers

Like the common source and common drain amplifiers, the analysis of MOSFET common gate amplifiers follows that of the JFET counterpart. The main point to note is that some biasing forms are not allowed as they ground the source terminal, which would short out the input signal, rendering the amplifier useless. These biasing forms include zero bias and voltage divider bias for the DE-MOSFET, and both the voltage divider and drain feedback bias forms for the E-MOSFET. For the most part, this leaves self and combination biasing with DE-MOSFETs. The relevant AC equations are reproduced below, for convenience.

$$A_v = g_m r_L \tag{13.9}$$
$$Z_{in} = R_S \parallel \frac{1}{g_m} \tag{13.10}$$
$$Z_{out} \approx R_D \tag{13.11}$$

Summary

DE- and E-MOSFET devices may be used to create both common source voltage amplifiers and common drain voltage followers. The common source amplifiers may be swamped or non-swamped, depending on the bias form used. If the bias type does not utilize a source resistor, swamping is not available. This includes zero bias for the DE-MOSFET and voltage divider bias for both the DE- and E-MOSFET.

Like their JFET counterparts, MOSFET common source amplifiers exhibit moderate inverting voltage gain, very high input impedance and moderate output impedance. The input impedance is a function of the biasing resistor configuration situated in front of the gate as the impedance looking into the gate itself is very, very high at low frequencies.

The MOSFET common drain followers also behave similarly to the JFET version. Again we see a non-inverting voltage gain approaching unity, a very high input impedance and a low output impedance. The higher the transconductance is, the closer the gain will be to unity and the lower the output impedance will be.

Review Questions

1. How well does the MOSFET voltage amplifier compare to its JFET counterpart?
2. How well does the MOSFET source follower compare to its JFET counterpart?
3. What are the practical differences between a voltage amplifier using a DE-MOSFET versus using an E-MOSFET ?
4. It has been stated that a source follower cannot be made using a standard zero biased DE-MOSFET. Why is this?
5. It has been stated that a swamped voltage amplifier cannot be made using a standard voltage divider biased E-MOSFET. Why is this?

Problems

Analysis Problems

1. For the amplifier of Figure 13.14, determine Z_{in} and A_v. V_{in} = 20 mV, I_{DSS} = 10 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 20 V, R_G = 750 kΩ, R_D = 2 kΩ, R_L = 4 kΩ, R_S = 1 kΩ, R_{SW} = 200 Ω.

2. For the amplifier of Figure 13.14, determine Z_{in} and V_{out}. V_{in} = 25 mV, I_{DSS} = 15 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 22 V, R_G = 330 kΩ, R_D = 2 kΩ, R_L = 6 kΩ, R_S = 510 Ω, R_{SW} = 220 Ω.

3. For the circuit of Figure 13.15, determine Z_{in} and A_v. V_{in} = 10 mV, I_{DSS} = 12 mA, $V_{GS(off)}$ = −2.5 V, V_{DD} = 26 V, R_G = 510 kΩ, R_D = 1.2 kΩ, R_L = 25 kΩ.

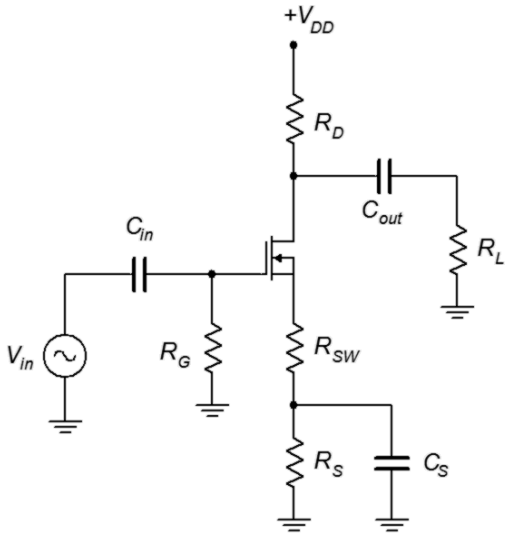


Figure 13.14

4. For the circuit of Figure 13.15, determine Z_{in} and V_{out}. V_{in} = 25 mV, I_{DSS} = 15 mA, $V_{GS(off)}$ = −1.5 V, V_{DD} = 24 V, R_G = 820 kΩ, R_D = 1 kΩ, R_L = 12 kΩ.

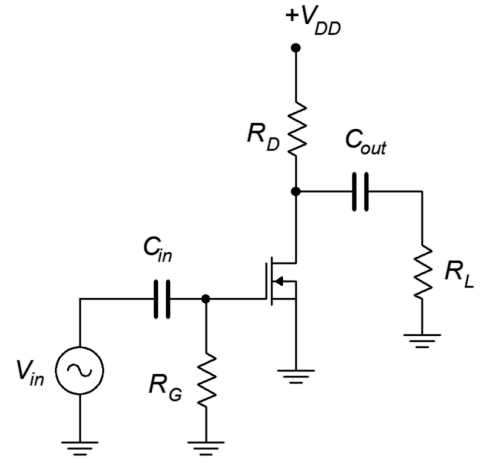


Figure 13.15

5. For the circuit of Figure 13.16, determine Z_{in} and V_{out}. V_{in} = 25 mV, I_{DSS} = 8 mA, $V_{GS(off)}$ = −3.5 V, V_{DD} = 24 V, R_1 = 1 MΩ, R_2 = 100 kΩ, R_D = 800 Ω, R_L = 10 kΩ.

6. For the circuit of Figure 13.16, determine Z_{in} and A_v. V_{in} = 10 mV, I_{DSS} = 6 mA, $V_{GS(off)}$ = −4 V, V_{DD} = 26 V, R_1 = 2 MΩ, R_2 = 120 kΩ, R_D = 1.2 kΩ, R_L = 15 kΩ.

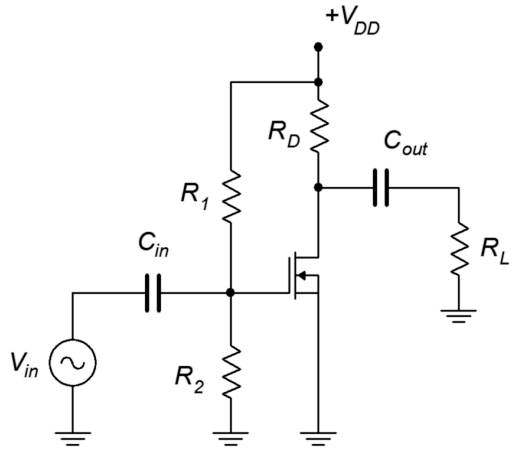


Figure 13.16

7. For the circuit of Figure 13.17, determine Z_{in} and V_{out}. V_{in} = 20 mV, $I_{D(on)}$ = 6 mA at $V_{DS(on)}$ = 3 V, $V_{GS(th)}$ = 2.5 V, V_{DD} = 34 V, R_1 = 1 MΩ, R_2 = 100 kΩ, R_D = 1 kΩ, R_L = 10 kΩ.

8. For the circuit of Figure 13.17, determine Z_{in} and A_v. V_{in} = 15 mV, $I_{D(on)}$ = 10 mA at $V_{DS(on)}$ = 4 V, $V_{GS(th)}$ = 2 V, V_{DD} = 30 V, R_1 = 2 MΩ, R_2 = 180 kΩ, R_D = 1.2 kΩ, R_L = 15 kΩ.

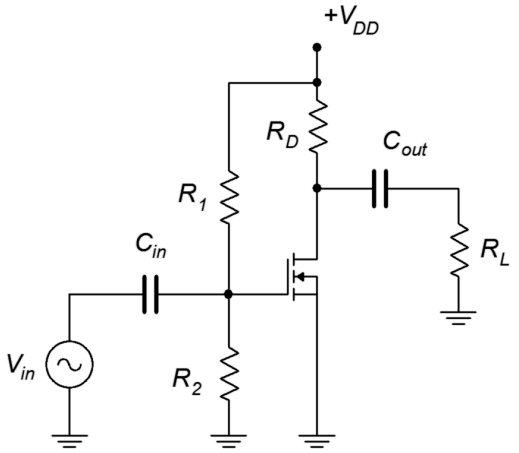


Figure 13.17

9. For the circuit of Figure 13.18, determine Z_{in} and V_{out}. V_{in} = 200 mV, I_{DSS} = 15 mA, $V_{GS(off)}$ = −3 V, V_{DD} = 15 V, R_G = 910 kΩ, R_L = 10 kΩ, R_S = 330 Ω.

10. For the circuit of Figure 13.18, determine Z_{in} and V_{out}. V_{in} = 200 mV, I_{DSS} = 20 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 12 V, R_G = 1 MΩ, R_L = 1.8 kΩ, R_S = 220 Ω.

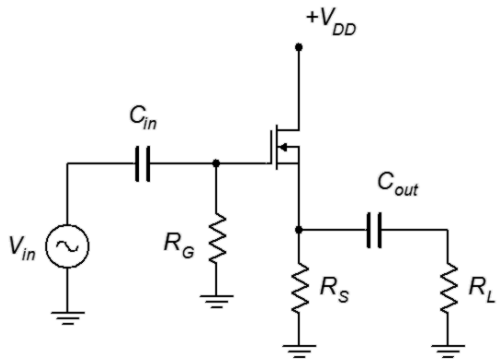


Figure 13.18

11. For the circuit of Figure 13.19, determine Z_{in} and A_v. I_{DSS} = 18 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 12 V, V_{SS} = −4 V, R_G = 680 kΩ, R_L = 10 kΩ, R_S = 1 kΩ.

12. For the circuit of Figure 13.19, determine Z_{in} and A_v. I_{DSS} = 20 mA, $V_{GS(off)}$ = −2 V, V_{DD} = 10 V, V_{SS} = −6 V, R_G = 2.2 MΩ, R_L = 5 kΩ, R_S = 510 Ω.

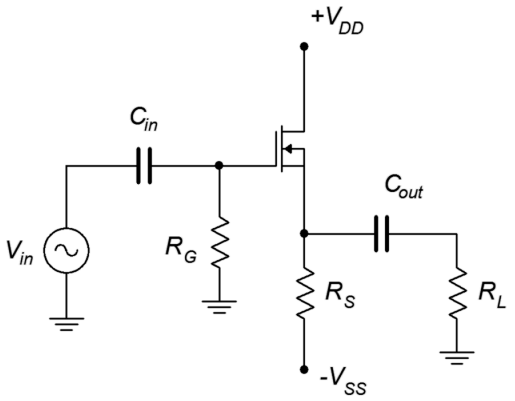


Figure 13.19

Design Problems

13. Following the circuit of Figure 13.14, design an amplifier with a gain of at least 5 and an input impedance of at least 500 kΩ. R_L = 10 kΩ. The MOSFET has the following parameters: $V_{GS(off)}$ = −2 V, I_{DSS} = 25 mA. Try to use standard resistor values.

14. Using the circuit of Figure 13.18, design a follower with a gain of at least 0.75 and an input impedance of at least 1 MΩ. R_L = 2 kΩ. The MOSFET has the following parameters: $V_{GS(off)}$ = −1.5 V, I_{DSS} = 40 mA. Try to use standard resistor values.

Challenge Problems

15. Using Figure 13.20, find Z_{in} and A_v. I_{DSS} = 15 mA, $V_{GS(off)}$ = −2 V.

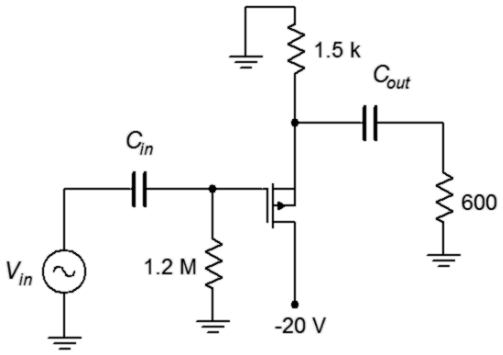


Figure 13.20

16. For the circuit of Figure 13.21, determine Z_{in} and A_v. I_{DSS} = 12 mA, $V_{GS(off)}$ = −1.5 V.

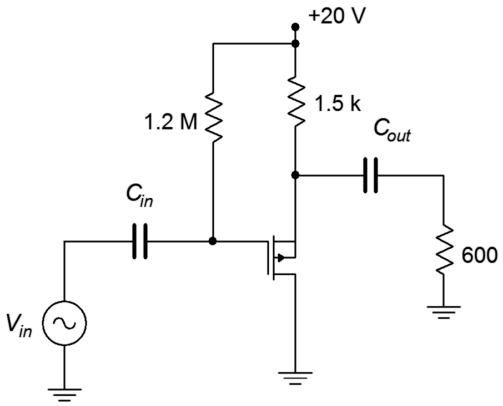


Figure 13.21

Computer Simulation Problems

17. Utilizing manufacturer's data sheets, find devices with the following specifications (typical) and verify them using the measurement techniques presented in the prior chapter. Device 1: $V_{GS(off)}$ = −2 V, I_{DSS} = 25 mA. Device 2: $V_{GS(off)}$ = −1.5 V, I_{DSS} = 40 mA.

18. Using the device model from the preceding problem, verify the design of Problem 13.

19. Using the device model from Problem 17, verify the design of Problem 14.

14 Class D Power Amplifiers

14.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Outline the building blocks of a class D amplifier.
- Discuss the advantages and disadvantages of the class D amplifier compared to class A and class B amplifiers.
- Explain the concepts of *pulse width modulation*, *shoot-through* and *dead time*.

14.1 Introduction

Much has been written in this text regarding the efficiency of various power amplifier topologies. While class A is known for its circuit layout simplicity, it is also known for its very low efficiency. Class B and class AB, while more complex than class A, present serious improvements in efficiency. In spite of these improvements, the family of class B amplifiers can hardly be considered as exhibiting high efficiency. Although not explicitly covered in this text, class G and H topologies are variations on class B and attempt to increase efficiency through the use of multiple sets of power supply rails or output devices, and in the process, tick the complexity up to another level.

The class D amplifier is perhaps the last word in amplifier efficiency. Theoretically with ideal devices, the efficiency of the output stage approaches 100%. Unlike the other amplifier forms, the transistors used in class D amplifiers never operate in the linear region; the output devices only operate as a switch, in either saturation or cutoff. High switching speed turns out to be a huge plus as it plays a major role in efficiency.

The increase in efficiency comes at a considerable increase in circuit complexity, however, for some applications this turns out to be a very good trade-off. As odd as it might at first seem, the two areas where class D topologies have taken root are at the opposite ends of the power output spectrum. The first area is perhaps the most obvious, mainly, very high output power amplifiers. An example might be an amplifier used as part of a large public address system and capable of delivering in excess of 1000 watts into a loudspeaker. High efficiency does two things here: First, it reduces the waste heat in the amplifier itself, and second, it reduces the current draw from the AC mains. Both of these are serious issues in a PA system used to fill a stadium or large concert hall as there may be dozens of such amplifiers comprising the system. As a bonus, improved efficiency also leads to a lighter and small enclosure because the need for heat sink area and mass will be reduced, as will the size of the AC power supply transformer. These traits will also reduce production costs and help offset the design complexity cost. The advantages have become so great that, in recent years, class D designs dominate the high end professional audio power amplifier market as well as the very high power automotive audio market (here there is another system limitation working in favor of class D, and that's the limited current capacity of the

vehicle's alternator to deliver current).

The second area where class D has found acceptance is for low power portable devices. Examples include personal music devices, cell phones and hearing aids. Output powers for these applications might range from tens of milliwatts up to a few watts, so excess heat is generally not a big problem except in the most compact of enclosures. What is a problem, though, is the energy budget. Unlike a large PA amplifier that might pump out in excess of two horsepower, these portable devices do not have the luxury of running off of the AC mains with tens or even hundreds of amps of current capacity. Instead, these devices are restricted to battery power and batteries can only store so much energy. For a given battery capacity, higher efficiency directly translates into longer battery life. Another way of thinking about this is that, given a higher efficiency, a smaller battery can be used to achieve the same battery life, and this means that the unit can be both smaller and less expensive. Of course, nothing says that we can't opt for a little of each.

14.2 Class D Basics

The key to the high efficiency of class D operation is to only operate the output devices as switches. That is, they are operated at the extreme ends of the load line, in either cutoff or saturation. The only exception to this rule is when the output device is transitioning from one state to another. Either BJTs or E-MOSFETs can be used, although for reasons that we shall examine, E-MOSFETs tend to be preferred in many applications and therefore we shall use them here as a general rule.

To understand the power advantage of class D switching, consider the circuit of Figure 14.1. Here we have an E-MOSFET being driven by a square wave at its gate terminal. The square wave runs from zero to some voltage well above $V_{GS(th)}$, sufficient to fully turn-on the MOSFET.

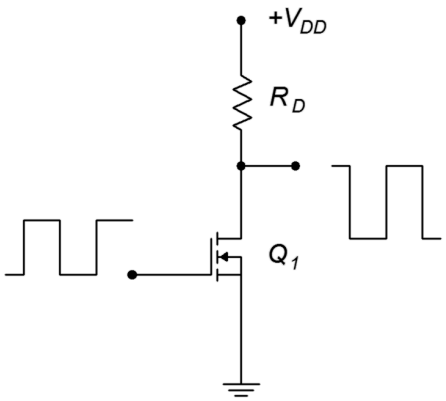


Figure 14.1
A simple MOSFET switch.

At the drain, an inverted square wave appears that ideally runs from $+V_{DD}$ down to zero volts. When the MOSFET is off, no drain current flows. Therefore, the drop

across the drain resistor, R_D, is zero. Consequently, by KVL, V_{DS} must equal V_{DD}. When the gate signal goes high, it turns on the MOSFET causing a large current. Ideally, this current equals V_{DD}/R_D and $V_{DS} = 0$ V. In reality, the MOSFET will present some resistance ($r_{DS(on)}$) and this essentially creates a voltage divider between the device and R_D. Obviously, if $r_{DS(on)} << R_D$ then we can approximate the low state as zero volts.

In the ideal case, the transistor dissipates no power. Here's why: When the gate is low, the device is off so no current flows. Although the device voltage is very high, the product of the device's voltage and current is zero. When the gate is high, the transistor turns on and conducts maximum current, however, the voltage across the device is zero, and the resulting product is zero once again. Therefore, the device dissipates no power.

The reality of the situation reveals that some power is indeed wasted by the output device. There are two chief culprits: non-zero turn-on voltage (caused by $r_{DS(on)}$ for example) and state transitions that are not instantaneous (i.e., the rise and fall times of the waveforms are not zero). These effects are illustrated in Figure 14.2.

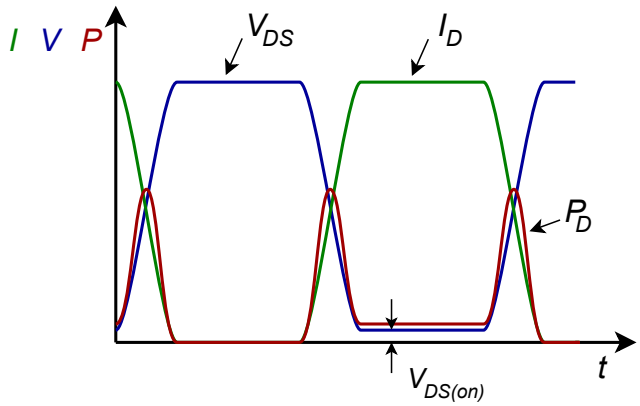


Figure 14.2
Switching losses.

The rise and fall times are exaggerated for clarity as is the voltage drop across the transistor. The green curve represents the device current while the blue curve indicates the device voltage. Note that during the on-state the transistor voltage is not zero, although it is very small. The product of these two curves is the power dissipated by the transistor and is shown in red. In the off-state, no power is dissipated. In the on-state, a small amount of power is dissipated due to the device's on-voltage. Power spikes also occur at the state transitions as neither the current nor the voltage are at zero.

The area under the red curve represents the total power dissipated by the output transistor. In order to minimize this loss and maximize the efficiency, we would like as low of an $r_{DS(on)}$ as possible along with very fast switching speeds. As power E-MOSFETs tend to switch faster than their BJT counterparts, this presents a compelling argument for their use. Further, large power E-MOSFETs may exhibit

$r_{DS(on)}$ values of just a few milliohms, guaranteeing relatively small on-state losses (compared to load power) even with drain currents of several tens of amps.

14.3 Pulse Width Modulation

Clearly, class D presents the possibility of minimal wasted power and high efficiency. We are now left with the problem of how to turn a series of pulses into a continuous, smoothly varying waveform, such as a voice or music signal. There are a few ways to accomplish this; it's a matter of encoding the amplitude of the original signal into the pulse train that drives the output devices. Theoretically, as long as the "area under the curve" for a segment of input signal is identical to the area represented by the pulse train, we will have successfully encoded the signal and we then should be able to decode it, turning it back into a smoothly varying output signal. For this to work properly, the pulse train will have to be at much higher frequency than the input signal in order to follow its changes over time. One way to do this is through *pulse density modulation*, or PDM. The idea is to produce a number of narrow pulses to represent the area. If the input amplitude is large, we create a large number of pulses and if the amplitude is small, we produce a small number of pulses. While this technique can work, it is somewhat challenging to turn this pulse train back into the desired signal at the load.

Another technique to encode the input is *pulse width modulation*, or PWM. Instead of altering the number of pulses in a given period of time, we keep the frequency constant and adjust the width of the pulses. If the input amplitude is high, the width of the corresponding pulse will be wide and if the amplitude is low, the pulse width will be narrow. The decoding of PWM is easier than that of PDM and is generally the preferred route.

Generating PWM is a relatively straightforward affair. All we need is a triangle wave and a comparator. The comparator has two input terminals: the signal to be encoded (input signal) and the reference wave (triangle wave). It has a two-state, logical output. The output will be high if the signal is more positive than the reference and it will be low if the reference is more positive than the signal. This is shown in block diagram form in Figure 14.3.

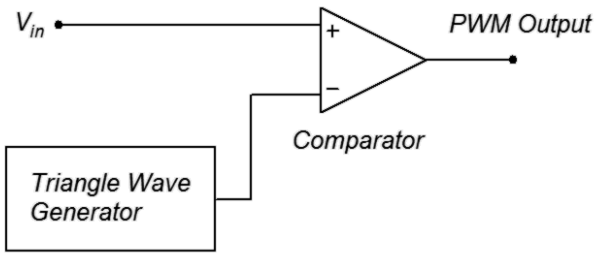


Figure 14.3
PWM encoder.

As mentioned, the triangle wave needs to be at a much higher frequency than the signal that is being encoded. As a general rule, this frequency should be at least ten times the highest input signal frequency.

$$f_{triangle} \geq 10 f_2 \tag{14.1}$$

A simulation showing the PWM waveforms is presented in Figure 14.4.

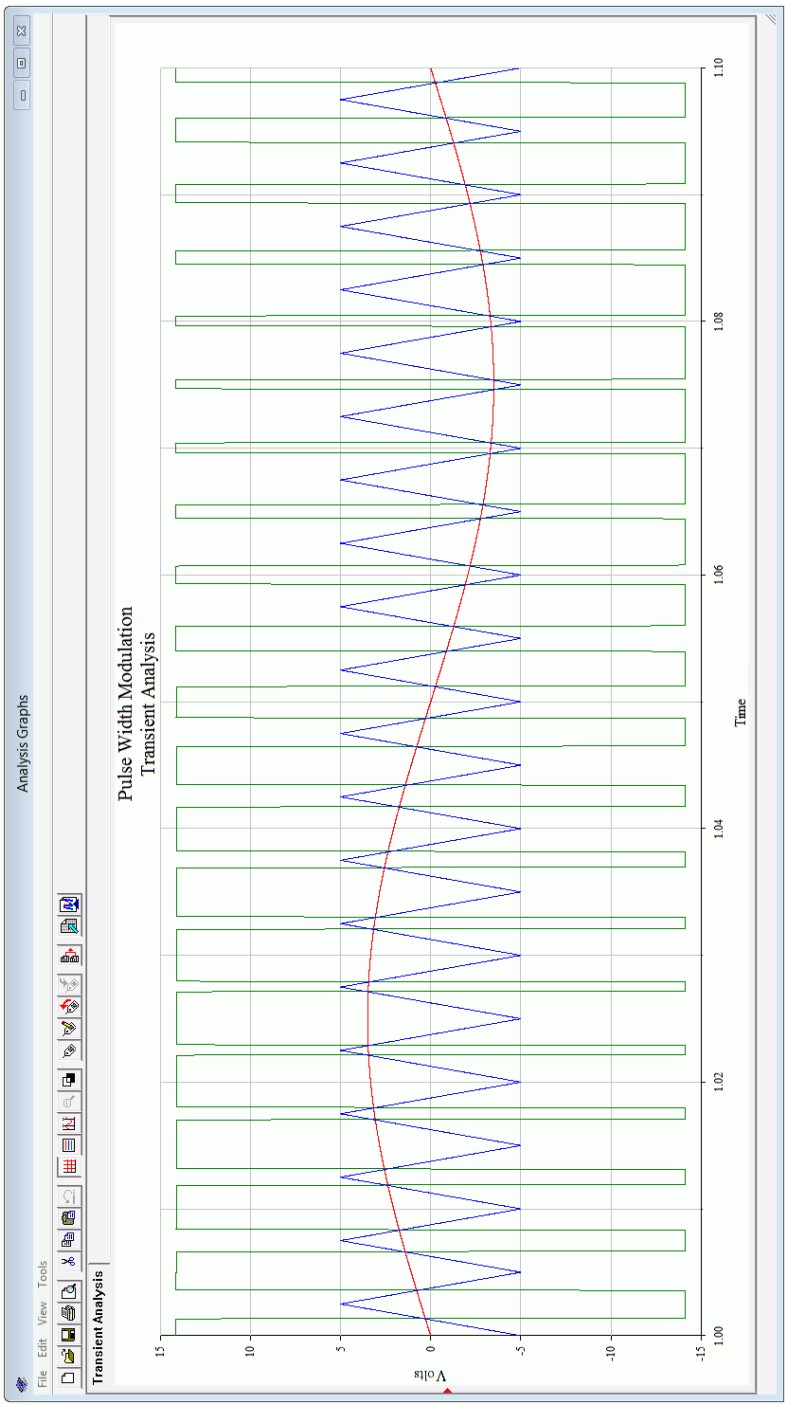


Figure 14.4
PWM waveforms.

The input signal is the red sine wave. The blue triangle wave is the reference and is approximately 20 times higher in frequency. The green wave is the PWM output. Note that when the red input signal climbs above the blue reference triangle, the green output goes high, otherwise, the output is low. Thus, the duty cycle of the pulses correlates with the input signal amplitude. The input signal should not exceed the amplitude of the triangle wave otherwise accuracy will impaired. Also, the accuracy of the encoding process is dependent on the linearity of the triangle wave, so a high quality triangle wave generator is needed. Lastly, for accuracy and ease of decoding, the output pulses should not be allowed to become too thin, so the input signal should be limited to about 75% of the amplitude of the triangle wave.

Reconstituting the Output

Although we have successfully encoded the input signal into a series of pulses, we still need to decode them, that is, reconstitute the original continuously variable input signal. Mathematically, the PWM signal contains all of the original input signal frequency components and amplitudes, it just has added a large of number of new frequency components. These new components are multiples (harmonics) of the fundamental PWM frequency, and consequently, they are all higher than the input signal frequencies. As such, they may be removed with a low-pass filter. This will effectively reconstitute the original signal (but at a much higher amplitude, of course). A simple passive *LC* filter would be appropriate in this instance as it must pass large currents and voltages. An example is shown in Figure 14.5.

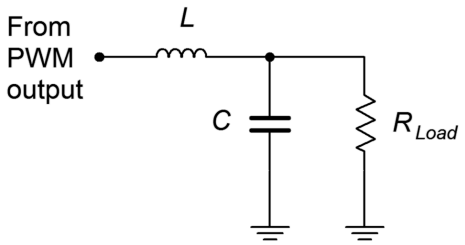


Figure 14.5
Passive low-pass filter.

At low frequencies, X_L will be very small while X_C will be very high, thus virtually all of the input signal frequencies will pass on to the load. At high frequencies, such as the harmonics of the PWM pulses, the situation is reversed: X_L is large and X_C is small, creating a large loss so that these components do not reach the load.[44] The critical frequency of the network is set to the highest input signal frequency (e.g., for high fidelity audio, slightly above 20 kHz).

44 For an audio amplifier, it is important that these components do not reach the loudspeaker. Even though they are beyond the range of human hearing, they can damage loudspeaker sub-components and, at the very least, present an extra power dissipation burden to them. Other kinds of loads may not be effected by the harmonics and filtering may not be needed.

We now have a complete outline for the class D amplifier, as shown in Figure 14.6.

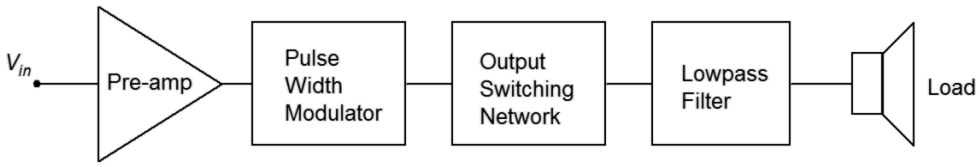


Figure 14.6
Class D amplifier block diagram.

The pre-amp can be comprised of any of the linear amplifier outlines presented in earlier chapters, whether they use BJTs or FETs. What remains then, is further investigation into the output switching network.

14.4 Output Configurations

If we apply the switching concept to a dual supply, push-pull topology, we arrive at the generic circuit of Figure 14.7.

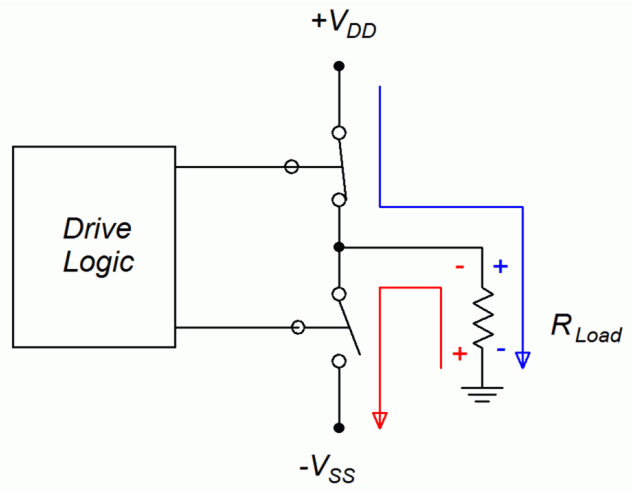


Figure 14.7
Generic push-pull switching.

The two output devices are alternately switched on and off. When the upper device is on, the lower device is off, and current flows from the positive supply to the load (blue path). Alternately, when the lower device is on, the upper device is off, and current flows through the load via V_{SS} (red path). We could use either BJTs or E-MOSFETs for these devices.

Two obvious variations exist of the generic output circuit. The first version, shown in Figure 14.8, appears to be a direct take-off of a class B output. It is shown with E-MOSFETs but could be made with BJTs. Biasing details are not shown, instead a generic "driver" circuit block will prove sufficient for our discussion.

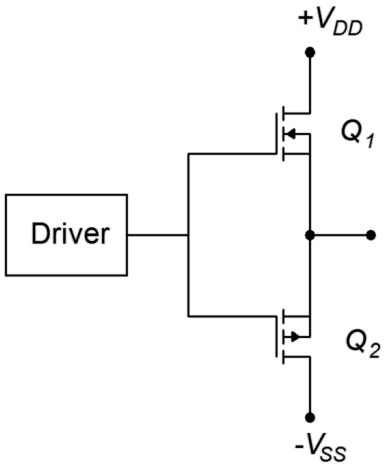


Figure 14.8
Basic push-pull switching.

In this circuit, the driver produces a bipolar pulse train that swings from negative to positive rather than from ground to positive. A positive level from the driver will turn on the upper N-channel device, allowing current flow to the load. In contrast, a negative level will turn on the lower P-channel device, allowing load current flow via V_{SS}.[45] It is worth noting that the gate drive signal must swing higher and lower than the two power supplies. This is because when a device is on, V_{DS} will be nearly zero, meaning that the source will be at the power rail. As V_{GS} must be greater than $V_{GS(th)}$, this means that V_G must be greater than the power supply.

An alternate connection scheme is shown in Figure 14.9. Here, the N- and P-channel devices have switched positions.

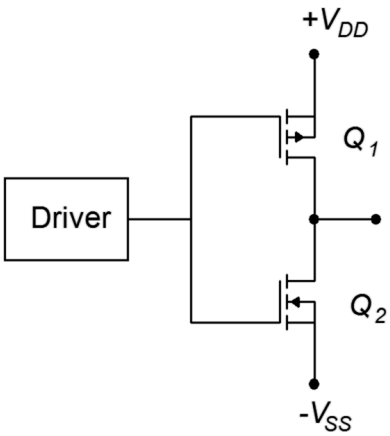


Figure 14.9
Alternate push-pull switching.

The logic here is reversed: the negative pulse turns on the upper device and the positive pulse turns on the lower device. The gate swing is lessened a little compared

45 Yes, it is labeled V_{SS} in spite of the fact that it's connected to the drain of the P-channel device. It's a matter of consistency with other circuits. “A rose by any other name” and all that...

to the earlier circuit but it suffers from a flaw common to both configurations, namely, asymmetry between the N- and P-channel device characteristics. This includes variations between internal device capacitances and $r_{DS(on)}$ values. For the best possible matching, and thus the lowest distortion and highest performance, it would be better to configure the output using identical devices. This is shown in Figure 14.10.

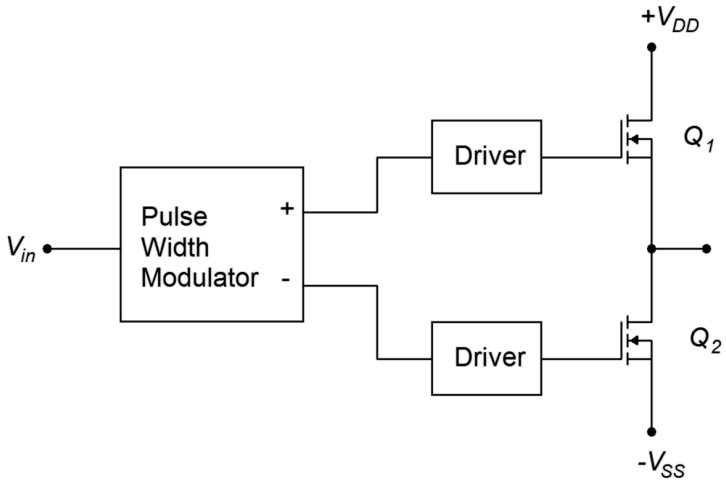
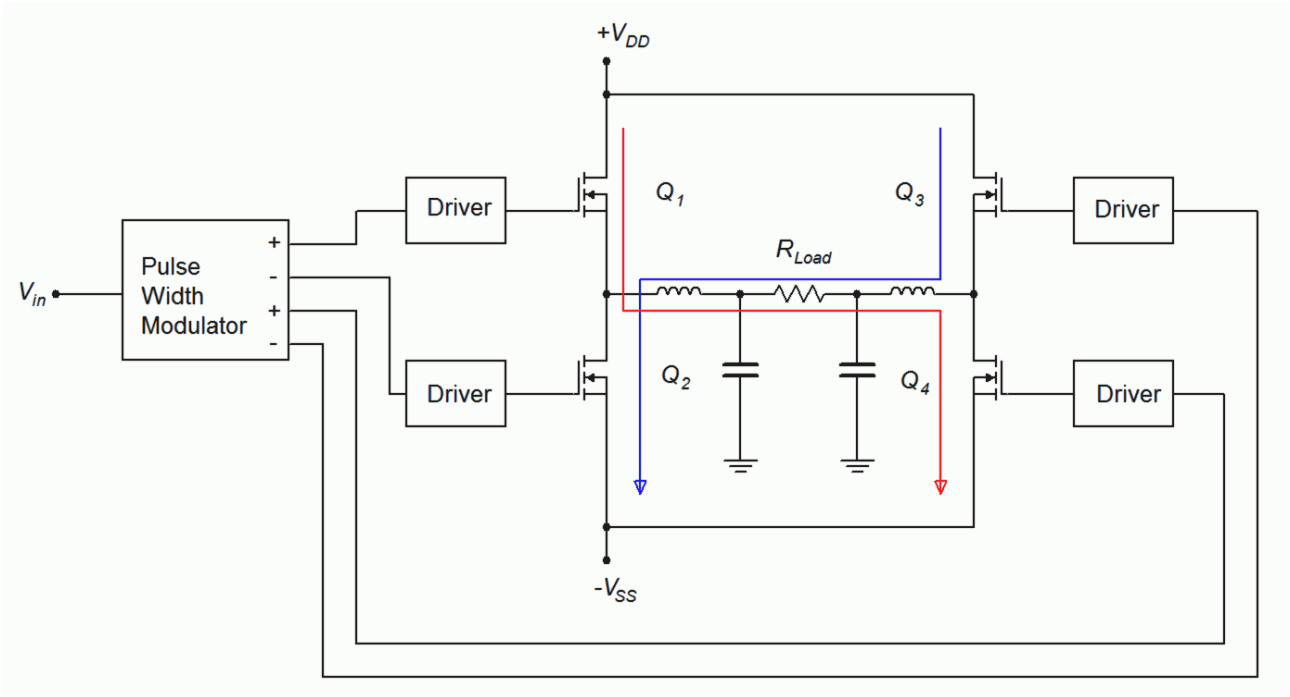


Figure 14.10
Output switching using identical devices.

This configuration complicates the drive signal in that we can no longer drive both gates with the same signal; instead, unique signals must be presented to each gate terminal. This circuit is known as a *half bridge*. Our final step will be to drive both ends of the load in differential fashion using a *full bridge*, or *H bridge* as it is sometimes known. This is shown in Figure 14.11.

Figure 14.11
Full bridge output.



The output devices are controlled as diagonal pairs. When Q_1 is on, Q_4 is on, creating a path for load current from left to right (red trace). In contrast, when Q_2 is on, Q_3 will also be on, thus creating a load current path from right to left (blue trace). This effectively doubles the current amplitude which quadruples the load power (because power varies as the square of current). This is the same technique discussed in Chapter 9 with class B amplifiers. A dual *LC* filter is included in this diagram to remove unwanted frequency components.

Example 14.1

A pair of E-MOSFETs are configured to drive an 8 Ω load as in Figure 14.10. Assuming that ±50 volt sources are used and that each device has an $r_{DS(on)}$ of 0.03 Ω, determine the peak load current and V_{DS} for the MOSFETs.

At any given time, one MOSFET will be on, creating a path between one supply, itself, the load and ground. The total resistance to limit the current will be the load plus $r_{DS(on)}$.

$$i_{load} = \frac{V_{DD}}{r_{load} + r_{DS(on)}}$$
$$i_{load} = \frac{50\,\text{V}}{8\,\Omega + 0.03\,\Omega}$$
$$i_{load} = 6.227\,\text{A}$$

The device voltage is found via Ohms law as load and drain current are identical.

$$v_{DS} = i_{load}\, r_{DS(on)}$$
$$v_{DS} = 6.227\,\text{A} \times 0.03\,\Omega$$
$$v_{DS} = 0.19\,\text{V}$$

Practical Concerns

There are a few details left that should not be overlooked. Two of them are related to the edge transition areas, another concerns the complexity of the drive circuits, and the final issue deals with the power supplies themselves

The first item of concern is precisely what happens during the transition. All of the output forms we have examined utilize two active devices configured in series between two power sources. There is nothing in that path to limit current. If both devices were to be simultaneously triggered to the on-state, a huge and possibly

damaging current would flow. While it would be foolish to turn both devices on intentionally, the rise and fall times of the pulses effectively do this. As one device is turning on and the other is turning off, both devices are in a conducting state, even if it's not maximum conduction. Essentially, we have two low impedance devices in series between two sources. This results in a large current pulse known as *shoot-through*. This situation is depicted in Figure 14.12.

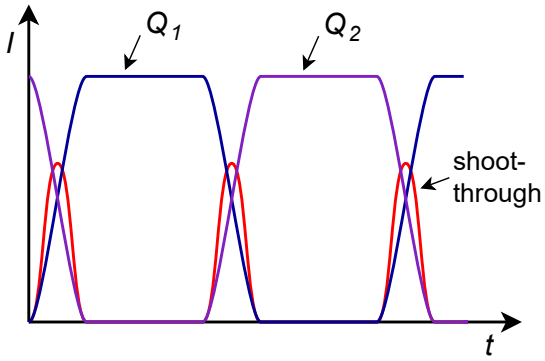


Figure 14.12
Shoot-through.

The current pulses for the two devices are shown in blue and violet. The maximal currents are directed to the load, but during the transition, a pulse of current, shown in red, "shoots through" the two devices, from one power supply directly to the other.

The solution to shoot-through is to adjust the turn-on and turn-off pulse timing so as to create a *dead time*, that is, a time span when neither device is directed to turn on. This is illustrated in Figure 14.13.

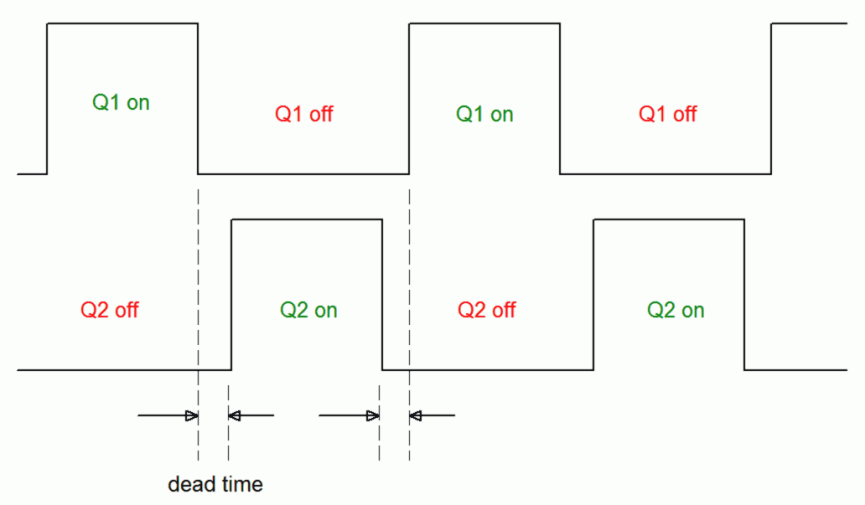


Figure 14.13
Dead time.

Dead time is adjusted to correspond with the rise and fall times of the output devices. Basically, a device is not allowed to turn on until the other device is, indeed, fully off. The inclusion of dead time alters the width of the pulse and consequently can introduce waveform distortion. A minimum amount of dead time should be used

to avoid this. This is another reason to use very fast output devices as they will require shorter dead times.

The second issue regarding timing is one of device capacitance. Power MOSFETs exhibit relatively high device capacitances. For example, the FDMS86180 examined in Chapter 12 exhibits input and output capacitances of roughly 4.4 nF and 2.7 nF, respectively. Although the extremely high gate input resistance might seem to indicate that very little drive current is needed to turn these devices on, the capacitance tells a different story.

The rate of change of voltage across a capacitor is a function of the capacitance and the current driving it:

$$\frac{dv_C}{dt} = \frac{i_C}{C}$$

The larger the current, or the smaller the capacitor, the greater the rate of change of voltage. This can place a serious limit on how quickly a device may be controlled. For example, suppose the drive circuit can pump out up to 10 mA. At first glance that may seem like an enormous amount of current to drive a MOSFET. Now, consider what happens if the input capacitance is 2 nF:

$$\frac{dv_C}{dt} = \frac{i_C}{C}$$
$$\frac{dv_C}{dt} = \frac{10\,\text{mA}}{2\,\text{nF}}$$
$$\frac{dv_C}{dt} = 5\text{E}6\,\text{V/s}$$

While a 5 million volt-per-second slope might sound fast, it's only 5 volts per microsecond. Compared to the requirements of, say, a 200 kHz to 300 kHz switching frequency, that is horribly slow.

Computer Simulation

To see the effect of input capacitance, a two-stage amplifier is captured in a simulator, as shown in Figure 14.14. The circuit consists of a relatively standard small signal amplifier feeding a medium power E-MOSFET, the IRF7201. A 10 kHz square wave is used to drive the circuit. The input capacitance of the MOSFET is 550 pF, certainly larger than a small signal FET but not an extremely large value. A single capacitor is placed across the gate that will be used to simulate a much larger device and the associated increased input capacitance.

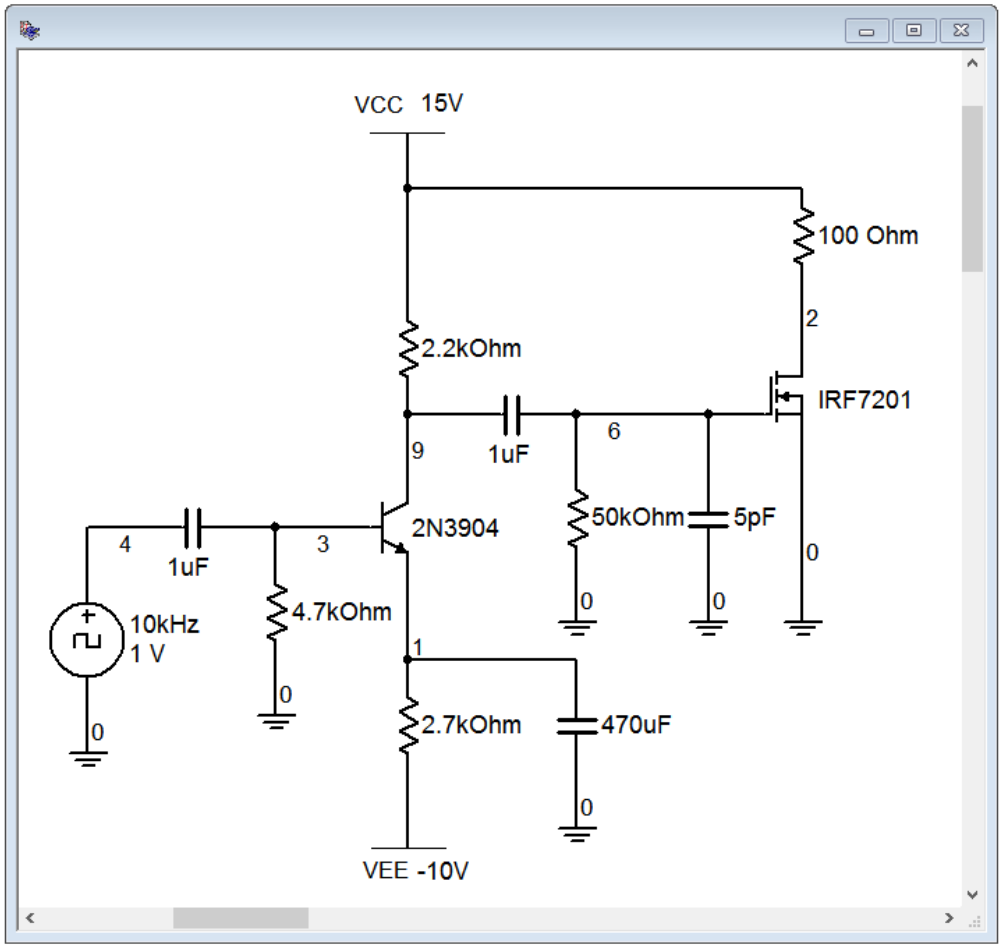


Figure 14.14
Circuit for input capacitance testing.

The initial transient analysis is run using a 5 pF gate capacitance which has no appreciable effect on the outcome. The result of the simulation is shown in Figure 14.15.

The red trace is the gate voltage at node 6 while the blue trace is the final output at node 2. The gate drive signal is suffering somewhat at the upper portion and the rise and fall times are evident. The output signal is swinging from the power supply of +15 volts down to ground, as expected. The rise time is somewhat quicker than the fall time but, in general, the output presents a decent square pulse at close to 50% duty cycle. The simulation is repeated but this time the gate capacitance is increased from 5 pF to 5 nF, a value more typical of a large power FET. The result is shown in Figure 14.16. The red gate drive signal has taken a serious hit and is no longer square in shape. The output pulse in blue still runs to and from the expected voltage levels, but the rising and falling edges are noticeably slowed. Also, the positive pulse width has been stretched due to the slowing of the gate signal which retards the turn on of the MOSFET. The result is a duty cycle that is greater than 50%.

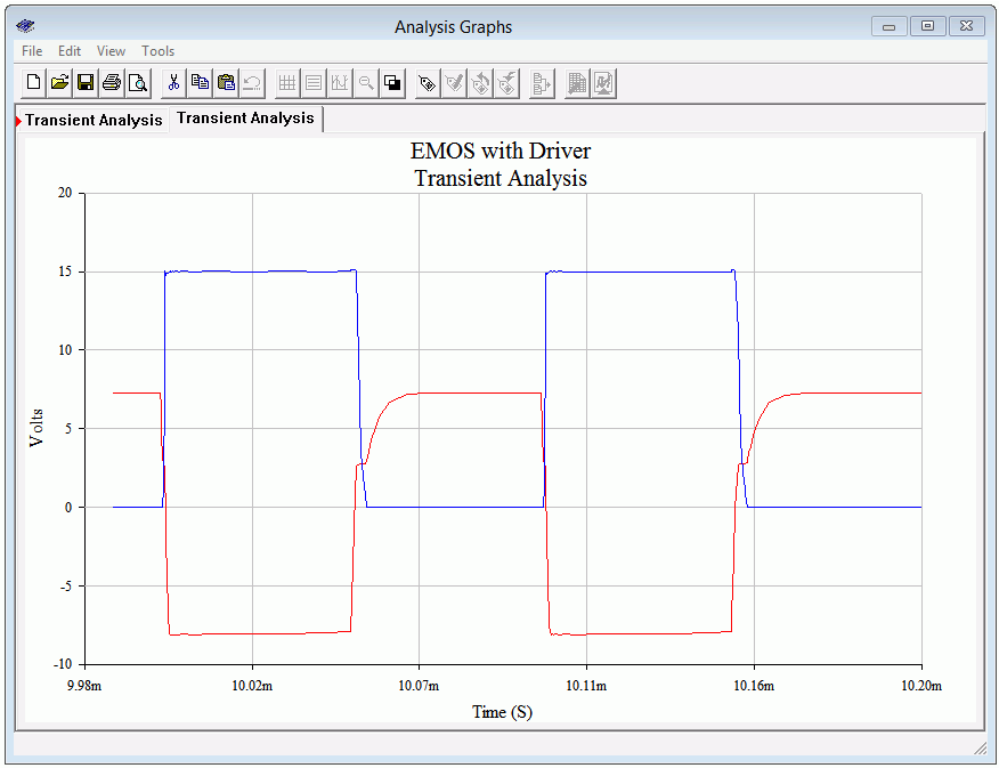


Figure 14.15
Waveforms for normal circuit.

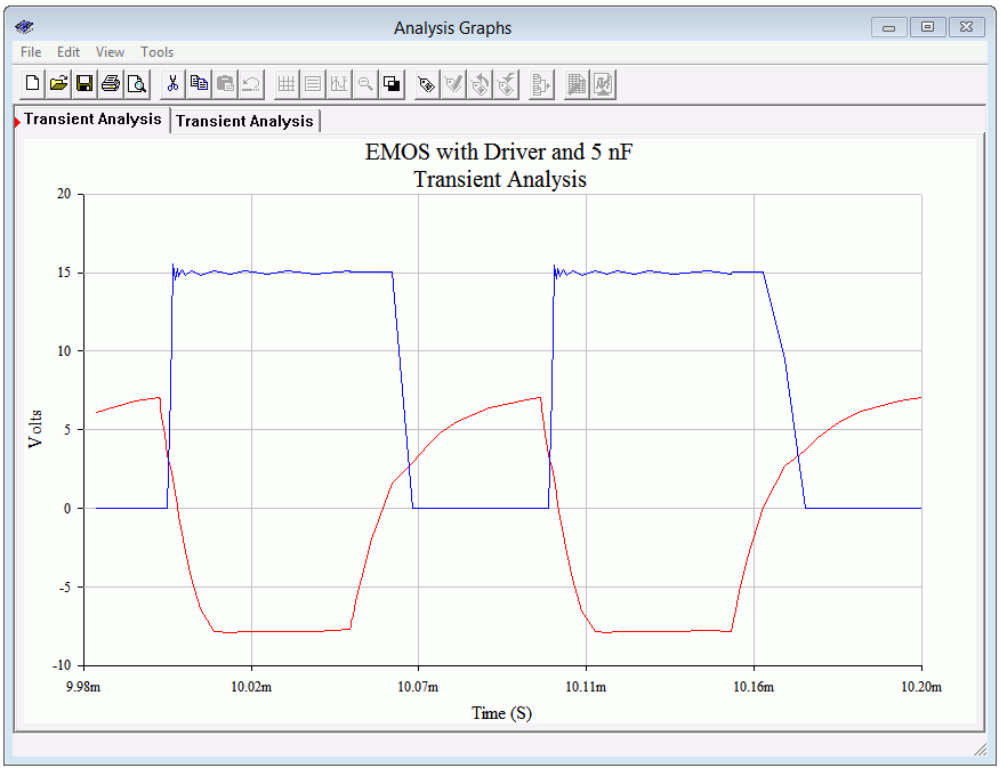


Figure 14.15
Waveforms for circuit with increased gate capacitance.

The bottom line is that, in order maximize speed, care must be taken to minimize capacitance, decrease the output impedance of the driver circuit and increase the drive current. Fortunately, this issue has been largely solved by IC manufacturers who offer FET driver integrated circuits designed specifically for these applications.

The final item of practical concern is the power supply itself, or more precisely, the quality of the supply voltage. Remember, the output device is being used as a switch. When the device is on, the power supply is directly connected to the output (with the exception of the small voltage drop across the output device). This means that any noise or ripple on the power supply will make its way to the output filter. Whatever noise components fall within the desired input signal range will not be filtered out, and thus are delivered to the load. In short, the output devices will "leak" the power supply noise into the output, so care must be taken to have as clean of a power supply voltage as possible.

Summary

The class D amplifier boasts very high efficiency, theoretically approaching 100%. The amplifier operates its output devices as switches; they are either fully on or fully off. The power losses are mostly relegated to switching edge losses so it is important to not switch the output devices at too high of a frequency.

The input signal is encoded as a series of pulses, typically via pulse width modulation. The pulse frequency is much higher than the highest input signal frequency, typically by an order of magnitude. The width of the pulses is a function of the amplitude of the input signal. That is, the higher the input signal, the greater the pulse width. The pulses are amplified by applying them to the output devices which then act as switches to alternately connect and disconnect the power supplies to the output terminal. The sequence of much larger amplitude pulses are then fed to a low-pass filter, typically, an *LC* filter, to remove the high frequency components of the pulse train. The reconstituted input signal is what remains, but at a much higher amplitude.

The output can be configured using either two-device half-bridge or four-device full-bridge arrangements. The full-bridge is preferred for higher performance. The input capacitance of the output devices can be relatively high, so care must be taken to ensure that sufficient capacity is available from the driver circuit.

Review Questions

1. What is PWM?
2. How does PWM differ from PDM (pulse density modulation)?
3. What is *shoot-through*? What causes it?
4. What is *dead time*? What is its purpose?
5. What is the function of the output *LC* filter?
6. What effect does device input (e.g., gate) capacitance have on the operation of the amplifier?
7. Sketch a half-bridge configuration.
8. Sketch a full-bridge configuration.
9. What is the effect of $r_{DS(on)}$ regarding load current and power losses?
10. What is the effect of drive current capacity on the output devices?
11. What effect does power supply noise have on the output?

Problems

Analysis Problems

1. A telephony system has a frequency range from 200 Hz to 3.5 kHz. Determine the minimum acceptable PWM frequency.

2. A background music and paging system has a frequency range from 50 Hz to 10 kHz. Determine the minimum acceptable PWM frequency.

3. Determine the maximum rate of change of input voltage for a driver circuit capable of producing 50 mA with a load consisting of a 1.5 nF gate capacitance.

4. Determine the maximum rate of change of input voltage for a driver circuit capable of producing 60 mA with a load consisting of a 3.5 nF gate capacitance.

5. A power E-MOSFET has an $r_{DS(on)}$ of 0.012 Ω and switches a 100 volt source to an 8 Ω load. Determine the maximum load current and V_{DS}.

6. Four power E-MOSFETs drive a 4 Ω load via a full bridge network. If each device has an $r_{DS(on)}$ of 0.02 Ω and they switch ±75 volt sources to the load, determine the load current and V_{DS} of the devices.

15 Insulated Gate Bipolar Transistors (IGBTs)

15.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Discuss the basic operation of the IGBT.
- Describe the internal structure of the IGBT.
- Describe the differences between PT and NPT IGBTs.
- Compare and contrast the IGBT to the power BJT.
- Compare and contrast the IGBT to the power MOSFET.
- Interpret important parameters found on an IGBT data sheet.
- Describe basic circuits using the IGBT as a power control device.

15.1 Introduction

The *Insulated Gate Bipolar Transistor*, or *IGBT*, is a power semiconductor that first became available to the commercial market during the 1980s. Initial devices had certain performance issues but these problems largely have been taken care of with subsequent generations. Today, the IGBT is in wide use, competing with power BJTs and power E-MOSFETs across a range of applications. The IGBT is designed to be used as a high voltage/high current switch and typically is not used for linear applications such as an audio class B power amplifier. The IGBT has also overtaken the older thyristor devices (e.g., SCR) in many areas due to its speed and the relative simplicity of the driving circuits used to control it.

The IGBT offers a mix of performance characteristics of both the power BJT and the power E-MOSFET. Like the BJT, the IGBT offers low on-state power loss and the ability to handle large currents and voltages. Like the power E-MOSFET, it is relatively easy to drive, being a voltage-controlled device rather than a current-controlled device. On the down-side, it is not as fast as the current generation of power E-MOSFETs and tends to be more costly than both the power BJT and power E-MOSFET. Consequently, the choice of which of these three devices should be used for a given power switching application will depend on the specifics of the design. For example, a medium to high power design that focuses on lowest cost may favor the BJT, a low to medium power application that requires very high switching speeds may be best solved with a power E-MOSFET, while an IGBT might be ideal for a very high power application utilizing low to medium speed clocking. This chapter will examine a number of power switching applications, and while they will all revolve around using the IGBT, please bear in mind that, depending on the specifics, power BJTs and E-MOSFETs might also be used.

The IGBT is available in two variants: the *punch through*, or *PT*; and the *non-punch through*, or *NPT*, versions. We shall look at both.

15.2 IGBT Internals

The IGBT is a multi-layer device. The cutaway shown in Figure 15.1 uses an N-channel, although P-channel is possible. This device has many features in common with the power E-MOSFET discussed in Chapter 12.

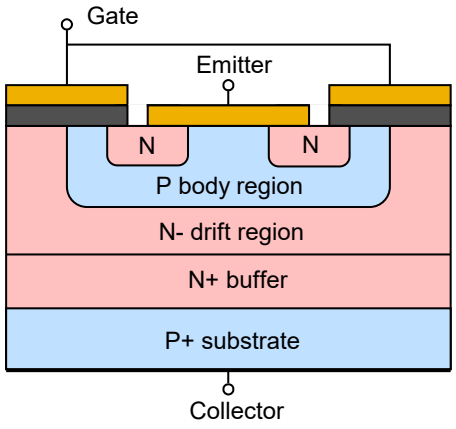


Figure 15.1
Cutaway view of IGBT.

To begin, note the labeling of the three terminals: gate, emitter and collector. The gate is isolated just as it is in MOSFETs. When a positive gate-emitter potential is applied, an N-type inversion layer will develop in the the P body region, allowing current to flow. The middle N region is split into two sections, the main N− drift region and the N+ buffer layer (the +/− signs indicate heavy and light doping, respectively). The device shown is the punch through (PT) variety. The non-punch through (NPT) omits the buffer layer. The functional differences between the two are that splitting the N material to include the buffer layer improves speed and lowers on-state voltage. Consequently, the PT version exhibits lower switching and conduction losses.

The upper three layers of Figure 15.1 form an E-MOSFET while the lower section (P body, N drift/buffer and P substrate) form a PNP transistor. Thus, we can make a simplified model of the IGBT using these other devices, as shown in Figure 15.2.

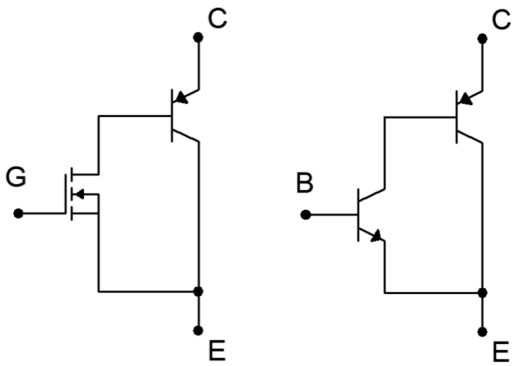


Figure 15.2
Simple model of an IGBT (left) compared to a Sziklai pair (right).

This simplified model is reminiscent of the NPN version of the Sziklai pair examined in Chapter 9. The input device has been replaced by an E-MOSFET. Therefore we expect the very small gate current and relatively simple drive requirements of the E-MOSFET with the power handling of a BJT.

The operational device curves are, unsurprisingly, also reminiscent of these two components. A set of collector curves is presented in Figure 15.3 using representative values for voltage and current.

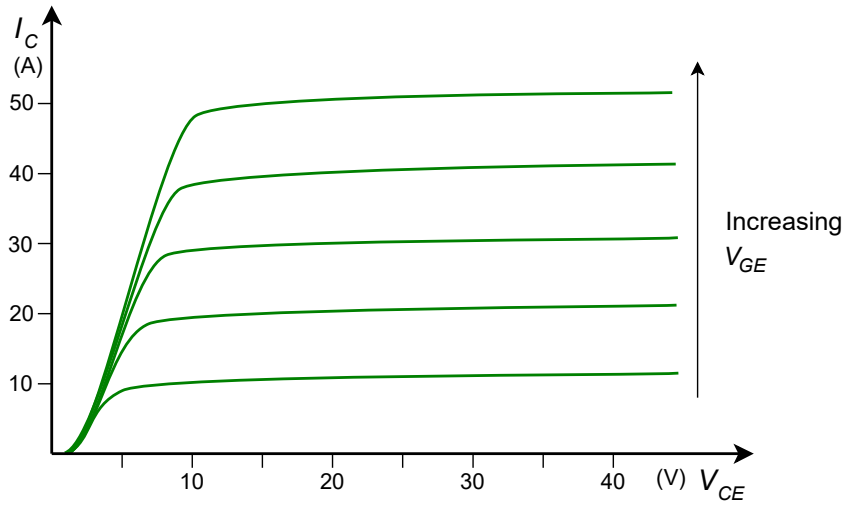


Figure 15.3
IGBT collector curve family.

This set of curves appears as a cross between the collector curve family seen in Chapter 4 and the drain curve family seen in Chapter 10, with one slight twist: the entire set of curves is displaced positively from the origin by about one volt. As with any E-MOSFET, channel current does not begin to flow until the gate threshold voltage is reached, here referred to as $V_{GE(th)}$. Increasing values of V_{GE} cause increased conduction and current flow. Finally, note that the curves do not begin to "flatten" until V_{CE} has reached several volts, unlike the saturation voltage of a single BJT which might be only tenths of a volt.

The forward current-voltage characteristic curve reflects the E-MOSFET portion of the model. This is shown in Figure 15.4.

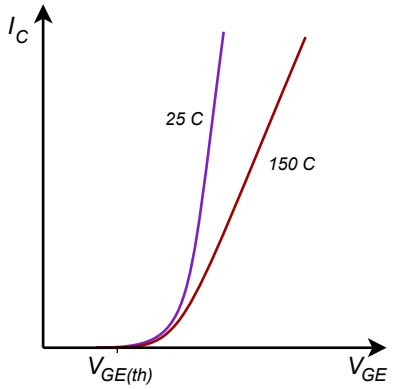


Figure 15.4
IGBT characteristic curve and variation with temperature.

Two traces are plotted for two different temperatures. The curves are essentially the same as the characteristic curve of the E-MOSFET, as seen in the introductory MOSFET chapter, Figure 12.11 (here the current axis has been expanded, making the traces appear steeper, in order to more clearly show the variation with temperature). Conduction begins at the threshold voltage, $V_{GE(th)}$, and then rises rapidly, following a square law trajectory. Once a sufficient current level is reached, the curve can be approximated as a straight line.

Of particular interest here is how the curve varies with temperature. As temperature increases (red trace), the slope decreases. Recalling that the slope of the current-voltage characteristic curve is the device transconductance, this means that the transconductance decreases with temperature. In other words, the IGBT exhibits a negative temperature coefficient of transconductance, just like power E-MOSFETs, and thus are also less inclined to suffer from current hogging and thermal runaway problems than BJTs.

Unfortunately, there isn't a single, standardized schematic symbol for the IGBT. Two versions are shown in Figure 15.5.

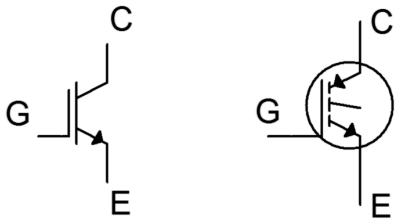


Figure 15.5
IGBT schematic symbols.

Both symbols attempt to reflect the dual E-MOSFET/BJT character of the device. The symbol on the left appears to have wider use currently, although there is also a third variation based on it where the gate connection is drawn toward the center rather than closer to the emitter.

15.3 IGBT Data Sheet Interpretation

A portion of the data sheet for the Fairchild/ON Semiconductor FGH50T65SQD IGBT is shown in Figure 15.6. This is a fourth generation IGBT featuring trench construction. It is rated for 650 volts and 50 amps. The device includes an anti-parallel diode. This is useful for bridge applications that drive inductive loads (recall that the current through an inductor cannot change instantaneously, thus, when devices are switched on/off in a bridge, the diode serves to create a path around the IGBT for this current).

Figure 15.6a
FGH50T65SQD data sheet.
Used with permission from SCILLC dba ON Semiconductor.

FAIRCHILD

April 2016

FGH50T65SQD

650 V, 50 A Field Stop Trench IGBT

Features

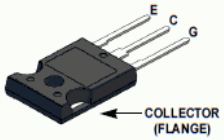
- Maximum Junction Temperature : $T_J = 175^oC$
- Positive Temperaure Co-efficient for Easy Parallel Operating
- High Current Capability
- Low Saturation Voltage: $V_{CE(sat)} = 1.6$ V(Typ.) @ $I_C = 50$ A
- 100% of the Parts Tested for $I_{LM}(1)$
- High Input Impedance
- Fast Switching
- Tighten Parameter Distribution
- RoHS Compliant

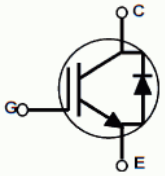
General Description

Using novel field stop IGBT technology, Fairchild's new series of field stop 4th generation IGBTs offer the optimum performance for solar inverter, UPS, welder, telecom, ESS and PFC applications where low conduction and switching losses are essential.

Applications

- Solar Inverter, UPS, Welder, Telecom, ESS, PFC





Absolute Maximum Ratings

Symbol	Description		FGH50T65SQD_F155	Unit
V_{CES}	Collector to Emitter Voltage		650	V
V_{GES}	Gate to Emitter Voltage		±20	V
	Transient Gate to Emitter Voltage		± 30	V
I_C	Collector Current	@ $T_C = 25^oC$	100	A
	Collector Current	@ $T_C = 100^oC$	50	A
$I_{LM\ (1)}$	Pulsed Collector Current	@ $T_C = 25^oC$	200	A
$I_{CM\ (2)}$	Pulsed Collector Current		200	A
I_F	Diode Forward Current	@ $T_C = 25^oC$	50	A
	Diode Forward Current	@ $T_C = 100^oC$	30	A
I_{FM}	Pulsed Diode Maximum Forward Current		200	A
P_D	Maximum Power Dissipation	@ $T_C = 25^oC$	268	W
	Maximum Power Dissipation	@ $T_C = 100^oC$	134	W
T_J	Operating Junction Temperature		-55 to +175	°C
T_{stg}	Storage Temperature Range		-55 to +175	°C
T_L	Maximum Lead Temp. for soldering Purposes, 1/8" from case for 5 seconds		300	°C

Notes:
1. $V_{CC} = 400$ V, $V_{GE} = 15$ V, $I_C = 200$ A, $R_G = 3\ \Omega$, Inductive Load
2. Repetitive rating: Pulse width limited by max. junction temperature

Figure 15.6b
FGH50T65SQD data sheet (cont).

Thermal Characteristics

Symbol	Parameter	FGH50T65SQD_F155	Unit
$R_{\theta JC}$(IGBT)	Thermal Resistance, Junction to Case, Max.	0.56	°C/W
$R_{\theta JC}$(Diode)	Thermal Resistance, Junction to Case, Max.	1.25	°C/W
$R_{\theta JA}$	Thermal Resistance, Junction to Ambient, Max.	40	°C/W

Package Marking and Ordering Information

Part Number	Top Mark	Package	Packing Method	Reel Size	Tape Width	Qty per Tube
FGH50T65SQD_F155	FGH50T65SQD	TO-247 G03	Tube	-	-	30

Electrical Characteristics of the IGBT T_C = 25°C unless otherwise noted

Symbol	Parameter	Test Conditions	Min.	Typ.	Max.	Unit
Off Characteristics						
BV_{CES}	Collector to Emitter Breakdown Voltage	V_{GE} = 0V, I_C = 1 mA	650	-	-	V
ΔBV_{CES} / ΔT_J	Temperature Coefficient of Breakdown Voltage	I_C = 1 mA, Reference to 25°C	-	0.6	-	V/°C
I_{CES}	Collector Cut-Off Current	V_{CE} = V_{CES}, V_{GE} = 0 V	-	-	250	µA
I_{GES}	G-E Leakage Current	V_{GE} = V_{GES}, V_{CE} = 0 V	-	-	±400	nA
On Characteristics						
$V_{GE(th)}$	G-E Threshold Voltage	I_C = 50 mA, V_{CE} = V_{GE}	2.6	4.5	6.4	V
$V_{CE(sat)}$	Collector to Emitter Saturation Voltage	I_C = 50 A, V_{GE} = 15 V, T_C = 25°C	-	1.6	2.1	V
		I_C = 50 A, V_{GE} = 15 V, T_C = 175°C	-	1.92	-	V
Dynamic Characteristics						
C_{ies}	Input Capacitance	V_{CE} = 30 V, V_{GE} = 0 V, f = 1MHz	-	3275	-	pF
C_{oes}	Output Capacitance		-	84	-	pF
C_{res}	Reverse Transfer Capacitance		-	12	-	pF
Switching Characteristics						
$t_{d(on)}$	Turn-On Delay Time	V_{CC} = 400 V, I_C = 12.5 A, R_G = 4.7 Ω, V_{GE} = 15 V, Inductive Load, T_C = 25°C	-	22	-	ns
t_r	Rise Time		-	8.7	-	ns
$t_{d(off)}$	Turn-Off Delay Time		-	105	-	ns
t_f	Fall Time		-	2.5	-	ns
E_{on}	Turn-On Switching Loss		-	180	-	uJ
E_{off}	Turn-Off Switching Loss		-	45	-	uJ
E_{ts}	Total Switching Loss		-	225	-	uJ
$t_{d(on)}$	Turn-On Delay Time	V_{CC} = 400 V, I_C = 25 A, R_G = 4.7 Ω, V_{GE} = 15 V, Inductive Load, T_C = 25°C	-	19	-	ns
t_r	Rise Time		-	13	-	ns
$t_{d(off)}$	Turn-Off Delay Time		-	93	-	ns
t_f	Fall Time		-	6.4	-	ns
E_{on}	Turn-On Switching Loss		-	410	-	uJ
E_{off}	Turn-Off Switching Loss		-	88	-	uJ
E_{ts}	Total Switching Loss		-	498	-	uJ

The main features are the 650 volt rating for V_{CE}, 100 amp continuous collector current and 268 watt dissipation at 25°C. The current and power ratings are essentially halved at the more practical temperature of 100°C. The threshold voltage, $V_{GE(th)}$, is specified as 4.5 volts with a ±1.9 volt spread. The saturation voltage typically is 1.6 volts at room temperature with 50 amps of collector current. This compares favorably to basic power BJTs. Like power MOSFETs, the input capacitance is relatively high at 3275 pF, so the same gate drive precautions must be followed. Finally, note the asymmetry in switching times. At room temperature and 12.5 amps of collector current, the turn-on delay plus rise time is specified as approximately 31 nanoseconds while the turn-off delay and fall time is nearly 110 nanoseconds. This relative slowing of the off-state transition is typical of IGBTs. Further, as both current and temperature increase (Figure 15.6c), these times increase by a few percent. By comparison, the FDMS86180 power E-MOSFET examined in Chapter 12 exhibited symmetrical values in the mid-30 nanosecond region at a drain current of 67 amps.

Figure 15.6c
FGH50T65SQD data sheet (cont).

Electrical Characteristics of the IGBT (Continued)

Symbol	Parameter	Test Conditions	Min.	Typ.	Max	Unit
$t_{d(on)}$	Turn-On Delay Time	V_{CC} = 400 V, I_C = 12.5 A, R_G = 4.7 Ω, V_{GE} = 15 V, Inductive Load, T_C = 175°C	-	20	-	ns
t_r	Rise Time		-	9.8	-	ns
$t_{d(off)}$	Turn-Off Delay Time		-	116	-	ns
t_f	Fall Time		-	3.5	-	ns
E_{on}	Turn-On Switching Loss		-	402	-	uJ
E_{off}	Turn-Off Switching Loss		-	110	-	uJ
E_{ts}	Total Switching Loss		-	512	-	uJ
$t_{d(on)}$	Turn-On Delay Time	V_{CC} = 400 V, I_C = 25 A, R_G = 4.7 Ω, V_{GE} = 15 V, Inductive Load, T_C = 175°C	-	18	-	ns
t_r	Rise Time		-	15	-	ns
$t_{d(off)}$	Turn-Off Delay Time		-	102	-	ns
t_f	Fall Time		-	8	-	ns
E_{on}	Turn-On Switching Loss		-	641	-	uJ
E_{off}	Turn-Off Switching Loss		-	203	-	uJ
E_{ts}	Total Switching Loss		-	844	-	uJ
Q_g	Total Gate Charge	V_{CE} = 400 V, I_C = 50 A, V_{GE} = 15 V	-	99	-	nC
Q_{ge}	Gate to Emitter Charge		-	17	-	nC
Q_{gc}	Gate to Collector Charge		-	23	-	nC

Electrical Characteristics of the Diode T_C = 25°C unless otherwise noted

Symbol	Parameter	Test Conditions		Min.	Typ.	Max	Unit
V_{FM}	Diode Forward Voltage	I_F = 30 A	T_C = 25°C	-	2.2	2.6	V
			T_C = 175°C	-	1.9	-	
E_{rec}	Reverse Recovery Energy	I_F =30 A, dI_F/dt = 200 A/µs	T_C = 175°C	-	40	-	uJ
t_{rr}	Diode Reverse Recovery Time		T_C = 25°C	-	31	-	ns
			T_C = 175°C	-	207	-	
Q_{rr}	Diode Reverse Recovery Charge		T_C = 25°C	-	48	-	nC
			T_C = 175°C	-	820	-	

Finally, consider the graphical data presented in Figure 15.6d.

Figure 15.6d
FGH50T65SQD data sheet (cont).

Typical Performance Characteristics

Figure 1. Typical Output Characteristics

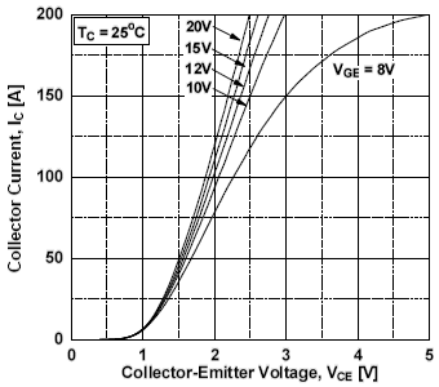


Figure 2. Typical Output Characteristics

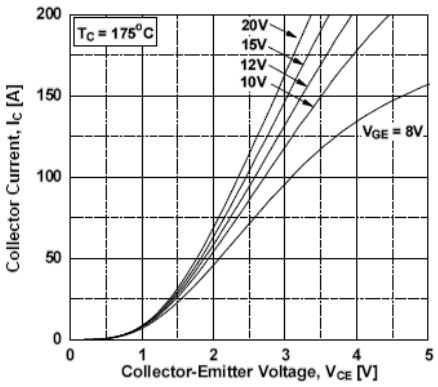


Figure 3. Typical Saturation Voltage Characteristics

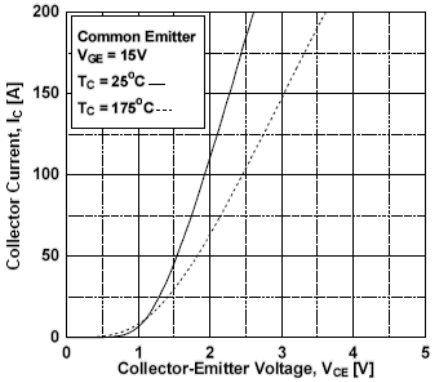


Figure 4. Saturation Voltage vs. Case Temperature at Variant Current Level

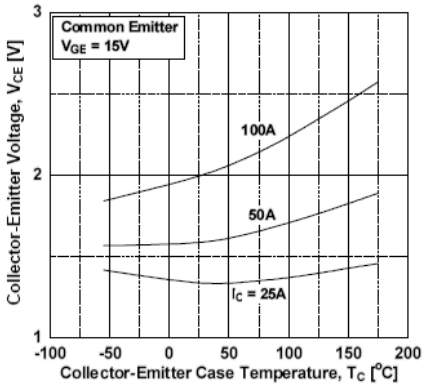


Figure 5. Saturation Voltage vs. V_{GE}

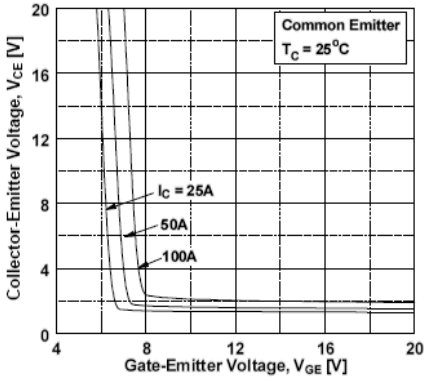
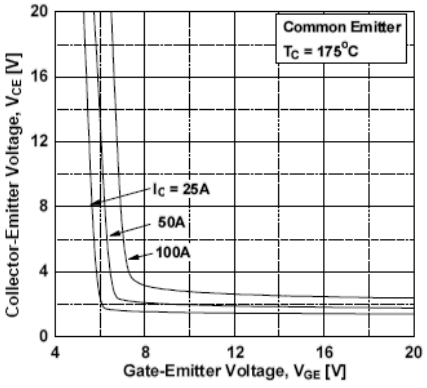


Figure 6. Saturation Voltage vs. V_{GE}



The top two graphs show the collector curves in the low voltage region. Compare these with Figure 15.3. Similarly, the middle-left graph shows the current-voltage characteristic with temperature variation. This reflects Figure 15.4. Finally, the bottom two graphs show the collector-emitter saturation voltage with respect to gate-emitter drive voltage for three different collector currents. Note that for the highest current at room temperature, saturation voltage is around 2 volts for a gate drive of at least 8 volts. This rises to about 3 volts at 175°C.

Computer Simulation

To highlight the performance of the IGBT, a simple series load switch is simulated. The circuit is shown in Figure 15.7.

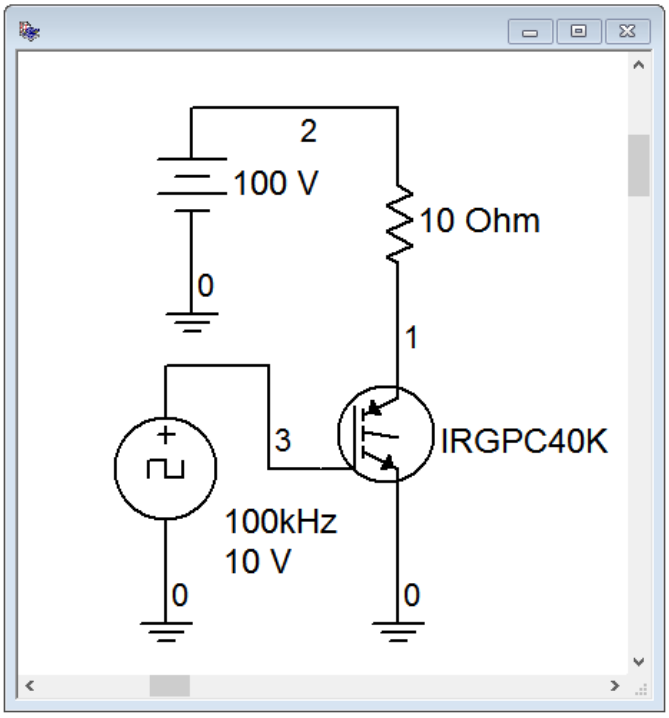


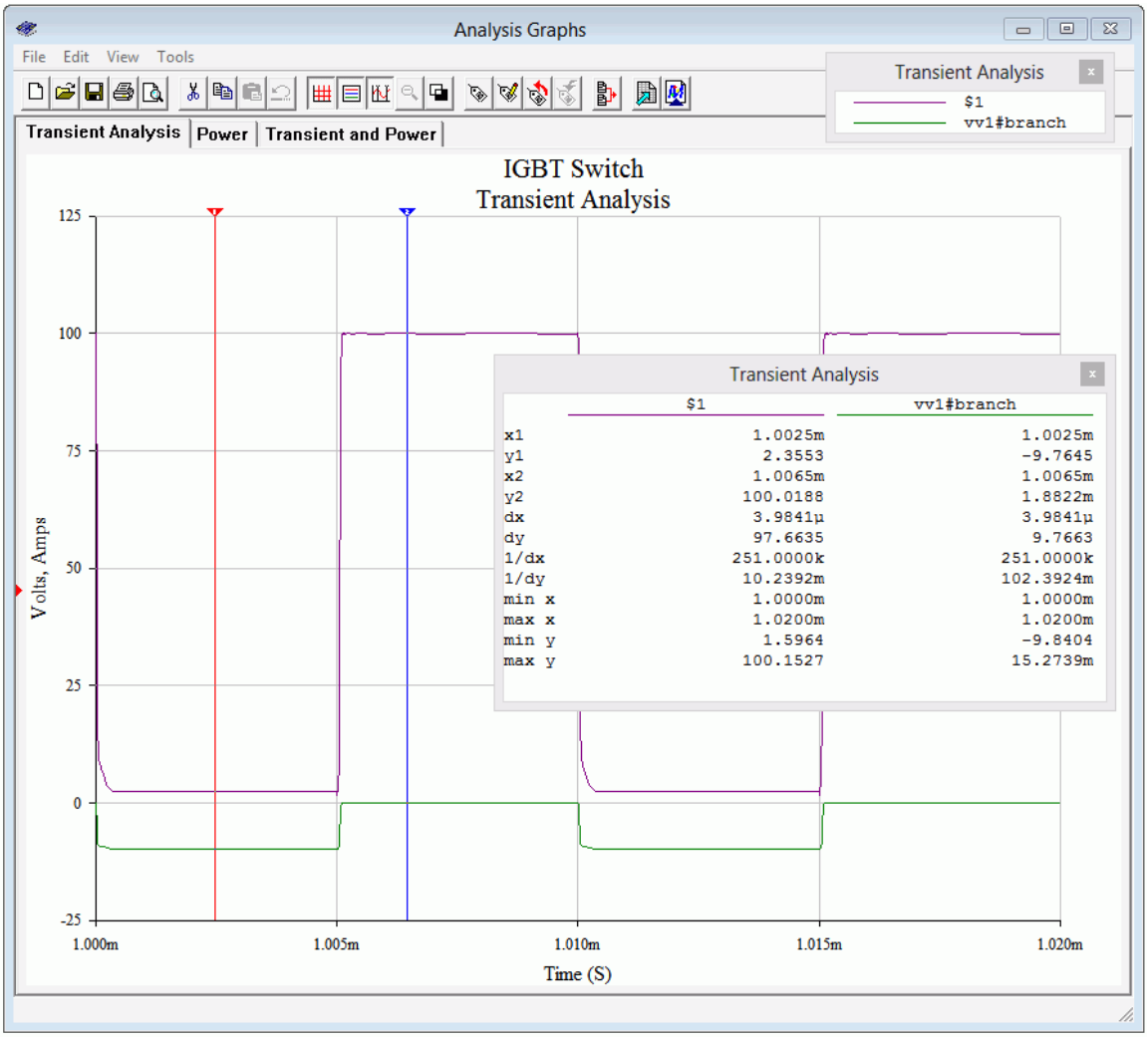
Figure 15.7
IGBT simulation schematic.

A 10 ohm load is switched from a 100 volt DC power supply via an International Rectifier IRGPC40K IGBT. The gate is driven from a 10 volt peak square wave running at 100 kHz.

Ideally, if the IGBT produced no losses, the full 100 volt source would drop across the load, producing 10 amp current pulses. According to the device data sheet, collector-emitter saturation voltage typically is 2.1 volts (3.2 volts maximum) with a 25 amp collector current. The total turn-on/rise time and turn-off/fall time values are 62 nanoseconds and 290 nanoseconds typically, at 25 amps and room temperature.

A transient analysis is run, with the results shown in Figure 15.8.

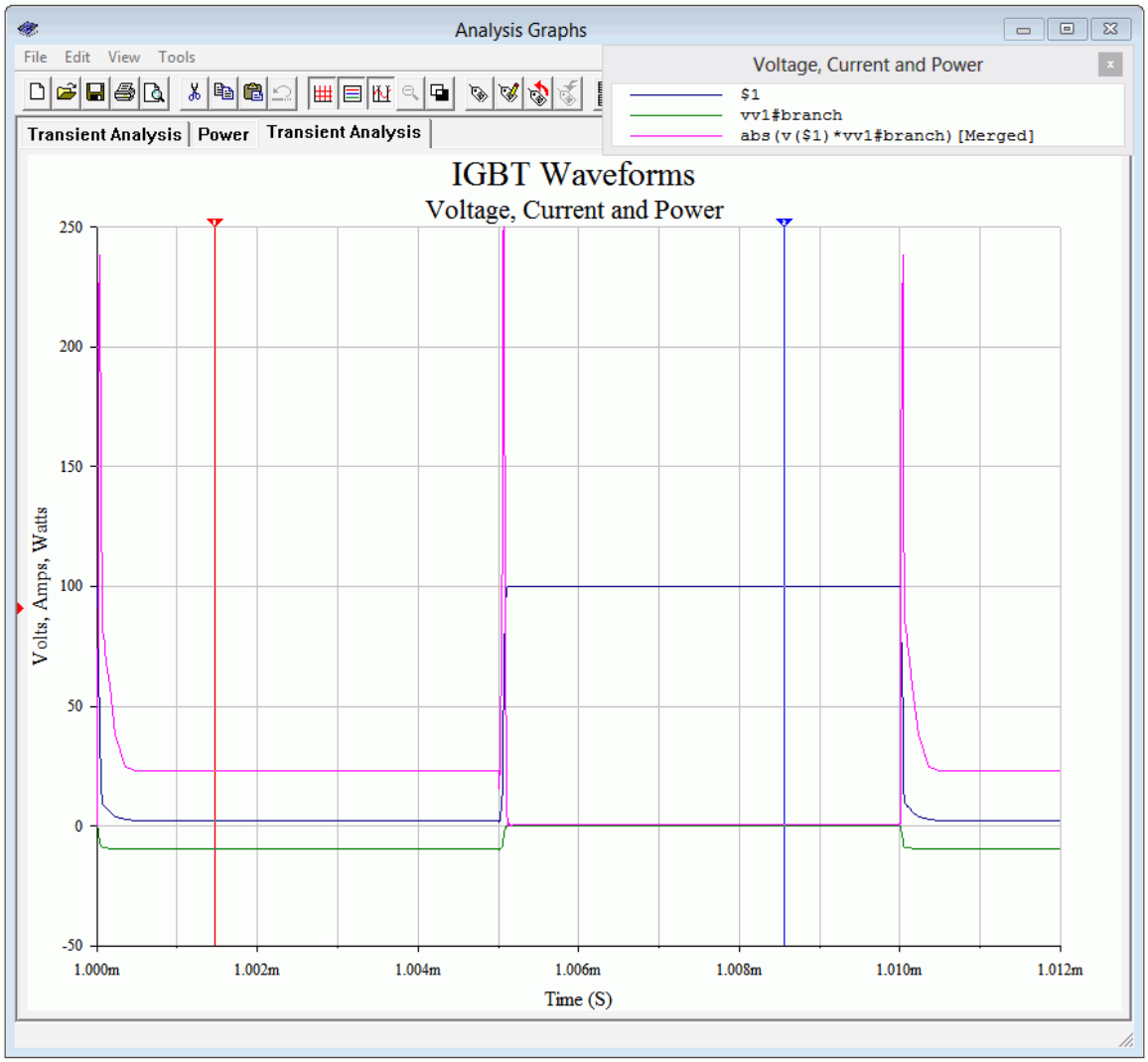
Figure 15.8
Transient analysis of simple IGBT switch.



The drain voltage (node 1) is shown as the purple trace and the drain current is shown in green (reversed in polarity to see it more clearly). The two cursors (red is #1, blue is #2) show the various levels. As expected, the drain voltage peaks at 100 volts in the off-state, at which time the current is nil. During the on-state, the drain voltage drops to about 2.36 volts, very close to the data sheet's typical saturation potential. This small drop reduces the voltage across the load to about 97.64 volts. This is verified by cursor #1 showing a current of 9.765 amps. Also, although it is not possible to determine the edge timings with great precision from this plot, the asymmetry between rise and fall is apparent, and the edges are generally consistent with the numeric values from the data sheet.

Figure 15.9 shows the result of multiplying the current and voltage waveforms. This new waveform represents the power dissipation of the IGBT.

Figure 15.9
Power waveform of simple IGBT switch.



The power trace is shown in fuchsia, or magenta, or screaming purple-pink, or whatever-you-want-to-call-it. The edge spikes dominate but the on-state power is apparent as well. Remember, during the on-state, the load is dissipating close to 1000 watts. The cursor output window is shown separately in Figure 15.10. The on-state power dissipation is approximately 23 watts which represents less than 2.5 percent of the load power. In contrast, the edge spikes are maxing out at around 250 watts. Of course, the time duration is very short, being only a few percent of the cycle period, but it cannot be ignored.

Figure 15.10
Numeric values at waveform cursors for the IGBT switch.

Voltage, Current and Power

	$1	vv1#branch	abs(v($1)*vv1#branch) [Mer
x1	1.0015m	1.0015m	1.0015m
y1	2.3509	-9.7649	22.9559
x2	1.0086m	1.0086m	1.0086m
y2	100.0292	2.9179m	417.2293m
dx	7.0916µ	7.0916µ	7.0916µ
dy	97.6783	9.7678	-22.5387
1/dx	141.0112k	141.0112k	141.0112k
1/dy	10.2377m	102.3769m	-44.3681m
min x	1.0000m	1.0000m	1.0000m
max x	1.0200m	1.0200m	1.0200m
min y	1.5964	-9.8404	28.1911m
max y	100.1527	15.2739m	249.7386

15.4 IGBT Applications

IGBTs lend themselves to a variety of high power switching applications. In this section, we shall look at four of them. Bear in mind that power BJTs and power E-MOSFETs might also be used for these applications, depending on the specifics of the design. In general, power E-MOSFETs will be preferred when using high switching frequencies at medium to low powers and voltages, while IGBTs are favored at higher voltages, currents and powers.

Induction Heating

Although thermal conduction is the first method commonly thought of when it comes to heating something, magnetic induction may also be used. Magnetic induction creates heat through the Joule effect and can be used for large scale industrial processes, such as creating metal alloys via an induction furnace, to much smaller scale consumer applications, such as an inductive cooktop. Induction heating is efficient because the vessel itself is heated directly and less heat is lost to the immediate environment. Also, control of heating can be very precise. The basic idea is to create a rapidly changing magnetic field placed next to the container to be heated. If this vessel is ferromagnetic, eddy currents will be induced in the vessel, creating heat. Thus, if we were to place a cast iron pot within the field, the pot itself would heat up as eddy currents are induced within it, thus heating the pot's contents. There are no open flames or surface heating elements involved. The only downside to this process is that the vessel must be made of ferromagnetic material. For a cooktop, that means that pots and pans must be made of iron or certain steel alloys. An aluminum sauté pan or ceramic container will not work with this system.

As an example, let's consider an inductive cooktop. A sophisticated design could feature a full- or half-bridge arrangement of IGBTs, but for illustrative purposes we'll focus on a simple single-ended system using just one IGBT.

The system features four main components: the rectifier/EMI filter, the control/drive circuit, the IGBT switch and the *LC* resonant tank sub-circuit that generates the field. This is illustrated in Figure 15.11.

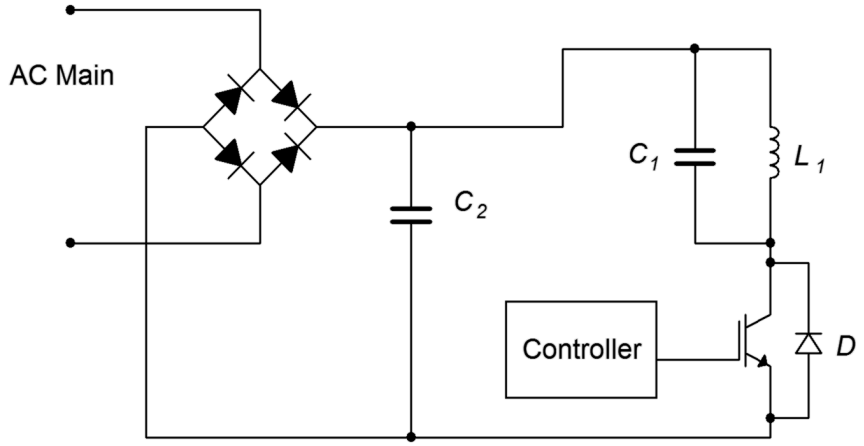


Figure 15.11
Simple inductive heater.

The rectifier produces full-wave pulsating DC, and the associated capacitor, C_2, is used to help minimize EMI (electromagnetic interference) and also provide a return path for the tank. The control circuit produces a variable duty cycle pulse train to drive the gate of the IGBT. The greater the duty cycle, the longer the on-state of the IGBT, and ultimately, the greater the heating. Between the IGBT and the rectified power signal is a parallel resonant tank circuit comprised of C_1 and L_1. The inductor is comprised of a series of loops of large gauge wiring or copper tubing embedded in the cook surface, typically under a glass or ceramic top. The resonant frequency of the tank is tuned to the frequency of the controlling pulse train. This will maximize the tank current and thus produce a more powerful magnetic field. The switching frequency is usually placed just above the range of human hearing to avoid audible microphonics.[46] Values in the range of 20 kHz to 30 kHz are typical, and the base frequency may change as the heat demand changes. For example, to minimize switching losses, the controlling frequency might start at 30 kHz for modest heating and decrease to 20 kHz for maximum heating.

From the cook's perspective there is no change between using the inductive cooktop and an ordinary electric cooktop using resistive heating elements: The cook places the pot or pan on the surface, under which lies the coil. A heat level control knob is provided for them to adjust the heat intensity. To their advantage, when they remove the pot or pan, the cooking surface itself will not be as hot as an ordinary cooktop.

46 Granted, it may still fall within the hearing range of your dog, so don't be surprised if your border collie prefers a gas cooktop to make a balsamic reduction.

From the designer's perspective, the heat control knob simply changes the duty cycle of the controlling pulse train (and optionally, its frequency, as mentioned previously). Other refinements might include sensing whether or not a vessel is on the cooktop and throttling back control if nothing is detected. Finally, an even simpler system could switch the IGBT on and off at a much slower rate (think in terms of seconds) to greatly reduce switching losses but this runs the risk of heat cycling if the pots and pans used are of very light gauge construction (i.e., their thermal time constant will be faster).

DC-to-AC Inversion

There are many instances where we wish to derive an AC voltage from an existing DC voltage. Examples include an uninterruptible power supply (UPS) that would draw current from a battery and deliver standard AC power when there is a disruption in the power grid, and the need to operate electronic devices designed for the home in a remote location. This process is known as *DC-to-AC inversion*.

The simplest method to create AC from DC is to just "chop" the DC at the desired line frequency and then scale it, that is, feed the DC into a simple IGBT switch which will produce a square wave and then feed the square wave into a transformer to arrive at the desired voltage. The obvious problem with this technique is that the AC signal will not be a nice, smooth sine wave, but rather, a distorted square wave. Unfortunately, for many electronic components this will present a challenge to their power supply circuits. A possible refinement involves making a step-wise approximation of a sine wave but this is still not ideal.

A more accurate scheme involves synthesizing a sine wave via PWM. We still chop the DC, but now the action is performed at a higher frequency and with a varying duty cycle such that, when the output is averaged, we arrive at a sine wave. A block diagram of this scheme is shown in Figure 15.12.

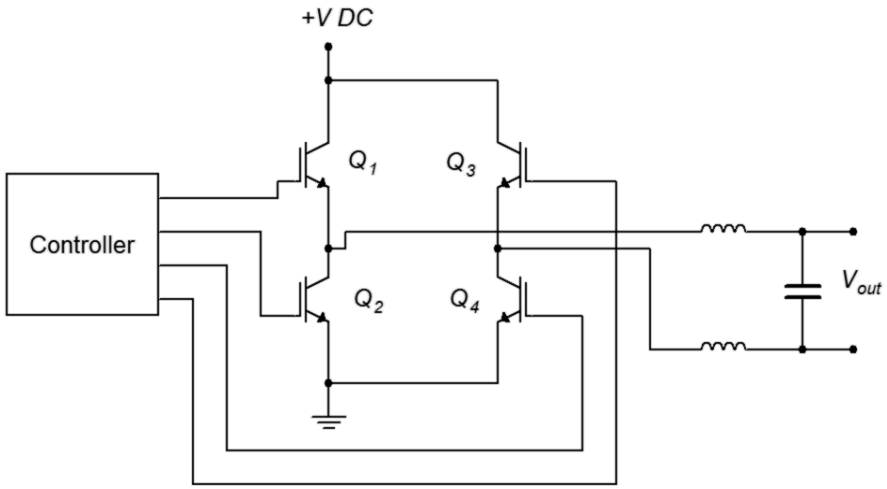


Figure 15.12
DC-to-AC inverter (anti-parallel diodes not shown).

The controller generates a set of PWM signals to drive the four IGBTs configured as a full-bridge. The bridge output is then fed to a balanced *LC* filter that removes the high frequency PWM components, leaving a smoothed sine wave. This signal can be used as is, or fed into a step-up transformer if the desired AC voltage is greater than the starting DC voltage. An example of this would be the need to supply a device designed to run on 120 VAC from the nominal 12 VDC system found in a car.

There is a side item worthy of mention here. Some devices derive timing signals from the line frequency (a classic example is an electronic alarm clock/radio). This is possible because the power generation utility monitors this frequency with great accuracy. If the controller shown in Figure 15.12 does not produce an accurate base frequency, then that clock/radio will not tell time accurately.

Motor Control

IGBTs can be used to control the speed of electric motors. The configuration of the control circuit will depend on the kind of motor being controlled. In simplest terms, the speed of a DC motor is controlled by the voltage applied to it: the higher the voltage, the higher the speed. In contrast, the speed of an AC motor depends on the frequency of the applied source (they are proportional).

Controlling a DC motor is a straightforward situation. If all we need to do is start and stop the motor, the IGBT can be inserted in series with the motor and used as a switch to open and close the circuit. Being solid-state, the IGBT has numerous advantages over a mechanical switch or relay including long-term reliability and simplicity of the control circuit. To alter the speed, the IGBT can be controlled via PWM. This is illustrated in Figure 15.13.

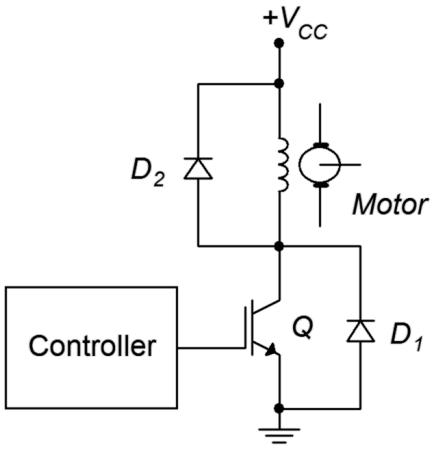


Figure 15.13
DC motor control via IGBT.

D_1 and D_2 are *flyback* or *snubber* diodes used for protection from inductive current spikes caused by the motor's current being switched on and off. As noted earlier, some IGBTs are co-packaged with an anti-parallel diode (D_1).

To vary the motor's speed, the controller produces a PWM drive signal. The smaller the duty cycle of this pulse train, the lower the average applied voltage to the motor, and therefore the slower its speed. The base frequency of the PWM signal does not have to be particularly high in this scenario; a few hundred hertz may prove sufficient. This will help minimize switching losses.

For an AC motor the situation is a little more complex. One approach is to use the PWM technique explained under the DC-to-AC converter section. The difference is that the power source would not be DC, but rather, AC. Consequently, we would need to transform the AC power source into a more usable signal and then apply the circuit depicted in Figure 15.12 to power the motor. The controller itself will need to be considerably more sophisticated. In Figure 15.12, the duty cycle is continuously changed such that the "area under the curve" approximates a sine wave. Eventually, the pattern will repeat itself for subsequent cycles of the sine wave. In other words, the rate at which the pattern repeats is the sine wave's period. In the DC-to-AC inverter application, this rate never changes because we need a constant output frequency (e.g., 60 Hz). In the AC motor control application, such is not the case. This repetition rate needs to be adjustable because that's what controls the motor's speed. One way to do this is to simply increase the base frequency of the PWM pulse train. This method is simple and direct but has the disadvantage of creating more transient edges per unit time and therefore tends to increase switching losses. An alternate approach is to keep the base frequency constant and instead alter the duty cycle pattern. This helps minimize the switching loss issues but has the disadvantage of requiring a more complex control circuit and possibly producing a lower quality sine wave at higher output frequencies.

DC-to-DC Conversion

Our final application is DC-to-DC up-conversion, that is, producing a new DC voltage that is higher than the original source and also capable of high output current (voltage doublers and triplers can be made from diode/capacitor lattices but they are not designed to deliver high, continuous currents). Applications requiring up-conversion include photovoltaic systems (i.e, combining the outputs of several solar panels and tying them into the power system) and high output car audio systems. Specifically, the nominal 12 volt automotive power system is insufficient to supply an amplifier intended to deliver hundreds or even thousands of watts to a subwoofer. The 12 volt source will need to be increased, perhaps by a factor of ten, to achieve the desired output levels.[47]

47 In the process, the current demand will be increased greatly as well, perhaps beyond the capabilities of the vehicle's alternator (which will also require upgrading), but these are the prices one must pay if one desires very high sound pressure levels in what is arguably the worst acoustical environment in which to listen to music. Of course, we should also admit that the act of critically listening to and enjoying music may not be the point of such an exercise.

As we have already seen how DC can be translated into AC via an inverter, it is possible to simply rectify and filter the new AC, producing a higher DC level. This scheme is illustrated in Figure 15.14.

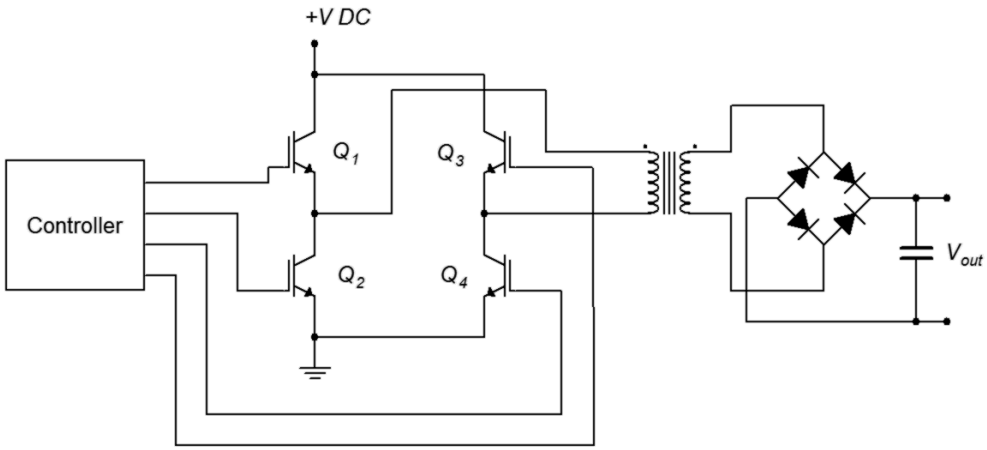


Figure 15.14
DC-to-DC conversion via transformer (anti-parallel diodes not shown).

The transformer pictured here will need to be a step-up variety in order to achieve the desired output voltage level. Note that the AC generation side does not have to produce a particularly nice sine wave nor does it have to be at the usual line frequency. In fact, increasing the frequency will likely result in reduced sizes for the transformer and filter capacitors.

A completely different approach is to use a switching regulator. Switching regulators use a feedback control system to generate a very stable output voltage by comparing it to a reference voltage. They can be configured in step-down, step-up or polarity inversion forms.[48] In this case, we can use the step-up, or *boost*, form. An example is shown in Figure 15.15.

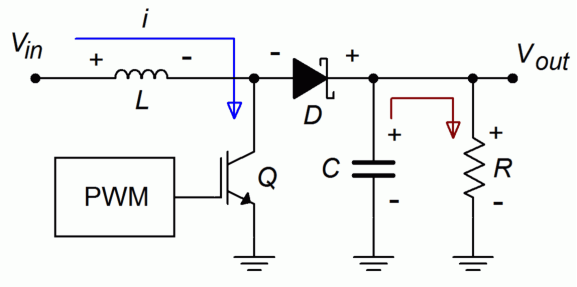


Figure 15.15a
Step-up switching regulator, on-state.

As with the other applications presented so far, the IGBT is being used as a switch. Figure 15.15a illustrates the on-state of the IGBT. During this phase, current is drawn through the inductor, *L*, storing energy in the associated magnetic field. The reverse-biased Schottky diode, *D*, isolates this section from the output section, where *C* is delivering the load voltage and current.

48 For details on switching regulators, see Fiore, J, *Operational Amplifiers and Linear Integrated Circuits: Theory and Application*, another free OER text.

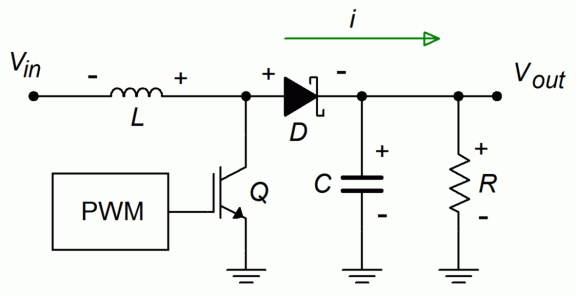


Figure 15.15b
Step-up switching regulator, off-state.

The off-state is depicted in Figure 15.15b. During this portion, the inductor discharges and appears as a source. As it is in series with the input voltage, the voltage at the output will equal the inductor's voltage plus the input voltage. Also, during this phase the capacitor is recharged, ready for the next on-state of the IGBT. Note that a Schottky diode is used here because it exhibits fast switching times and a low forward voltage drop. Switching frequencies for these circuits tend to be high (100 kHz and up are common) as that minimizes the sizes of the inductor and capacitor.

Summary

The Insulated Gate Bipolar Transistor, or IGBT, can be thought of as a combination of a power BJT and a power E-MOSFET. As such, it combines the low on-state conduction losses of the BJT with the relatively easy drive requirements of the E-MOSFET. The IGBT is available in two variants; the PT, or punch through, and the NPT, or non-punch through types. The PT type includes an N+ buffer layer in its construction and this endows the device with faster switching speed and lower on-state losses.

The IGBT's characteristic curves tend to echo that of the E-MOSFET. Conduction does not begin until the gate-emitter voltage exceeds a threshold voltage, $V_{GE(th)}$. From there, the current-voltage characteristic follows a square-law trajectory, and at sufficiently high current levels it can be approximated as a straight line. The IGBT exhibits a negative temperature coefficient of transconductance, like the MOSFET, making it less prone to thermal runaway and current hogging issues. A family of collector curves (i.e., V_{CE} vs. I_C) shares attributes with BJT collector curves and MOSFET drain curves. The curves echo the same overall shape, starting with a section where current rises rapidly compared to voltage, and then leveling out into a constant current region. The initial region of rapid change is somewhat drawn out as it is in the MOSFET. Also, the entire set of curves is displaced positively by about a volt, rather than current increasing immediately from the origin.

In general, the IGBT offers higher voltage, current and power capability than the power E-MOSFET although it lags behind in switching speed. Further, switching times for the on- and off-state are asymmetrical. Compared to the power BJT, the

IGBT tends to be more expensive. Consequently, power E-MOSFETs tend to be favored at low and moderate power levels when high switching speeds are needed and BJTs tend to be preferred when cost is a major component in more modest designs. As such, IGBTs find use across a range of applications including power inverters, uninterruptible power supplies, induction heaters, solar power systems, motor controllers and so forth.

Review Questions

1. What are the advantages and disadvantages of IGBTs compared to power BJTs?
2. What are the advantages and disadvantages of IGBTs compared to power E-MOSFETs?
3. What are the differences between NPT and PT IGBTs?
4. Compare the simple IGBT model to that of a Sziklai pair.
5. Explain the basic operation of an induction heater and how an IGBT is used to control the generation of heat.
6. Explain how pulse width modulation might be used to control the speed of a DC motor via an IGBT.
7. Explain how an IGBT can be used to translate a DC power source into an AC power source.

Problems

Analysis Problems

1. From the FGH50T65SQD data sheet, determine the collector-emitter saturation voltage at 25°C for V_{GE} = 7 volts and I_C = 50 amps.

2. From the FGH50T65SQD data sheet, determine the change in collector-emitter saturation voltage from 25°C to 175°C for V_{GE} = 8 volts and I_C = 100 amps.

3. From the FGH50T65SQD data sheet, determine the rise time for a 50 amp collector current at 25°C for V_{GE} = 15 volts.

4. From the IRGPC40K data sheet, determine the combined turn-on and rise time at 150°C.

5. From the IRGPC40K data sheet, determine the combined turn-off and fall time at 150°C.

6. Repeat the simple switch simulation from this chapter using the FGH50T65SQD IGBT and compare the resulting voltage, current and power waveforms.

7. Repeat the simple switch simulation from this chapter using the FDMS86180 power E-MOSFET from Chapter 12 and compare the resulting voltage, current and power waveforms.

16 Decibels and Bode Plots

16.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Convert between ordinary and decibel based power and voltage gains.
- Utilize decibel-based voltage and power measurements during circuit analysis.
- Define and graph a general Bode plot.
- Detail the differences between lead and lag networks, and graph Bode plots for each.
- Combine the effects of several lead and lag networks together in order to determine a system Bode plot.

16.1 Introduction

This chapter introduces two related items; the decibel and Bode plots. The decibel measurement scheme is in wide use, particularly in the fields of audio and communications. We will be examining its advantages over the ordinary system of measurement used up to now and how to convert values of one form into the other. One of the more important parameters of a circuit is its frequency response, that is, the way in which the circuit responds to input signals over a range of frequencies. While we have investigated this to some extent in prior work, in this chapter we shall take it to its logical conclusion, namely the Bode plot. A Bode plot is, in fact, a pair of plots; one of relative signal magnitude or gain with respect to frequency and a second detailing the phase response with respect to frequency. Bode plots are of particular importance in the study of circuits such as amplifiers and filters, as well as in systems that make use of negative feedback. The gain magnitude plot makes use of a decibel scale and thus it makes sense to begin our study looking at the decibel system: specifically how it is defined and its practical use.

16.2 The Decibel

Most people are familiar with the term “decibel” in reference to sound pressure. It’s not uncommon to hear someone say something such as “It was 110 decibels at the concert last night and my ears are still ringing.” This popular use is somewhat inaccurate, but does show that decibels indicate some sort of quantity or relative level; in this case, sound pressure level.

Decibel Representation of Power and Voltage Gains

In its simplest form, the decibel is used to measure system gain, such as power or voltage gain, where gain is simply the ratio of an output signal to an input signal. For an amplifying circuit, the gain would be greater than one, but for purely passive systems it will likely be fractional (i.e., the output quantity is smaller than the input

quantity). For example, a simple voltage divider might be said to have a “gain” of 0.2, or some such, meaning that the output signal is only 20% of the input signal. Unlike the ordinary gain measurements, the decibel form is logarithmic. Because of this, it can be very useful for showing ratios of change, as well as absolute change. The base unit is the Bel, named after Alexander Graham Bell, the noted American scientist and inventor. To convert an ordinary gain to its Bel counterpart, just take the common log (base 10) of the gain. In equation form:

$$\textit{Bel gain} = \log_{10}(\textit{ ordinary gain }) \qquad (16.1)$$

Note that on most hand calculators common log is denoted as “log” while the natural log is given as “ln”. Unfortunately, some programming languages use “log” to indicate natural log and “log10” for common log. More than one student has been bitten by this bug, so be forewarned! As an example, if an amplifier circuit produces an output power of 200 milliwatts for an input of 10 milliwatts, we would normally say that it has a power gain of:

$$G = \frac{P_{out}}{P_{in}}$$
$$G = \frac{200\,\text{mW}}{10\,\text{mW}}$$
$$G = 20$$

For the Bel version, just take the log of this result.

$$G' = \log_{10} G$$
$$G' = \log_{10} 20$$
$$G' = 1.301$$

The Bel gain is 1.3 Bels. The term “Bels” is not a unit in the strict sense of the word (as in “watts”), but is simply used to indicate that this is not an ordinary gain. In contrast, ordinary power and voltage gains are sometimes given units of W/W and V/V to distinguish them from Bel gains. Also, note that the symbol for Bel power gain is G' and not G. All Bel gains are denoted with the following prime (’) notation to avoid confusion. Because Bels tend to be rather large, we typically use one-tenth of a Bel as the norm. The result is the decibel (one-tenth Bel). To convert to decibels, simply multiply the number of Bels by 10. Our gain of 1.3 Bels is equivalent to 13 decibels. The units are commonly shortened to dB. Consequently, we may say:

$$G' = 10 \log_{10} G \qquad (16.2)$$

Where the result is in dB.

At this point, you may be wondering what the big advantage of the decibel system is. To answer this, recall a few log identities. Normal multiplication becomes

addition in the log system, and division becomes subtraction. Likewise, powers and roots become multiplication and division. Because of this, two important things show up. First, ratios of change become constant offsets in the decibel system, and second, the entire range of values diminishes in size. The result is that a very wide range of gains may be represented within a fairly small scope of values, and the corresponding calculations can become quicker.

There are a couple of dB values that are useful to remember, and are illustrated in Figure 16.1. With the aid of your calculator, it is very easy to show the following:

Factor	dB Value using $G' = 10 \log_{10} G$
1	0 dB
2	3.01 dB
4	6.02 dB
8	9.03 dB
10	10 dB

Figure 16.1
Positive dB factors.

We can also look at fractional factors (i.e., losses instead of gains, Figure 16.2):

Factor	dB Value
0.5	−3.01 dB
0.25	−6.02 dB
0.125	−9.03 dB
0.1	−10 dB

Figure 16.2
Negative dB factors.

If you look carefully, you will notice that a doubling is represented by an increase of approximately 3 dB. A factor of 4 is in essence, two doublings. Therefore, it is equivalent to 3 dB + 3 dB, or 6 dB. Remember, because we are using logs, multiplication turns into simple addition. In a similar manner, a halving is represented by approximately −3 dB. The negative sign indicates a reduction. To simplify things a bit, think of factors of 2 as ±3 dB, the sign indicating whether you are increasing (multiplying), or decreasing (dividing). As you can see, factors of 10 work out to a very convenient 10 dB. By remembering these two factors, you can often estimate a dB conversion without the use of your calculator. For instance, we could rework our initial conversion problem as follows:

- The amplifier has a gain of 20.
- 20 can be written as 2 times 10.
- The factor of 2 is 3 dB, the factor of 10 is 10 dB.
- The answer must be 3 dB + 10 dB, or 13 dB.

Time for a few examples.

Example 16.1

An amplifier has a power gain of 800. What is the decibel power gain?

$$G' = 10 \log_{10} G$$
$$G' = 10 \log_{10} 800$$
$$G' = 10 \times 2.903$$
$$G' = 29.03\,\text{dB}$$

We could also use our estimation technique:

- G = 800 = $8 \cdot 10^2$
- 8 is equivalent to 3 factors of 2, or 2·2·2, and can be expressed as 3 dB + 3 dB + 3 dB, which is, of course, 9 dB
- 10^2 is equivalent to 2 factors of 10, or 10 dB + 10 dB = 20 dB. Alternately, the power of 2 literally represents 2 Bels, and thus 20 dB.
- The result is 9 dB + 20 dB, or 29 dB

Note that if the leading digit is not a power of 2, the estimation will not be as precise. For example, if the gain is 850, you know that the decibel gain is just a bit over 29 dB. You also know that it must be less than 30 dB ($1000=10^3$ which is 3 factors of 10, or 30 dB.) As you can see, by using the dB form, you tend to concentrate on the magnitude of gain, and not so much on trailing digits.

Example 16.2

An attenuator reduces signal power by a factor of 10,000. What is this loss expressed in dB?

$$G' = 10 \log_{10} \frac{1}{10{,}000}$$
$$G' = 10 \times (-4)$$
$$G' = -40\,\text{dB}$$

By using the approximation, we can say,

$$\frac{1}{10{,}000} = 10^{-4}$$

The negative exponent tells us we have a loss (negative dB value), and 4 factors of 10 (i.e., 4 Bels).

$$G' = -10\,\text{dB} - 10\,\text{dB} - 10\,\text{dB} - 10\,\text{dB}$$
$$G' = -40\,\text{dB}$$

Remember, if an increase in signal is produced, the result will be a positive dB value. A decrease in signal will always result in a negative dB value. A signal that is unchanged indicates a gain of unity, or 0 dB.

To convert from dB to ordinary form, just invert the steps; that is, divide by ten and then take the antilog.

$$G = \log_{10}^{-1} \frac{G'}{10} \tag{16.3}$$

On most hand calculators, base 10 antilog is denoted as 10^x. In most computer languages, you just raise 10 to the appropriate power, as in G = 10.0**(Gprime / 10.0) (Python), or use an exponent function, as in pow(10.0, Gprime / 10.0) (C).

Example 16.3

An amplifier has a power gain of 23 dB. If the input is 1 mW, what is the output?

In order find the output power, we need to find the ordinary power gain, *G*.

$$G = \log_{10}^{-1} \frac{G'}{10}$$
$$G = \log_{10}^{-1} \frac{23}{10}$$
$$G = 199.5$$

Therefore, P_{out} =199.5 · 1 mW, or 199.5 mW

You could also use the approximation technique in reverse. To do this, break up the dB gain in 10 dB and 3 dB chunks:

$$23\,\text{dB} = 3\,\text{dB} + 10\,\text{dB} + 10\,\text{dB}$$

Now replace each chunk with the appropriate factor, and multiply them together (remember, when going from log to ordinary form, addition turns into multiplication.)

$3\text{dB} = 2\text{X}$, $10\text{dB} = 10\text{X}$, so,

$$G = 2 \times 10 \times 10$$
$$G = 200$$

While the approximation technique appears to be slower than the calculator, practice will show otherwise. Being able to quickly estimate dB values can prove to be a very handy skill in the electronics field. This is particularly true in larger, multi-stage designs.

Example 16.4

A three-stage amplifier has gains of 10 dB, 16 dB, and 14 dB per section. What is the total dB gain?

Because dB gains are a log form, just add the individual stage gains to arrive at the system gain.

$$G'_{total} = G'_1 + G'_2 + G'_3$$
$$G'_{total} = 10\,\text{dB} + 16\,\text{dB} + 14\,\text{dB}$$
$$G'_{total} = 40\,\text{dB}$$

As you may have noticed, all of the examples up to this point have used power gain and not voltage gain. You may be tempted to use the same equations for voltage gain. In a word, **don't**. If you think back for a moment, you will recall that power varies as the square of voltage. In other words, a doubling of voltage will produce a quadrupling of power. If you were to use the same dB conversions, a doubling of voltage would be 3 dB, yet, because the power has quadrupled, this would indicate a 6 dB rise. Consequently, voltage gain (and current gain as well) are treated in a slightly different fashion. We would rather have our doubling of voltage work out to 6 dB, so that it matches the power calculation. The correction factor is very simple. Because power varies as the second power of voltage, the dB form should be twice as large for voltage (remember, exponentiation turns into multiplication when using logs). Applying this factor to equation 16.2 yields:

$$A'_v = 20 \log_{10} A_v \tag{16.4}$$

Be careful though, the Bel voltage gain only equals the Bel power gain if the input and output impedances of the system are matched (you may recall from your other work that it is quite possible to design a circuit with vastly different voltage and power gains). If we were to recalculate our earlier table of common factors, we would find that a doubling of voltage gain is equivalent to a 6 dB rise, and a ten fold increase is equivalent to a 20 dB rise, twice the number of decibels of their power gain counterparts.

Note that current gain may be treated in the same manner as voltage gain (although this is less commonly done in practice).

Example 16.5

A circuit has an output signal of 2 V for an input of 50 mV. What is A'_v?
First find the ordinary gain.

$$A_v = \frac{2}{0.05} = 40$$

Now convert to dB form.

$$A'_v = 20 \log_{10} 40$$
$$A'_v = 20 \times 1.602$$
$$A'_v = 32.04 \text{ dB}$$

The approximation technique yields 40=2·2·10, or 6 dB + 6 dB + 20 dB, or 32 dB.

To convert A'_v to A, reverse the process.

$$A_v = \log_{10}^{-1} \frac{A'_v}{20} \tag{16.5}$$

Example 16.6

An amplifier has a gain of 26 dB. If the input signal is 10 mV, what is the output?

$$A_v = \log_{10}^{-1} \frac{A'_v}{20}$$
$$A_v = \log_{10}^{-1} \frac{26}{20}$$
$$A_v = 19.95$$
$$V_{out} = A_v V_{in}$$
$$V_{out} = 19.95 \times 10 \text{ mV}$$

The final point to note in this section is that, as in the case of power gain, a negative decibel value indicates a loss. Therefore, a 2:1 voltage divider would have a gain of −6 dB.

Signal Representation in dBW and dBV

As you can see from the preceding section, it is possible to spend considerable time converting between decibel gains and ordinary voltages and powers. Because the decibel form does offer advantages for gain measurement, it would make sense to use a decibel form for power and voltage levels as well. This is a relatively straightforward process. There is no reason why we can't express a power or voltage in a logarithmic form. Because a dB value just indicates a ratio, all we need to do is decide on a reference (i.e., a comparative base for the ratio). For power measurements, a likely choice would be 1 watt. In other words, we can describe a power as being a certain number of dB above or below 1 watt. Positive values will indicate powers greater than 1 watt, while negative values will indicate powers less than 1 watt. In general equation form:

$$P' = 10\log_{10}\frac{P}{reference} \tag{16.6}$$

The answer will have units of dBW, that is, decibels relative to 1 watt.

Example 16.7

A power amplifier has a maximum output of 120 W. What is this power in dBW?

$$P' = 10\log_{10}\frac{P}{1\text{ Watt}}$$
$$P' = 10\log_{10}\frac{120\text{ W}}{1\text{ W}}$$
$$P' = 20.8\text{ dBW}$$

There is nothing sacred about the 1 watt reference, short of its convenience. We could just as easily choose a different reference. Other common reference points are 1 milliwatt (dBm) and 1 femtowatt (dBf). Obviously, dBf is used for very low signal levels, such as those coming from an antenna. dBm is in very wide use in the communications industry. To use these other references, just divide the given power by the new reference.

Example 16.8

A small personal audio music player delivers 200 mW to its headphones. What is this output power in dBW, and in dBm?

For an answer in units of dBW, use the 1 watt reference.

$$P' = 10 \log_{10} \frac{P}{1\,\text{Watt}}$$
$$P' = 10 \log_{10} \frac{200\,\text{mW}}{1\,\text{W}}$$
$$P' = -7\,\text{dBW}$$

For units of dBm, use a 1 milliwatt reference.

$$P' = 10 \log_{10} \frac{P}{1\,\text{Watt}}$$
$$P' = 10 \log_{10} \frac{200\,\text{mW}}{1\,\text{mW}}$$
$$P' = 23\,\text{dBm}$$

200 mW, −7 dBW, and 23 dBm are three ways of saying the same thing. Note that the dBW and dBm values are 30 dB apart. This will always be true, because the references are a factor of 1000 (30 dB) apart.

In order to transfer a dBW or similar value into watts, reverse the process.

$$P = \log_{10}^{-1} \frac{P'}{10} \times \textit{reference} \tag{16.7}$$

Example 16.9

A studio microphone produces a 12 dBm signal while recording normal speech. What is the output power in watts?

$$P = \log_{10}^{-1} \frac{P'}{10} \times \textit{reference}$$
$$P = \log_{10}^{-1} \frac{12\,\text{dBm}}{10} \times 1\,\text{mW}$$
$$P = 15.8\,\text{mW} = 0.0158\,\text{W}$$

For voltages, we can use a similar system. A logical reference is 1 V, with the resulting units being dBV. As before, these voltage measurements will use a multiplier of 20 instead of 10.

$$V' = 20 \log_{10} \frac{V}{\textit{reference}} \tag{16.8}$$

Example 16.10

A test oscillator produces a 2 volt signal. What is this value in dBV?

$$V' = 20\log_{10}\frac{V}{reference}$$
$$V' = 20\log_{10}\frac{2\,\text{V}}{1\,\text{V}}$$
$$V' = 6.02\,\text{dB}$$

When both circuit gains and signal levels are specified in decibel form, analysis can be very quick. Given an input level, simply add the gain to it in order to find the output level. Given input and output levels, subtract them in order to find the gain.

Example 16.11

A computer hard drive read/write amplifier exhibits a gain of 35 dB. If the input signal is −42 dBV, what is the output signal?

$$V'_{out} = V'_{in} + A'_v$$
$$V'_{out} = -42\,\text{dBV} + 35\,\text{dB}$$
$$V'_{out} = -7\,\text{dBV}$$

Note that the final units are dBV and not dB, thus indicating a voltage and not merely a gain.

Example 16.12

A guitar power amp needs an input of 20 dBm to achieve an output of 25 dBW. What is the gain of the amplifier in dB?

First, it is necessary to convert the power readings so that they share the same reference unit. Because dBm represents a reference 30 dB smaller than the dBW reference, just subtract 30 dB to compensate.

$$20\,\text{dBm} = -10\,\text{dBW}$$

$$G' = P'_{out} - P'_{in}$$
$$G' = 25\,\text{dBW} - (-10\,\text{dBW})$$
$$G' = 35\,\text{dB}$$

Note that the units are dB and not dBW. This is very important! Saying that the gain is "so many" dBW is the same as saying the gain is "so many" watts. Obviously, gains are "pure" numbers and do not carry units such as watts or volts.

The usage of a dB-based system is shown graphically in Figure 16.3. Note how the stage gains are added to the input signal to form the output. Even large circuits can be quickly analyzed in this form. To make life in the lab even easier, it is possible to take measurements directly in dB form. By doing this, you need never convert while troubleshooting a design. For general-purpose work, voltage measurements are the norm, and therefore a dBV scale is often used.

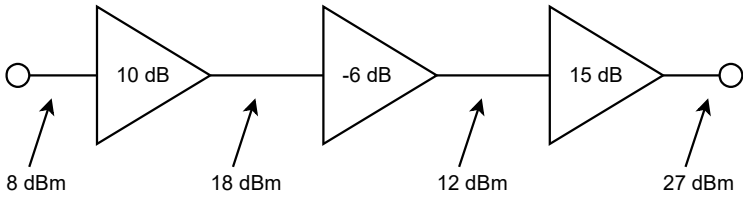


Figure 16.3
Multistage dB application.

Items Of Interest In The Laboratory

When using a digital meter on a dBV scale it is possible to "underflow" the meter if the signal is too weak. This will happen if you try to measure around zero volts, for example. If you attempt to calculate the corresponding dBV value, your calculator will probably show "error". The effective value is negative infinite dBV. The meter will certainly have a hard time showing this value! Another item of interest revolves around the use of dBm measurements. It is common to use a voltmeter to make dBm measurements, in lieu of a wattmeter. While the connections are considerably simpler, a voltmeter cannot measure power. How is this accomplished then? Well, as long as the circuit impedance is known, power can be derived from a voltage measurement. A common impedance in communication systems (such as recording studios) is 600 Ω, so a meter can be calibrated to give correct dBm readings by using Power Law. If this meter is used on a non-600 Ω circuit, the readings will no longer reflect accurate dBm values (but will still properly reflect relative changes in dB).

Finally, recalling the chapter introduction regarding "110 dB" concert levels, properly, that would read "110 dB-SPL", referring to "*S*ound *P*ressure *L*evel". The reference level corresponding to 0 dB-SPL is the quietest sound the average person can hear; the threshold of hearing (20 micropascals for young healthy humans). Thus, 110 dB-SPL refers to a sound pressure that is 110 dB greater than the threshold of hearing. Typically, 1 dB represents a "just noticeable difference" in loudness for humans, although this depends on the precise frequency and sound pressure.

16.3 Bode Plots

The Bode plot is a graphical response prediction technique that is useful for both circuit design and analysis. It is named after Hendrik Wade Bode, an American engineer known for his work in control systems theory and telecommunications. A Bode plot is, in actuality, a pair of plots: One graphs the signal gain or loss of a system versus frequency, while the other details the circuit phase versus frequency. Both of these items are very important in the design of well-behaved, optimal circuits.

Generally, Bode plots are drawn with logarithmic frequency axes, a decibel gain axis, and a phase axis in degrees. First, let's take a look at the gain plot. A typical gain plot is shown Figure 16.4. Remember, "gains" can be fractional, as with a voltage divider.

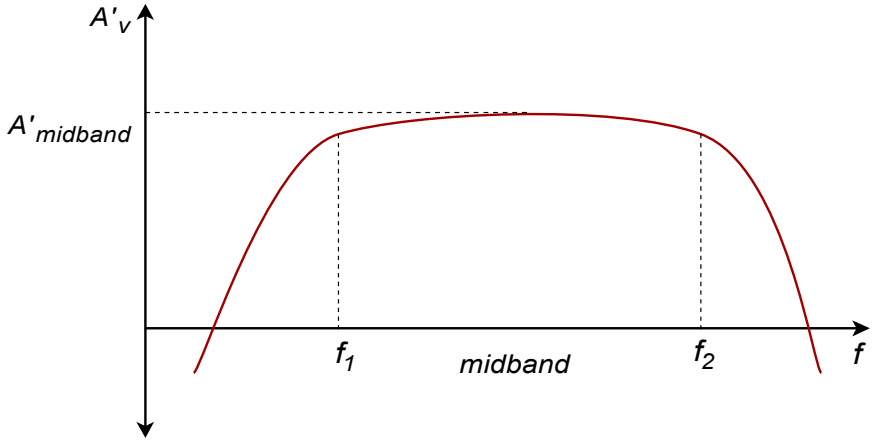


Figure 16.4
Generic gain plot.

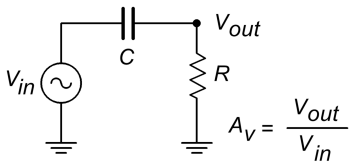
Note how the plot is relatively flat in the middle, or midband, region. The gain value in this region is known as the midband gain. In purely passive circuits this value may be fractional (i.e., a negative dB value). At either extreme of the midband region, the gain begins to decrease. The gain plot shows two important frequencies, f_1 and f_2. f_1 is the lower break frequency while f_2 is the upper break frequency. The gain at the break frequencies is 3 dB less than the midband gain. These frequencies are also known as the half-power points, or corner frequencies. Normally, amplifiers are only used for signals between f_1 and f_2. The exact shape of the rolloff regions will depend on the design of the circuit. For example, it is possible to design amplifiers with no lower break frequency (i.e., a DC amplifier), however, all amplifiers will exhibit an upper break. The break points are caused by the presence of circuit reactances, typically coupling and stray capacitances. The gain plot is a summation of the midband response with the upper and lower frequency limiting networks. Let's take a look at the lower break, f_1.

Lead Network Gain Response

Reduction in low frequency gain is caused by lead networks. A generic lead network is shown in Figure 16.5. It gets its name from the fact that the output voltage developed across R leads the input. At very high frequencies the circuit will be essentially resistive. Conceptually, think of this as a simple voltage divider. The divider ratio depends on the reactance of C. As the input frequency drops, X_c increases. This makes V_{out} decrease. At very high frequencies, where $X_c << R$, V_{out} is approximately equal to V_{in}. This can be seen graphically in Figure 16.6, where the frequency axis is normalized to f_c. The break frequency (i.e., the frequency at which the signal has decreased by 3 dB) is found via the standard equation,

$$f_c = \frac{1}{2\pi RC}$$

Figure 16.5
Lead network.



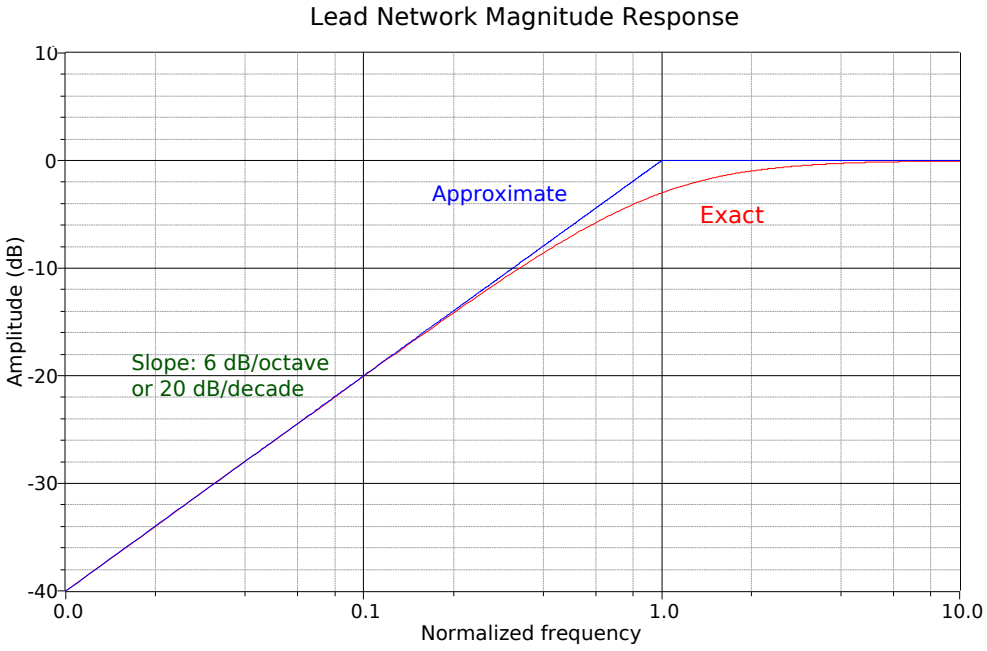


Figure 16.6
Lead network gain plot.

The response below f_c will be a straight line if a decibel gain axis and a logarithmic frequency axis are used. This makes for very quick and convenient sketching of circuit response. The slope of this line is 6 dB per octave (an octave is a doubling or halving of frequency, e.g., 800 Hz is 3 octaves above 100 Hz).[49] This range covers a factor of two in frequency. This slope may also be expressed as 20 dB per decade, where a decade is a factor of 10 in frequency. With reasonable accuracy, this curve may be approximated as two line segments, called *asymptotes*, shown in Figure 16.6 (blue). The shape of this curve is the same for any lead network. Because of this, it is very easy to find the approximate gain at any given frequency as long as f_c is known.

49 The term *octave* is borrowed from the field of music. It gets its name from the fact that there are eight notes in the standard western scale: do-re-mi-fa-sol-la-ti-do.

It is not necessary to go through reactance and phasor calculations. To create a general response equation, start with the voltage divider rule to find the gain:

$$\frac{V_{out}}{V_{in}} = \frac{R}{R - jX_c}$$
$$\frac{V_{out}}{V_{in}} = \frac{R\angle 0}{\sqrt{R^2 + X_c^2}\angle -\arctan\frac{X_c}{R}}$$

The magnitude of this is,

$$|A_v| = \frac{R}{\sqrt{R^2 + X_c^2}}$$
$$|A_v| = \frac{1}{\sqrt{1 + \frac{X_c^2}{R^2}}} \qquad (16.9)$$

Recalling that,

$$f_c = \frac{1}{2\pi RC}$$

we may say,

$$R = \frac{1}{2\pi f_c C}$$

For any frequency of interest, *f*,

$$X_c = \frac{1}{2\pi fC}$$

Equating the two preceding equations yields,

$$\frac{f_c}{f} = \frac{X_c}{R} \qquad (16.10)$$

Substituting equation 16.10 in equation 16.9 gives,

$$A_v = \frac{1}{\sqrt{1 + \frac{f_c^2}{f^2}}} \qquad (16.11)$$

To express A_v in dB, substitute equation 16.11 into equation 16.5.

$$A'_v = 20 \log_{10} \frac{1}{\sqrt{1 + \frac{f_c^2}{f^2}}}$$

After simplification, the final result is:

$$A'_v = -10 \log_{10} \left(1 + \frac{f_c^2}{f^2} \right) \tag{16.12}$$

Where
f_c is the critical frequency,
f is the frequency of interest,
A'_v is the decibel gain at the frequency of interest.

Example 16.13

A circuit has a lower break frequency of 40 Hz. How much signal is lost at 10 Hz?

$$A'_v = -10 \log_{10} \left(1 + \frac{f_c^2}{f^2} \right)$$

$$A'_v = -10 \log_{10} \left(1 + \frac{40^2}{10^2} \right)$$

$$A'_v = -12.3 \text{ dB}$$

In other words, the signal level is 12.3 dB lower than it is in the midband. Note that 10 Hz is 2 octaves below the break frequency. Because the cutoff slope is 6 dB per octave, each octave loses 6 dB. Therefore, the approximate result is −12 dB, which double-checks the exact result. Without the lead network, the gain would stay at 0 dB all the way down to DC (0 Hz.)

Lead Network Phase Response

At very low frequencies, the circuit of Figure 16.5 is largely capacitive. Because of this, the output voltage developed across R leads by 90 degrees. At very high frequencies the circuit will be largely resistive. At this point V_{out} will be in phase with V_{in}. At the critical frequency, V_{out} will lead by 45 degrees. A general phase graph is shown in Figure 16.7. As with the gain plot, the phase plot shape is the same for any lead network. The general phase equation may be obtained from the voltage divider:

$$\frac{V_{out}}{V_{in}} = \frac{R}{R - jX_c}$$
$$\frac{V_{out}}{V_{in}} = \frac{R\angle 0}{\sqrt{R^2 + X_c^2}\angle -\arctan\frac{X_c}{R}}$$

The phase portion of this is,

$$\theta = \arctan\frac{X_c}{R}$$

By using equation 10.6, this simplifies to,

$$\theta = \arctan\frac{f_c}{f} \tag{16.13}$$

Where
f_c is the critical frequency,
f is the frequency of interest,
θ is the phase angle at the frequency of interest.

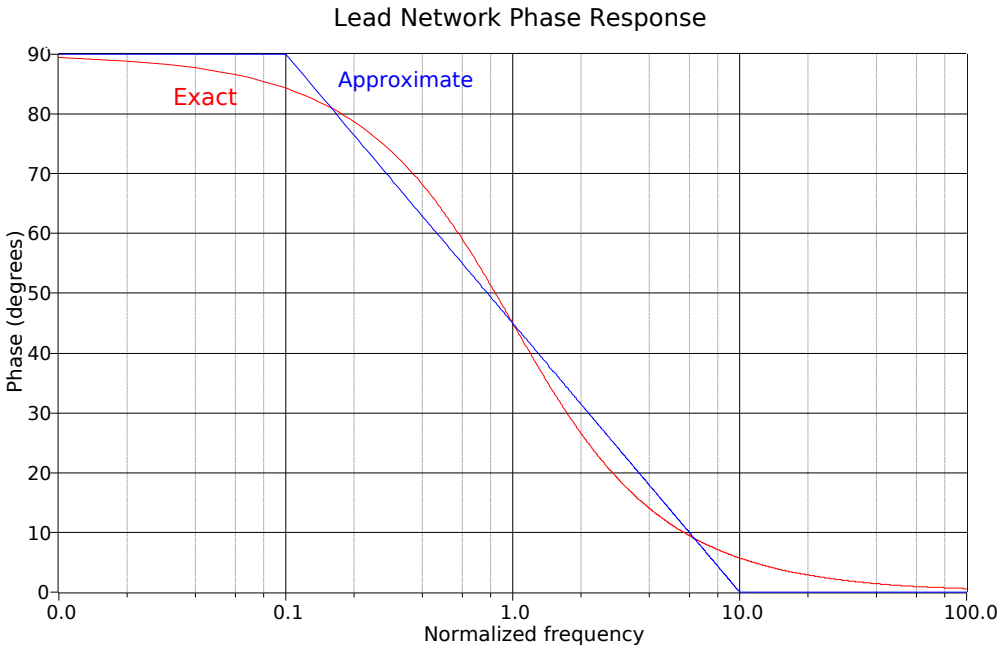


Figure 16.7
Lead network phase response.

Often, an approximation, such as the blue line in Figure 16.7, is sufficient. By using Equation 16.13, you can show that this approximation is off by no more than 6 degrees at the corners.

Example 16.14

A telephone amplifier has a lower break frequency of 120 Hz. What is the phase response one decade below and one decade above?

One decade below 120 Hz is 12 Hz, while one decade above is 1.2 kHz.

$$\theta = \arctan \frac{f_c}{f}$$

$$\theta = \arctan \frac{120\,\text{Hz}}{12\,\text{Hz}}$$

$\theta = 84.3$ degrees one decade below f_c (i.e, approaching 90 degrees)

$$\theta = \arctan \frac{120\,\text{Hz}}{1.2\,\text{kHz}}$$

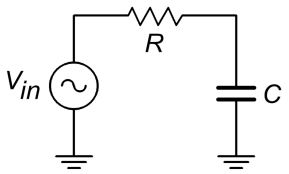
$\theta = 5.71$ degrees one decade above f_c (i.e., approaching 0 degrees)

Remember, if a circuit or amplifier is direct-coupled, and has no lead networks, the phase will remain at 0 degrees right back to 0 Hz (DC).

Lag Network Response

Unlike its lead network counterpart, all systems will contain lag networks. In essence, it is little more than an inverted lead network. As you can see from Figure 16.8, it simply transposes the *R* and *C* locations. Because of this, the response tends to be inverted as well. In terms of gain, X_c is very large at low frequencies, and thus V_{out} equals V_{in}. At high frequencies, X_c decreases, and V_{out} falls. The break point occurs when X_c equals *R*. The general gain plot is shown in Figure 16.9. Like the lead network response, the slope of this curve is −6 dB per octave (or −20 dB per decade.) Note that the slope is negative instead of positive. We can derive a general gain equation for this circuit in virtually the same manner as we did for the lead network. The derivation is left as an exercise.

Figure 16.8
Lag network.



$$A'_v = -10 \log_{10}\left(1 + \frac{f^2}{f_c^2}\right) \tag{16.14}$$

Where
f_c is the critical frequency,
f is the frequency of interest,
A'_v is the decibel gain at the frequency of interest.

Note that this equation is almost the same as Equation 16.12. The only difference is that f and f_c have been transposed.

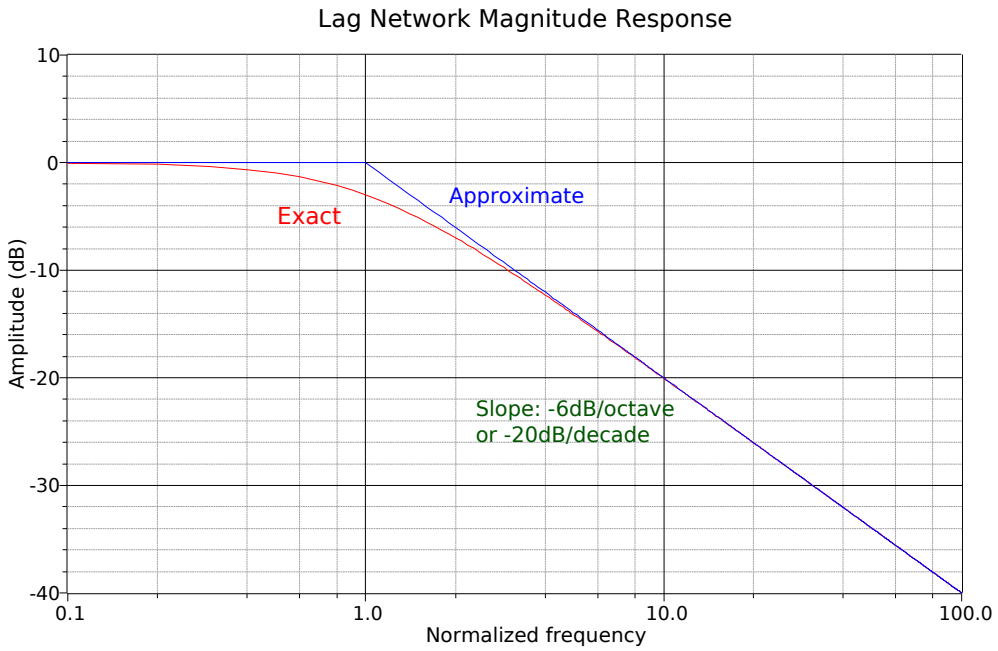


Figure 16.9
Lag network gain plot.

In a similar vein, we may examine the phase response. At very low frequencies, the circuit is basically capacitive. Because the output is taken across C, V_{out} will be in phase with V_{in}. At very high frequencies, the circuit is essentially resistive. Consequently, the output voltage across C will lag by 90 degrees. At the break frequency the phase will be −45 degrees. A general phase plot is shown in Figure 16.10. As with the lead network,we may derive a phase equation. Again, the exact steps are very similar, and left as an exercise.

$$\theta = -90 + \arctan\frac{f_c}{f} \tag{16.15}$$

Where
f_c is the critical frequency,
f is the frequency of interest,
θ is the phase angle at the frequency of interest.

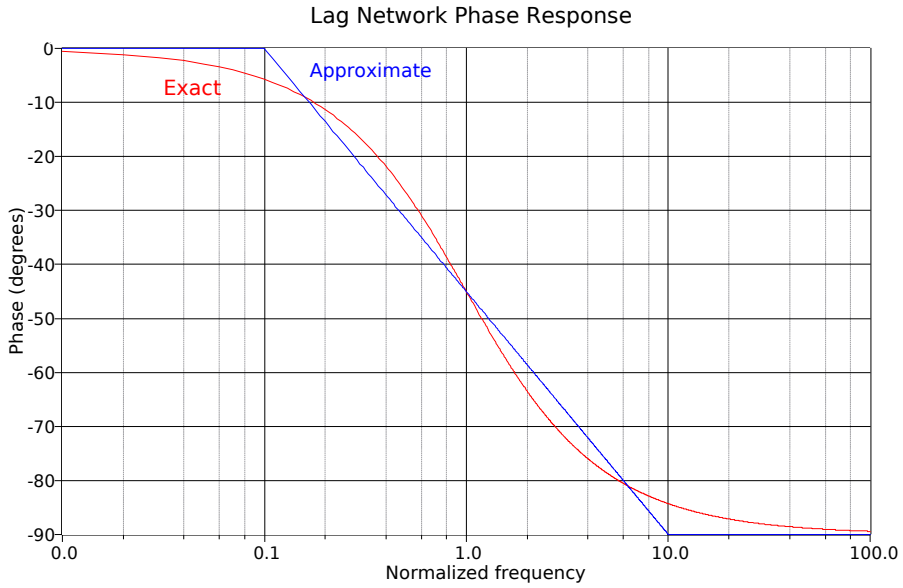


Figure 16.10
Lag network phase response.

Example 16.15

A medical ultra sound transducer feeds a lag network with an upper break frequency of 150 kHz. What are the gain and phase values at 1.6 MHz?

Because this represents a little more than a 1 decade increase, the approximate values are −20 dB and −90 degrees, from Figures 16.7 and 16.8, respectively. The exact values are:

$$A'_v = -10\log_{10}\left(1+\frac{f^2}{f_c^2}\right)$$

$$A'_v = -10\log_{10}\left(1+\frac{1.6\,\text{MHz}^2}{150\,\text{kHz}^2}\right)$$

$$A'_v = -20.6\,\text{dB}$$

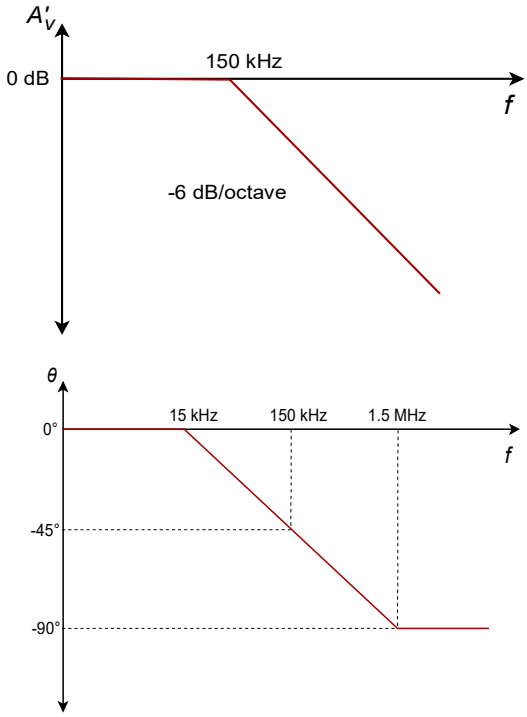
$$\theta = -90 + \arctan\frac{f_c}{f}$$

$$\theta = -90 + \arctan\frac{150\,\text{kHz}}{1.6\,\text{MHz}}$$

$$\theta = -84.6\,\text{degrees}$$

The complete Bode plot for this network is shown in Figure 16.11. It is very useful to examine both plots simultaneously. In this manner you can find the exact phase change for a given gain quite easily. For example, if you look carefully at the plots of Figure 16.11, you will note that at the critical frequency of 150 kHz, the total phase change is −45 degrees.

Figure 16.11
Bode plot for 150 kHz lag.



Because this circuit involved the use of a single lag network, this is exactly what you would expect.

Rise Time versus Bandwidth

For pulse-type signals, the "speed" of a circuit is often expressed in terms of its *rise time*. If a square pulse such as Figure 16.12a is passed into a simple lag network, the capacitor charging effect will produce a rounded variation, as seen in Figure 16.12b. This effect places an upper limit on the duration of pulses that a given circuit can handle without producing excessive distortion.

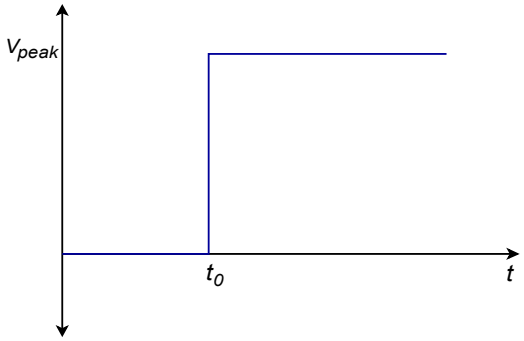


Figure 16.12a
Pulse rise time effect: Input to network.

By definition, rise time is the amount of time it takes for the signal to traverse from 10% to 90% of the peak value of the pulse. The shape of this pulse is defined by the standard capacitor charge equation examined in earlier course work, and is valid for any system with a single clearly dominant lag network.

$$V_{out} = V_{peak}\left(1 - \epsilon^{\frac{-t}{RC}}\right) \tag{16.16}$$

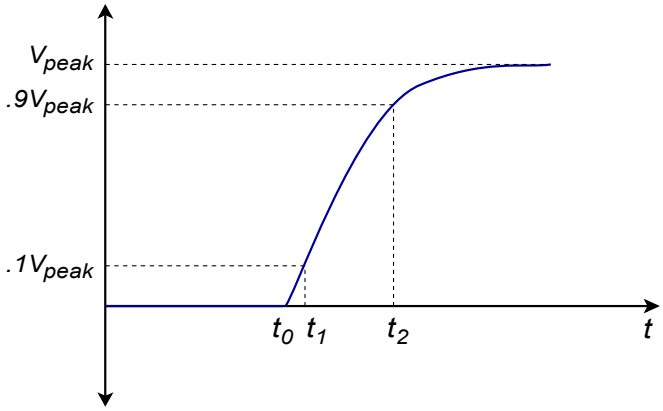


Figure 16.12b
Pulse rise time effect: Output of network.

In order to find the time internal from the initial starting point to the 10% point, set V_{out} to $0.1V_{peak}$ in Equation 16.16 and solve for t_1.

$$
\begin{aligned}
0.1\,V_{peak} &= V_{peak}\left(1-\epsilon^{\frac{-t_1}{RC}}\right)\\
0.1\,V_{peak} &= V_{peak}-V_{peak}\,\epsilon^{\frac{-t_1}{RC}}\\
0.9\,V_{peak} &= V_{peak}\,\epsilon^{\frac{-t_1}{RC}}\\
0.9 &= \epsilon^{\frac{-t_1}{RC}}\\
\log 0.9 &= \frac{-t_1}{RC}\\
t_1 &= 0.105\,RC
\end{aligned}
\qquad (16.17)
$$

To find the interval up to the 90% point, follow the same technique using $0.9V_{peak}$. Doing so yields:

$$t_2 = 2.303\,RC \qquad (16.18)$$

The rise time, T_r, is the difference between t_1 and t_2

$$
\begin{aligned}
T_r &= t_1 - t_2\\
T_r &= 2.303\,RC - 0.105\,RC\\
T_r &\approx 2.2\,RC
\end{aligned}
\qquad (16.19)
$$

Equation 16.19 ties the rise time to the lag network's R and C values. These same values also set the critical frequency f_2. By combining equation 16.15 with the basic critical frequency relationship, we can derive an equation relating f_2 to T_r.

$$f_2 = \frac{1}{2\pi RC}$$

Solving 16.19 in terms of RC, and substituting yields

$$
\begin{aligned}
f_2 &= \frac{2.2}{2\pi T_r}\\
f_2 &= \frac{0.35}{T_r}
\end{aligned}
\qquad (16.20)
$$

Where
f_2 is the upper critical frequency,
T_r is the rise time of the output pulse.

Example 16.16

Determine the rise time for a lag network critical at 100 kHz.

$$f_2 = \frac{0.35}{T_r}$$
$$T_r = \frac{0.35}{f_2}$$
$$T_r = \frac{0.35}{100\,\text{kHz}}$$
$$T_r = 3.5\,\mu\text{s}$$

16.4 Combining the Elements - Multi-Stage Effects

A complete gain or phase plot combines three elements: (1) the midband response, (2) the lead response, and (3) the lag response. Normally, a particular design will contain multiple lead and lag networks. The complete response is the summation of the individual responses. For this reason, it is useful to find the dominant lead and lag networks. These are the networks that affect the midband response first. For lead networks, the dominant one will be the one with the highest f_c. Conversely, the dominant lag network will be the one with the lowest f_c. It is very common to approximate the complete system response by drawing straight-line segments such as those given in Figures 16.5 and 16.7. The process goes something like this:

- Locate all f_cs on the frequency axis.
- Draw a straight line between the dominant lag and lead f_cs at the midband gain. If the system does not contain any lead networks, continue the midband gain line down to DC.
- Draw a 6 dB per octave slope between the dominant lead and the next lower lead network.
- Because the effects of the networks are cumulative, draw a 12 dB per octave slope between the second lead f_c and the third f_c. After the third f_c, the slope should be 18 dB per octave, after the fourth, 24 dB per octave, and so on.
- Draw a −6 dB per octave slope between the dominant lag f_c and the next highest f_c. Again, the effects are cumulative, so increase the slope by −6 dB at every new f_c.

Example 16.17

Draw the Bode gain plot for the following amplifier: A'_v midband = 26 dB, one lead network critical at 200 Hz, one lag network critical at 10 kHz, and another lag network critical at 30 kHz.

The dominant lag network is 10 kHz. There is only one lead network, so it's dominant by default.

- Draw a straight line between 200 Hz and 10 kHz at an amplitude of 26 dB.
- Draw a 6 dB per octave slope below 200 Hz. To do this, drop down one octave (100 Hz) and subtract 6 dB from the present gain (26 dB − 6dB = 20 dB.) The line will start at the point 200 Hz/26 dB, and pass through the point 100 Hz/20 dB. Because there are no other lead networks, this line may be extended to the left edge of the graph.
- Draw a −6 dB per octave slope between 10 kHz and 30 kHz. The construction point will be 20 kHz/20 dB. Continue this line to 30 kHz. The gain at the 30 kHz intersection should be around 16 dB. The slope above this second f_c will be −12 dB per octave. Therefore, the second construction point should be at 60 kHz/4 dB (one octave above 30 kHz, and 12 dB down from the 30 kHz gain). Because this is the final lag network, this line may be extended to the right edge of the graph.

A completed graph is shown in Figure 16.13.

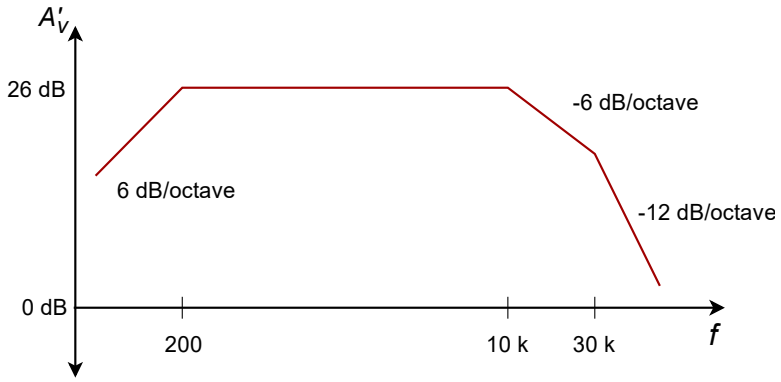


Figure 16.13
Gain plot of complete amplifier.

There is one item that should be noted before we leave this section, and that is the concept of *narrowing*. Narrowing occurs when two or more networks share similar critical frequencies, and one of them is a dominant network. The result is that the true −3 dB breakpoints may be altered. Here is an extreme example. Assume that a circuit has two lag networks, both critical at 1 MHz. A Bode plot would indicate that the breakpoint is 1 MHz. This is not really true. Remember, the effects of lead and lag networks are cumulative. Because each network produces a 3 dB loss at 1 MHz, the net loss at this frequency is actually 6 dB. The true −3 dB point will have been shifted down. The Bode plot only gives you the approximate shape of the response.

Summary

We have seen how to convert gains and signals into a decibel form for both powers and voltages. This is convenient because what would require multiplication and division under the ordinary scheme only requires simple addition and subtraction in the dB scheme. Along with this, dB measurement is used almost exclusively for Bode gain plots. A Bode plot details a system's gain magnitude and phase response. For gain, the amplitude is measured in dB, while the frequency is normally presented in log form. For a phase plot, phase is measured in degrees, and again, the frequency axis is logarithmic. The changes in gain and phase at the frequency extremes are caused by lead and lag networks. Lead networks cause the low frequency gain to roll off. The roll off rate is 6 dB per octave per network. The phase will change from +90 degrees to 0 degrees per network. Lag networks cause the high frequency gain to roll off at a rate of −6 dB per octave per network. The phase change per lag network is from 0 degrees to −90 degrees.

Review Questions

1. What are the advantages of using decibels over the ordinary scheme?
2. How do decibel power and voltage gain calculations differ?
3. Define the differences between dB, dBW, dBm, dBV and dBu.
4. Describe a Bode plot.
5. What is a lead network? What general response does it yield?
6. What is a lag network? What general response does it yield?
7. What do the terms f_1 and f_2 indicate about a system's response?
8. What are the rolloff slopes for lead and lag networks?
9. What are the phase changes produced by individual lead and lag networks?
10. How is rise time related to upper break frequency?
11. How do multiple lead or lag networks interact to form an overall system response?
12. How does the decibel measurement scheme differ from the ordinary method of indicating gains and signal level?

Problems

Analysis Problems - dB emphasis

1. Convert the following power gains into dB form: a) 10 b) 80 c) 500 d) 1 e) 0.2 f) 0.03.
2. Convert the following dB power gains into ordinary form: a) 0 dB b) 12 dB c) 33.1 dB d) 0.2 dB e) −5.4 dB f) −20 dB.
3. An amplifier has an input signal of 1 mW, and produces a 2 W output. What is the power gain in dB?
4. A Hi-Fi power amplifier has a maximum output of 50 W and a power gain of 19 dB. What is the maximum input signal power?
5. An amplifier with a power gain of 27 dB is driven by a 25 mW source. Assuming the amplifier doesn't clip, what is the output signal in watts?
6. Convert the following voltage gains into dB form: a) 10 b) 40 c) 250 d) 1 e) 0.5 f) 0.004
7. Convert the following dB voltage gains into ordinary form: a) 0.5 dB b) 0 dB c) 46 dB d) 10.7 dB e) −8 dB f) −14.5 dB
8. A guitar pre-amp has a gain of 44 dB. If the input signal is 12 mV, what is the output signal?
9. A video amplifier has a 140 mV input and a 1.2 V output. What is the voltage gain in dB?
10. The pre-amp in a particular tape deck can output a maximum signal of 4 V. If this amplifier has a gain of 18 dB, what is the maximum input signal?
11. Convert the following powers into dBW: a) 1 W b) 23 W c) 6.5 W d) 0.2W e) 2.3 mW f) 1.2 kW g) 0.045 mW h) 0.3 μW i) 5.6E−18 W.
12. Repeat Problem 11 for units of dBm.
13. Repeat Problem 11 for units of dBf.
14. Convert the following voltages into dBV: a) 12.4 V b) 1 V c) 0.25 V d) 1.414 V e) 0.1 V f) 10.6 kV g) 13 mV h) 2.78 μV.
15. A two stage power amplifier has power gains of 12 dB and 16 dB. What is the total gain in dB and in ordinary form?
16. If the amplifier of Problem 15 has an input of −18 dBW, what is the final output in dBW? in dBm? in watts?
17. Referring to Figure 16.3, what are the various stages' outputs if the input is changed to −4 dBm? to −34 dBW?
18. Which amplifier has the greatest power output? a) 50 watts b) 18 dBW c) 50 dBm.
19. Which amplifier has the greatest power output? a) 200 mW b) −10 dBW c) 22 dBm.

20. A three stage amplifier has voltage gains of 20 dB, 5 dB, and 12 dB respectively. What is the total voltage gain in dB and in ordinary form?
21. If the circuit of Problem 20 has an input voltage of −16 dBV, what are the outputs of the various stages in dBV? In volts?
22. Repeat Problem 21 for an input of 12 mV.
23. Which amplifier produces the largest output voltage? a) 15 V b) 16 dBV

Analysis Problems - Bode plot emphasis

24. Given a lead network critical at 3 kHz, what are the gain and phase values at 100 Hz, 3 kHz, and 40 kHz?
25. Given a lag network tuned to 700 kHz, what are the gain and phase values at 50 kHz, 700 kHz, and 10 MHz? What is the rise time?
26. A noninverting amplifier has a midband voltage gain of 18 dB and a single lag network at 200 kHz. What are the gain and phase values at 30 kHz, 200 kHz, and 1 MHz. What is the rise time?
27. Repeat Problem 26 for an inverting (−180 degrees) amplifier.
28. Draw the Bode plot for the circuit of Problem 26.
29. Draw the Bode plot for the circuit of Problem 27.
30. An inverting (−180 degrees) amplifier has a midband gain of 32 dB and a single lead network critical at 20 Hz (assume the lag network f_c is high enough to ignore for low frequency calculations). What are the gain and phase values at 4 Hz, 20Hz, and 100 Hz?
31. Repeat Problem 29 with a noninverting amplifier.
32. Draw the Bode plot for the circuit of Problem 30.
33. Draw the Bode plot for the circuit of Problem 31.
34. A noninverting amplifier used for ultrasonic applications has a midband gain of 41 dB, a lag network critical at 250 kHz, and a lead network critical at 30 kHz. Draw its gain Bode plot.
35. Find the gain and phase at 20 kHz, 100 kHz, and 800 kHz for the circuit of Problem 34.
36. If the circuit of Problem 34 has a second lag network added at 300 kHz, What are the new gain and phase values at 20 kHz, 100 kHz, and 800 kHz?
37. Draw the gain Bode plot for the circuit of Problem 36.
38. What are the maximum and minimum phase shifts across the entire frequency spectrum for the circuit of Problem 36?
39. A noninverting DC amplifier has a midband gain of 36 dB, and lag networks at 100 kHz, 750 kHz, and 1.2 MHz. Draw its gain Bode plot.

40. What are the maximum and minimum phase shifts across the entire frequency spectrum for the circuit of Problem 39?
41. What is the maximum rate of high frequency attenuation for the circuit of Problem 39 in dB/Decade?
42. If an amplifier has two lead networks, what is the maximum rate of low frequency attenuation in dB/Octave?

Challenge Problems

43. You would like to use a voltmeter to take dBm readings in a 600 Ω system. What voltage should produce 0 dBm?
44. Assuming that it takes about an 8 dB increase in sound pressure level in order to produce a sound that is subjectively "twice as loud" to the human ear, can a Hi-Fi using a 100 W amplifier sound twice as loud as one with a 40 W amplifier (assuming the same loudspeakers)?
45. Hi-Fi amplifiers are often rated with a "headroom factor" in dB. This indicates how much extra power the amplifier can produce for short periods of time, over and above its nominal rating. What is the maximum output power of a 250 W amplifier with 1.6 dB headroom?
46. If the amplifier of Problem 34 picks up an extraneous signal that is a −10 dBV sine wave at 15 kHz, what is the output?
47. If the amplifier of Problem 39 picks up a high frequency interference signal at 30 MHz, how much is it attenuated over a normal signal? If this input signal is measured at 2 dBV, what should the output be?
48. If an amplifier has two lag networks, and both are critical at 2 MHz, is the resulting f_2 less than, equal to, or greater than 2 MHz?
49. If an amplifier has two lead networks, and both are critical at 30 Hz, is the resulting f_1 less than, equal to, or greater than 30 Hz?

Simulation Problems

50. Use a simulator to plot the Bode gain response of the circuit in Problem 39.
51. Use a simulator to plot the Bode phase response of the circuit in Problem 34.
52. Use a simulation program to generate a Bode plot for a lead network comprised of a 1 kΩ resistor and a 100 nF capacitor.

17 Frequency Limits

17.0 Chapter Learning Objectives

After completing this chapter, you should be able to:

- Recognize circuit components which impact the low frequency performance of BJT and FET amplifiers.
- Recognize circuit components which impact the high frequency performance of BJT and FET amplifiers.
- Analyze and compute the frequency limits (f_1 and f_2) of BJT and FET amplifiers.
- Design or modify the frequency limits (f_1 and f_2) of BJT and FET amplifiers.
- Understand how Miller's Theorem can impact the high frequency limit (f_2) of certain amplifiers.

17.1 Introduction

All amplifiers, regardless of design, exhibit frequency limits. That is to say, there is a range of frequencies over which an amplifier will work effectively. This concept was introduced in Chapter 6 and will be expanded greatly here. The primary goal of this chapter is to offer methods of determining this range for both BJT and FET amplifiers, as well as offering insight in how to design amplifiers to meet certain frequency criteria. As a reminder, the lower and upper limit frequencies are denoted as f_1 and f_2, respectively. They represent the frequencies at which the midband response has fallen by three decibels, or half power. They are also referred to as the *3 dB down frequencies* or *corner frequencies*. Without exception, all amplifiers will exhibit an upper frequency limit, however, amplifiers can be designed to have no lower frequency limit. In other words, the f_1 is DC or 0 Hertz. Normally, these amplifiers are referred to as being *DC* or *direct coupled*.

In order to analyze the frequency range of amplifiers, we shall introduce new AC models for the circuits; one set for low frequencies, and a second set for high frequencies. Certain assumptions that we took for granted, such as coupling capacitors being treated as short circuits, may no longer be valid. Consequently, in order to obtain a full frequency analysis of a given amplifier, we may need to examine three separate equivalent circuits: one for the midband response, a second for the low frequency response, and a third for the high frequency response. The low frequency equivalent will contain lead networks while the high frequency equivalent will contain lag networks. The end result will be a Bode plot of the amplifier, much like the generic plots created in the prior chapter.

We shall also make use of Miller's Theorem in this analysis. This will prove to be particularly important when considering the high frequency performance of common emitter BJT and common source FET amplifiers. Miller's Theorem also can be of great use when designing an amplifier to meet a specific high frequency limit as it effectively multiplies the size of the associated capacitor.

17.2 Low Frequency Response

We begin our discussion at the low end of the frequency spectrum. Here, our focus is on obtaining a value for f_l, or designing for a desired f_l.

Lead Networks

The system response at lower frequencies is dependent on the lead networks in the circuit. Fortunately, these are easy to spot as generally they are caused by the presence of coupling and/or bypass capacitors. Thus, a typical single stage amplifier, whether BJT or FET, will have three such networks: one at the input, a second at the output, and a third in the emitter or source bypass position. In multi-stage systems, the output network of one stage will combine with the input network of the following stage to create a single, combined network. Technically, the bypass network is not a true lead network like the input and output networks, but the practical difference is usually trivial. In any case, we will examine the difference a little later in this section.

BJT Amplifiers

Common Emitter

Let us begin with a fairly generic swamped common emitter amplifier, as shown in Figure 17.1.

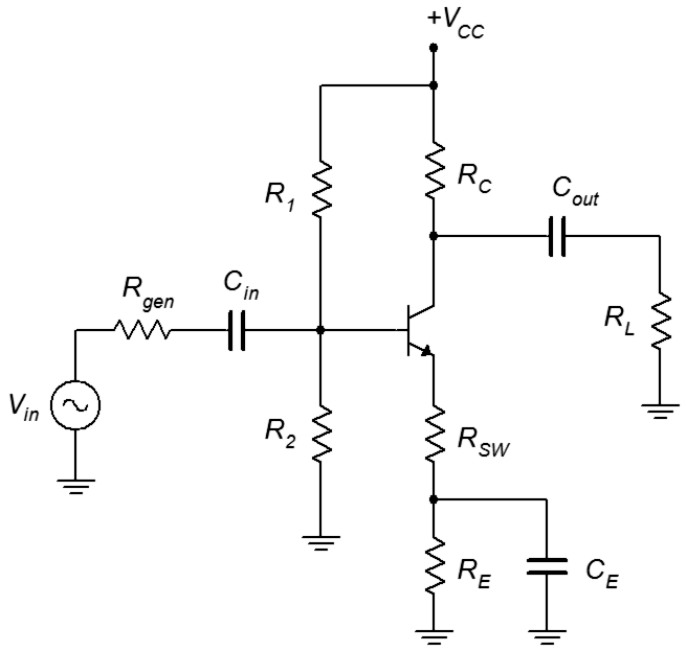


Figure 17.1
BJT common emitter amplifier.

The low frequency equivalent circuit is similar to the midband equivalent, except that we do not assume that the coupling and bypass capacitors are shorted. This is shown in Figure 17.2, redrawn for clarity.

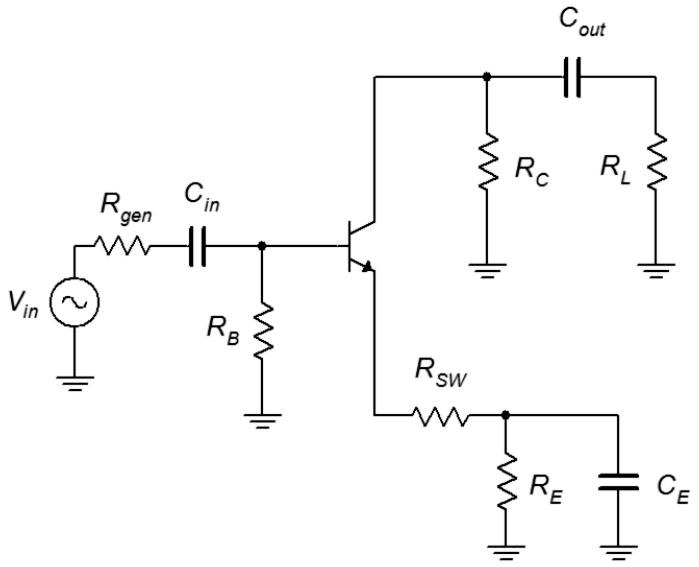
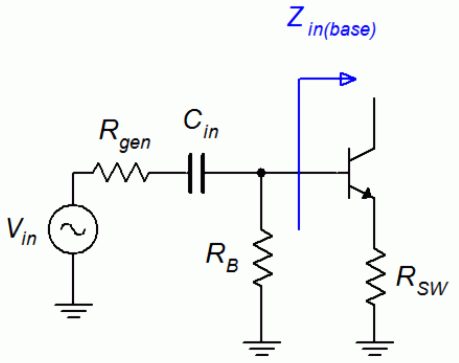


Figure 17.2
Low frequency equivalent of CE amplifier.

In general, a lead network exists wherever the signal current flows through a capacitor on its way to the load. The capacitor creates a frequency dependent voltage divider with the associated impedance. The capacitive reactance increases with decreasing frequency, and thus the voltage divider effect worsens as the frequency is lowered. That is, the percentage of the signal that gets to the load is progressively reduced as the frequency decreases. This is the hallmark of a lead network. This effect should be obvious with the input and output networks. The emitter bypass network operates a little differently, as we shall see in a moment.

Figure 17.3
Input lead network of CE amplifier.



First, let's consider the input network. This section is shown in Figure 17.3. In order to find the critical frequency, we need values for the capacitance and resistance. The capacitance is just C_{in}, the input coupling capacitor. To find the resistance, find the Thevenin equivalent resistance seen from the capacitor. To the left, we have the internal resistance of the signal source, R_{gen}. The source itself is shorted, leaving that branch at ground. To the right, we see the input impedance of the stage which is the base biasing resistor, R_B, in parallel with $Z_{in(base)}$. $Z_{in(base)}$ is found by the usual equation, $\beta(R_{SW} + r'_e)$, assuming the reactance of C_E is negligible. For lack of a better term, we can call this total resistance $R_{in(lead)}$.

$$R_{in(lead)} = R_{gen} + R_B \,||\, \beta(R_{SW} + r'_e) \qquad (17.1)$$

If the circuit is not swamped, we can just set the swamping resistor, R_{SW}, to zero.

The general equation for critical frequency is

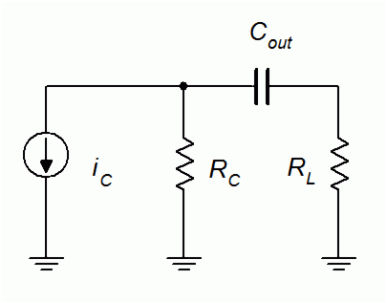
$$f_c = \frac{1}{2\pi RC}$$

Substituting the values for the input network, we find

$$f_{c(input\,lead)} = \frac{1}{2\pi R_{in(lead)} C_{in}} \tag{17.2}$$

The analysis for the output network is similar. The equivalent circuit of the output network is shown in Figure 17.4. Finding the Thevenin resistance, which we shall call $R_{out(lead)}$, is a simple matter of opening the current source. The resistance is the series combination of R_C and R_L.

Figure 17.4
Output lead network of CE amplifier.



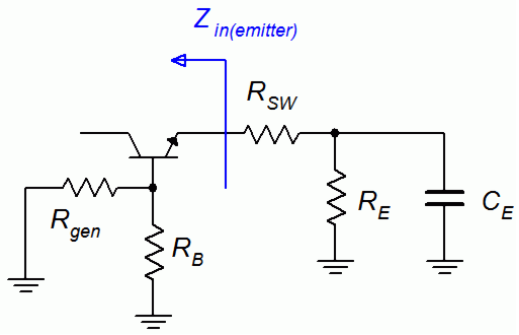
$$R_{out(lead)} = R_C + R_L \tag{17.3}$$

Thus, we find,

$$f_{c(output\,lead)} = \frac{1}{2\pi R_{out(lead)} C_{out}} \tag{17.4}$$

Finally, we need to consider the emitter bypass network. The associated capacitance is C_E, And the low frequency equivalent circuit is shown in Figure 17.5. By inspection, we can see that the resistance that the capacitor "sees" is the emitter bias resistor, R_E, in parallel with the series combination of the swamping resistor, R_{SW}, and the impedance looking into the emitter, or $Z_{in(emitter)}$. From prior work with common collector amplifiers in Chapter 7, it was discovered that the impedance looking into the emitter is equal to r'_e in series with the base resistance divided by beta, which gives us

Figure 17.5
Bypass lead network of CE amplifier.



$$R_{bypass(lead)} = R_E \parallel \left(R_{SW} + r'_e + \frac{R_B \parallel R_{gen}}{\beta} \right) \tag{17.5}$$

The swamping resistor tends to dominate the result as the final term usually is rather small. We now have the third and final lead critical frequency formula for our BJT common emitter amplifier:

$$f_{c(bypass\,lead)} = \frac{1}{2\pi R_{bypass(lead)} C_E} \tag{17.6}$$

As mentioned previously, the bypass network is not a true lead network. By “true”, it is meant that the response does not roll off continuously as frequency is decreased. In fact, at some point, the response will level off with no further loss incurred. This can be understood by looking at precisely how the bypass capacitor controls the gain. The job of this capacitor is to “short out” the emitter bias resistor, R_E. This occurs because X_C is much smaller than the associated resistance at midband frequencies. As the frequency decreases, the reactance increases, and the effective AC emitter resistance is no longer zero. At this point we can think of the capacitor as only partially bypassing the resistor. Thus, the voltage gain starts to decrease. It's as if the swamping is increasing. With further decreases in frequency, the reactance continues to increase. At some point, the reactance magnitude will equal the value of the biasing resistor and further decreases in the frequency will have little effect on the total AC emitter impedance. Thus, the gain will decrease no further. The result is a “stepped” response. That is, after the initial rolloff, the gain flattens again, producing the appearance of a step. From this we can deduce that this second frequency is equal to

$$f_{c(bypass\,step)} = \frac{1}{2\pi R_E C_E} \tag{17.7}$$

A quick approximation of this frequency can be found by simply dividing the bypass critical frequency by the ratio of the bias and swamping resistors. For example, if the bias resistor is 5 kΩ and the swamping resistor is 250 Ω, the step frequency will be roughly 20 times lower than the bypass critical frequency. Usually, this frequency is low enough that it will not impact greatly the overall response of the amplifier, and we need not bother with it.

Once the lead frequencies are computed, an appropriate Bode plot for the low frequency portion of the spectrum may be generated for the amplifier, as outlined in the prior chapter. Remember that the highest of these frequencies is the dominant frequency, that is, the frequency at which midband response is first affected. Also, if two or more of the frequencies are in close proximity and one of them is dominant, there will be an interaction between them. This may lead to the system f_l being somewhat higher than the dominant f_c.

One final consideration is the absence of a swamping resistor. If the amplifier is not swamped, the expressions developed for the input and bypass networks (i.e., equations 17.1 and 17.5) can still be used; just insert zero for the value of R_{SW}. It is worth noting, though, that the elimination of the swamping resistor will decrease the associated effective resistance in these networks, resulting in higher required capacitance values for a given critical frequency. In the case of the bypass network, the effect can be dramatic.

Time for a few illustrative examples.

For the circuit of Figure 17.6, determine f_l. Assume β = 165.

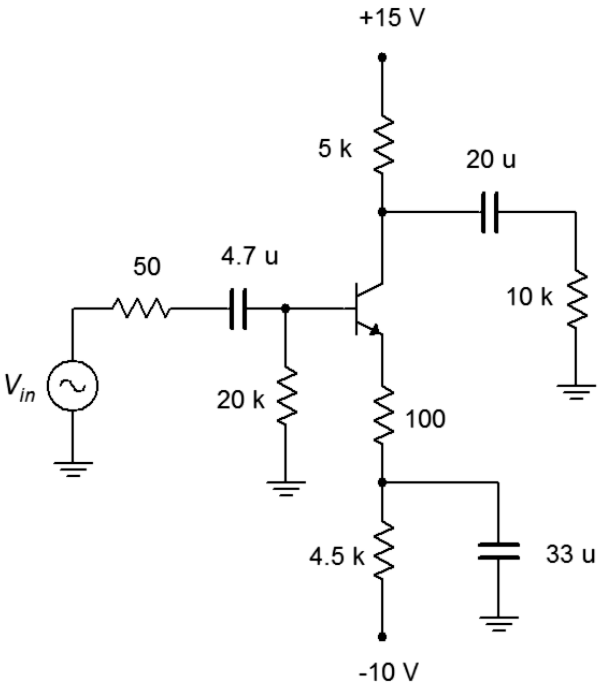


Figure 17.6
Circuit for Example 17.1.

First, we can approximate the DC emitter current to be about 2 mA which sets r'_e to 13 Ω. Starting with the input network, we find the total impedance using equation 17.1, and then find the input critical frequency.

$$R_{in(lead)} = R_{gen} + R_B \parallel \beta(R_{SW} + r'_e)$$
$$R_{in(lead)} = 50\,\Omega + 20\ \text{k}\Omega \parallel 165(100\,\Omega + 13\,\Omega)$$
$$R_{in(lead)} = 9.7\,\text{k}\Omega$$
$$f_{c(input\ lead)} = \frac{1}{2\pi R_{in(lead)} C_{in}}$$
$$f_{c(input\ lead)} = \frac{1}{2\pi 9.7\,\text{k}\Omega 4.7\,\mu\text{F}}$$
$$f_{c(input\ lead)} = 3.5\,\text{Hz}$$

Turning to the output network, we see that the total resistance is the load plus R_C, or 15 kΩ.

$$f_{c(output\ lead)} = \frac{1}{2\pi R_{out(lead)} C_{out}}$$
$$f_{c(output\ lead)} = \frac{1}{2\pi 15\,\text{k}\Omega 20\,\mu\text{F}}$$
$$f_{c(output\ lead)} = 0.53\,\text{Hz}$$

Finally, we consider the bypass network. First, we find the effective resistance using equation 17.5.

$$R_{bypass(lead)} = R_E \parallel \left(R_{SW} + r'_e + \frac{R_B \parallel R_{gen}}{\beta} \right)$$
$$R_{bypass(lead)} = 4.5\,\text{k}\Omega \parallel \left(100\,\Omega + 13\,\Omega + \frac{20\,\text{k}\Omega \parallel 50\,\Omega}{165} \right)$$
$$R_{bypass(lead)} = 113\,\Omega$$

$$f_{c(bypass\ lead)} = \frac{1}{2\pi R_{bypass(lead)} C_E}$$
$$f_{c(bypass\ lead)} = \frac{1}{2\pi 113\,\Omega 33\,\mu\text{F}}$$
$$f_{c(bypass\ lead)} = 42.7\,\text{Hz}$$

The dominant network is the highest of the lead networks. This is the bypass network at 42.7 Hz. The other two networks are reasonably well below this point, thus, f_l is approximately 42.7 Hz.

Computer Simulation

To crosscheck the calculations, the circuit of Example 17.1 is entered into a simulator, as shown in Figure 17.7.

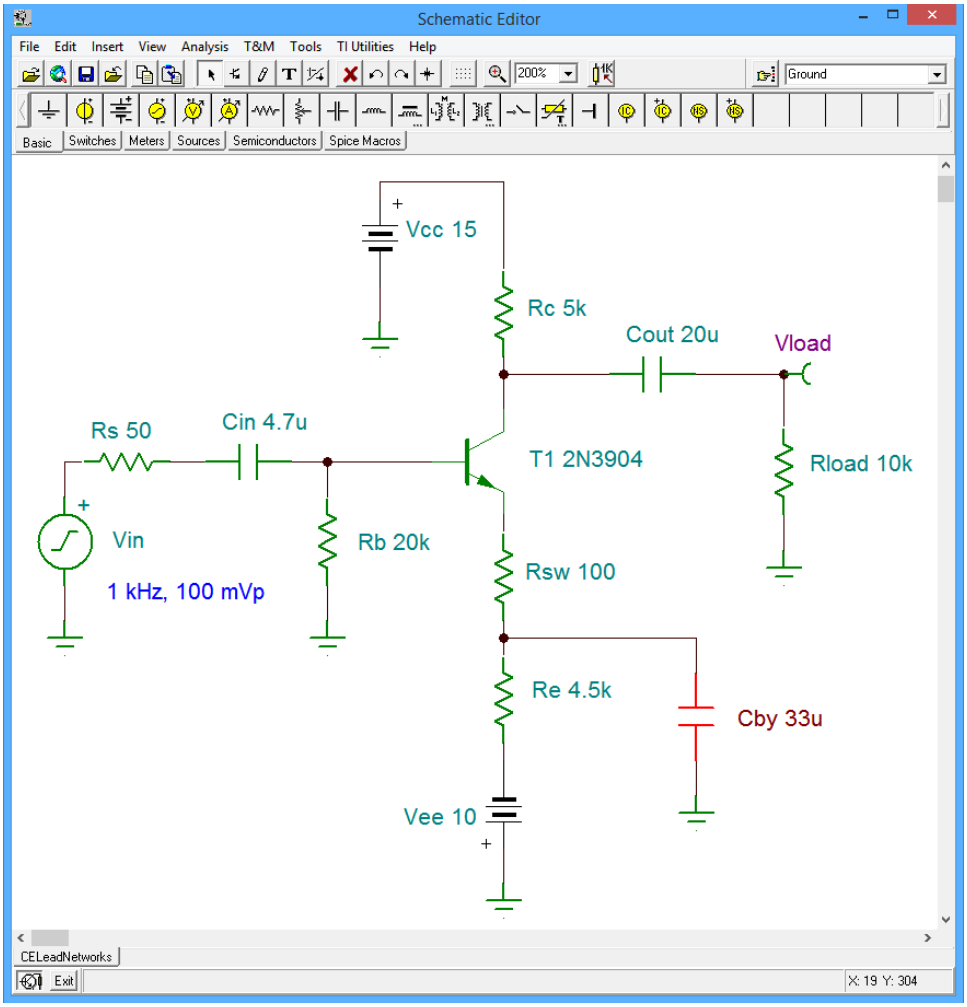


Figure 17.7
BJT low frequency simulation schematic.

An AC analysis is performed on the circuit and a Bode plot is generated. The result is shown in Figure 17.8. The amplifier's midband gain is just below 29 dB. The "3 dB down" corner frequency is at approximately 42.8 Hz, verifying our computed value.

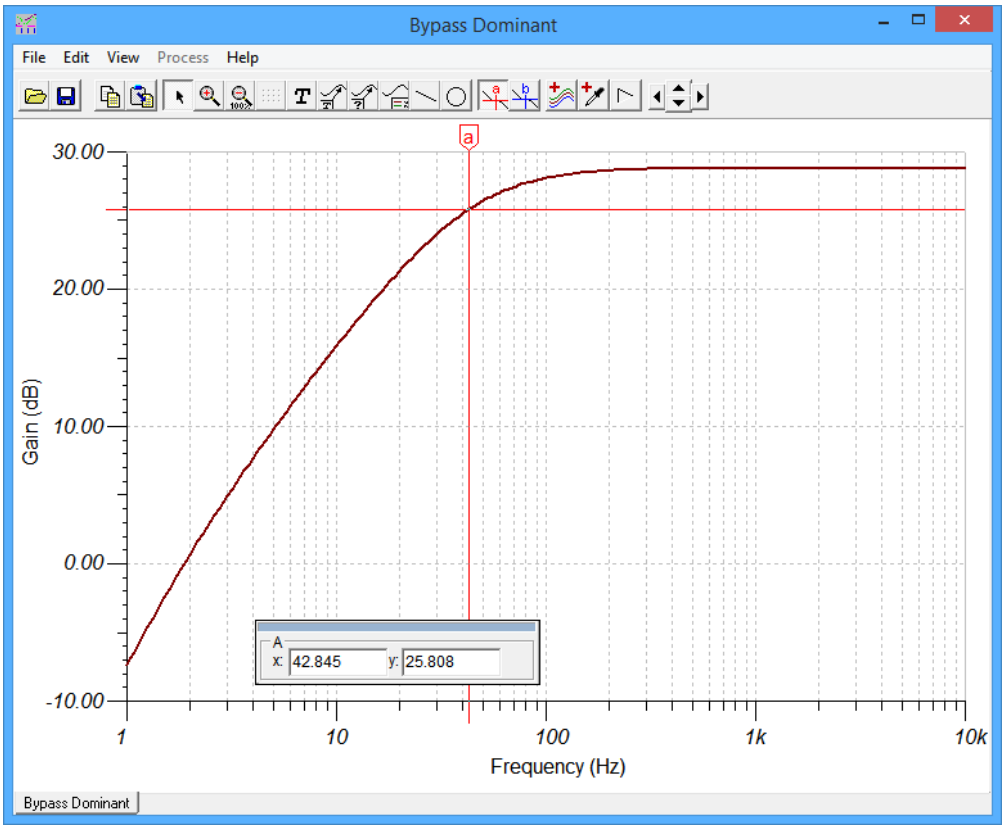


Figure 17.8
BJT low frequency simulation results for original circuit (C_{by} dominant).

The next question is, how do we verify that the analyses of the other two networks are correct? This turns out to be quite easy. As each network is dependent on a single capacitor, all we have to do is change the capacitor value(s) such that the remaining networks become dominant in turn. For example, in order to make the input network dominant, we can decrease C_{in} by a large amount which raises the associated frequency. We also increase C_{by}, dropping its frequency, and thus effectively swapping their places in the spectrum.

To see the input network, C_{in} is reduced by a factor of ten to 470 nF. This raises its expected critical frequency by a factor of ten, to 35 Hz. At the same time, the bypass capacitor is increased by a factor of ten to 330 µF, reducing its frequency to 4.27 Hz. The results of the modified circuit are shown in Figure 17.9. The cursor indicates a frequency of approximately 34 Hz, once again confirming our calculations. It is worth noting that the interaction with the two lower frequency networks is apparent by the increasing steepness of the rolloff in the vicinity of 3 to 5 Hz.

To use the same technique on the output network, we can decrease C_{out} to 200 nF, or a factor of 100. This will push up its critical frequency by the same factor, to 53 Hz. To avoid interaction, C_{in} is changed back to 4.7 µF. The results can be seen in Figure 17.10, producing a frequency of about 54.5 Hz. Again, the agreement is good.

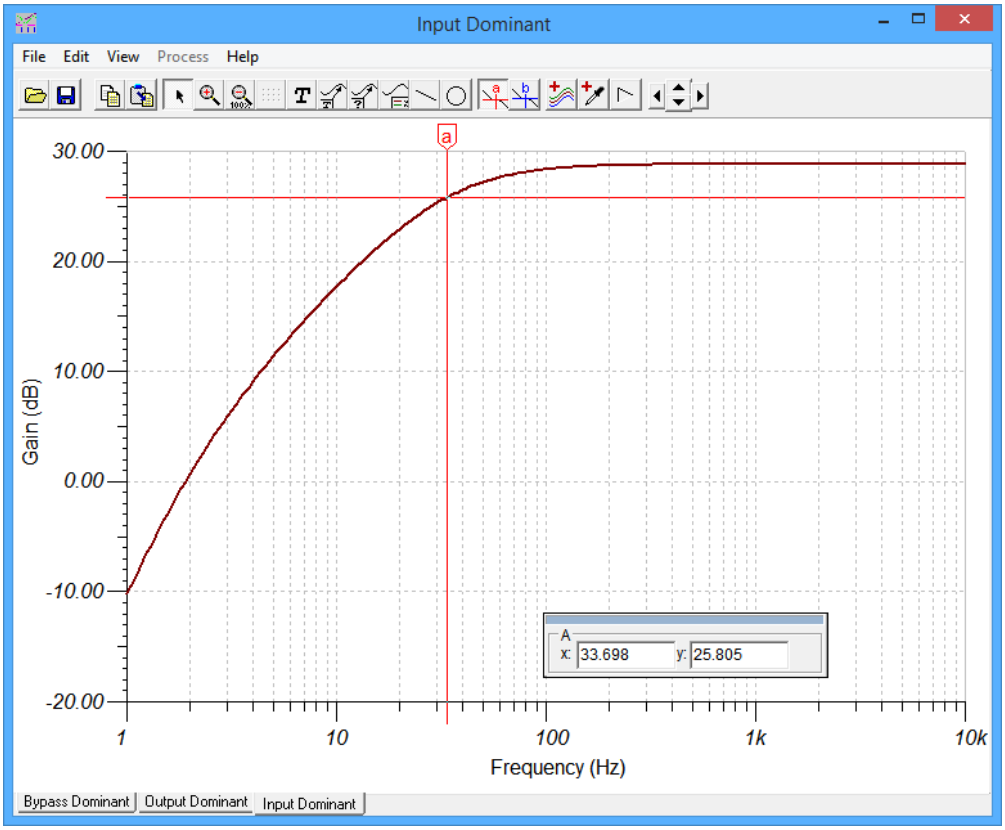


Figure 17.9

BJT low frequency simulation results for second capacitor configuration (C_{in} dominant).

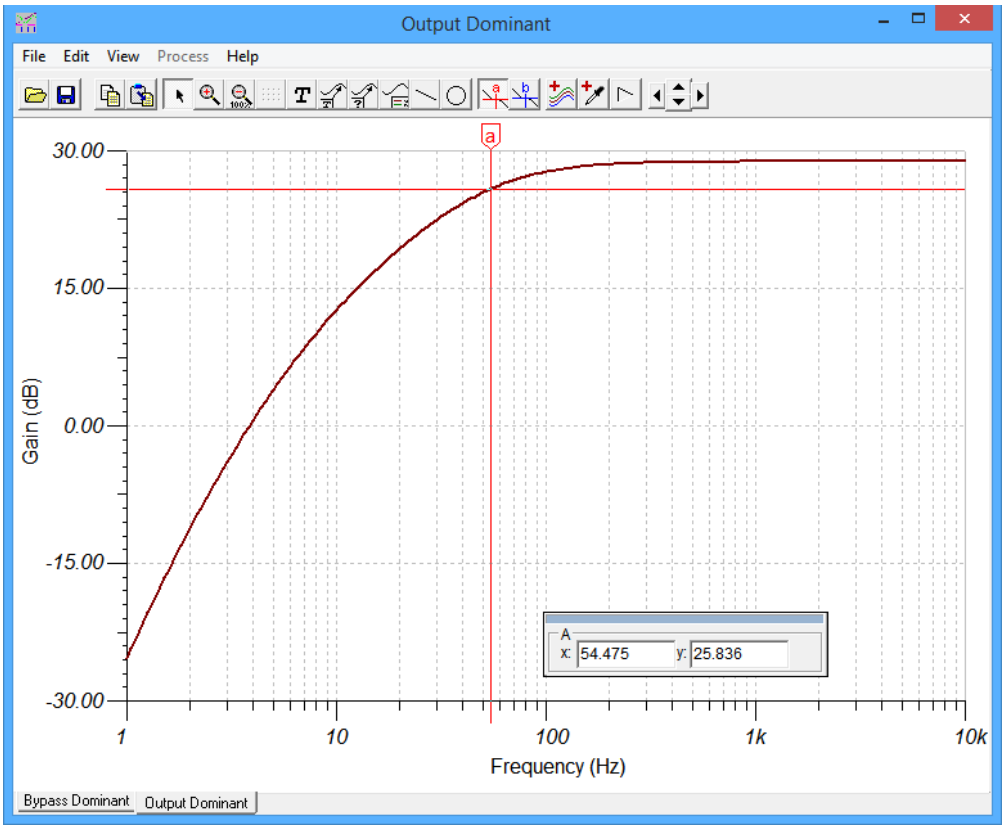


Figure 17.10

BJT low frequency simulation results for third capacitor configuration (C_{out} dominant).

The final item to check for is the "stepped" response of the bypass network that was mentioned earlier. This is a little trickier in that the secondary frequency (i.e., where the rolloff begins to revert to horizontal), is easiest to see if it occurs above the remaining critical frequencies. To see this, we will need to reduce the bypass capacitor by a considerable amount. By dropping to a 330 nF, the critical frequency will be translated up to 4.27 kHz. The "step" frequency will be lower by a factor just a little less than the ratio of R_E to R_{SW}, or a factor of about 40 here. The results of this simulation can be seen in Figure 17.11. In this case, the frequency of 110 Hz is found by moving *above* the flat stepped portion by 3 dB.

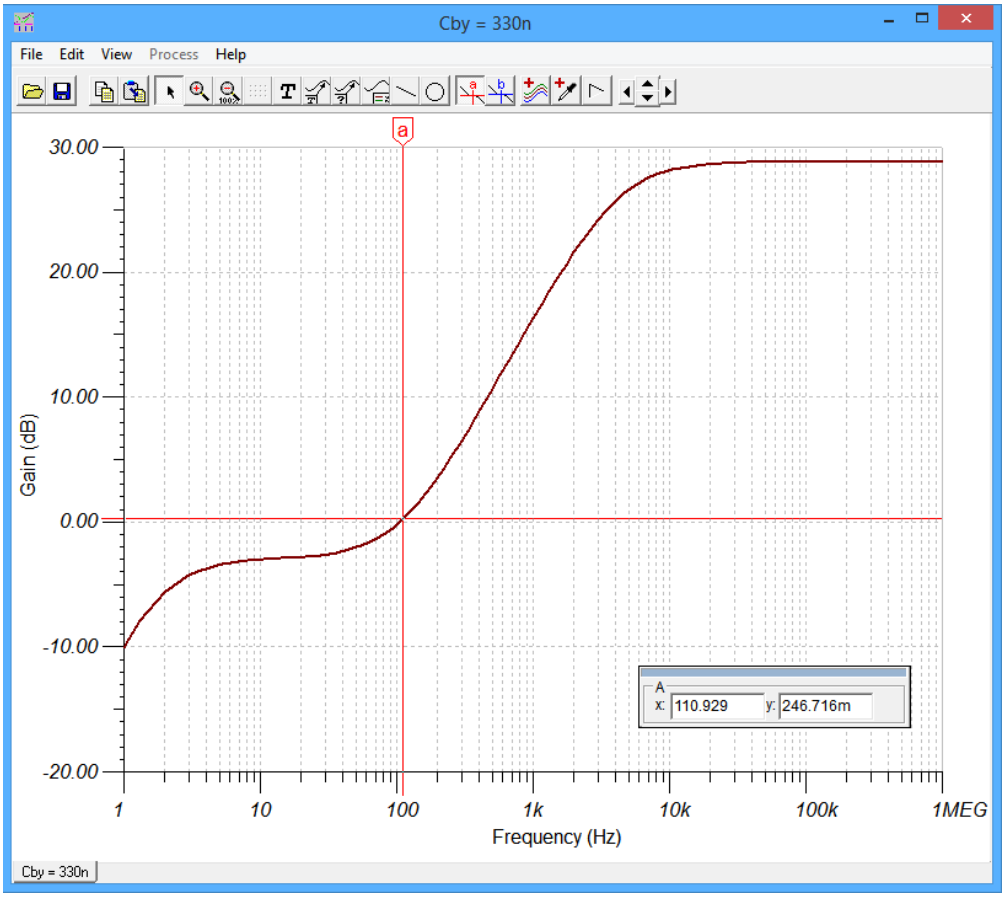


Figure 17.11
BJT low frequency simulation results showing stepped response of bypass network with C_{by} set to 330 nF.

Common Collector (Emitter Follower)

We turn our attention next to the emitter follower configuration. An example circuit is shown in Figure 17.12. The first thing to notice is that this circuit has only two lead networks; one at the input and another at the output. Consequently, unlike the common emitter amplifier, the rolloff rate will be no steeper than 12 dB per octave (40 dB per decade).

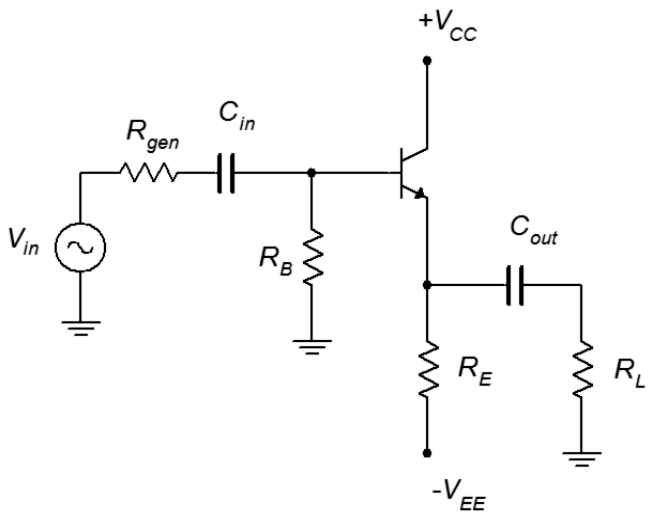


Figure 17.12
BJT common collector or emitter follower circuit.

The analysis of this circuit is similar to that of the common emitter amplifier. The input network formulas are virtually unchanged, except for a minor variation on the computation of $Z_{in(base)}$, namely, that R_L must be placed in parallel with R_E, and this combination serves the place of the former swamping resistor, R_{SW}, in the equation. This assumes that the reactance of C_{out} is negligible.

$$R_{in(lead)} = R_{gen} + R_B \,||\, \beta(R_E \,||\, R_L + r'_e) \tag{17.8}$$

The critical frequency equation is unchanged. As far as the output network is concerned, the effective resistance to the right is R_L, and to the left is R_E in parallel with the impedance looking into the emitter. In other words, it is similar to the expression used for the common emitter's bypass network. The critical frequency equation remains unchanged here as well.

$$R_{output\,(lead\,)} = R_L + R_E \,||\, \left(r'_e + \frac{R_B \,||\, R_{gen}}{\beta} \right) \tag{17.9}$$

Example 17.2

For the circuit of Figure 17.13, determine f_l. Assume $\beta = 100$.

If we assume that the DC base voltage is close to zero, we can approximate the DC emitter current to be approximately 1 mA. This sets r'_e to 26 Ω. For the input network, we find the total impedance using equation 17.8, and then the input critical frequency. The generator impedance is assumed to be zero.

$$R_{in(lead)} = R_{gen} + R_B \parallel \beta(R_E \| R_L + r'_e)$$
$$R_{in(lead)} = 0\,\Omega + 15\ \text{k}\Omega \parallel 100(10\,\text{k}\Omega \parallel 5.3\,\text{k}\Omega + 26\,\Omega)$$
$$R_{in(lead)} = 14.4\,\text{k}\Omega$$
$$f_{c(input\,lead)} = \frac{1}{2\pi R_{in(lead)} C_{in}}$$
$$f_{c(input\,lead)} = \frac{1}{2\pi\, 14.4\,\text{k}\Omega\, 3.3\,\mu\text{F}}$$
$$f_{c(input\,lead)} = 3.35\,\text{Hz}$$

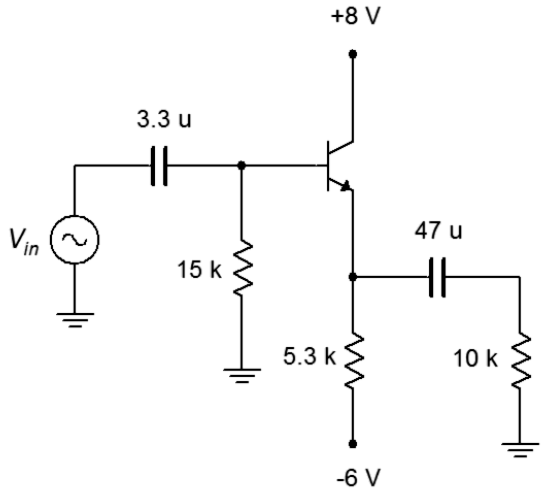


Figure 17.13
Circuit for Example 17.2.

For the output network, the resistance looking into the emitter is the r'_e of 26 Ω, due to the fact that the source shorts the base to ground. This value is placed in parallel with the 5.3 kΩ biasing resistor, for a result of approximately 26 Ω. Finally, the 10 kΩ load is added, leaving us with approximately 10 kΩ total. From here we can find the critical frequency.

$$f_{c(output\,lead)} = \frac{1}{2\pi R_{output(lead)} C_{out}}$$
$$f_{c(output\,lead)} = \frac{1}{2\pi\, 10\,\text{k}\Omega\, 47\,\mu\text{F}}$$
$$f_{c(output\,lead)} = 0.34\,\text{Hz}$$

Clearly, the input network is dominant, and therefore, f_l is approximately 3.35 Hz.

Common Base

Analysis of the common base configuration follows that of the common emitter and common source configurations. Like the emitter follower, it has only two lead networks; one at the input, and a second at the output. The output network is virtually identical to that of the common emitter (essentially a current source feeding two resistors that are separated by an in-line capacitor). As such, the equations for equivalent resistance and critical frequency are unchanged.

For the common base amplifier, the signal source drives the emitter. Thus, the input impedance is found in essentially the same manner as we find the impedance looking toward the emitter of the emitter follower. This is equal to the emitter bias resistor, R_E, in parallel with $Z_{in(emitter)}$. Further, as the base terminal of the common base amplifier is at AC ground, $Z_{in(emitter)}$ is equal to r'_e. Finally, the value of the signal source's internal impedance (i.e., R_{gen}) is added to the total to find the effective resistance. That might sound like a bit of work, but with typical values, this can be approximated as r'_e plus R_{gen}. Being relatively small, this will require a large capacitance if a low critical frequency is desired.

FET Amplifiers

When it comes to lead network response, FET-based circuits have much in common with their BJT counterparts. This is due, at least in part, to the fact that both devices are modeled as controlled current sources. Also, due to the fact that the gate of a FET exhibits a very high input impedance, finding f_l for FET circuits tends to be somewhat easier than for BJTs. Finally, the low frequency analysis for MOSFETs is virtually the same as that for JFETs, further easing our analysis efforts.

Common Source

To start, consider the generic common source amplifier shown in Figure 17.14. To keep this as general as possible, this amplifier includes a swamping resistor.

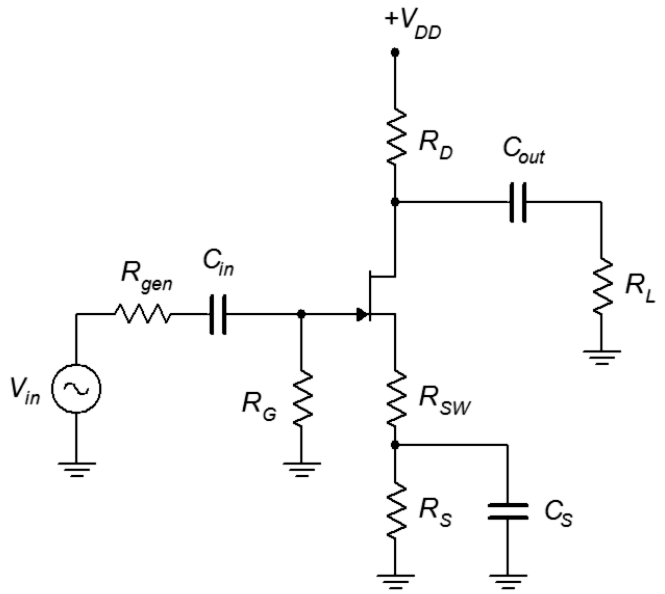


Figure 17.14
FET common source amplifier.

Like the bipolar version, this amplifier has three lead networks; one at the input due to C_{in}, a second at the output due to C_{out}, and a third due to the bypass capacitor, C_S. Once again, the bypass network is not a true lead network, and ultimately will produce a stepped-style response.

In order to determine the effective resistance values for these three networks, we shall use a low frequency AC equivalent of the circuit, as shown in Figure 17.15. This is just an ordinary AC midband equivalent, except that we keep the coupling and bypass capacitors. The objective will be to find the Thevenin equivalent resistance as seen from the perspective of each of the three capacitors.

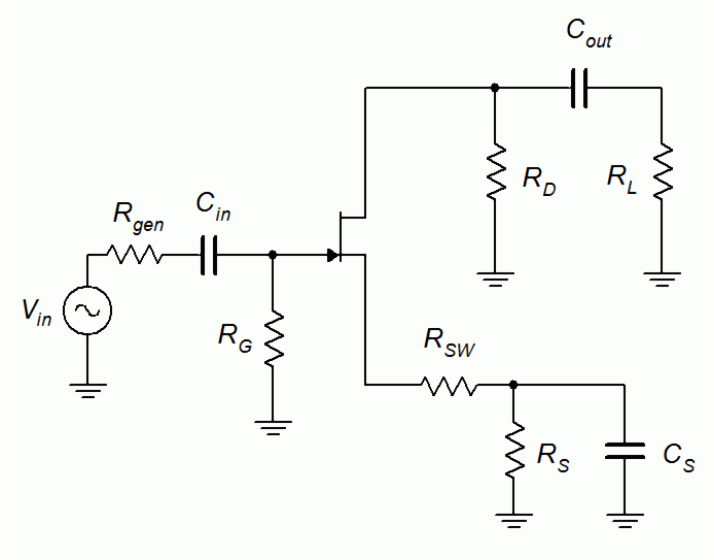


Figure 17.15
Low frequency equivalent of FET common source amplifier.

First, let's consider the input network. To the left of C_{in} we have R_{gen}, with the signal source shorting back to ground. To the right we see the gate biasing resistor, R_G, which is in parallel with $Z_{in(gate)}$. As $Z_{in(gate)}$ is assumed to be infinite at low frequencies, this reduces to just R_G. Thus, the effective resistance is

$$R_{in(lead)} = R_{gen} + R_G \tag{17.10}$$

Therefore, the input lead critical frequency is

$$f_{c(input\,lead)} = \frac{1}{2\pi R_{in(lead)} C_{in}} \tag{17.11}$$

For the output network, we see the same sort of connection that we had with the BJT version. The FET is modeled as a controlled current source and is opened, leaving one resistor on either side of C_{out}. The effective resistance is

$$R_{out(lead)} = R_D + R_L \tag{17.12}$$

The corresponding critical frequency is

$$f_{c(output\,lead)} = \frac{1}{2\pi R_{out(lead)} C_{out}} \tag{17.13}$$

We now turn our attention to the third and final network, that of the source bypass capacitor, C_S. To help clarify the analysis, this network is redrawn in Figure 17.16.

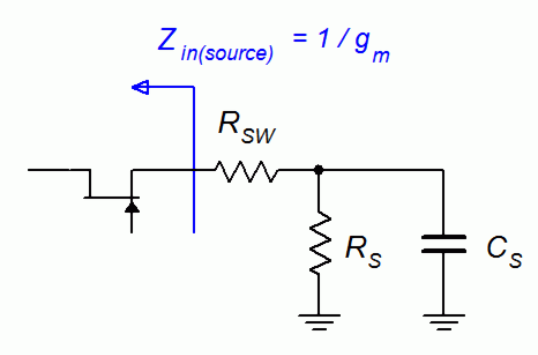


Figure 17.16
Low frequency equivalent of FET bypass network.

Unsurprisingly, this network is not much different from the BJT bypass network. The sole difference is in the computation for the impedance seen looking into the transistor's terminal. The FET version is simpler, being the same as the input impedance of a common gate amplifier, or the reciprocal of the transconductance. The resulting equivalent value as seen by the capacitor is

$$R_{bypass\,(lead)} = R_S \parallel \left(R_{SW} + \frac{1}{g_m} \right) \tag{17.14}$$

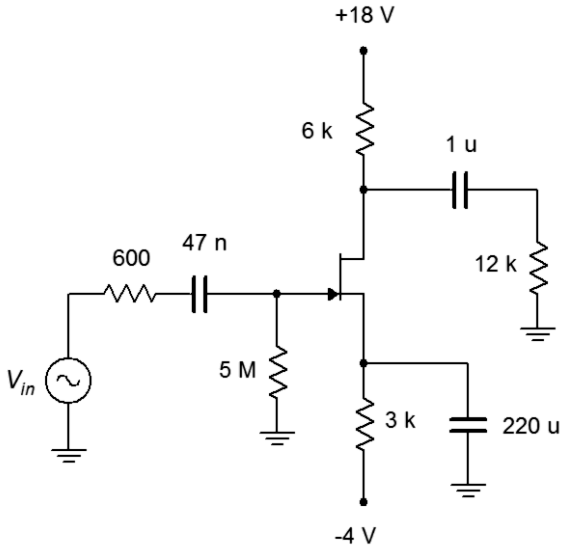
The associated critical frequency is

$$f_{c\,(bypass\,lead)} = \frac{1}{2\pi R_{bypass\,(lead)} C_S} \tag{17.15}$$

If the amplifier is not swamped, simply set R_{SW} to 0. Now for an example:

Example 17.3

Figure 17.17
Circuit for Example 17.3.



Given the circuit of Figure 17.17, determine f_l. Assume g_m is 2 mS.

Beginning with the input network, to the right of the 47 nF capacitor we have 5 MΩ in parallel with $Z_{in(gate)}$ (which approaches infinity), leaving just 5 MΩ. To this we add the 600 Ω generator impedance for a total of virtually 5 MΩ.

$$f_{c\,(input\,lead)} = \frac{1}{2\pi R_{in\,(lead)} C_{in}}$$

$$f_{c\,(input\,lead)} = \frac{1}{2\pi\, 5\,\mathrm{M\Omega}\, 47\,\mathrm{nF}}$$

$$f_{c\,(input\,lead)} = 0.677\,\mathrm{Hz}$$

At the output network, the 1 μF capacitor "sees" the drain biasing resistor and the load, for a total of 18 kΩ.

$$f_{c(\text{output lead})} = \frac{1}{2\pi R_{output(lead)} C_{out}}$$
$$f_{c(\text{output lead})} = \frac{1}{2\pi\, 18\,\text{k}\Omega\, 1\,\mu\text{F}}$$
$$f_{c(\text{output lead})} = 8.84\,\text{Hz}$$

For the unswamped bypass network, the effective resistance is

$$R_{bypass(lead)} = R_S \parallel \frac{1}{g_m}$$
$$R_{bypass(lead)} = 3\,\text{k}\Omega \parallel \frac{1}{2\ \text{mS}}$$
$$R_{bypass(lead)} = 429\,\Omega$$

Therefore, the critical frequency for the bypass network is

$$f_{c(\text{bypass lead})} = \frac{1}{2\pi R_{bypass(lead)} C_S}$$
$$f_{c(\text{bypass lead})} = \frac{1}{2\pi\, 429\,\Omega\, 220\,\mu\text{F}}$$
$$f_{c(\text{bypass lead})} = 1.69\,\text{Hz}$$

The output network is dominant and thus the expected f_l is approximately 8.84 Hz. The other networks are not that far off, so some interaction can be expected. This would push the true f_l to a slightly higher frequency.

Common Drain (Source Follower)

The lead network analysis for the common drain configuration is similar to that of the BJT common collector follower. Further, some of the equations are identical to those used for a non-swamped common source amplifier. To investigate, we shall make use of a simple source follower, as seen in Figure 17.18. To begin with, the input network is identical to the common source configuration. This yields an equivalent resistance of

$$R_{in(lead)} = R_{gen} + R_G \tag{17.16}$$

Consequently, the input lead critical frequency must be

$$f_{c(\text{input lead})} = \frac{1}{2\pi R_{in(lead)} C_{in}} \tag{17.17}$$

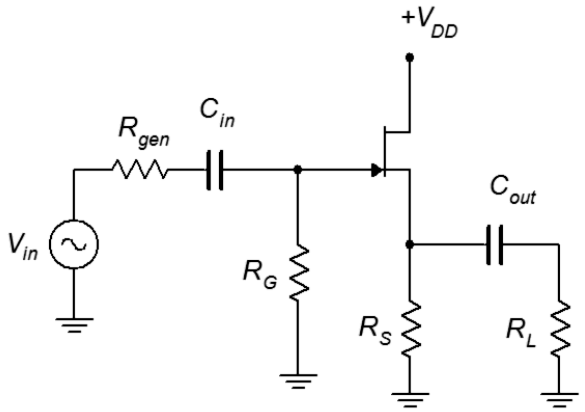


Figure 17.18
Simple common drain follower.

At the output network, the capacitor "sees" the load to the right. To the left, we find the source biasing resistor, R_S, in parallel with the value looking into the source, which we know is equal to the reciprocal of the transconductance.

$$R_{bypass(lead)} = R_L + R_S \parallel \frac{1}{g_m} \tag{17.18}$$

Therefore, the output lead critical frequency is

$$f_{c(output\,lead)} = \frac{1}{2\pi R_{output(lead)} C_{out}} \tag{17.19}$$

And now for something slightly different: an example using a design target frequency.

Example 17.4

Determine values for the capacitors in the circuit of Figure 17.19 to achieve a lower break frequency, f_l, of 20 Hz. Assume g_m is 4 mS.

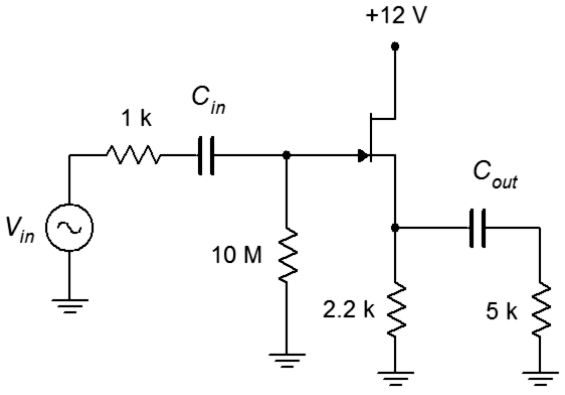


Figure 17.19
Circuit for Example 17.4.

Unlike the earlier analysis examples, here we have a known value for the lower corner frequency. Thus, we will need to find the effective resistance values for the input and output networks, and then solve for the capacitances. It is important to note that we cannot make both networks critical at 20 Hz. Doing so would create a 6 dB loss rather than a 3 dB loss at that frequency. This means that the real 3 dB frequency would be higher than 20 Hz. What we will do instead is find the capacitor values that will each achieve 20 Hz, and then increase the size of one of them in order to push its critical frequency down to the point

where there is no interference. Usually, this will be the smaller of the two as this will tend to minimize the size of the components.

First, the input network. Given the size differential between the generator and gate resistances, we can approximate the total value as 10 MΩ. For a critical frequency of 20 Hz, we find

$$C_{in} = \frac{1}{2\pi R_{in(lead)} f_{c(input\,lead)}}$$
$$C_{in} = \frac{1}{2\pi 10\,\mathrm{M\Omega} 20\,\mathrm{Hz}}$$
$$C_{in} = 796\ \mathrm{pF}$$

Moving to the output network, the effective resistance is

$$R_{bypass(lead)} = R_L + R_S \parallel \frac{1}{g_m}$$
$$R_{bypass(lead)} = 5\ \mathrm{k\Omega} + 2.2\ \mathrm{k\Omega} \parallel \frac{1}{4\ \mathrm{mS}}$$
$$R_{bypass(lead)} = 5225\,\Omega$$

The corresponding capacitance for 20 Hz is

$$C_{out} = \frac{1}{2\pi R_{out(lead)} f_{c(output\,lead)}}$$
$$C_{out} = \frac{1}{2\pi 5225\,\Omega 20\,\mathrm{Hz}}$$
$$C_{out} = 1.52\,\mu\mathrm{F}$$

For the final step, the obvious choice here would be to increase the value of C_{in}, perhaps by a factor of ten, leaving us with just shy 8 nF. The nearest standard value would be 8.2 nF. C_{out} would not change; the nearest standard value being 1.5 μF.

Common Gate

Finally, our tour of lead network analysis arrives at the common gate configuration. As you might have guessed, there is much in common here with the common base configuration as well as the common source and common drain configurations. Like the source follower, the common gate has only two lead networks; one at the input and one at the output. First off, the output network is identical to that of the common source amplifier and uses the same equations. The effective resistance is R_D in parallel with R_L. For the input network, the analysis is very similar to the output network of the source follower. In both cases we need to determine the impedance

looking toward the source terminal. Recall that the impedance looking toward the FET's source is the source bias resistor in parallel with the reciprocal of the transconductance. To this we add the generator impedance to arrive at the effective ohmic value.

$$R_{input(lead)} = R_{gen} + R_S \parallel \frac{1}{g_m} \tag{17.20}$$

Before leaving this section, there are two items worth repeating: First, not all amplifiers have lead networks, and second, design is a relatively straightforward process of picking the right values for coupling and bypass capacitors. If you want to extend the frequency response to lower and lower frequencies, just increase the sizes of these capacitors. The frequency will drop by the same ratio as the capacitor increase. Similarly, to limit the low frequency response to a higher value, decrease the sizes of the capacitors. As we shall see shortly, designing is not always quite so easy at the other end of the spectrum.

17.3 High Frequency Response

In this section we shall examine the performance of amplifiers at the high end of their range. We are interested in what causes the limits and how to predict the various critical frequencies involved. Further, we would like to determine ways of controlling those limits. It is important to remember that, unlike the low frequency response, all amplifiers without exception exhibit high frequency limits. In other words, while it is possible to design amplifiers without an f_1, all amplifiers will have an f_2. Our discussion mostly will focus on common emitter and common source amplifiers. As these are inverting configurations, Miller's Theorem will be of considerable use.

Lag Networks

The high frequency response of amplifiers entails lag networks. Remember, a lag network can be considered to be a frequency-dependent voltage divider with the resistor in-line with the signal and the capacitor situated across the load. As frequency increases, the reactance decreases, and thus, the voltage available to the load is reduced. It turns out that a single stage amplifier will have two lag networks; one at the input and a second at the output. For multi-stage amplifiers, the output of one stage will combine with the input of the following stage to create a single, interstage lag network.

Unlike the low end of the spectrum, the high end limits are caused by capacitors that you can't see. These are undesirable parasitic capacitances found within the devices themselves, and due to printed circuit board traces and cabling. Transistor capacitances can range from just a few picofarads up into the nanofarad range,

depending on transistor type and size. Individual carbon or metal film resistors are around one picofarad or less, although it is possible to find surface mount versions that are below a tenth of a picofarad.

Regarding the wiring capacitance, this can vary greatly. Simple circuit board traces may be just a few picofarads or less, but loads can also include cabling that is measured in tens of picofarads per foot (perhaps 100 picofarads per meter). The extra capacitance from resistors, PCB traces, and the like all appear in parallel, and can be lumped together into a single value for each lag network that we will generically refer to as $C_{(stray)}$.

BJT Amplifiers

Let us begin with a typical common emitter amplifier, as shown in Figure 17.20.

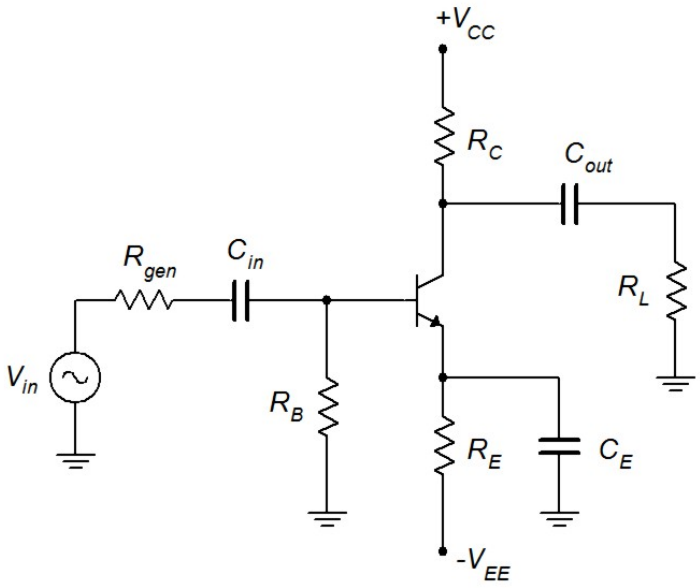
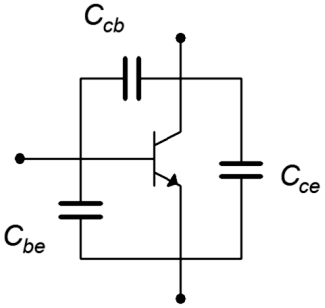


Figure 17.20
A typical common emitter amplifier.

For the midband analysis, we assumed that the coupling and bypass capacitors are effective shorts. At the even higher frequencies that we are now considering, this must still be the case. Before we attempt to redraw the circuit for high frequency analysis, we must remember that there are small parasitic capacitances found between the transistor's terminals, as shown in Figure 17.21.

The parasitic capacitances of a small signal BJT are quite small when compared to the values used for coupling and bypass capacitors. Typical values are in the single digit picofarad range. In fact, the collector–emitter capacitance, C_{ce}, is usually small enough to ignore.

Figure 17.21
BJT parasitic capacitances.



With this model, the amplifier can now be redrawn for the high frequency case, as shown in Figure 17.22. Of primary importance, we have added the relevant device capacitances. Note that C_{cb} is in the Miller position; that is, it straddles the amplifier from input to output. Remember, Miller's Theorem only applies to inverting amplifiers, like this one, not to non-inverting amplifiers. The theorem allows us to split this capacitor into two capacitors, one across the input and another across the output, that will have the same impact as the original.

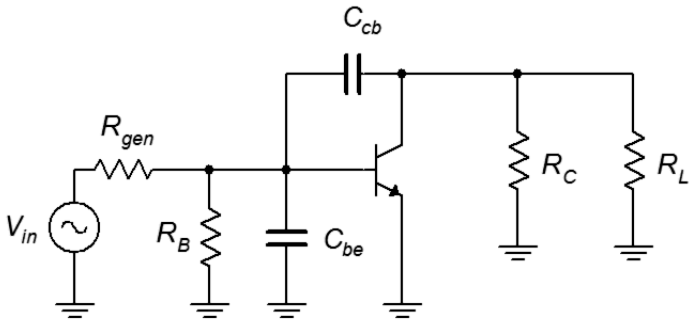


Figure 17.22
High frequency equivalent of CE amplifier with transistor capacitances.

By applying Miller's Theorem, we can create an equivalent input network and an equivalent output network. The input network is shown in Figure 17.23.

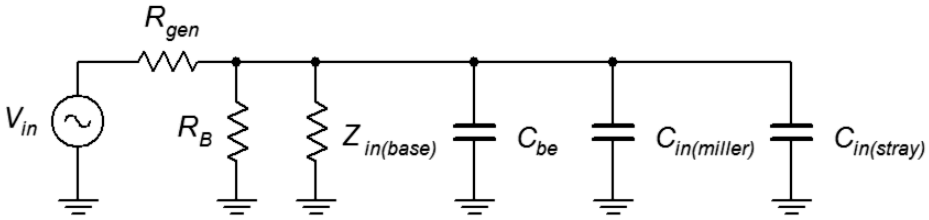


Figure 17.23
High frequency equivalent of BJT input network.

For this network, we have moved $Z_{in(base)}$ to the left to avoid visual clutter. This is not a problem as all of the elements are in parallel. We have also included a lumped $C_{in(stray)}$ value along with $C_{in(miller)}$, the Millerized value of C_{cb}. As a reminder, this increases the capacitance by the magnitude of the amplifier's gain plus 1. Thus,

$$C_{in(miller)} = C_{cb} (|A_v| + 1) \tag{17.21}$$

In many amplifiers, $C_{in(miller)}$ is the largest capacitance in the group by a large margin. As all of the capacitances are in parallel, the total is just the sum of their values. Also, all of the resistances are in parallel when looking back from the capacitors.

$$C_{input(lag)} = C_{be} + C_{in(miller)} + C_{in(stray)} \tag{17.22}$$

$$R_{input(lag)} = R_{gen} \,||\, R_B \,||\, Z_{in(base)} \tag{17.23}$$

And, as expected, we find

$$f_{c(input\,lag)} = \frac{1}{2\pi R_{input(lag)} C_{input(lag)}} \tag{17.24}$$

The situation at the output network is similar except that the network is driven by a parallel current source instead of series voltage source. The Thevenin equivalents will be the same, though. Refer to Figure 17.24.

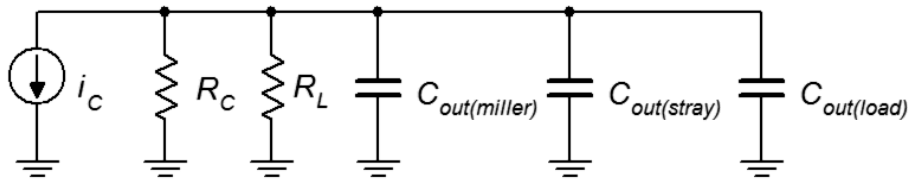


Figure 17.24
High frequency equivalent of BJT output network.

A reminder for $C_{out(miller)}$ is useful at this point.

$$C_{out(miller)} = C_{cb} \frac{|A_v| + 1}{|A_v|} \tag{17.25}$$

By inspection, we find

$$C_{output(lag)} = C_{out(load)} + C_{out(miller)} + C_{out(stray)} \tag{17.26}$$

$$R_{output(lag)} = R_C \parallel R_L \tag{17.27}$$

And finally, we reach

$$f_{c(output\,lag)} = \frac{1}{2\pi R_{output(lag)} C_{output(lag)}} \tag{17.28}$$

The lower result of equations 17.24 and 17.28 will be the dominant frequency, and thus, is the system f_2.

At this point, there is a practical wrinkle. Unfortunately, manufacturer's data sheets do not list the values of C_{cb} and C_{be} directly. Instead, they list other values that are easier to measure. Fortunately, it's easy enough to convert said data to what we need. Data sheets usually list values for C_{ibo} and C_{obo} (and sometimes, these values go by different names, so be forewarned). The measurement circuit for these parameters is shown in Figure 17.25.

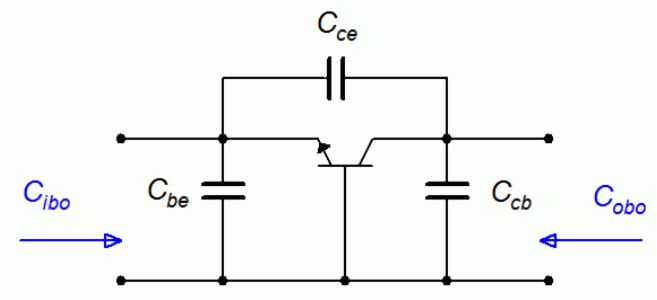


Figure 17.25
Measuring C_{ibo} and C_{obo}.

The subscript *ibo* stands for "**i**nput in common **b**ase configuration with an **o**pen output". Similarly, *obo* refers to "**o**utput in common **b**ase configuration with an **o**pen input". By observation, C_{ibo} is equal to C_{be} in parallel with the series combination of C_{ce} and C_{cb}. Recalling that capacitors in series combine in the same manner as resistors in parallel, and that C_{ce} is usually very small, we can see that

$$C_{be} \approx C_{ibo} \tag{17.29}$$

$$C_{cb} \approx C_{obo} \tag{17.30}$$

To make things just a little more fun, these two capacitances are not fixed values. In fact their precise values depend on the transistor's bias voltages. For example, consider C_{cb}. This parasitic capacitance is caused by the reverse biased collector-base junction. You can think of the two terminals as the plates of a capacitor and the depletion region between them as the capacitor's dielectric. As the reverse bias potential increases, this widens the depletion region, which is akin to increasing the plate separation distance. The result is that the capacitance decreases with increasing reverse voltage. Fortunately, manufacturers will publish curves for C_{ibo} and C_{obo} in their device data sheets. An example of such is shown in Figure 17.26 for the popular 2N3904.

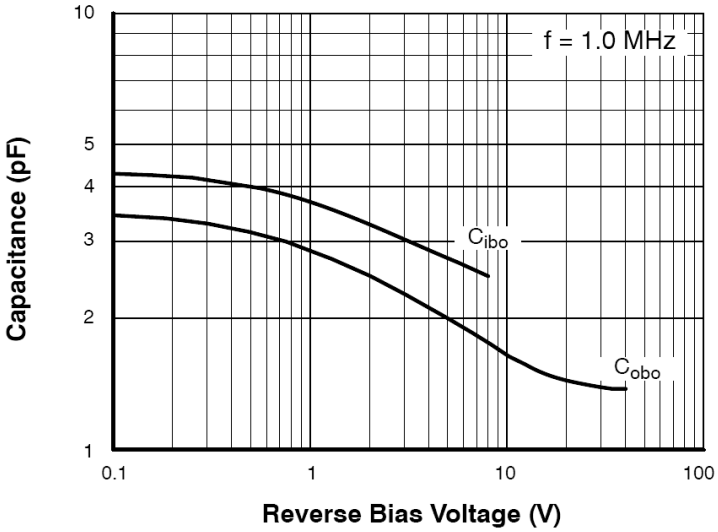


Figure 17.26
C_{ibo} and C_{obo} curves.
Used with permission from SCILLC dba ON Semiconductor.

For instance, in this curve we can see that C_{obo} is about 3.5 picofarads at very low reverse voltages and decreases to less than 2 picofarads above 5 volts. Data sheets will also list values for these capacitances but these are usually maximums measured at zero volts. For highest accuracy, it is best to perform a DC bias calculation to determine the reverse voltage, and then use a graph like this to obtain the capacitance value.

A couple of final twists: first, for amplifiers being driven by a particularly low impedance (i.e., R_{gen}), it is advisable to use the more accurate transistor model that includes the base spreading resistance, r'_b. This value appears in series with R_{gen}, so the two values should be added together before placing the sum in parallel with the other input resistors. The second twist is the issue of a swamped amplifier. A reexamination of the input network for the swamped case shows that $C_{in(miller)}$ is not in parallel with C_{be} due to R_{SW}. In practice, $C_{in(miller)}$ is usually much larger than C_{be}, and placing them in parallel generally will yield acceptable results.

Time for an illustrative example.

Example 17.5

For the circuit of Figure 17.27, determine the input and output lag critical frequencies, and the system f_2. Assume $\beta = 100$, $C_{obo} = 3$ pF, $C_{ibo} = 5$ pF. and ignore any stray or load capacitances.

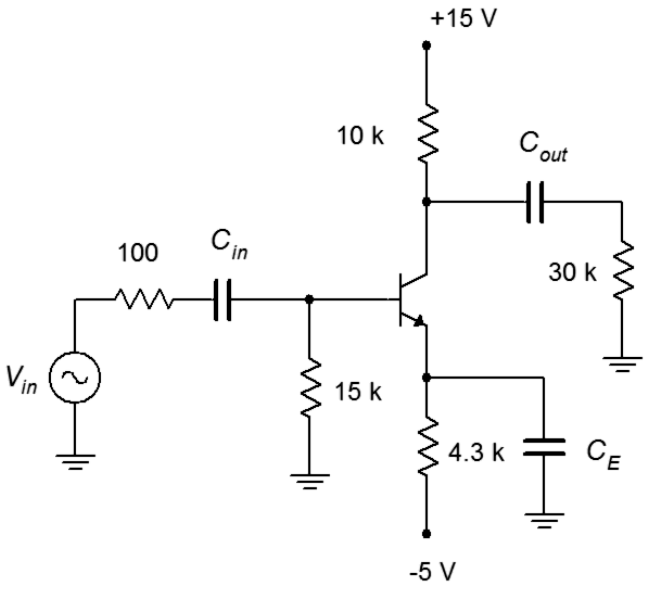


Figure 17.27
Circuit for Example 17.5.

First, we need to determine the system voltage gain so that we can find the input and output Miller capacitances. To do that, we need r'_e, and thus, I_E.

If we assume that the DC base voltage is at nearly ground, then approximately 4.3 VDC drops across R_E, leaving $I_E = 1$ mA, and thus, $r'_e = 26\ \Omega$. The effective AC collector impedance is 10 k Ω in parallel with 30 kΩ , or 7.5 kΩ. This yields a voltage gain of 288, inverting.

C_{cb} is approximately equal to C_{obo}, and C_{be} is approximately equal to $C_{ibo.}$ Therefore, the input and output Miller capacitances are:

$$C_{in(miller)} = C_{cb}(|A_v|+1)$$
$$C_{in(miller)} = 3\text{ pF}(|-288|+1)$$
$$C_{in(miller)} = 867\text{ pF}$$

$$C_{out(miller)} = C_{cb}\frac{|A_v|+1}{|A_v|}$$
$$C_{out(miller)} = 3\,\text{pF}\frac{|-288|+1}{|-288|}$$
$$C_{out(miller)} \approx 3\,\text{pF}$$

For the input network's resistance, we will need $Z_{in(base)}$. In this circuit, that's $\beta\, r'_e$, or 2.6 kΩ. Thus, the input lag resistance is:

$$R_{input(lag)} = R_{gen} \parallel R_B \parallel Z_{in(base)}$$
$$R_{input(lag)} = 100\,\Omega \parallel 15\,\text{k}\,\Omega \parallel 2.6\,\text{k}\,\Omega$$
$$R_{input(lag)} = 95.7\,\Omega$$

Ignoring stray input capacitance, the total input lag capacitance is:

$$C_{input(lag)} = C_{be} + C_{in(miller)}$$
$$C_{input(lag)} = 5\,\text{pF} + 867\,\text{pF}$$
$$C_{input(lag)} = 872\,\text{pF}$$

Thus, the input lag f_c is:

$$f_{c(input\,lag)} = \frac{1}{2\pi R_{input(lag)} C_{input(lag)}}$$
$$f_{c(input\,lag)} = \frac{1}{2\pi 95.7\,\Omega 872\,\text{pF}}$$
$$f_{c(input\,lag)} = 1.91\,\text{MHz}$$

Clearly, the Miller input capacitance dominates. This is typical in high gain amplifiers of this type.

The output network analysis is similar. The output lag resistance is:

$$R_{output(lag)} = R_C \parallel R_L$$
$$R_{output(lag)} = 10\,\text{k}\,\Omega \parallel 30\,\text{k}\,\Omega$$
$$R_{output(lag)} = 7.5\,\text{k}\,\Omega$$

Once again, ignoring any stray or load capacitance, the total output lag capacitance is:

$$C_{output(lag)} = C_{out(miller)}$$
$$C_{output(lag)} = 3\,\text{pF}$$

The output lag f_c is:

$$f_{c(output\,lag)} = \frac{1}{2\pi R_{output\,(lag)} C_{output\,(lag)}}$$
$$f_{c(output\,lag)} = \frac{1}{2\pi 7.5\,\mathrm{k\Omega}\,3\,\mathrm{pF}}$$
$$f_{c(output\,lag)} = 7.07\,\mathrm{MHz}$$

Here we see that, even though the capacitance is much lower than it is in the input network, the increased resistance places the critical frequency in the same neighborhood. Between the two, the input network is lower, and therefore, dominant. Thus, the system f_2 is approximately 1.91 MHz. In reality, the true "3 dB down frequency" will be somewhat lower than this due to the fact that the two critical frequencies are relatively close. Consequently, there will be some non-zero loss for the output network at 1.91 MHz.

The analysis for common collector followers and common base amplifiers is similar to the preceding, but with one important caveat: these amplifiers are non-inverting, and thus, Miller's Theorem does not apply. In the common collector follower, C_{bc} winds up in parallel with C_{be} across the input, while C_{ce} appears across the output. For the common base amplifier, C_{be} is across the input, C_{cb} is across the output, and the bridged C_{ce} usually is small enough to ignore.

FET Amplifiers

To start our discussion of FET high frequency response, let's consider a typical common source amplifier, such as the one shown in Figure 17.28. This circuit features a JFET, but for the most part, the analysis for MOSFETs is the same. To be clear, there are some differences that arise from the availability of alternate biasing schemes, and power MOSFETs can exhibit considerably higher values for gate capacitance than small signal JFETs, but these differences are still compatible with the analysis. For consistency, we'll stick with JFETs in the following discussion and examples.

The amplifier of Figure 17.28 features three coupling and bypass capacitors that shape the circuit's low frequency response. Like the BJT amplifiers examined in the prior section, the high frequency response of this circuit is controlled by capacitances that, in general, you don't see on a schematic. Here, we deal with the same set of capacitances; namely, the parasitic device capacitances of the FET, those of associated circuit components such as resistors, stray wiring capacitances, and load capacitance.

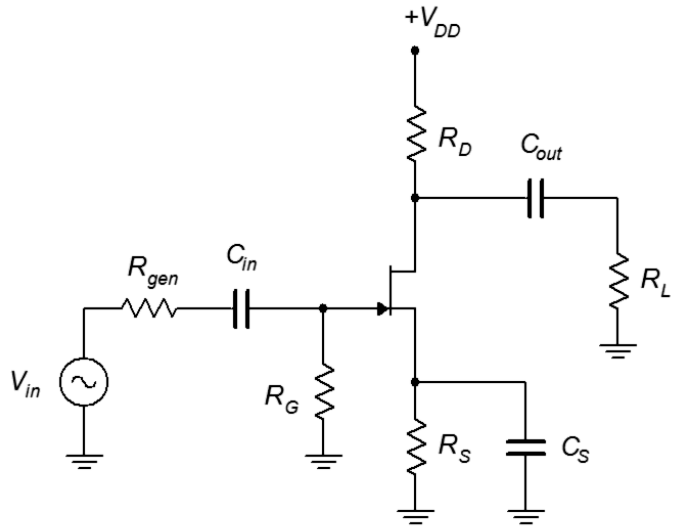
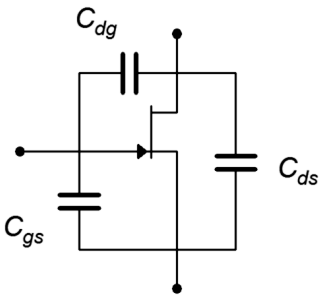


Figure 17.28
A typical common source amplifier.

Like the BJT amplifier, this circuit will also have two lag networks; one at the input and a second at the output. The FET itself exhibits parasitic capacitances between its terminals, as shown in Figure 17.29. The drain–source capacitance, C_{ds}, tends to be small enough to ignore in most amplifiers. For common source amplifiers, C_{dg} is in the Miller position, and C_{gs} is at the input. For typical small signal devices, these parasitics are usually measured in the single digit picofarad range, although for large signal power devices, C_{gs} can be in the nanofarad range.

Figure 17.29
FET parasitic capacitances.



Using this transistor model, we can now create a high frequency model for the amplifier. We begin by shorting the coupling and bypass capacitors, replacing the DC source with its ideal internal resistance (a short), and then redrawing. Perhaps unsurprisingly, it is similar to its BJT counterpart. The resulting circuit is shown in Figure 17.30.

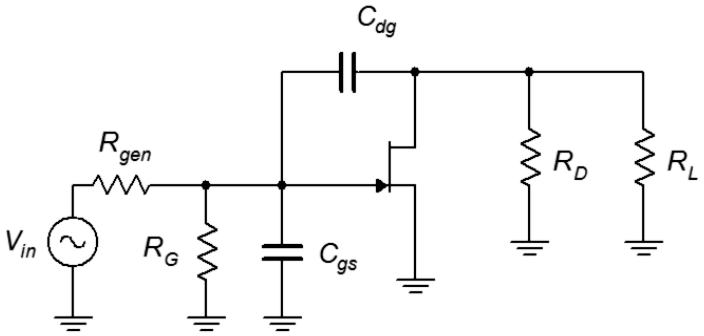


Figure 17.30
High frequency equivalent of CS amplifier with transistor capacitances.

Continuing the process, we can Millerize C_{dg}, which splits into input and output equivalents.

$$C_{in(miller)} = C_{dg}\,(|A_v| + 1) \tag{17.31}$$

$$C_{out\,(miller)} = C_{dg}\frac{|A_v|+1}{|A_v|} \tag{17.32}$$

Focusing on the input network, the equivalent circuit is shown in Figure 17.31.

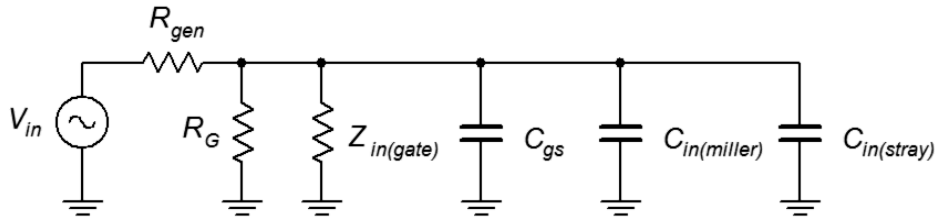


Figure 17.31
High frequency equivalent of FET input network.

By inspection, we can determine the equivalent lag capacitance and resistance.

$$C_{input(lag)} = C_{gs} + C_{in(miller)} + C_{in(stray)} \tag{17.33}$$

$$R_{input(lag)} = R_{gen} \,||\, R_G \,||\, Z_{in(gate)} \tag{17.34}$$

Given that the values of $Z_{in(gate)}$ and R_G are often very large in comparison to R_{gen}, in many circuits we can approximate Equation 17.34 as $R_{input(lag)} = R_{gen}$.

Finally, we arrive at the formula for critical frequency of the input network:

$$f_{c(input\,lag)} = \frac{1}{2\pi R_{input\,(lag)} C_{input\,(lag)}} \tag{17.35}$$

For the output network, the high frequency equivalent is shown in Figure 17.32.

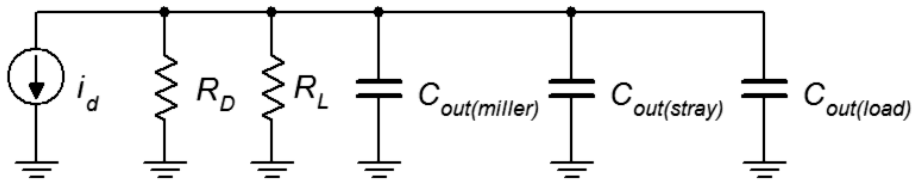


Figure 17.32
High frequency equivalent of FET output network.

The effective capacitance is the parallel combination of the three capacitances. Looking back from that combination while opening the current source shows that the effective resistance is the parallel combination of R_D and R_L.

$$C_{output(lag)} = C_{out(miller)} + C_{out(stray)} + C_{out(load)} \tag{17.36}$$

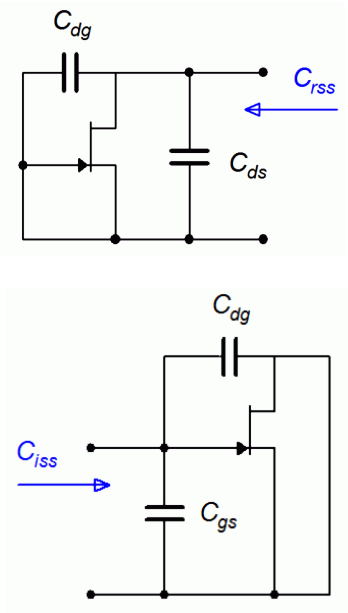
$$R_{output(lag)} = R_D \parallel R_L \tag{17.37}$$

And this leads us to,

$$f_{c(output\,lag)} = \frac{1}{2\pi R_{output\,(lag)} C_{output\,(lag)}} \tag{17.38}$$

Manufacturer's datasheets do not always list values for our idealized device capacitances. Instead, the often listed values are C_{iss} and C_{rss}, due to ease of measurement. These measurement circuits are shown in Figure 17.33.

Figure 17.33
Measurement of C_{rss} and C_{iss}.



The *rss* subscript stands for "**r**everse in common **s**ource configuration with **s**horted input", while the subscript *iss* stands for "**i**nput in common **s**ource configuration with **s**horted output". For the top circuit, shorting the input removes C_{gs}. This leaves C_{ds} in parallel with C_{dg}. Typically, C_{ds} is small enough to ignore, and thus

$$C_{dg} \approx C_{rss} \tag{17.39}$$

In the lower circuit, note that shorting the output removes C_{ds} and also places C_{dg} in parallel with C_{gs}. From this, we can deduce that

$$C_{gs} \approx C_{iss} - C_{rss} \tag{17.40}$$

At this point, astute readers might observe that the FET's parasitic device capacitances will vary with the applied reverse bias, just as was the case with the BJT. As such, the manufacturer may give graphical data in place of simple numerics, as shown in Figure 17.34. In this case, the manufacture has conveniently (for us, anyway) plotted C_{gs} and C_{gd} instead of C_{iss} and C_{rss}. Note also that the graph explicitly states that C_{ds} is negligible. As expected, the increasing reverse bias potential widens the depletion region which results in reduced capacitance. For example, C_{gs} is just over 10 picofarads in the region near a zero volt potential but drops to 4 picofarads at 10 volts, and just over 3 picofarads at 30 volts. A similar situation can be seen for C_{gd}.

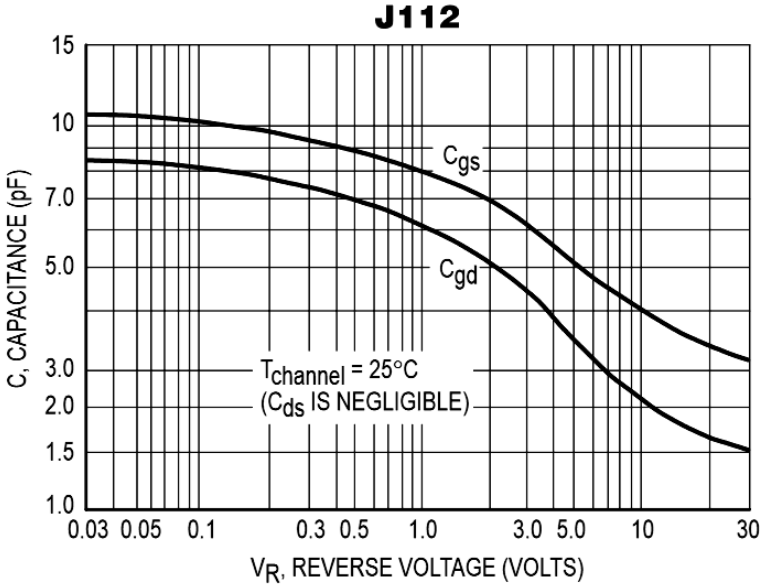


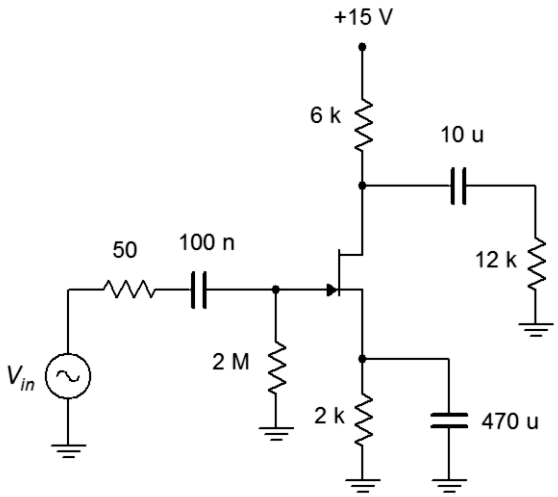
Figure 17.34
C_{gs} and C_{gd} curves.
Used with permission from SCILLC dba ON Semiconductor.

In this discussion we have focused on unswamped amplifiers. For swamped amplifiers, an inspection of the equivalent circuit reveals that C_{gs} is no longer in parallel with $C_{in(miller)}$ and $C_{in(stray)}$. In many circuits, the loss of accuracy in treating them as though they are in parallel will be acceptable. This is more likely to be a problem when gains are particularly low or when C_{gs} is much larger than C_{dg}. Finally, for common gate and common drain followers, the issues for FETs are similar to those of BJT common base and common collector followers, namely that these configurations are non-inverting and Miller's Theorem does not apply.

Once again, it's time for a few illustrative examples.

Example 17.6

Figure 17.35
Circuit for Example 17.6.



For the circuit of Figure 17.35, determine the input and output lag critical frequencies, and the amplifier's f_2. Assume g_m = 1.25 mS, C_{rss} = 1.5 pF, and C_{iss} = 4.5 pF. Ignore stray and load capacitance.

First, we need to translate the device capacitances into a more convenient form.

$$C_{gd} = C_{rss} = 1.5 \text{ pF}$$

$$C_{gs} = C_{iss} - C_{rss}$$
$$C_{gs} = 4.5\,\text{pF} - 1.5\,\text{pF}$$
$$C_{gs} = 3\,\text{pF}$$

Next, we can determine the Miller input and output capacitances. To do this, we'll need to find the voltage gain.

$$
\begin{aligned}
A_v &= -g_m r_L \\
A_v &= -1.25\,\text{mS}\left(6\,\text{k}\Omega \parallel 12\,\text{k}\Omega\right) \\
A_v &= -5
\end{aligned}
$$

$$
\begin{aligned}
C_{in(miller)} &= C_{dg}\left(|A_v|+1\right) \\
C_{in(miller)} &= 1.5\,\text{pF}\left(|-5|+1\right) \\
C_{in(miller)} &= 9\,\text{pF}
\end{aligned}
$$

$$
\begin{aligned}
C_{out(miller)} &= C_{gd}\frac{|A_v|+1}{|A_v|} \\
C_{out(miller)} &= 1.5\,\text{pF}\,\frac{|-5|+1}{|-5|} \\
C_{out(miller)} &= 1.8\,\text{pF}
\end{aligned}
$$

Now we can move to the input lag network's resistance value. We can assume that $Z_{in(gate)}$ is large enough to ignore.

$$
\begin{aligned}
R_{input(lag)} &= R_{gen} \parallel R_G \parallel Z_{in(gate)} \\
R_{input(lag)} &= 50\,\Omega \parallel 2\,\text{M}\Omega \parallel \infty\,\Omega \\
R_{input(lag)} &= 50\,\Omega
\end{aligned}
$$

Ignoring input stray capacitance, the total input lag capacitance is:

$$
\begin{aligned}
C_{input(lag)} &= C_{gs}+C_{in(miller)} \\
C_{input(lag)} &= 3\,\text{pF}+9\,\text{pF} \\
C_{input(lag)} &= 12\,\text{pF}
\end{aligned}
$$

Thus, the input lag f_c is:

$$
\begin{aligned}
f_{c(input\,lag)} &= \frac{1}{2\,\pi\,R_{input(lag)}C_{input(lag)}} \\
f_{c(input\,lag)} &= \frac{1}{2\,\pi\,50\,\Omega\,12\,\text{pF}} \\
f_{c(input\,lag)} &= 265\,\text{MHz}
\end{aligned}
$$

Continuing with the output network analysis, the output lag resistance is:

$$R_{output(lag)} = R_D \parallel R_L$$
$$R_{output(lag)} = 6\,\text{k}\Omega \parallel 12\,\text{k}\Omega$$
$$R_{output(lag)} = 4\,\text{k}\Omega$$

Ignoring any output stray and load capacitance, the total output lag capacitance is:

$$C_{output(lag)} = C_{out(miller)}$$
$$C_{output(lag)} = 1.8\,\text{pF}$$

The output lag f_c is:

$$f_{c(output\,lag)} = \frac{1}{2\pi R_{output(lag)} C_{output(lag)}}$$
$$f_{c(output\,lag)} = \frac{1}{2\pi\, 4\,\text{k}\Omega\, 1.8\,\text{pF}}$$
$$f_{c(output\,lag)} = 22.1\,\text{MHz}$$

Obviously, the output network is dominant and the two critical frequencies are over a decade apart, thus the system f_2 is 22.1 MHz.

Computer Simulation

In order to crosscheck the results of Example 17.6, the circuit is captured in a simulator, as shown in Figure 17.36. A 2N5458 JFET was chosen for the simulation as the device capacitance and transconductance values match those given in the example. A simple transient analysis (not shown) was performed using an input sine wave of 100 millivolts peak amplitude at 1 kHz. The output signal showed a slight asymmetry, which is to be expected in such an amplifier, but the amplitude averaged to 500 millivolts peak, verifying the expected gain of 5, inverting.

An AC analysis follows, as shown in Figure 17.37. This Bode plot provides further verification of the midband gain, with a value of approximately 14 dB. The system f_2 is the point where the gain drops by 3 dB, or in this case, down to 11 dB. A measurement cursor pinpoints this frequency at 25.25 MHz. This is a little over ten percent higher than our computed value. This deviation may be due to slight variances between our assumed device capacitances and those of the simulator's transistor model. Naturally, in the real world, there will be variances from device to device, as well as the parasitic capacitances due to resistors, wiring, and the like. Consequently, the simulation results are within expected norms.

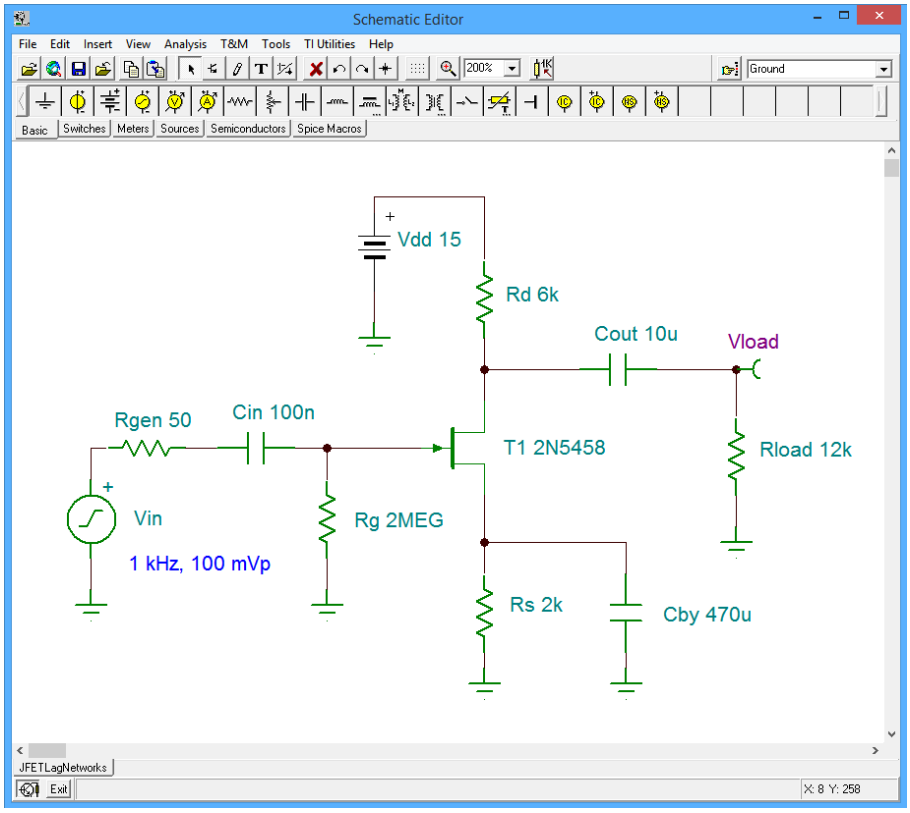


Figure 17.36
JFET high frequency simulation schematic.

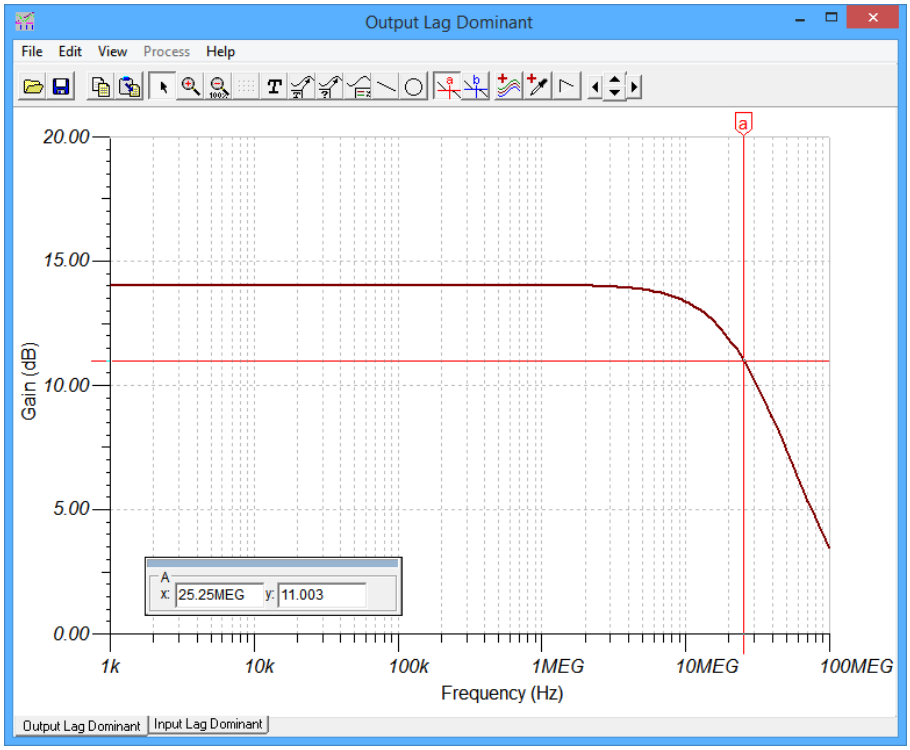


Figure 17.37
JFET high frequency simulation results of original circuit.

Of course, this simulation only verifies the output network. In order to check the input network, we can perform a scaling trick, similar to the one performed with lead networks. Instead of scaling the capacitors, though, this time we can scale the internal resistance of the signal generator, R_{gen}. This is possible because R_{gen} doesn't affect the gain (which would effect the Miller capacitances) and is the primary determiner of the input lag resistance; all without affecting the output network.

The value of R_{gen} was increased by a factor of 100 to 5 kΩ. This translates the input critical frequency down by a factor of 100, to 2.65 MHz. The network is now dominant, although there may be a slight interaction with the output network, pushing f_2 a little lower. The resulting Bode plot is shown in Figure 17.38.

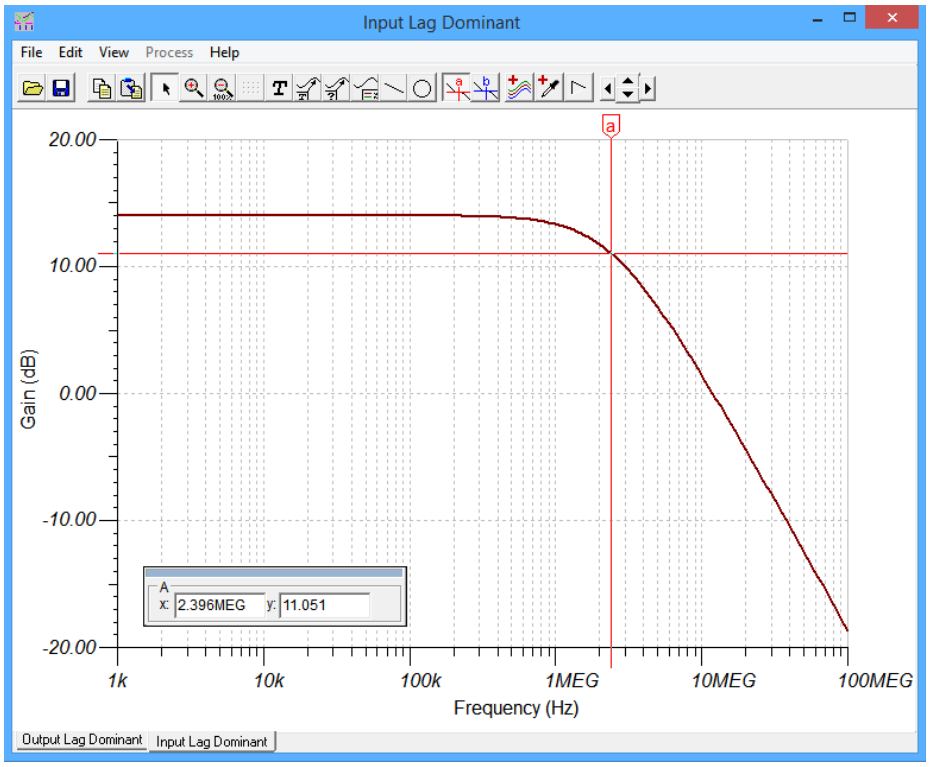


Figure 17.38
JFET high frequency simulation results of modified circuit.

The measurement cursor is again moved to 11 dB, showing a system 3 dB down frequency of approximately 2.4 MHz. Scaled back up by a factor of 100, this is in line with the computed value, although a little on the low side.

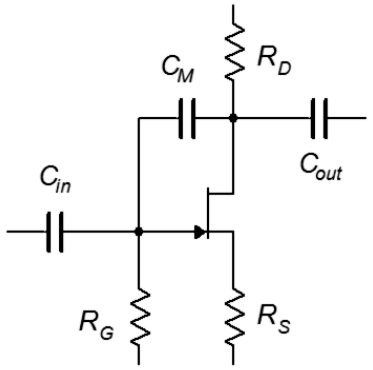
Granted, changing the generator's internal impedance is a nice enough trick for the simulator, but is often not possible or desirable in real world circuits. This brings up the question of how we might alter the high frequency response of an amplifier. The reasons for increasing f_2 might be obvious, but decreasing it can also be useful, particularly when it comes to avoiding high frequency interference. Unfortunately, raising the frequency turns out to be a bit of work as it requires a decrease in either the device capacitances or the surrounding resistances. Indeed, this may require a

different transistor model or even a complete redesign of the circuit. On the other hand, decreasing f_2 is relatively easy. This only requires an increase in capacitance, and it is a simple matter to add shunt capacitance to the input and/or output network(s). A more efficient approach is to add a capacitor in the Miller position (i.e., from drain to gate in a FET, or collector to base in a BJT). Such a capacitor will affect both the input and output networks, so it's almost like getting a second capacitor for free. This idea is explored in the following example.

Example 17.7

Using the circuit and results from Example 17.6, add a Miller capacitor, C_M, like the one shown in Figure 17.39, to shift f_2 down to 100 kHz.

Figure 17.39
Circuit for Example 17.7.



We will be changing nothing else in the circuit, so the prior values for gain and network resistance remain unchanged. All we have to do is determine the total capacitance for each network that will bring its critical frequency down to 100 kHz. We then subtract off the the associated device, stray, and load capacitances (if any). This results in the required $C_{in(miller)}$ and $C_{out(miller)}$. From there it is a simple matter to find C_M by using the Miller equations in reverse. The smaller result is the one we'll use.

Starting with the input network, the input lag resistance is 50 Ω. The required total input lag capacitance is:

$$C_{input(lag)} = \frac{1}{2\pi R_{input(lag)} f_{input(lag)}}$$
$$C_{input(lag)} = \frac{1}{2\pi 50\Omega\, 100\,\text{kHz}}$$
$$C_{input(lag)} = 31.8\,\text{nF}$$

The input already has 12 pF from C_{gs} and the Millerized C_{dg}, but this is small enough to ignore by comparison. Consequently, C_M will be 31.8 nF divided by the gain magnitude plus one.

$$C_{M(in)} = \frac{C_{input(lag)}}{|A_v|+1}$$
$$C_{M(in)} = \frac{31.8\,\text{nF}}{|-5|+1}$$
$$C_{M(in)} = 5.3\,\text{nF}$$

We follow a similar process for the output network. The output lag resistance is 4 kΩ and required capacitance is:

$$C_{output\,(lag)} = \frac{1}{2\,\pi\,R_{output\,(lag)}\,f_{output\,(lag)}}$$

$$C_{output\,(lag)} = \frac{1}{2\,\pi\,4\,\text{k}\Omega\,100\,\text{kHz}}$$

$$C_{output\,(lag)} = 398\,\text{pF}$$

Once again, the existing output capacitance is small enough to ignore. Now we apply Miller's Theorem in reverse to find the required C_M.

$$C_{M\,(out)} = C_{output\,(lag)}\frac{|A_v|}{|A_v|+1}$$

$$C_{M\,(out)} = 398\,\text{pF}\frac{|-5|}{|-5|+1}$$

$$C_{M\,(out)} = 332\,\text{pF}$$

The output network version of C_M is lower, so that's what we'll use. Note this will also lower the input network to approximately 1.6 MHz. If we had chosen the input network version of C_M, the output network critical frequency would have dropped below the target of 100 kHz.

If we add a 332 pF capacitor to the simulator circuit used in the prior example and then rerun the simulation, we get the Bode plot shown in Figure 17.40.

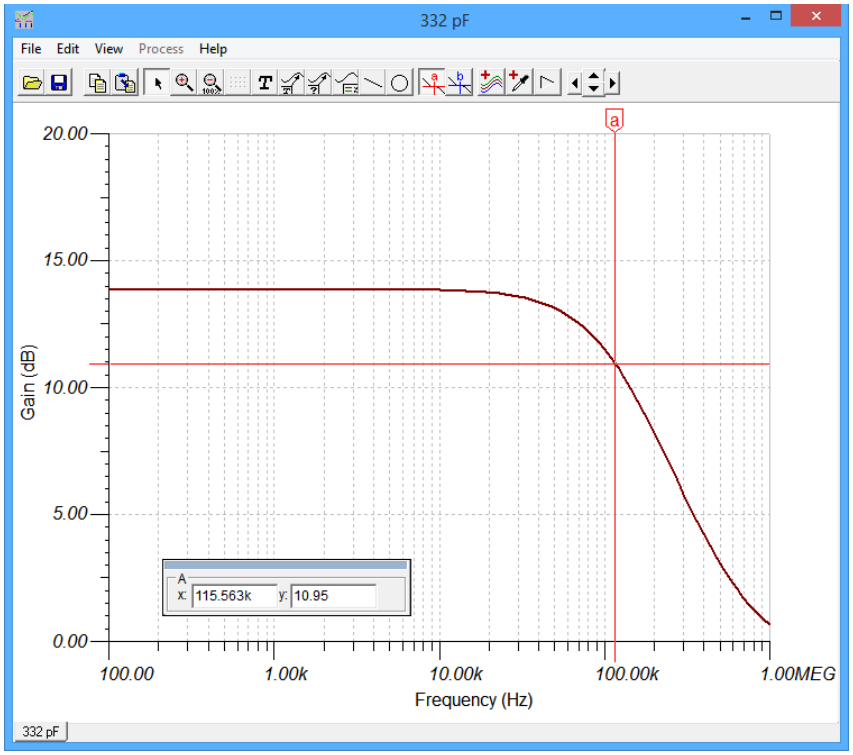


Figure 17.40
JFET high frequency simulation results using a Miller position capacitor.

The simulation shows a critical frequency of 115 kHz, or about 15% above target. If desired, this can be remedied by increasing C_M to about 380 pF.

Summary

All transistor amplifiers exhibit limits that define the frequency range over which they remain effective. The limits are the half power points, meaning the frequencies at which the voltage gain drops to 70.7% of its midband value. The lower and upper limits are denoted as f_1 and f_2, respectively. All transistor amplifiers have an upper limit, but it is possible to design amplifiers that do not have a lower limit, meaning that the gain is maintained down to 0 Hertz (DC). The lower and upper frequency limits of amplifiers are caused by lead networks and lag networks, respectively.

Lead networks involve coupling and bypass capacitors. A typical single stage inverting amplifier, whether based on a BJT or FET, exhibits three such networks: one at the input, a second at the output, and a third for the emitter or source bypass. The standard critical frequency equation is used for each network with the resistance value being equal to the Thevenin resistance seen from the position of the associated capacitor. The highest of these frequencies is dominant and will set the system f_1, although the actual value may be somewhat higher if any of the other network's critical frequency is relatively close to the dominant frequency.

Lag networks are caused generally by small parasitic capacitances associated with the transistor, other components, wiring, and the like. A typical single stage inverting amplifier will exhibit two lag networks: one at the input and another at the output. Any capacitance that bridges an inverting amplifier, running from input to output across the transistor, must be converted into equivalent input and output capacitances using Miller's Theorem. These capacitances appear in parallel with the other device and stray capacitances to form the equivalent input and output lag capacitances. The equivalent resistances are found by Thevenizing the surrounding elements from the perspective of the lag capacitors. The standard critical frequency equation can then be used to determine the input and output network frequencies. The lower of two is dominant and sets the system f_2. Again, if the other network is close in terms of its frequency, the interaction will shift f_2. In this case f_2 will shift to a lower frequency.

It is relatively straightforward to design an amplifier for a specific value of f_1. This is because individual coupling and bypass capacitors can be specified to reach said value. Manipulating f_2 is somewhat trickier. Reducing f_2 can be achieved most efficiently by adding a capacitor in the Miller position. In contrast, increasing f_2 may require selection of a different transistor, or even a major redesign of the circuit.

Review Questions

1. What circuit elements define the low frequency response of amplifiers?
2. What circuit elements define the high frequency response of amplifiers?
3. Do all amplifiers exhibit a lower frequency limit, f_1? Why/why not?
4. Do all amplifiers exhibit a higher frequency limit, f_2? Why/why not?
5. Compare and contrast the low frequency analysis of BJT versus FET amplifiers.
6. Compare and contrast the high frequency analysis of BJT versus FET amplifiers.
7. Of what use is Miller's theorem when analyzing the frequency limits of amplifiers?

Problems

Analysis Problems

1. For the circuit of Figure 17.41, determine the lead critical frequencies and the system f_1. $V_{CC} = 32$ V, $V_{EE} = -10$ V, $\beta = 100$, $R_L = 8$ kΩ, $R_C = 22$ kΩ, $R_E = 9.3$ kΩ, $R_B = 20$ kΩ, $R_{gen} = 75$ Ω, $C_{in} = 1$ μF, $C_{out} = 2$ μF, $C_E = 100$ μF.

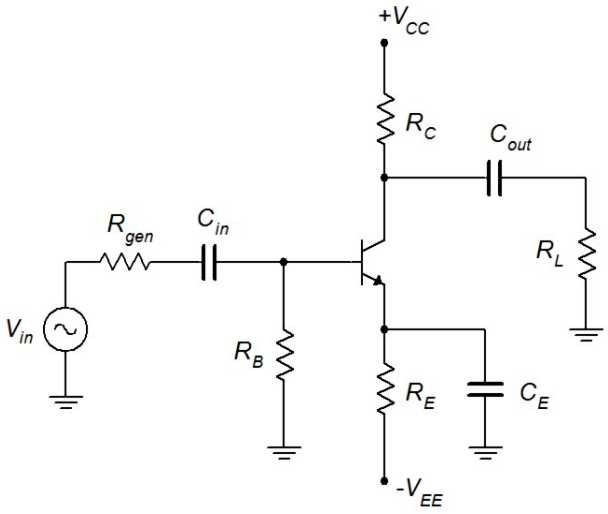


Figure 17.41

2. For the circuit of Figure 17.41, determine the lead critical frequencies and the system f_1. $V_{CC} = 20$ V, $V_{EE} = -10$ V, $\beta = 150$, $R_L = 4.7$ kΩ, $R_C = 12$ kΩ, $R_E = 4$ kΩ, $R_B = 10$ kΩ, $R_{gen} = 50$ Ω, $C_{in} = 2$ μF, $C_{out} = 10$ μF, $C_E = 470$ μF.

3. In the circuit of Figure 17.42, find the lead critical frequencies and system f_1. $V_{CC} = 20$ V, $V_{EE} = -8$ V, $\beta = 100$, $R_{gen} = 50$ Ω, $R_B = 22$ kΩ, $R_E = 36$ kΩ, $R_{SW} = 500$ Ω, $R_C = 40$ kΩ, $R_L = 60$ kΩ, $C_{in} = 1$ μF, $C_{out} = 2$ μF, $C_E = 100$ μF.

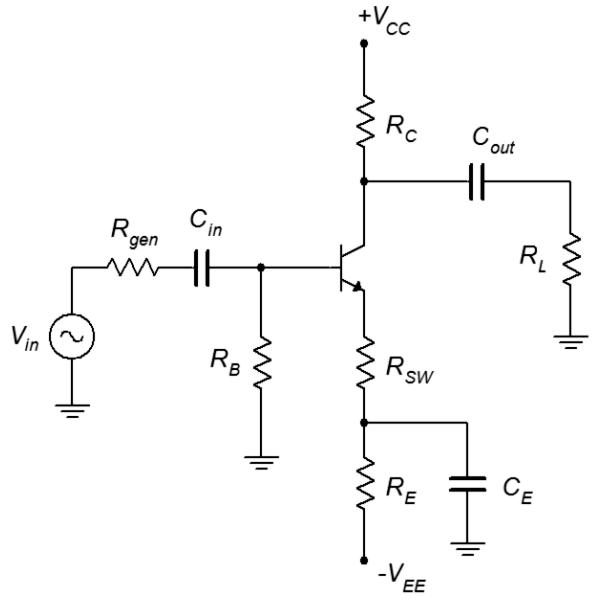


Figure 17.42

4. In the circuit of Figure 17.42, find the lead critical frequencies and system f_l. $V_{CC} = 18$ V, $V_{EE} = -12$ V, $\beta = 160$, $R_{gen} = 600\ \Omega$, $R_B = 8.2$ kΩ, $R_E = 7.5$ kΩ, $R_{SW} = 100\ \Omega$, $R_C = 6$ kΩ, $R_L = 10$ kΩ, $C_{in} = 2.2$ μF, $C_{out} = 10$ μF, $C_E = 330$ μF.

5. For the circuit of Figure 17.43, determine the lead critical frequencies and the system f_l. $\beta = 100$.

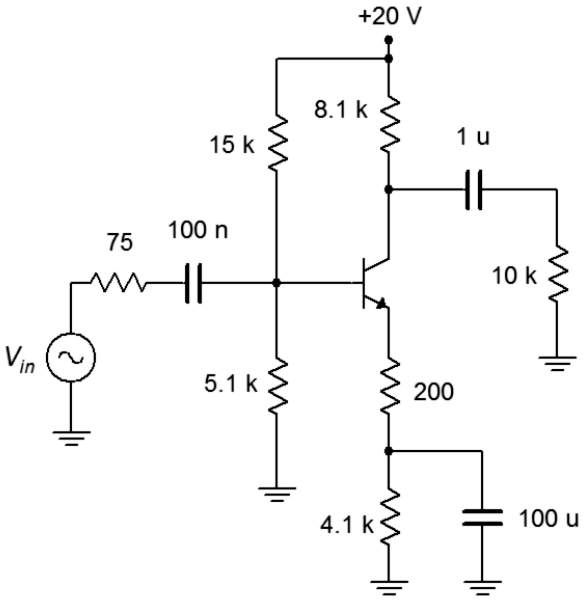


Figure 17.43

6. For the circuit of Figure 17.43, determine the lead critical frequencies and the system f_l. $\beta = 200$.

7. Given the circuit of Figure 17.44, determine the lead critical frequencies and the system f_l. V_{CC} = 25 V, β = 100, R_{gen} = 100 Ω, R_1 = 10 kΩ, R_2 = 2.5 kΩ, R_E = 1 kΩ, R_C = 3 kΩ, R_L = 6 kΩ, C_{in} = 10 μF, C_{out} = 6.8 μF, C_E = 330 μF.

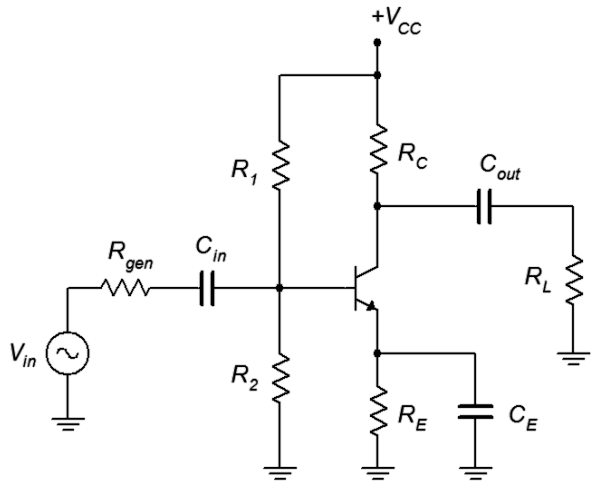


Figure 17.44

8. Given the circuit of Figure 17.44, determine the lead critical frequencies and the system f_l. V_{CC} = 15 V, β = 150, R_{gen} = 500 Ω, R_1 = 4.7 kΩ, R_2 = 2.2 kΩ, R_E = 3.3 kΩ, R_C = 3.9 kΩ, R_L = 5 kΩ, C_{in} = 4.7 μF, C_{out} = 10 μF, C_E = 680 μF.

9. For the circuit of Figure 17.45, determine the lead critical frequencies and the system f_l. β = 100, C_{in} =10 μF, C_{out} = 22 μF.

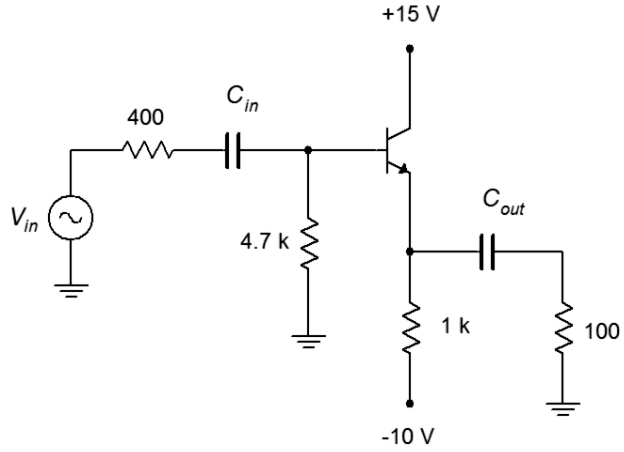


Figure 17.45

10. For the circuit of Figure 17.45, determine the lead critical frequencies and the system f_l. β = 200, C_{in} = 1 μF, C_{out} = 470 μF.

11. Given the circuit of Figure 17.46, determine the lead critical frequencies and the system f_l. V_{DD} = 20 V, g_m = 2 mS, R_{gen} = 50 Ω, R_G = 4.7 MΩ, R_D = 12 kΩ, R_S = 4 kΩ, R_L = 10 kΩ, C_{in} = 100 nF, C_{out} = 10 μF, C_S = 22 μF.

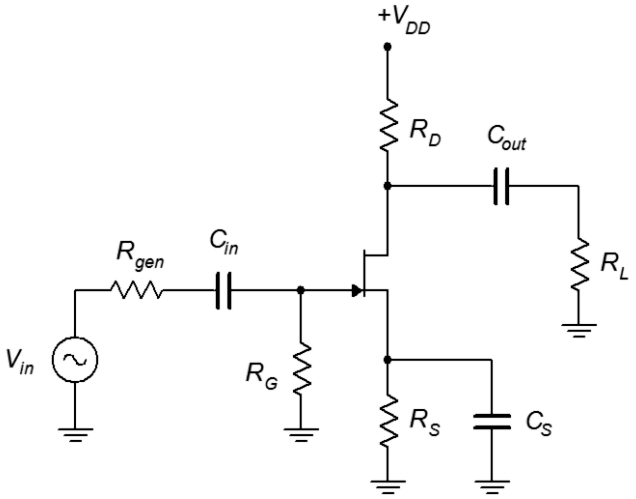


Figure 17.46

12. Given the circuit of Figure 17.46, determine the lead critical frequencies and the system f_l. $V_{DD} = 15$ V, $g_m = 3$ mS, $R_{gen} = 100\ \Omega$, $R_G = 1$ MΩ, $R_D = 6.8$ kΩ, $R_S = 3.3$ kΩ, $R_L = 15$ kΩ, $C_{in} = 330$ nF, $C_{out} = 2.2$ μF, $C_S = 680$ μF.

13. In the circuit of Figure 17.47, determine the lead critical frequencies and the system f_l. $V_{DD} = 22$ V, $V_{SS} = -4$ V, $g_m = 2.5$ mS, $R_G = 1$ MΩ, $R_D = 5$ kΩ, $R_S = 4$ kΩ, $R_L = 10$ kΩ, $C_{in} = 100$ nF, $C_{out} = 3.3$ μF, $C_S = 560$ μF.

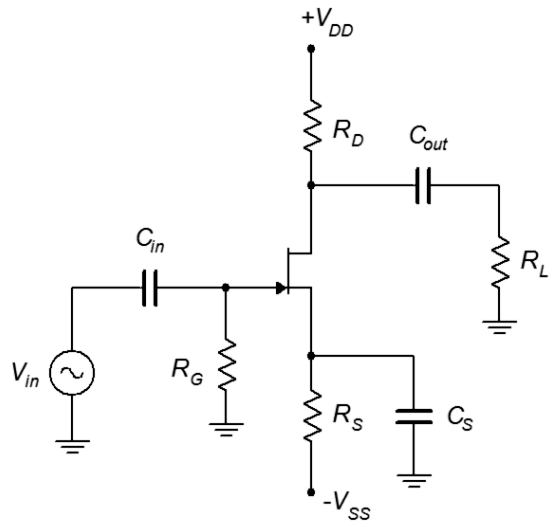


Figure 17.47

14. In the circuit of Figure 17.47, find the lead critical frequencies and the system f_l. $V_{DD} = 25$ V, $V_{SS} = -3$ V, $g_m = 5$ mS, $R_G = 2.2$ MΩ, $R_D = 10$ kΩ, $R_S = 4.7$ kΩ, $R_L = 20$ kΩ, $C_{in} = 470$ nF, $C_{out} = 100$ μF, $C_S = 33$ μF.

15. Given the circuit of Figure 17.48, determine the lead critical frequencies and the system f_1. $g_m = 8$ mS, $C_{in} = 10$ nF, $C_{out} = 330$ μF.

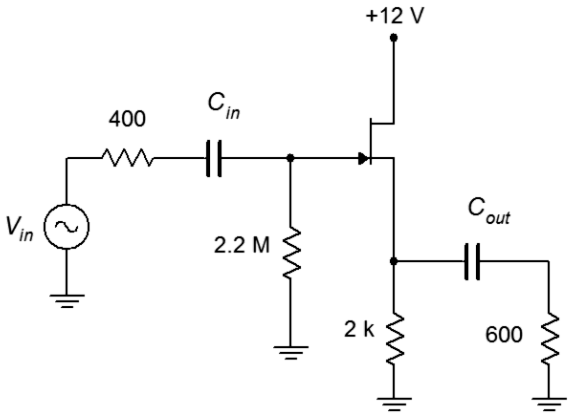


Figure 17.48

16. Given the circuit of Figure 17.48, determine the lead critical frequencies and the system f_1. $g_m = 10$ mS, $C_{in} = 150$ nF, $C_{out} = 47$ μF.

17. For the circuit of Figure 17.41, determine the lag critical frequencies and the system f_2. $V_{CC} = 32$ V, $V_{EE} = -10$ V, $\beta = 100$, $R_L = 8$ kΩ, $R_C = 22$ kΩ, $R_E = 9.3$ kΩ, $R_B = 20$ kΩ, $R_{gen} = 75$ Ω, $C_{ibo} = 5$ pF, $C_{obo} = 2$ pF.

18. For the circuit of Figure 17.41, determine the lag critical frequencies and the system f_2. $V_{CC} = 20$ V, $V_{EE} = -10$ V, $\beta = 150$, $R_L = 4.7$ kΩ, $R_C = 12$ kΩ, $R_E = 4$ kΩ, $R_B = 10$ kΩ, $R_{gen} = 50$ Ω, $C_{ibo} = 15$ pF, $C_{obo} = 3$ pF.

19. In the circuit of Figure 17.42, find the lag critical frequencies and system f_2. $V_{CC} = 20$ V, $V_{EE} = -8$ V, $\beta = 100$, $R_{gen} = 50$ Ω, $R_B = 22$ kΩ, $R_E = 36$ kΩ, $R_{SW} = 500$ Ω, $R_C = 40$ kΩ, $R_L = 60$ kΩ, $C_{in(stray)} = 20$ pF, $C_{out(stray)} = 50$ pF, $C_{ibo} = 3$ pF, $C_{obo} = 2$ pF.

20. In the circuit of Figure 17.42, find the lag critical frequencies and system f_2. $V_{CC} = 18$ V, $V_{EE} = -12$ V, $\beta = 160$, $R_{gen} = 600$ Ω, $R_B = 8.2$ kΩ, $R_E = 7.5$ kΩ, $R_{SW} = 100$ Ω, $R_C = 6$ kΩ, $R_L = 10$ kΩ, $C_{in(stray)} = 10$ pF, $C_{out(stray)} = 100$ pF, $C_{ibo} = 6$ pF, $C_{obo} = 2$ pF.

21. For the circuit of Figure 17.43, determine the lag critical frequencies and the system f_2. $\beta = 100$, $C_{ibo} = 5$ pF, $C_{obo} = 1$ pF.

22. For the circuit of Figure 17.43, determine the lag critical frequencies and the system f_2. $\beta = 200$, $C_{ibo} = 2$ pF, $C_{obo} = 4$ pF.

23. Given the circuit of Figure 17.44, determine the lag critical frequencies and the system f_2. $V_{CC} = 25$ V, $\beta = 100$, $R_{gen} = 100$ Ω, $R_1 = 10$ kΩ, $R_2 = 2.5$ kΩ, $R_E = 1$ kΩ, $R_C = 3$ kΩ, $R_L = 6$ kΩ, $C_{ibo} = 10$ pF, $C_{obo} = 1$ pF.

24. Given the circuit of Figure 17.44, determine the lag critical frequencies and the system f_2. V_{CC} = 15 V, β = 150, R_{gen} = 500 Ω, R_1 = 4.7 kΩ, R_2 = 2.2 kΩ, R_E = 3.3 kΩ, R_C = 3.9 kΩ, R_L = 5 kΩ, C_{ibo} = 20 pF, C_{obo} = 1.5 pF.

25. Given the circuit of Figure 17.46, determine the lag critical frequencies and the system f_2. V_{DD} = 20 V, g_m = 2 mS, R_{gen} = 50 Ω, R_G = 4.7 MΩ, R_D = 12 kΩ, R_S = 4 kΩ, R_L = 10 kΩ, R_{gen} = 50 Ω, C_{gs} = 3 pF, C_{dg} = 2 pF.

26. Given the circuit of Figure 17.46, determine the lag critical frequencies and the system f_2. V_{DD} = 15 V, g_m = 3 mS, R_{gen} = 100 Ω, R_G = 1 MΩ, R_D = 6.8 kΩ, R_S = 3.3 kΩ, R_L = 15 kΩ, C_{gs} = 5 pF, C_{dg} = 2.5 pF.

27. Given the circuit of Figure 17.49, determine the lag critical frequencies and the system f_2. g_m = 3 mS, C_{gs} = 5 pF, C_{dg} = 4 pF.

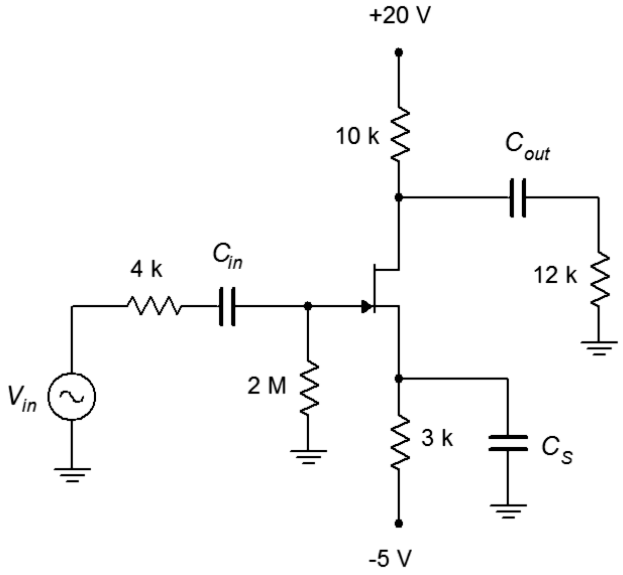


Figure 17.49

28. Given the circuit of Figure 17.49, determine the lag critical frequencies and the system f_2. g_m = 8 mS, C_{gs} = 20 pF, C_{dg} = 3 pF.

29. Given the circuit of Figure 17.49, determine the lag critical frequencies and the system f_2. g_m = 3 mS, C_{iss} = 5 pF, C_{rss} = 4 pF.

30. Given the circuit of Figure 17.49, determine the lag critical frequencies and the system f_2. g_m = 8 mS, C_{iss} = 20 pF, C_{rss} = 3 pF.

Design Problems

31. Alter the coupling and bypass capacitors of Figure 17.43 so that the input f_c is 200 Hz, the output f_c is 20 Hz, and the bypass f_c is 2 Hz. Assume $\beta = 100$.

32. Alter the coupling and/or bypass capacitors of Figure 17.43 so that the system f_l is 100 Hz . Assume $\beta = 200$.

33. Determine values for the coupling capacitors of Figure 17.45 so that the system f_l is 50 Hz. Assume $\beta = 100$.

34. Determine values for the coupling capacitors of Figure 17.45 so that the input f_c is 100 Hz and the output f_c is 10 Hz. Assume $\beta = 200$.

35. Determine values for the coupling capacitors of Figure 17.48 so that the system f_l is 50 Hz. Assume $g_m = 8$ mS.

36. Determine values for the coupling capacitors of Figure 17.48 so that the input f_c is 100 Hz and the output f_c is 10 Hz. Assume $g_m = 10$ mS.

37. Determine values for the coupling capacitors of Figure 17.49 so that the system f_l is 1 kHz. Assume $g_m = 3$ mS.

38. Determine values for the coupling capacitors of Figure 17.49 so that both the input and output f_c values are 500 Hz. Assume $g_m = 8$ mS.

Challenge Problems

39. Given the circuit of Figure 17.50, determine the lead critical frequencies and the system f_l. $\beta = 100$, $C_{in} = 4.7$ μF, $C_{out} = 10$ μF, $C_E = 33$ μF.

40. Given the circuit of Figure 17.50, determine the lag critical frequencies and the system f_2. $\beta = 150$, $C_{ibo} = 8$ pF, $C_{obo} = 3.5$ pF.

41. Given the circuit of Figure 17.51, determine the lead critical frequencies and the system f_l. $g_m = 5$ mS, $C_{in} = 56$ nF, $C_{out} = 15$ μF, $C_G = 33$ pF.

42. Given the circuit of Figure 17.51, determine the lag critical frequencies and the system f_2. $g_m = 12$ mS, $C_{iss} = 30$ pF, $C_{rss} = 5$ pF, $C_G = 36$ pF.

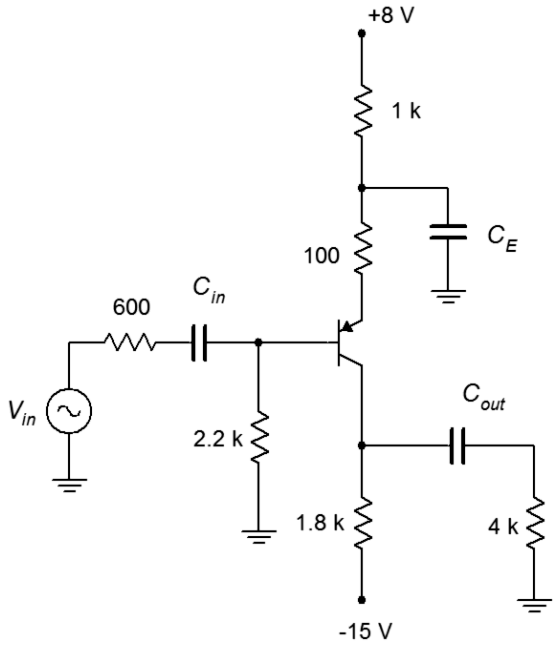


Figure 17.50

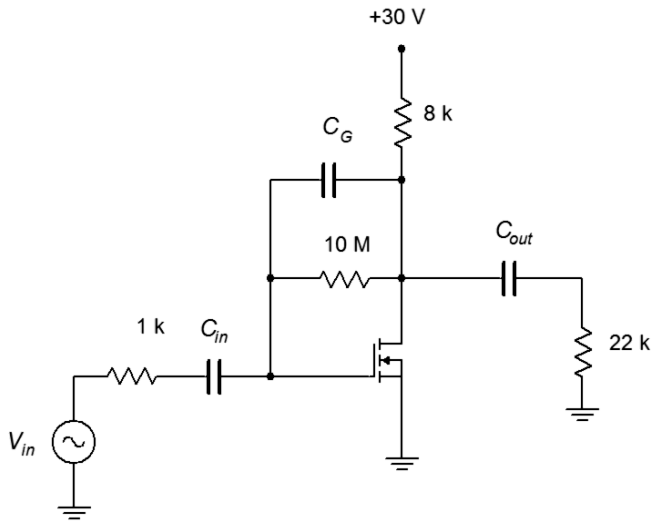


Figure 17.51

Computer Simulation Problems

43. Run a simulation to verify the results of Design Problem 32. Verify all critical frequencies.

44. Run a simulation to verify the results of Design Problem 34. Verify both critical frequencies.

45. Run a simulation to verify the results of Challenge Problem 39. Verify all critical frequencies.

Appendix A

Manufacturer's Data Sheet Links

These links are current as of Fall, 2022 and are listed here for convenience. Always check the manufacturer's site for the latest information before beginning any new design. Note that some devices are available from several different manufacturers.

Diodes

1N4002 rectifier: https://www.onsemi.com/pub/Collateral/1N4001-D.PDF
1N4148/1N914 switching: https://www.onsemi.com/pdf/datasheet/1n914-d.pdf
NZX5V1B Zener: https://assets.nexperia.com/documents/data-sheet/NZX_SER.pdf

LED

LED: https://www.cree.com/led-components/media/documents/ds-C566D-1275.pdf

Transistors

2N3904 NPN BJT: https://www.onsemi.com/pub/Collateral/2N3903-D.PDF
2N306 PNP BJT: https://www.onsemi.com/pub/Collateral/2N3906-D.PDF
J112 N JFET: https://www.onsemi.com/pub/Collateral/J111-D.PDF
2N3055 Power NPN BJT: http://www.onsemi.com/pub_link/Collateral/2N3055-D.PDF
FDMS86180 Power E-MOSFET: https://www.onsemi.com/pub/Collateral/FDMS86180-D.pdf
IRF7201 Power E-MOSFET: http://www.infineon.com/dgdl/irf7201pbf.pdf?fileId=5546d462533600a4015355f13efb1ada
FGH50T65SQD IGBT: https://www.onsemi.com/pub/Collateral/FGH50T65SQD-D.PDF
IRGPC40K IGBT: http://www.irf.com/product-info/datasheets/data/irgpc40k.pdf

Appendix B

Standard Component Sizes

Passive components (resistors, capacitors and inductors) are available in standard sizes. The tables below are for resistors. The same digits are used in subsequent decades up to at least 1 Meg ohm (higher decades are not shown). Capacitors and inductors are generally not available in as many standard values as are resistors. Capacitors below 10 nF (.01 μF) are usually available at the 5% standard digits while larger capacitances tend to be available at the 20% standards.

5% and 10% standard values, EIA E24 and EIA E12

10% values (EIA E12) are **bold**
20% values (seldom used) are every fourth value starting from 10
(i.e., every other 10% value)

10	11	**12**	13	**15**	16	**18**	20	**22**	24	**27**	30
33	36	**39**	43	**47**	51	**56**	62	**68**	75	**82**	91

1% and 2% standard values, EIA E96 and EIA E48

2% values (EIA E48) are **bold**

10.0	10.2	**10.5**	10.7	**11.0**	11.3	**11.5**	11.8	**12.1**	12.4	**12.7**	13.0
13.3	13.7	**14.0**	14.3	**14.7**	15.0	**15.4**	15.8	**16.2**	16.5	**16.9**	17.4
17.8	18.2	**18.7**	19.1	**19.6**	20.0	**20.5**	21.0	**21.5**	22.1	**22.6**	23.2
23.7	24.3	**24.9**	25.5	**26.1**	26.7	**27.4**	28.0	**28.7**	29.4	**30.1**	30.9
31.6	32.4	**33.2**	34.0	**34.8**	35.7	**36.5**	37.4	**38.3**	39.2	**40.2**	41.2
42.2	43.2	**44.2**	45.3	**46.4**	47.5	**48.7**	49.9	**51.1**	52.3	**53.6**	54.9
56.2	57.6	**59.0**	60.4	**61.9**	63.4	**64.9**	66.5	**68.1**	69.8	**71.5**	73.2
75.0	76.8	**78.7**	80.6	**82.5**	84.5	**86.6**	88.7	**90.9**	93.1	**95.3**	97.6

Appendix C

Answers to Selected Numbered Problems

Chapter 2

1. .53 mA
3. 4.37 V, 6.23 V
5. 8.79 mA
7. 1.18 mA for both
9. 4.7 V
11. 8.3 V
13. 430 Ω

Chapter 3

1. 16.3 V peak
3. 12 V peak
5. 24 V peak
7. 10sin2π100t, positive clipped at 8.7 V
9. 12sin2π100t, positive clipped at 6.7 V, negative clipped at −4.7 V
11. 8sin2π500t + 9.3 VDC
13. Use biased clipper (see Problem 11, Figure 3.37)

Chapter 4

1. 99
3. I_B = 21.5 μA, I_C = 2.15 mA, I_E = 2.175 mA
5. 10.74 mA
7. 0 mA
9. 10.74 mA
11. .2 V
13. 47% of former value
15. 3.37 kΩ
17. 150 Ω

Chapter 5

1. $V_{CE(cutoff)}$ = 28 V, $I_{C(sat)}$ = 1.27 mA, I_{CQ} = 0.73 mA, V_{CEQ} = 11.94 V
3. $V_{CE(cutoff)}$ = 23 V, $I_{C(sat)}$ = 8.52 mA, I_{CQ} = 3.03 mA, V_{CEQ} = 14.8 V
5. $V_{CE(cutoff)}$ = 20 V, $I_{C(sat)}$ = 6.49 mA, I_{CQ} = 1 mA, V_{CEQ} = 6.6 V
7. $V_{CE(cutoff)}$ = 16 V, $I_{C(sat)}$ = 0.988 mA, I_{CQ} = 0.613 mA, V_{CEQ} = 6.06 V
9. $V_{CE(cutoff)}$ = 12 V, $I_{C(sat)}$ = 3.64 mA, I_{CQ} = 1.27 mA, V_{CEQ} = 7.81 V
11. $V_{CE(cutoff)}$ = 18 V, $I_{C(sat)}$ = 2.4 mA, I_{CQ} = 1.6 mA, V_{CEQ} = 5.99 V
13. $V_{CE(cutoff)}$ = 15 V, $I_{C(sat)}$ = 3.89 mA, I_{CQ} = 2.72 mA, V_{CEQ} = 4.51 V
15. $V_{CE(cutoff)}$ = 18 V, $I_{C(sat)}$ = 2.3 mA, I_{CQ} = 1.96 mA, V_{CEQ} = 2.66 V
17. $V_{CE(cutoff)}$ = 18 V, $I_{C(sat)}$ = 4.15 mA, I_{CQ} = 1.58 mA, V_{CEQ} = 11.4 V
19. 3.65 kΩ
21. 7.58 kΩ

Chapter 6

1. 465 mV
3. No, yes
5. 50,000:1
7. 2.86 kΩ

Chapter 7

1. Z_{in} = 1.53 kΩ, Z_{out} = 8.1 kΩ
3. Z_{in} = 9.08 kΩ, Z_{out} = 22 kΩ, v_{load} = 4 V (inverted)
5. Z_{in} = 1.8 kΩ, Z_{out} = 1.82 kΩ, v_{load} = 224 mV (inverted)
7. Z_{in} = 3.25 kΩ, Z_{out} = 8.1 kΩ, v_{load} = 1.19 V (inverted)
9. Z_{in} = 90.4 kΩ, Z_{out} = 4.1 Ω, v_{load} = 182 mV
11. Z_{in} = 236 kΩ, Z_{out} = 6.6 Ω, v_{load} = 279 mV
13. Z_{in} = 42.7 Ω, Z_{out} = 22 kΩ, v_{load} = 594 mV
15. Set R_{SW} = 300 Ω to maintain A_v (note r'_e change), increase 22 kΩ several times (220 kΩ is fine)
17. Set R_{SW} = 49 Ω to double A_v (including effect of r'_e), increase R_E to 1049 Ω to maintain Q

Chapter 8

1. I_{CQ} = 24 mA, $i_{c(sat)}$ = 127.6 mA, V_{CEQ} = 6.7 V, $v_{ce(cutoff)}$ = 8.25 V, Compliance = 1.55 Vp, $P_{l(max)}$ = 16 mW, $P_{D(max)}$ = 161 mW, η = 3.7%
3. Not centered
5. I_{CQ} = 186 mA, $i_{c(sat)}$ = 1.374 A, V_{CEQ} = 16.4 V, $v_{ce(cutoff)}$ = 19 V, Compliance = 2.57 Vp, $P_{l(max)}$ = 206 mW, $P_{D(max)}$ = 3.05 W, η = 3.16%
7. I_{CQ} = 30.4 mA, $i_{c(sat)}$ = 554 mA, V_{CEQ} = 15.7 V, $v_{ce(cutoff)}$ = 16.6 V, Compliance = 0.912 Vp, $P_{l(max)}$ = 13 mW, $P_{D(max)}$ = 477 mW, η = 1.16%
9. I_{CQ} = 51.7 mA, $i_{c(sat)}$ = 391 mA, V_{CEQ} = 8.65 V, $v_{ce(cutoff)}$ = 9.79 V, Compliance = 0.79 Vp, $P_{l(max)}$ = 19.5 mW, $P_{D(max)}$ = 449 mW, η = 0.94%
11. Yes
13. 4.07 C°/W
15. 98 VDC (not particularly practical, changing R_E as well would be more flexible)

Chapter 9

1. BV_{CEO} = 30 V, $I_{C(max)}$ = 0.9375 A, $P_{l(max)}$ = 7.03 W, $P_{D(max)}$ = 1.4 W
3. 187 Ω
5. BV_{CEO} = 15 V, $I_{C(max)}$ = 0.469 A, $P_{l(max)}$ = 1.76 W, $P_{D(max)}$ = 351 mW
7. 187 Ω
9. BV_{CEO} = 48 V, $I_{C(max)}$ = 3 A, $P_{l(max)}$ = 36 W, $P_{D(max)}$ = 7.2 W
11. BV_{CEO} = 48 V, $I_{C(max)}$ = 1.5 A, $P_{l(max)}$ = 18 W, $P_{D(max)}$ = 3.6 W
13. 3.5 A
15. R_1:R_2 ratio is 8.55:1, e.g., if R_2 = 500 Ω then R_1 = 4.275 kΩ

Chapter 10

1. I_D = 10 mA, V_{DS} = 14 V
3. I_D = 2.11 mA, V_G = 0 V, V_D = 32.2 V
5. I_D = 3.77 mA, V_{DS} = 9.9 V
7. I_D = 4.15 mA, V_G = 0 V, V_{DS} = 13.8 V
9. R_S = 180 Ω
11. R_E = 2.825 kΩ

Chapter 11

1. Z_{in} = 220 kΩ, A_v = −7.5
3. Z_{in} = 270 kΩ, V_{out} = 173 mV (inverted)
5. Z_{in} = 390 kΩ, A_v = −12.5
7. Z_{in} = 220 kΩ, A_v = 0.643
9. Z_{in} = 470 kΩ, V_{out} = 83.8 mV
11. V_{out} = 99.3 mV
13. A_v = -12, V_{out} = 120 mV (inverted)
15. Z_{in} = 231 Ω, Z_{out} = 4 kΩ

Chapter 12

1. I_D = 3 mA, V_G = 0 V, V_D = 9.6 V
3. I_D = 5.1 mA, V_G = 0 V, V_D = 11.6 V
5. I_D = 12 mA, V_G = 0 V, V_D = 13.4 V
7. I_D = 9.1 mA, V_G = 1.17 V, V_D = 25.7 V
9. I_D = 3.125 mA, V_G = 4.25 V, V_D = 26.25 V
11. I_D = 1.56 mA, V_G = 0 V, V_D = −17.2 V
13. I_D = 3.67 mA, V_G = 0 V, V_D = −12.9 V
15. I_D = 11 mA, V_D = −4.8 V
17. 180 Ω
19. R_D = 1.29 kΩ, R_G is not tied to any particular value but probably > 1 MΩ

Chapter 13

1. Z_{in} = 750 kΩ, A_v = −2.58
3. Z_{in} = 510 kΩ, A_v = −8.86
5. Z_{in} = 90.9 kΩ, V_{out} = 142 mV (inverted)
7. Z_{in} = 90.9 kΩ, V_{out} = 515 mV (inverted)
9. Z_{in} = 910 kΩ, V_{out} = 125 mV
11. Z_{in} = 680 kΩ, A_v = 0.91
13. R_G = 510 kΩ or greater. The remainder of the design has nearly infinite possibilities. The first thing to notice is that the goal is achievable because g_{m0} = 25 mS and R_L = 10 kΩ, giving a ceiling A_v of 250 unswamped (but unrealistic). Using a standard 15 VDC supply and picking a mid-point bias (V_{GS} = 0.5 $V_{GS(off)}$ and I_D = 0.25 I_{DSS}) yields R_S = 1 V/6.25 mA = 160 Ω and g_m = 12.5 mS. Using a common 1 kΩ for R_D yields r_L = 909 Ω and V_D = 8.75 V (reasonable). The resulting unswamped gain is over 11 so the 160 Ω source resistance can be split into two parts, R_S and R_{SW}, with R_{SW} adding enough degeneration to position the gain at the desired level. For example, R_{SW} = 100 Ω with R_S = 60 Ω yields a gain of 5.

Chapter 14

1. 35 kHz
3. 33.3 V/μs
5. 12.48 A, 0.15 V

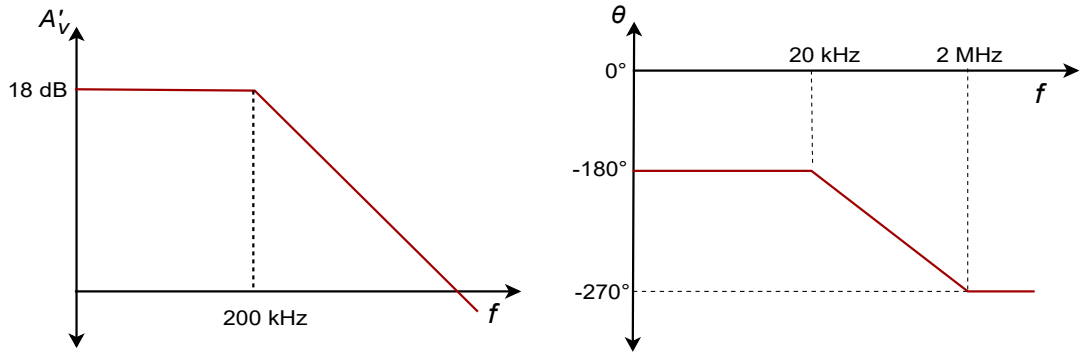
Chapter 15

1. 4 V
3. 30 ns
5. 610 ns

Chapter 16

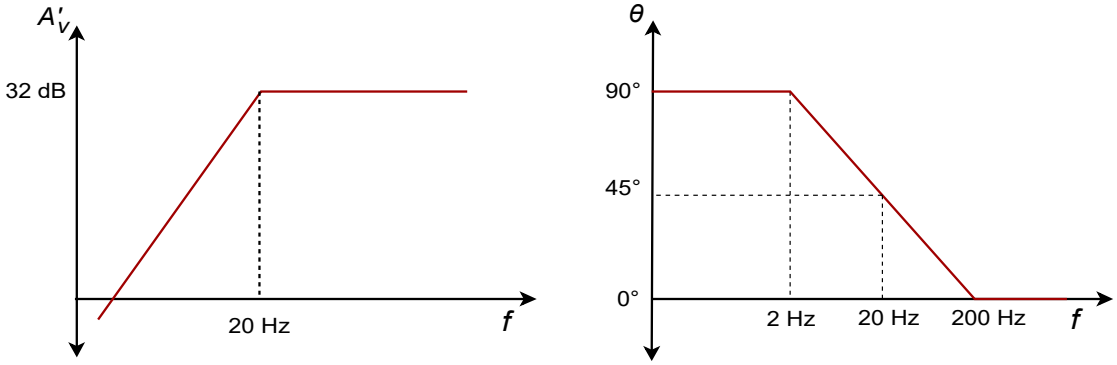
1. A) 10 dB B) 19 dB C) 26.99 dB D) 0 dB E) −6.99 dB F) −15.23 dB
3. 33 dB
5. G = 501, P_{out} = 12.53 W
7. A) 1.06 B) 1 C) 199.5 D) 3.43 E) 0.398 F) 0.188
9. A = 8.57 A' = 18.66 dB
11. A) 0 dBW B) 13.6 dBW C) 8.13 dBW D) −7 dBW E) −26.4 dBW F)30.8 dBW G) −43.5 dBW H) −65.2 dBW I) −172.5 dBW
13. A) 150 dBf B) 163.6 dBf C) 158.1 dBf D) 143 dBf E) 123.6 dBf F) 180.8 dBf G) 106.5 dBf H) 84.8 dBf I) −22.5 dBf
15. G'_{total} = 28 dB, G = 631
17. For P'_{in} = 4 dBm: output stage 1 = 6 dBm, stage 2 = 0 dBm, stage 3 = 15 dBm.
    For P'in = −34 dBm: output stage 1 = − 24 dBW, stage 2 = −30 dBW, stage 3 = −15 dBW.
19. a. 200 mW
21. V'_{out} = 21 dBV, 21 dBV = 11.2 V (final output)
    For stage 1: 4 dBV = 1.58 V. For stage 2: 9 dBV = 2.82 V
23. a. 15 V
25. At 50 kHz: −0.022 dB, −4.09 degrees. At 700 kHz: −3 dB, −45 degrees.
    At 10 MHz: −23.1 dB, −86 degrees. T_r = 500 μsec
27. The amplitude portion does not change. Phases are: At 30 kHz −188.5 degrees, at 200 kHz −225 degrees, at 1 MHz −258.7 degrees

29.



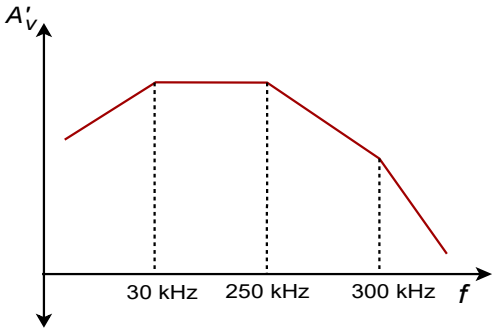
31. At 4 kHz: 78.7 degrees, at 20 Hz: 45 degrees, at 100 Hz: 11.3 degrees

33.

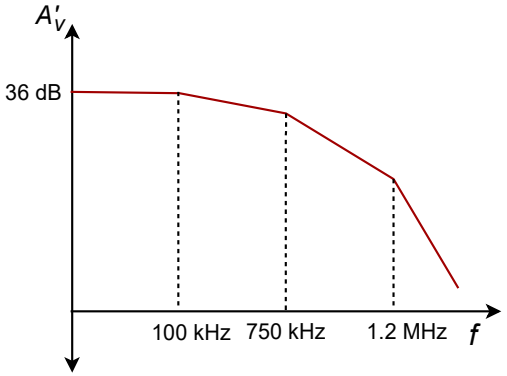


35. Net gain at 20 kHz = 35.87 dB. Net phase at 20 kHz = 51.7 degrees
At 100 kHz: phase = −5.1 degrees, A'_v = 40 dB. At 800 kHz: phase = −70.5 degrees, A'_v = 30.5 dB

37.



39.



41. Each lag network rolls off at 20 dB/decade for a 60 dB/decade total (i.e., above 1.2 MHz).
43. 0.775 V
45. 360 W
47. 71.5 dBV
49. Greater than 30 Hz.

Chapter 17

1. f_{in} = 67 Hz, f_{out} = 2.65 Hz, f_{bypass} = 61.2 Hz. f_1 > 67 Hz due to proximity.
3. f_{in} = 9.76 Hz, f_{out} = 0.796 Hz, f_{bypass} = 2.57 Hz. f_1 > 9.76 Hz due to proximity.
5. f_{in} = 477 Hz, f_{out} = 8.8 Hz, f_{bypass} = 7.42 Hz. $f_1 \approx$ 477 Hz.
7. f_{in} = 4.5 Hz, f_{out} = 2.6 Hz, f_{bypass} = 69.4 Hz. $f_1 \approx$ 69.4 Hz.
9. f_{in} = 4.5 Hz, f_{out} = 67.9 Hz. $f_1 \approx$ 67.9 Hz.
11. f_{in} = 0.34 Hz, f_{out} = 0.723 Hz, f_{bypass} = 16.3 Hz. $f_1 \approx$ 16.3 Hz.
13. f_{in} = 1.59 Hz, f_{out} = 3.22 Hz, f_{bypass} = 0.781 Hz. f_1 > 3.22 Hz due to proximity.
15. f_{in} = 7.23 Hz, f_{out} = 0.672 Hz. $f_1 \approx$ 7.23 Hz.
17. f_{in} = 4.78 MHz, f_{out} = 13.6 MHz. f_2 < 4.78 Hz due to proximity.
19. f_{in} = 40.8 MHz, f_{out} = 3.32 MHz. $f_2 \approx$ 3.32 MHz.
21. f_{in} = 84.1 MHz, f_{out} = 33.9 Hz. f_2 < 33.9 MHz due to proximity..
23. f_{in} = 5.63 MHz, f_{out} = 79.6 Hz. $f_2 \approx$ 5.63 MHz.
25. f_{in} = 128 MHz, f_{out} = 13.3 MHz. $f_2 \approx$ 13.3 MHz.
27. f_{in} = 535 kHz, f_{out} = 6.88 MHz. $f_2 \approx$ 535 kHz.
29. f_{in} = 566 kHz, f_{out} = 6.88 MHz. $f_2 \approx$ 566 kHz.
31. C_{in} = 239 nF, C_{out} = 440 nF, C_{bypass} = 371 µF.
33. For 50 Hz: C_{in} = 902 nF, C_{out} = 29.9 µF. Pick one and increase the other.
35. For 50 Hz: C_{in} = 1.45 nF, C_{out} = 4.43 µF. Pick one and increase the other.
37. For 1 kHz: C_{in} = 79.4 pF, C_{out} = 7.23 nF, C_{bypass} = 531 nF. Pick one and increase the others.

Appendix D

"There is life beyond the cymbals."

-Bill Bruford

After many years of observing the food ingestion habits of college students and faculty, it can be said that many do not cook, bake and eat food so much as they consume prepackaged *food-like substances* of dubious nutritional quality and limited flavor. There is no need for this as basic cooking is not difficult, and besides, it gives **you** control of what goes inside your body. Here is an excellent bit of snack-food: tasty, nutritious and relatively easy to make. And when you're done, you can feel good that you didn't drop a bunch of cash on some assembly-line manufactured excuse for sustenance that's loaded with saturated fat, refined sugar, salt and who-knows-what-else. Also, your cooking-challenged friends will think you're a genius.

Autumn Bread

Autumn bread is a variation on banana bread, but is very low in fat. If you add the optional nuts, the fat content goes up but it's not saturated fat or trans fat as in butter, margarine, or shortening. Walnuts in particular are a decent source of alpha-linolenic acid, a "good fat". Autumn bread is fairly dense and moist. It definitely *sticks with you*. For variations, you could swap one of the bananas for another apple if desired. In this case you might consider adding a little nutmeg. Also, if you like it less sweet, you can drop the honey back to 1/3 cup. For many people, there are sufficient raisins in the raisin bran but an extra half cup has been added "just because". In fact, extra raisins will help sweeten the flavor and that's another reason to cut back on the honey. Actually, don't bother to measure out the raisins and walnuts, just eyeball them and add what looks good at the moment. Cooking should be fun and a little bit experimental. "t" = teaspoon, "T" = tablespoon, "C" = 8 ounce cup. Yeow, I really do prefer metric units, but...

3 egg whites
3 overripe mashed bananas
1 medium apple, diced
3/4 C plain non-fat yogurt
2 T molasses
1/2 C honey
1 t cinnamon
1 t vanilla
1/2 C walnuts, almonds, or pecans; or combination thereof (optional)
1/2 C raisins
1 1/2 C raisin bran
2 1/2 C flour (try using 50/50 whole wheat/all-purpose)
1/2 t baking soda
2 1/2 t baking powder

Preheat oven to 350° and coat a 9x5 loaf pan with non-stick spray. Make sure that was 350°F and not 350°C; otherwise, prepare to eat carbon chunks when this comes out of the oven.

In a large bowl beat egg whites for about a minute by hand, then add next nine ingredients (through raisins), mixing well. Make sure you do this in a large bowl as doing it directly on the kitchen counter probably will be messy. Add raisin bran, mixing to make sure it's all wet. Let stand 10 minutes or so. While you're waiting, combine flours, baking soda and baking powder in a second bowl. Go back to first bowl and mix it around to make sure that the raisin bran is broken up. If it's not broken up, tell it a funny story and see if that breaks it up. Add the flour mixture to the wet bowl in batches, mixing thoroughly. If you don't know how to spell "thoroughly", just do the best you can. Now, grab the loaf pan. If you sprayed the *outside* of the pan instead of the inside, discard it, get another pan, and only spray the *inside* this time. Dump the contents of the bowl (not the empty one, because that would just be confusing) into the loaf pan and pop it in the oven (the loaf pan, not the bowl). Bake 60 – 65 minutes or until toothpick comes out clean when inserted in center of pan. Cool on rack for 15 minutes. You might wish to let the loaf cool and then cover it in plastic wrap for a couple of hours as this softens the crust.

This makes one big *gigundo* size loaf, enough to feed a family of four for a week, or one hungry distance runner for the afternoon.

Eventually, everything is history.