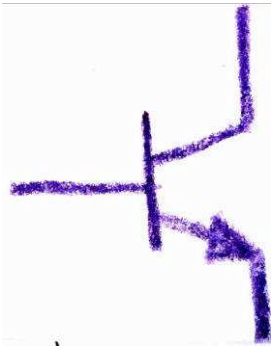
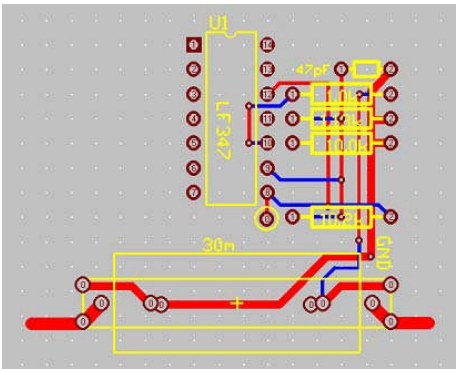
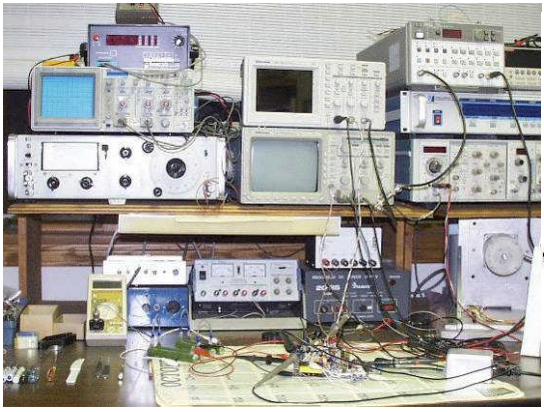
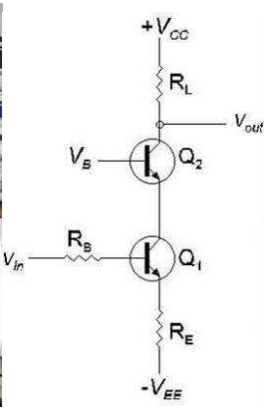
Transistor Amplifiers









Dennis L. Feucht

About the Book

Innovatia Laboratories

Revision date: APR 2014

This book is an Innovatia Laboratories publication and has no ISBN number. Innovatia publications are distributed for the purpose of encouraging a deeper yet simpler understanding of electronics, and to aid readers in improvement of their engineering skills and effectiveness on design projects.

Please distribute this PDF copy to your friends, colleagues, and associates who might benefit from its contents, and refer them to other books and booklets in the Innovatia series at

www.innovatia.com

Check for the availability of paper copies of the book at a nominal price at the above website.

If you recieve this copy free of charge and want to fiscally support Innovatia in publishing more books of this kind, please donate to the author through the above website or by postal mail at (complete address): Dennis Feucht, San Ignacio, Cayo, BELIZE, Central America. Thanks.

Please email from the above website any errors you find in this book.

The distribution form of this book is in PDF. The book was written in Microsoft Word 2003, Professional Edition. Graphics were created in CorelDraw and by digital oscilloscope and camera.

Contents

3 Matched-Transistor Circuits

4 Circuit Dynamics

5 Transistor Dynamics

6 Amplifier Design

7 Cascode Amplifiers

8 Differential and Cascade Amplifiers

9 Current- and Shunt-Feedback Amplifiers

10 A Wider View of Amplifiers

430 pages

Preface

Transistor amplifiers have been written about for well over a half century. Why another book on an old subject? I have three reasons. As a field develops, simplifications and refinements occur and misconceptions are corrected. This is still true for analog circuits. Some important concepts have yet to diffuse broadly, leaving engineers and technicians in the dark about puzzling kinds of circuit behavior. This book brings out some of what is either missing or crowded out of introductory active-circuits textbooks and corrects some misconceptions. It also presents refined ways of thinking about circuits that simplify understanding of them. A fuller engineering development of most of these themes is in my *Analog Circuit Design* book-set, and in this book, I offer fresh and different ways of thinking about neglected or confusing topics.

Second, this book attempts to address a very broad range of readership, from experienced engineers to those aspiring to be engineers, especially astute young pre-university students who are good at math and science. Consequently, the challenge is taken to develop circuits concepts with a minimum of advanced mathematics as such. The mathematical basis for Laplace transforms, differential equations, and most of differential and integral calculus (no less advanced calculus) is not actually needed to think in the complex-frequency domain, which mainly involves complex-number algebra. With a good grasp of algebra and trigonometry, it is amazing how far one can venture into analog circuit design, and in this book I set out to show that implicitly. Books containing advanced circuits engineering almost never present applied calculus needed to do circuit analysis. I have included just enough to bridge the gap for the pre-calculus student yet hopefully not annoy engineers too severely. Consequently, I hope that this book, in part, serves to bridge the gap between hobbyist electronics and engineering literature. The first part is about basic circuits concepts and the remainder is amplifier designs.

Third, my previous book-set, while presenting examples, does not do what most other books on analog circuits also do not do: “walk through” amplifier designs at an engineering level of detail as an actual design activity. About a dozen amplifiers are designed in this book, step-by-step, with explanation. By presenting a high level of detail, the reader is offered insight into the sequence of thinking of one engineer. The amount of detail is intended to leave the reader

with little or nothing to puzzle over or be left wondering how various major decisions or derivations were made.

Designs in a book cannot start with a blank screen (or piece of paper) if something specific is to be written. Circuits are presented, followed by details on how the parts values are determined. This involves a list of considerations and each is patiently gone through so that the resulting design is of industry-standard quality. Then a prototype is built and its behavior compared with design calculations. In engineering projects, often the two do not agree and for some designs in this book, they also do not and modification is required to refine the design after the incongruities have been pondered. At times the design is right and measurements on the bench are wrong because of instrumentation errors. Oscilloscope probes are capacitive loads and interconnecting cables sometimes require termination.

What is missing in the older books is discussion of "mixed-signal" A/D and D/A conversion that interfaces analog to digital circuits. Some of the difficulty in analyzing and designing systems with both is ameliorated with new concepts in how to envision ADC and DAC function, both at low speed and dynamically.

I introduce some simplifications and what I believe are improvements in notation and terminology. "AC" and "DC" are ambiguous and obsolete, and the awkward expression "low-frequency ac" was long ago avoided by the thermodynamicists, who call it *quasistatic*. Overall, the notation should be quite readable to anyone skilled in the art. I maintain the well-established use of upper- and lower-case letters and subscripts for static and incremental quantities, though I abandon the upper-lower case distinction for time and complex-frequency domains, preferring to let the domain variable identify the domain.

Finally, the last chapter, which logically would be first, expounds upon engineering and design in itself, topics that might be more interesting after one has acquired experience with them and is ripe for deeper reflection upon their wider meaning.

This book is intended to be complementary to my *Analog Circuit Design* book-set while also enabling newer designers to access it. Even the simpler circuits in this book raise issues that experienced electronics engineers can ponder. I hope that for them, a few useful insights will be added to their already considerable knowledge of transistor amplifiers. Writing this book has done that for me!

Circuit Concepts

Introduction

This book is about electronics for both those who are aspiring to be engineers and for engineers. Unlike many engineering textbooks, it includes more background material for aspiring youth, technicians, or technologists who have a serious interest in learning how to design electronic circuits. It is intended to enable the reader to apply electronics in the design and construction of working circuits. It begins with extended explanations of basic circuits concepts and continues with chapters that work through design of representative circuits in detail, well beyond what introductory engineering textbooks present.

With only pre-calculus math and science, it is possible to learn the basic principles of electronics and even design electronic circuits and systems of simple to moderate complexity. (The higher mathematics is, however, required to understand at engineering depth what more there is to know about circuit design.) By reducing complicated electronics theory to its essential concepts and then learning how to apply them, basic yet interesting electronic circuit design is achievable. Consequently, this book requires only pre-calculus knowledge of math and science, and in many cases, only elementary algebra and some basic physics. The reader is expected to know algebra, geometry, advanced algebra (complex numbers, logarithms and exponentials), trigonometry, and the algebra and geometry of vectors. Some elementary calculus is used and developed as needed. It is remarkable how little actual calculus is required to do circuit design, though the concept of limits is basic to the understanding of circuit behavior. Because of linearity assumptions, differential equations that describe circuits quickly transform to algebraic form, and by describing the behavior of circuit elements with this algebra, calculus and differential equations are avoided entirely. The reader should both like and be proficient with algebra, for it is used extensively in the analysis and design of circuits.

The reader should also have familiarity with basic electronics, know something about basic circuit laws such as Ohm's Law (ΩL) and Kirchhoff's Voltage and Current Laws (KVL, KCL), about

resistors, voltage and current sources, and maybe even Thevenin's theorem, should have some familiarity with how transistors work, and also have some practical experience "at the bench".

The electronics equipment recommended for the experiments and circuit designs in this book are a digital multimeter (DMM) that measures voltage, resistance, and transistor β (or h_{fe}) and a power supply that has +12 V and –12 V outputs and can deliver at least 50 mA from each. In addition, you will need a voltage source that you can vary from about 0 V to 5 V. (Or you can build the one given in the text.) As the book progresses, you will need a waveform source (preferably a function generator) and an oscilloscope. This equipment, along with a workbench dedicated to electronics work, constitutes the modest electronics laboratory setting in which to duplicate the analyses and designs in this book.

The laboratory requires basic electronics hand tools: a 15 to 25 watt soldering iron (18 W Antex G/3U with Antex iron tips: 2-IC recommended), a roll of rosin-core, 0.031 inch diameter solder, a soldering-iron stand (Antex), a metal-tine brush to clean off the iron tip, a wastebasket or metal recycling bucket in which to clean it off, a desoldering tool (or "solder sucker" Soldapullt model DS017) a needle-nose pliers (Diamond Utica LNB5317CG), diagonal cutter (Xcelite MS549J or Diamond Utica MS54-3J), wire stripper (Miller 102), small screwdriver (Xcelite R3323), scriber (General Tools # 83), and a component lead bender.

The parts you will need to stock in your laboratory might well be organized in Akron-Mills plastic cabinets with many transparent drawers per cabinet and handles at the top of the drawer so that you can place adhesive labels below them. Some parts needed for lab experiments are a 20 kΩ front-panel potentiometer, PN3904 NPN transistors, PN3906 PNP transistors, an assortment of 5 % and 1 %, ¼ W resistors; an assortment of low-voltage ceramic capacitors including 0.1 μF, 50 V. For active circuits, experiments use TL071 op-amps and CA3096 BJT arrays. Also needed is some prototyping circuit-board, with 0.1 inch hole spacings in both dimensions, on which to build the circuits. Sockets are recommended for integrated circuits (ICs) and transistors, for easy replacement of possibly failed parts or for comparing parts.

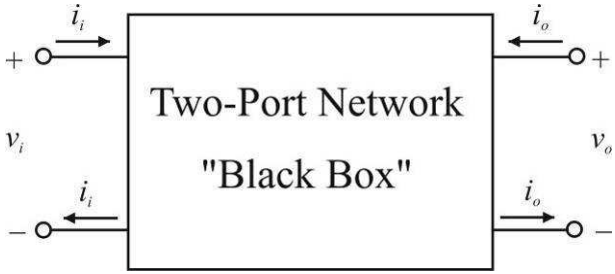
Electronic components have been shrinking in size to where it is now difficult to prototype with them; through-hole parts are recommended. However, if you prefer your laboratory inventory to be surface-mount parts, be prepared to invest some time in acquiring the

additional manual skills needed to work with them. In this case, tweezers are more useful than needle-nose pliers and a roll of solder wick for solder removal can be needed. If possible, avoid IC packages with the pinouts under the package. Side-extended pins that can be accessed with a soldering iron are preferred, though they are close in *pitch*. And to stay healthy, try to avoid breathing soldering fumes.

Waveforms, Ports, and Amplifiers

We need some words that label electronic circuit concepts. A *waveform* is an electrical quantity such as voltage or current that is a function of time. In the function notation of algebra, a voltage waveform can be expressed as $v(t)$ where v is voltage and t is time.

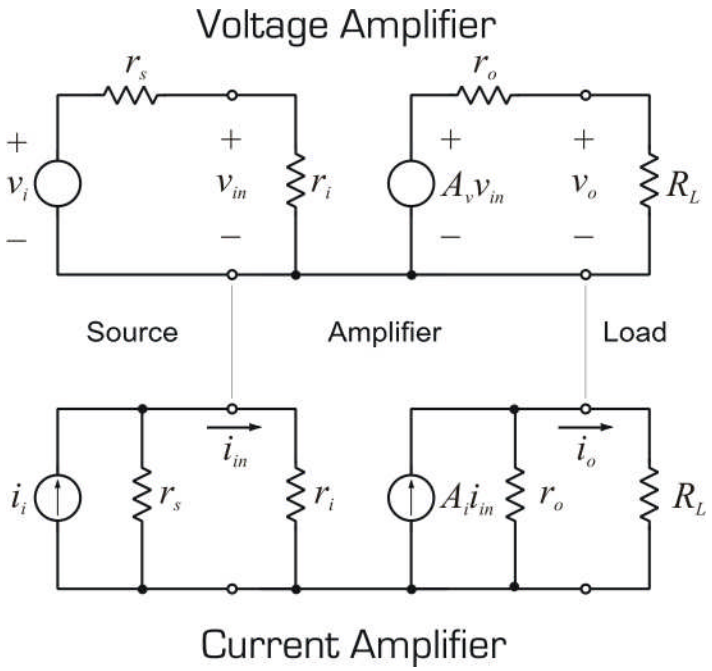
A *port* is associated with a pair of terminals of a circuit. Picture the circuit as being in a box with a couple pairs of terminals.



The box representing the circuit (or *network*) might be any circuit. It might even be inside a closed or "black" box where we cannot see what it is. Yet the circuit has two terminals coming out of the box to the left and two to the right. Each of these terminal pairs is a port. A *port* is defined by the voltage across its terminals and the current common to its terminals. The algebraic sign, or *polarity*, of the voltage is shown by the + and – signs at the terminals. If v_i (the input voltage) is 1 V, then the + terminal is one volt greater than the – terminal. If $v_i = -1$ V, then the voltage measured at the + terminal relative to the – terminal is negative. You can think of the – terminal as being the reference or 0 V terminal and v_i is then the voltage at the + terminal as measured *relative to* the – terminal.

The second defining feature of a port is the current, i_i. Positive current is defined as flowing *into* the + terminal of the port, as shown. The same current flows out of the – terminal, completing the port circuit loop. If the current flowing into the + terminal is not equal to the current flowing out of the – terminal, the terminals are not a port.

A second port, to the right (the output port) is similarly defined, with its port current and voltage variables. To refer to either voltage or current, we will use the more general variable symbol, *x*, which can be either *v* or *i*. The two-port circuit can be used to represent an *amplifier*. The input port waveform is made larger, or amplified, to become the output port waveform. The general form for two of the four basic kinds of amplifiers is shown below.



The *voltage amplifier* has a voltage-source input in series with a resistance, r_s, which need not be a resistor but can be a simplified *equivalent* resistance. It can represent resistive dividers and other complications but behaves the same as, or is equivalent to, a single resistance in series with a voltage source. This is called a *Thevenin equivalent circuit*. The input port of the amplifier "box" or *block* has a resistance across it, r_i. This also might consist of several resistors yet their equivalent resistance reduces to r_i. At the output port, the voltage from across the input port, v_{in}, is multiplied by the *voltage gain* of the amplifier, A_v; $v_{out} = A_v \cdot v_{in}$. The v_{out} source is dependent on the input port voltage. The amplifier output port also has series resistance r_o. External to the amplifier is *load resistance*, R_L, across the output port.

At the input port, the input source resistance, r_s, and the amplifier input resistance, r_i, form a voltage divider that reduces the amplifier input-port voltage to

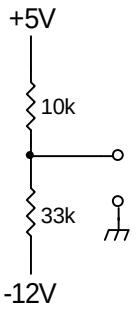
$$v_{in} = \frac{r_i}{r_i + r_s} \cdot v_i$$

A similar divider is at the output port;

$$v_o = \frac{R_L}{R_L + r_o} \cdot v_{out}$$

The dividers reduce the gain and can cause gain error. For the amplification or *gain* of the amplifier to be $v_o/v_i = A_v$, then it must be that $v_{in} = v_i$ and $v_o = v_{out}$, and the dividers must be eliminated. Consequently, an ideal voltage amplifier has infinite input resistance ($r_i \to \infty$) and zero output resistance, ($r_o = 0\ \Omega$).

The second kind of amplifier shown above is the *current amplifier*, the *dual* of the voltage amplifier. Dual circuits have their voltage and current variables interchanged. Instead of Thevenin voltage sources at input and output ports, current sources are in parallel with resistances instead. These are *Norton* equivalent circuits. Thevenin and Norton circuits can be converted to each other. Thevenin series resistance becomes Norton parallel (or *shunt*) resistance and the current source becomes $I_N = V_{TH}/R_{TH}$. The Norton source current is the short-circuit current of the Thevenin circuit port. The Thevenin source has the open-circuit voltage of the Norton circuit.



The current amplifier is ideal when it has no resistive current dividers. This occurs when $r_i = 0\ \Omega$ and $r_o \to \infty$. These conditions are the opposite, or the *dual*, of those for the ideal voltage amplifier.

The voltage divider shown here is drawn in an abbreviated form of circuit diagram. Both the +5 V and –12 V sources have their undrawn (–) terminal connected to the 0 V node, or *ground*, as shown by the symbol connected to the lower port terminal. The divider can be reduced to a Thevenin equivalent circuit by finding the open-circuit voltage (V_{OC}) and the short-circuit current (I_{SC}). To find I_{SC}, short the port (ground the divider output node). Then

$$-I_{SC} = \frac{5\text{ V}}{10\text{ k}\Omega} + \frac{-12\text{ V}}{33\text{ k}\Omega}$$
$$= 0.5\text{ mA} - 0.364\text{ mA} = 0.136\text{ mA}$$

The negative sign indicates that (according to the definition of a port) the short-circuit current is flowing out of the + divider terminal. (Unless indicated otherwise, it is assumed that the grounded terminal is the – port terminal.)

The open-circuit voltage can be found by *superposition*: the contributions to the divider voltage from each of the two voltage sources is found, one at a time, with the other source *nulled* (made zero for no contribution). Then the source contributions are added.

$$V_{OC} = \left(\frac{33\text{ k}\Omega}{33\text{ k}\Omega + 10\text{ k}\Omega}\right)\cdot 5\text{ V} + \left(\frac{10\text{ k}\Omega}{33\text{ k}\Omega + 10\text{ k}\Omega}\right)\cdot(-12\text{ V})$$
$$= 3.84\text{ V} + (-2.79\text{ V}) = 1.047\text{ V}$$

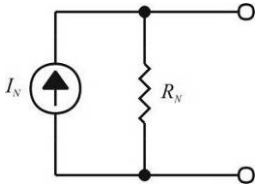
The Thevenin resistance is

$$R_{TH} = \frac{V_{OC}}{I_{SC}} = \frac{1.047\text{ V}}{0.136\text{ mA}} = 7.7\text{ k}\Omega$$

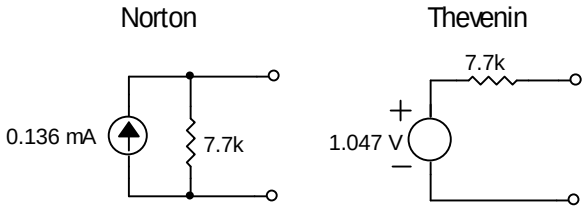
Voltage sources have zero or *short-circuit* resistance. Current sources have infinite or *open-circuit* resistance. From the divider port, the two divider resistances are in parallel through the zero-ohm voltage sources, and

$$R_{TH} = 10\text{ k}\Omega \| 33\text{ k}\Omega = \frac{10\text{ k}\Omega \cdot 33\text{ k}\Omega}{10\text{ k}\Omega + 33\text{ k}\Omega} = 7.7\text{ k}\Omega$$

The parallel resistance formula is often easier to use in finding the Thevenin equivalent resistance.



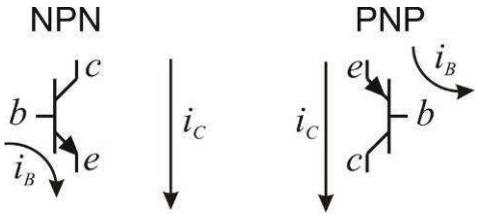
The Norton equivalent circuit is a Norton current source, I_N, in parallel with the Norton equivalent resistance, R_N. To find the Thevenin equivalent of the Norton circuit, first find $V_{OC} = I_N \cdot R_N$. Then $I_{SC} = I_N$, and $R_{TH} = V_{OC}/I_{SC} = R_N$. The Norton equivalent of the above circuit has $I_N = 0.136$ mA and $R_N = 7.7$ kΩ. The two circuits have the same port voltages and currents and are thus equivalent.



Bipolar Junction Transistors

A *transistor* as a circuit component is an *active device*, a device that can source power by amplifying voltage and current. There are two basic kinds of transistors: *bipolar junction transistors* (BJTs) which were commercially developed first, and *field-effect transistors* (FETs) which have two main categories: junction FETs (JFETs) and metal-oxide-semiconductor FETs (MOSFETs). This book will mainly present BJTs. Many of the principles carry over readily to FETs.

Both BJTs and FETs have two *polarities*. For BJTs, these are NPN and PNP. The circuit symbols for them are shown below.



A small amount of current flows between base (*b*) and emitter (*e*), i_B, causing a larger amount of current to flow between collector (*c*) and emitter, i_C. Base current flows into the base of an NPN BJT causing i_C to flow into the collector terminal. Base current flows out of the base of an PNP and i_C flows out of its collector. By inverting the PNP symbol on the page, all currents flow from top to bottom of the page. This often makes it easier to think about circuit behavior.

The gain in collector current over base current is defined as one of the two basic circuit parameters of BJTs, symbolized by the Greek letter *beta*:

$$\beta = \frac{i_C}{i_B}$$

For small changes in i_B and i_C, we denote these changes by lower-case letters and subscripts: $i_b \approx \Delta i_B$ and $i_c \approx \Delta i_C$. Then the *incremental* β is

$$\beta = \frac{i_c}{i_b} \cong \frac{\Delta i_C}{\Delta i_B}$$

The constant or unchanging component of a waveform is the *static* component, denoted by upper-case letters. For instance, for collector current, the *total* collector current,

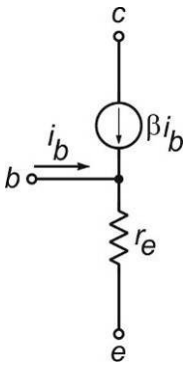
$$i_C = I_C + i_c$$

where the static current, I_C, is sometimes called the collector *bias* or *dc* current. Then i_c is the small variation of i_C around the I_C value. The static BJT voltages and currents set the *operating-point* for the transistor. When operating as an amplifier, the NPN has a base-to-emitter static voltage of $V_{BE} \approx 0.65$ V. The collector-to-base voltage is also positive for NPNs: $V_{CB} > 0$ V. Whenever $V_{CB} < 0$ V, the NPN BJT is in *saturation* and the diode that is the base-collector junction is *forward biased* and conducts. For the PNP, the polarities are reversed for normal operation: $V_{BE} \approx -0.65$ V and $V_{CB} < 0$ V.

The static and incremental (small varying) values of β are equal. We will use the PN3904 (NPN) and PN3906 (PNP) transistors for some example circuits. Their range of β is between 100 and 300 with a typical $\beta = 150$ in normal operation. In saturation, $\beta \approx 1$.

BJT T Model

Kirchhoff's Current Law (KCL) can be applied to the BJT terminals if we define positive emitter current as coming out of the NPN emitter terminal (the opposite of port convention). Then a positive change in i_E, or $\Delta i_E > 0$ A, is one in which the emitter current changes so that more current comes out of the emitter terminal.



Applying KCL,

$$i_e = i_c + i_b$$

The definition of β can be rearranged as

$$i_c = \beta \cdot i_b$$

When i_c is substituted into the first equation, then

$$i_e = \beta \cdot i_b + i_b = (\beta + 1) \cdot i_b$$

The BJT incremental (or *small-signal*) T model uses a *dependent* current source for collector current. The current-source

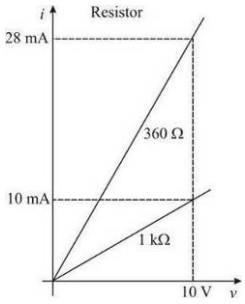
value depends on (and varies with) i_b as $i_c = \beta \cdot i_b$. This circuit model can be substituted for BJTs in a circuit diagram for incremental circuit analysis - that is, analysis of circuit behavior for small changes around a fixed operating-point.

The resistance, r_e, is the *incremental emitter resistance*. Its value can be calculated from the static operating-point emitter current;

$$r_e = \frac{26\,\text{mV}}{|\,I_E\,|}$$

At 1 mA of emitter current, $r_e = 26\ \Omega$. These two BJT *parameters* (constants during incremental analysis) define the T model equivalent circuit of a BJT biased at a linearly amplifying operating-point.

Incremental and Static Resistance



Incremental resistance is not the same as static resistance for nonlinear devices. Consider first a linear device, the resistor. Its current-voltage plots are shown for two values. The values of current for 10 V applied to 1 kΩ and 360 Ω resistors are found from Ohm's law. If the voltage varies across the resistor, the resulting change in current can also be found from Ohm's Law:

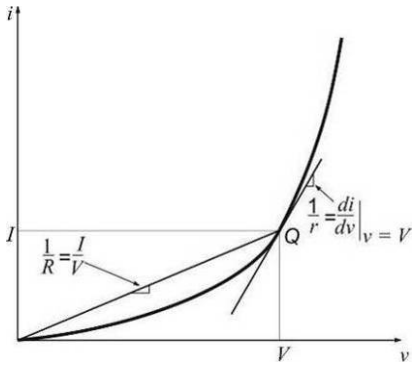
$$r = \frac{\Delta v}{\Delta i} = R = \frac{V}{I}$$

The static (R) and incremental (r) resistances are the same for linear devices.

Now consider a BJT base-emitter junction (which is a diode), a nonlinear device having an exponential function, $i(v)$, as shown below. The junction is operated at operating point Q which is at static voltage V and static current I. The static resistance is 1/slope of the line from the origin to Q and is

$$R = \frac{V}{I}$$

However, if the voltage is varied slightly around V by an amount Δv, then the change in current, Δi, around I is found from the slope of the line tangent to $i(v)$ at Q, which is written in calculus notation as di/dv.



The 1/slope of the tangent line at Q is the incremental resistance at operating-point Q;

$$r = \left.\frac{dv}{di}\right|_Q \cong \frac{\Delta v}{\Delta i}$$

$dx \approx \Delta x$ when the change in x is small and is exactly equal to it when Δx shrinks to become infinitesimal, or "infinitely small". Thus dx is an ideal infinitely-small change and any finite small change, Δx, is an approximation to it.

The slope is approximately the ratio of the changes in the total variables, Δv and Δi. The 1/slope of r is less than the 1/slope of R and $R > r$. For a typical *b-e* junction (2N3904), $V = 0.65$ V at $I = 1$ mA. Then

$$R = \frac{V}{I} = \frac{0.65\,\text{V}}{1\,\text{mA}} = 650\,\Omega$$

whereas the incremental resistance for a silicon p-n junction is

$$r \approx \frac{26\,\text{mV}}{1\,\text{mA}} = 26\,\Omega$$

The static and incremental resistances are obviously not the same. Transistors are thus nonlinear in that r_e depends on I_E. β is somewhat dependent on BJT static current but will be considered constant.

Static and Incremental Circuit Design

Circuit design consists of two different kinds of design, in the following sequence:

1. *Static* design sets the operating-point or *bias* for the (nonlinear) transistors and diodes.
2. *Incremental* design sets the incremental parameters such as gain.

Each circuit voltage and current is a total variable that is the sum of static and incremental variables:

$$x_{TOTAL} = X + x$$

Keep in mind that the incremental $x \approx \Delta x_{TOTAL}$. Static design is first because the static I_E affects the value of the incremental parameter r_e.

When incremental quantities, x, change slowly, they are called *quasistatic* changes. In some electronics literature, *quasistatic* is called "low-frequency ac". When reactances (capacitance, inductance) affect circuit behavior, then the changes are *dynamic*. *Static* is constant (unchanging in time), *quasistatic* refers to slow incremental changes for which reactances have no effect, and *dynamic* refers to incremental changes including the effects of circuit reactances. Mathematically, static and quasistatic circuit quantities are real numbers and dynamic circuits have imaginary numbers.

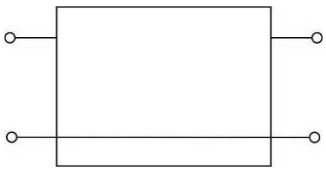
To show calculation of a static BJT quantity, a PN3904 BJT has the typical $\beta = 150$ and a static base current of $I_B = 10\ \mu A$. What is the collector current?

$$I_C = \beta \cdot I_B = (150) \cdot (10\,\mu A) = 1.5\,mA$$

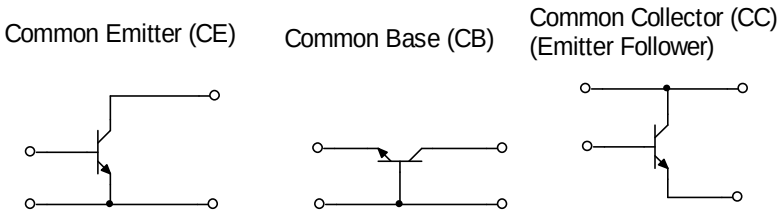
The static and quasistatic β are equal for BJTs.

BJT Configurations

Transistors have three terminals but two-port models of amplifiers have four. Two of the four terminals must have a common connection if the three transistor terminals are to correspond to two-port terminals. The common terminals are the negative input and output port terminals.



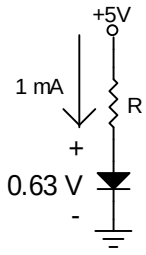
There are consequently three *configurations* of transistor two-port amplifiers, each having one of the transistor terminals as the common terminal. The three BJT configurations are shown below.



Input (left) and output (right) ports are shown as terminal pairs. The CC common (collector) node is usually not ground, yet is a *quasistatic ground*: it does not change in voltage. Additional resistors are required to connect the transistors to power supplies to set their operating point by *biasing* them.

BJT Biasing

Static circuit design biases the transistors. Diode biasing is considered first. Suppose we want to bias a diode to operate at 1 mA of static current: $I_D = 1$ mA. What is the value of R? The chosen diode is specified from its component data to have a 0.63 V drop across it at 1 mA of current. Then the voltage across the resistor is



$$5\text{ V} - 0.63\text{ V} = 4.37\text{ V}$$

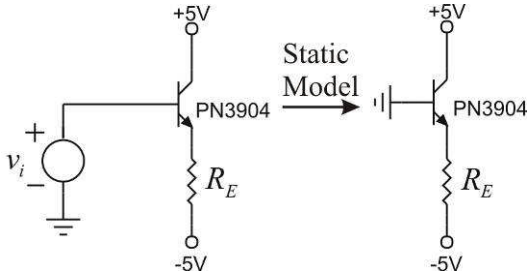
For 1 mA, then

$$R = \frac{4.37\text{ V}}{1\text{ mA}} = 4.37\,\text{k}\Omega \rightarrow 4.3\,\text{k}\Omega,\ 5\%$$

Static BJT design is illustrated by the following circuit. The CC BJT incremental voltage input source, v_i, is the varying waveform to be amplified. It is grounded in common with the +5V and –5V supplies through their supply ground terminals, as is the output voltage. It has a static voltage of 0 V. In the standard notation,

$$v_I = V_I + v_i = 0\text{ V} + v_i$$

(The average voltage of v_i is also 0 V.) As a voltage source, v_i has zero internal resistance and can be shorted to 0 V (ground), as shown in the static model of the circuit to the right.

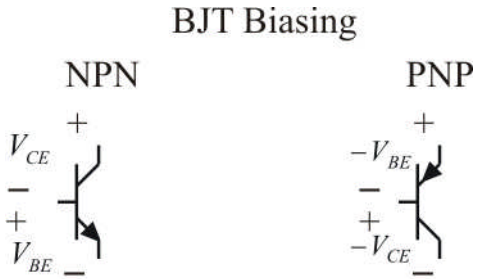


For a PN3904, a 1 mA static emitter current, or $I_E = 1$ mA, will result in a base-to-emitter voltage, $V_{BE} = 0.65$ V. Assuming we want the emitter static (or bias) current to be 1 mA, then what should we choose as the value of the emitter resistor, R_E?

The solution is like that of the diode circuit;

$$R_E = \frac{-0.65\text{ V} - (-5\text{ V})}{1\text{ mA}} = 4.35\text{ k}\Omega$$

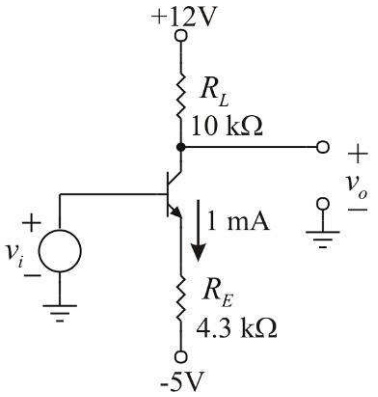
For the NPN BJT to behave linearly so that the BJT T model is valid, the base-emitter (*b*-*e*) junction must be forward biased (conducting) and the collector-base (*c*-*b*) junction reverse biased ($V_C > V_B$). For this circuit, $V_{BE} = 0.65$ V; the *b*-*e* junction is forward-biased. The *c*-*b* junction is reverse biased; $V_{CB} = 5\text{ V} > 0$ V. The BJT is in the *linear region* of operation. BJT biasing in general is shown below.



If $V_{CE} \leq 0$ V, then the BJT is *saturated* and the *b*-*c* junction is forward-biased. If V_{BE} is much less than 0.5 V, the emitter current is negligible and the BJT is *cut off*. In neither of these cases does it operate linearly and the incremental T model does not apply.

CE Amplifier

A common-emitter (CE) amplifier is shown below. The circuit has two current loops, input and output. The input current flows from the input source, v_i, into the base of the BJT, through the *b-e* junction to the emitter, through emitter resistor R_E to the –5 V source and out its ground to the input voltage source ground.



Once again, the shorthand convention for *supply* (or static biasing) sources is to not show their grounded terminal. Yet keep in mind that it exists and the supply current also flows to or from ground. (Note the different ground symbol; there are several.) The input source, v_i, is written in lower-case letters to indicate that it is an incremental source. Its voltage does not cause the input current to depart much from the operating-point current.

The –5 V supply in the input loop sets the bias current. For static analysis, v_i is set to zero; no changing voltage exists within the circuit for static (unchanging) analysis. Then the input circuit reduces to the *b-e* junction (a diode) in series with R_E and –5 V. We have analyzed this kind of circuit before. If we assume that V_{BE} will be about 0.65 V, then the emitter voltage is –0.65 V and the emitter current calculates to be the voltage across R_E divided by its resistance;

$$I_E = \frac{(-0.65\,\text{V}) - (-5\,\text{V})}{4.3\,\text{k}\Omega} = 1.01\,\text{mA}$$

The output loop has the +12 V supply in series with collector load resistor, R_L, the collector-emitter terminals of the BJT, and the emitter circuit that is shared in common with the input loop. The collector current can be calculated from the definition of β. The prior T model

terminal-current calculations with β assumed incremental (small changes in) currents but the same equations apply for constant currents because β is the same for both.

The following equations are equivalent to the T-model current equations, but they will be expressed somewhat differently to show how the polarity (the algebraic sign) of the currents is related to circuit port convention. We did not bother with ports in T-model development but instead defined the positive direction of currents as their actual directions. For a NPN BJT, emitter current comes out of the emitter terminal and we consider this the positive direction that makes I_E a positive number. However, BJT manufacturer data uses the two-port convention for polarities. By Kirchhoff's current law (KCL), the sum of the currents of the three BJT terminals must be zero, where positive current (by port definition) is current going into a terminal:

$$I_C + I_B + I_E = 0$$

If all currents flowing into a transistor are positive (by port convention), then for the sum to be zero, one or more currents must be negative. The equation can be solved for the common-terminal (emitter) current:

$$-I_E = I_C + I_B$$

The sum of the collector and base currents going into those terminals equals the emitter current coming out of its terminal. Solving instead for I_C,

$$I_C = -(I_B + I_E) = -I_B - I_E$$

On the right side of the equation, $-I_E$ is positive because emitter current flows out of the terminal, making $I_E < 0$ A. The negative sign before it makes $-I_E > 0$ A. The base current flows into the base terminal and is positive. The negative sign before it makes $-I_B$ negative. In other words, $I_B > 0$ A and thus $-I_B < 0$ A. It subtracts from the positive emitter current, $-I_E$, as it must.

If the signs of the currents are made explicit by using absolute value to denote their magnitudes (which are non-negative), then

$$|I_C| = |I_E| - |I_B|$$

From the definition of β, we know that

$$I_C = \beta \cdot I_B$$

Substituting for I_C, and continuing for now with port convention,

$$I_E = -(I_C + I_B) = -(\beta + 1) \cdot I_B$$

In other words, the emitter current is β + 1 times larger than the base current. Whereas the base current goes into the base terminal, the negative sign shows that it comes out of the emitter port terminal.

Furthermore,

$$I_B = -\frac{I_E}{\beta + 1} = -\frac{-1.01\text{mA}}{151} \approx 6.69\,\mu\text{A}$$

where a microampere (μA) is a mA/1000. I_B is positive, indicating that base current is flowing into the base. Substituting for I_B,

$$I_C = -(I_B + I_E) = -\left(-\frac{I_E}{\beta + 1} + I_E\right) = \left(\frac{\beta}{\beta + 1}\right) \cdot I_E = \alpha \cdot I_E$$

The $\beta/(\beta + 1)$ expression relates emitter to collector current and defines the new transistor parameter, *alpha*:

$$\alpha = \frac{\beta}{\beta + 1}$$

The circuit meaning of α is that it is the fraction of emitter current not lost to the base on the way from (or to in a PNP) the collector.

Returning to defined BJT currents (and leaving port convention) the CE circuit PN3904 has a typical β = 150. Thus its typical

$$\alpha = \frac{150}{151} \approx 0.993$$

The static collector current of the above circuit is then

$$I_C = \alpha \cdot I_E \approx (0.993) \cdot (1.01\,\text{mA}) \approx 1.00\ \text{mA}$$

Knowing I_C, we can calculate the collector voltage by subtracting the voltage drop across R_L from the +12 V supply:

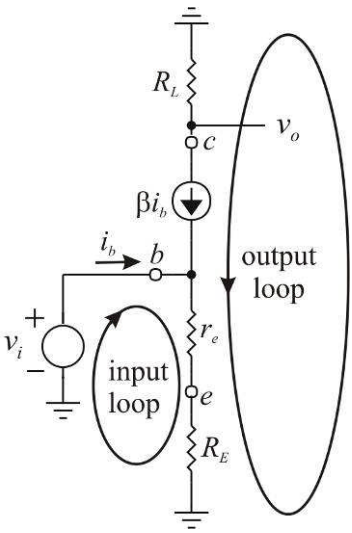
$$V_C = V_{CC} - I_C \cdot R_L = 12\,\text{V} - (1\,\text{mA}) \cdot (10\,\text{k}\Omega) = 2\,\text{V}$$

Then V_{CB} = +2 V and the transistor is not saturated. Because V_{BE} is also positive enough for significant emitter current, it is not cut off either and is thus operating in its linear region and can amplify.

From the static circuit analysis we know the constant voltages and currents in the circuit. We proceed to the incremental or "small-

signal" analysis by letting small changes in variables occur around the bias values. The source of changing values is the input source, v_i, which is $v_i = \Delta v_I = \Delta v_B = v_b$. For an increase ($\uparrow$) in v_B, then $\Delta v_B = v_b$ is positive. For a decrease ($\downarrow$) in v_B, v_b is a negative change. We can trace the sequence of effects from an increase in v_i: $v_i \uparrow \Rightarrow v_b \uparrow \Rightarrow v_e \uparrow \Rightarrow i_e \uparrow \Rightarrow v(R_L) \uparrow \Rightarrow v_c = v_o \downarrow$. A small increase in base voltage, v_i, causes an increase in emitter voltage (assuming V_{BE} is a constant 0.65 V) and this increases the emitter and collector currents. Then the voltage drop across R_L increases and subtracts from the +12 V supply to produce a decrease in collector voltage. Thus, v_o, which is Δv_C, changes with the opposite polarity of v_i and we say that the CE amplifier *inverts*. As a voltage amplifier, its voltage gain or amplification, A_v, is negative.

CE Amplifier Quasistatic Model



In the causal propagation of increases and decreases of variables through the circuit, we assumed that V_{BE} remains constant, or $v_{be} = 0$ V. However, from the T model, we know that the incremental resistance in the emitter circuit between base and emitter is r_e and though it is small, it is not zero but is

$$r_e \approx 26 \text{ mA}/|I_E| = 26 \ \Omega$$

for the above circuit. (The absolute value of I_E lets us apply the r_e formula to both NPN and PNP BJTs without regard to current polarity.) The resulting incremental quasistatic circuit model is shown. Static voltage sources are grounded.

A general method for finding the gain of BJT amplifier circuits is the *transresistance method*. It appeals to intuition about circuits because it follows the flow of cause and effect in the circuit:

$$x_i \rightarrow x_{common} \rightarrow x_o$$

For the CE circuit, this is $v_i \rightarrow i_e \rightarrow v_o$. The incremental circuit model uses the incremental T model for the transistor. The steps in solving for the gain are as follows:

1. $v_i \rightarrow i_e$: Start with the input loop and find the emitter current caused by v_i. It is a current developed across the emitter input-loop resistance, r_M, the *transresistance* of the circuit. Then

$$i_e = \frac{v_i}{r_M}$$

What is r_M? It is the resistance across which v_i develops the emitter current. In the input loop, the resistance across the v_i voltage source is

$$r_M = r_e + R_E \Rightarrow i_e = \frac{v_i}{r_M} = \frac{v_i}{r_e + R_E}$$

2. $i_e \rightarrow i_c$ (intermediate step): The emitter current is ideally the current common to both input and output loops, but in a BJT (and not a FET) some of it is lost to the base. In this intermediate step, we take this into account by introducing α. Then

$$i_c = \alpha \cdot i_e = \alpha \cdot \frac{v_{in}}{r_M}$$

For $\beta >> 1$, $\alpha \approx 1$ and can be omitted for approximate gain calculations.

3. $i_c \rightarrow v_o$: In the output loop, i_c develops the output voltage across R_L;

$$v_o = -i_c \cdot R_L$$

Then substituting for i_c from the previous steps,

$$v_o = -\left(\alpha \cdot \frac{v_i}{r_e + R_E} \right) \cdot R_L$$

Move v_i to the left side of the equation by dividing both sides by it.

$$A_v = \frac{v_o}{v_i} = -\alpha \cdot \frac{R_L}{r_e + R_E}$$

A_v is the *voltage amplification* or *voltage gain* of the CE amplifier. Besides the negative sign for polarity and α loss, the gain is a ratio of

resistances. In the numerator is the resistance across which the common current (i_e, modified by α) develops the output voltage. In the denominator is the transresistance, r_M, the resistance across which the input voltage, v_i, develops the common current, i_e.

The gain of the given CE amplifier can now be calculated;

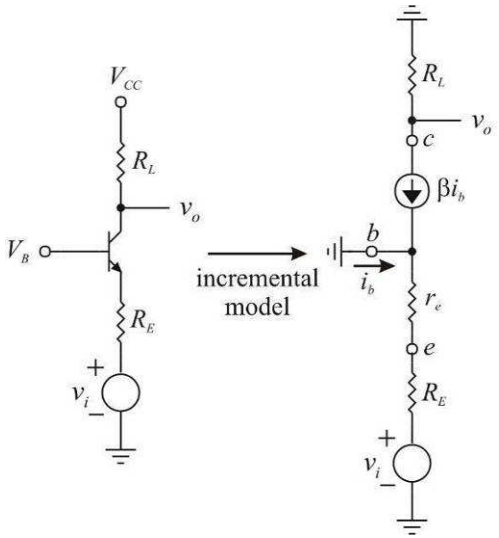
$$A_v = -\alpha \cdot \frac{R_L}{r_e + R_E} = -(0.993) \cdot \frac{10\,\text{k}\Omega}{26\,\Omega + 4.3\,\text{k}\Omega} = -2.3$$

When $R_E >> r_e$ then the effect of r_e on the gain becomes negligible and R_E *dominates*. This is an important design principle for both BJT and FET amplifiers. External resistors such as R_E can be made stable over temperature, T, while r_e changes with both T and I_E, causing the gain to change and be inaccurate. It is desirable in amplifier design that R_E dominate r_e for this reason.

With the above concepts and the equations useful for design (as design *formulas*), we are now able to calculate the values of the resistances in a CE amplifier to result in a given voltage gain and to also choose the static voltages and currents.

CB Amplifier

Shown below is a common-base (CB) amplifier and its incremental model using the T model of the BJT.



The input, v_i, now drives the emitter instead of the base. Applying the transresistance method to find the voltage gain, the first step is to find r_M. The common quantity, $x_{common} = i_e$ flows through

$$r_M = r_e + R_E$$

in the input loop, caused by v_i.

Second, base current is lost and α is included as a factor. Then

$$i_c = \alpha \cdot \left(-\frac{v_i}{r_e + R_E} \right)$$

v_i causes current to flow into the emitter; hence the negative sign for i_e which is flowing in the opposite direction through r_e in the T model. The static current flows from base to emitter as usual, but the incremental model shows only changes in quantities. As v_I increases a little, this causes a positive change that is v_i. It causes i_B to decrease a little and this causes i_E to decrease a little. Then $i_C = \beta \cdot i_B$ decreases a little. The flow of cause and effect is $v_i \uparrow \Rightarrow i_e \downarrow \Rightarrow i_c \downarrow$.

Finally, the (α-modified) common current is the output current, i_c. It develops the output voltage across R_L so that

$$v_o = -R_L \cdot i_c = -R_L \cdot \alpha \cdot \left(-\frac{v_i}{r_e + R_E} \right)$$

Then the voltage gain is

$$A_v = \frac{v_o}{v_i} = +\alpha \cdot \frac{R_L}{r_e + R_E}$$

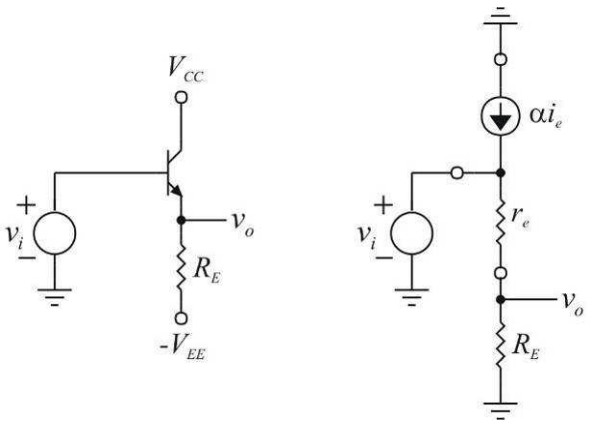
The only difference between the gain of the CE and CB amplifiers is that the CB amplifier does not invert. Whenever v_i increases, v_o also increases. Then v_i and v_o are said to be “in phase” because they change with the same polarity (increase or decrease) at the same time.

CC Amplifier

The last of the three basic BJT configurations is the *emitter-follower* or *common-collector* (CC), shown below with the quasistatic T model. Applying the transresistance method to find the gain, the resistance across which v_i develops the common (emitter) current is the same as in the previous configurations: $r_M = r_e + R_E$. What is

different is that the common current develops the output voltage across R_E without α loss. Then

$$A_v = \frac{v_o}{v_i} = \frac{R_E}{r_M} = + \frac{R_E}{r_e + R_E}$$



The gain equation is again a ratio of input and output resistances and has the form of a voltage divider with r_e the top resistance and R_E at bottom. Because of this, the CC has $A_v < 1$ but it has a current gain of

$$A_i = \frac{i_o}{i_i} = \frac{i_e}{i_b} = \beta + 1$$

for an open-circuit output. It is useful as a *buffer amplifier*, an amplifier which has an input resistance much larger than its output resistance. Referring back to the general voltage amplifier on page 4, a buffer amplifier causes r_i to be made much larger than it would be without the buffer and it reduces the gain error caused by the input voltage divider. Dividers have a gain less than one. A gain magnitude $|A| < 1$ is *attenuation*.

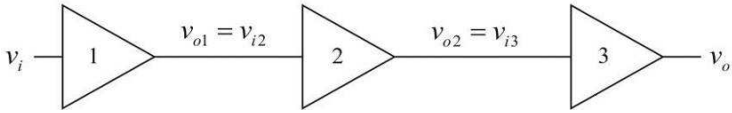
In contrast to the CC, the CB has a current gain of $\alpha < 1$, but can have $A_v > 1$. The CE is the most versatile configuration in that it can have both $|A_v|$ and $|A_i|$ greater than one.

The circuits for the three BJT configurations were kept simple to introduce the concept of the transresistance method. We have yet to derive the equations for input and output resistance of these configurations. We want to know them because of the input and output divider attenuation that can occur at the input and output of amplifiers that affects their overall gain.

Cascade Stages

When more gain is needed than a single transistor configuration can produce, the single-BJT amplifiers can be placed in *cascade*; the output of the first drives the input to the second, and so forth, as shown below in the form of a *block diagram*.

Cascade Stages



Each amplifier in the cascade connection is a *stage* of amplification. The voltage gain of the amplifier is

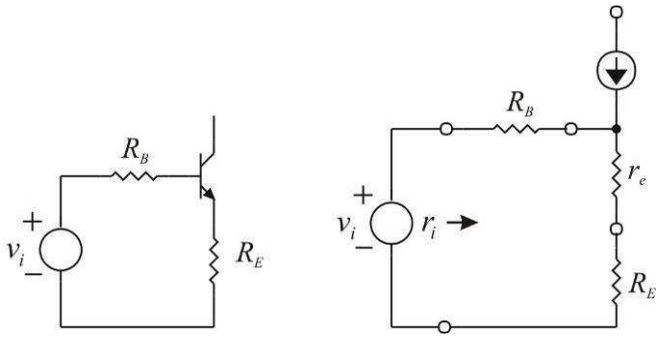
$$A_v = \frac{v_o}{v_i} = \frac{v_{o1}}{v_i} \cdot \frac{v_{o2}}{v_{i2}} \cdot \frac{v_o}{v_{i3}} = A_{v1} \cdot A_{v2} \cdot A_{v3}$$

Equate intermediate input and output voltages, $v_{o1} = v_{i2}$ and $v_{o2} = v_{i3}$ and cancel in numerator and denominator, leaving v_o/v_i.

If each stage has a gain of 5, then the overall gain is $5^3 = 125$. For a voltage amplifier, the input stage should have high input resistance and the output stage have low output resistance. The CC configuration in both cases can be a good choice.

β Transform

Another basic transistor concept is the β *transform.* It is derived from the BJT current equations we already have. It will let us include base resistance in gain formulas as needed for the following circuit.



Let $R_B = 0\ \Omega$ for now. Then the input resistance (using the lower-case r to indicate incremental resistance) is also the base port resistance,

$$r_i = \frac{v_i}{i_i} = r_B = \frac{v_i}{i_b} = \frac{v_i}{\left(\frac{i_e}{\beta+1}\right)} = (\beta+1)\cdot\frac{v_i}{\left(\frac{v_i}{r_e+R_E}\right)} = (\beta+1)\cdot(r_e+R_E)$$

The input resistance is $(\beta + 1)$ times larger than the resistance of the emitter circuit (which is r_M) because the base current is $1/(\beta + 1)$ times smaller than the current that flows through the emitter resistances. Then r_M is *referred to* (or it "appears from") the base as $\beta + 1$ times larger.

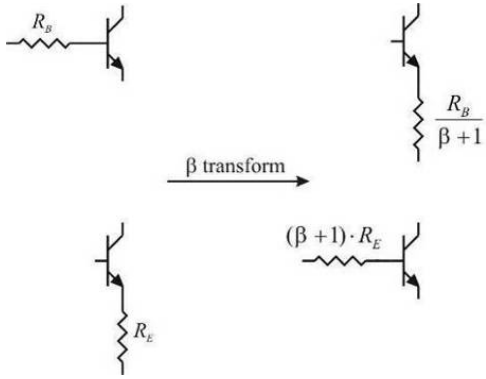
With R_B included, i_b flows through R_B. Because it is in the base circuit it is not β-transformed. Then

$$r_i = R_B + (\beta+1)\cdot r_M = R_B + r_B$$

Now consider how R_B "looks" from the emitter circuit. The voltage across R_B is v_b. When R_B is referred to the emitter circuit, an emitter-referred R_B has i_e of current (because it is in the emitter circuit) and v_b of voltage; then the base resistance referred to the emitter is

$$\frac{v_b}{i_e} = \frac{v_b}{(\beta+1)\cdot i_b} = \frac{(v_b / i_b)}{\beta+1} = \frac{R_B}{\beta+1}$$

From the emitter, resistance in the base circuit appears $1/(\beta + 1)$ times smaller. The β transform can be summarized by the following figures.



Base resistance can be moved to the emitter by dividing its value by $\beta + 1$; emitter resistance can be moved to the base by multiplying its value by $\beta + 1$.

With the β transform, R_B (and any other resistance in the input loop) can be added to the gain equations. For the CE amplifier,

$$A_v(\text{CE}) = -\alpha \cdot \frac{R_L}{r_e + R_E + \dfrac{R_B}{\beta + 1}}$$

Now the transresistance is

$$r_M = r_e + R_E + \frac{R_B}{\beta + 1} = R_E + r_E$$

where r_E is defined as the emitter port resistance, or $r_e + R_B/(\beta + 1)$. This is the resistance *into* the emitter port. The resistance looking out of the emitter terminal is connected external to it and is R_E.

Input and Output Resistance

The β transform also makes it possible to determine amplifier input and output resistances. Input resistance of the CE amplifier is

$$r_i(\text{CE}) = R_B + (\beta + 1) \cdot (r_e + R_E) = R_B + r_B$$

This follows immediately from the β transform and also applies to the CC amplifier:

$$r_i(\text{CC}) = R_B + (\beta + 1) \cdot (r_e + R_E) = R_B + r_B$$

The CB amplifier, with R_B, has an input resistance in the emitter circuit:

$$r_i(\text{CB}) = R_E + r_e + R_B/(\beta + 1) = r_M$$

and is $r_i/(\beta + 1)$ of the other two configurations.

The last of the three most important amplifier circuit design characteristics is incremental output resistance, r_o. The CE and CB stages both have for their output circuits R_L in parallel with (or *shunting*) a collector current source. Current sources have infinite resistance and are like open circuits, so that $R_L \parallel (\infty\ \Omega) = R_L$. Then

$$r_o(\text{CE}) = r_o(\text{CB}) = R_L$$

For the CC stage, R_E is in parallel with

$$r_E = r_e + R_B/(\beta + 1)$$

where r_E is the resistance of the emitter port - that is, looking into the emitter. Then the CC output resistance is

$$r_{out}(\text{CC}) = R_E \parallel r_E \approx r_E,\ R_E >> r_E$$

Just as the port for r_i (CB) was in the emitter circuit, r_o(CC) is also.

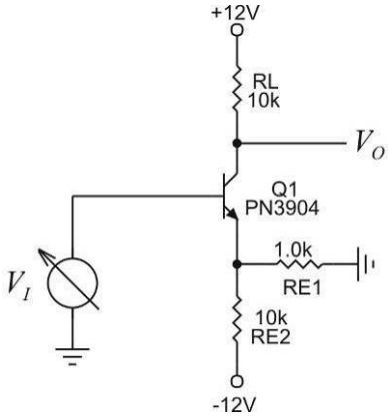
Input and output resistances are important because they can form voltage or current dividers at the input or output of a stage and attenuate the interstage voltage or current, thereby affecting the gain. When stages are cascaded, the input resistance of the second stage can form a divider with the output resistance of the first stage, and this must be included in gain calculations.

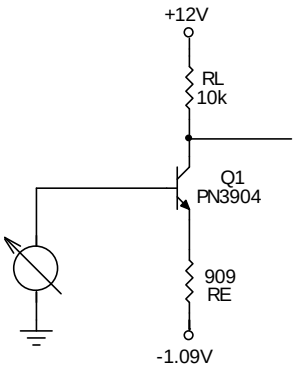
Lab Experiment: BJT CE Amplifier

The following amplifier power supplies should have 1 μF, 25 V (or higher) ceramic capacitors (not shown) placed across both of them (to ground). This should be a standard practice for all powered circuits and will be henceforth assumed. The capacitors provide a low "resistance" path for variations in current so that the amplifier power terminals are a low-resistance voltage source.

A variable-voltage laboratory power supply (V_I) is connected to the BJT base and set to 0 V. A digital multimeter (DMM) is used to measure V_I and V_O. Change V_I from 0 V to 0.1 V (100 mV), measure v_O and calculate the change in v_O: $\Delta v_O = v_o$. Then calculate the gain:

$$A_v = \frac{v_o}{v_i} = \frac{\Delta v_O}{100\,\text{mV}}$$





Now calculate from circuit theory what the gain should be;

$$A_v = -\alpha \cdot \frac{R_L}{r_e + R_{E1} \parallel R_{E2}}$$

$$= -(0.993) \cdot \frac{10\ \text{k}\Omega}{54\ \Omega + 909\ \Omega} = -10.3$$

To calculate r_M, thevenize the emitter circuit. The –12 V supply, R_{E1}, and R_{E2} have a Thevenin equivalent circuit as shown. The equivalent series resistance is

$$R_E = R_{E1} \parallel R_{E2} = \frac{R_{E1} \cdot R_{E2}}{R_{E1} + R_{E2}} = 909\ \Omega$$

Static analysis results in the Thevenin voltage of

$$V_{TH} = (-12\ \text{V}) \cdot \left(\frac{R_{E1}}{R_{E1} + R_{E2}} \right) = -1.09\ \text{V}$$

Then (more static analysis)

$$I_E = \frac{-0.65\ \text{V} - (-1.09\ \text{V})}{909\ \Omega} = 485\ \mu\text{A},\ I_C = \alpha \cdot I_E = 482\ \mu\text{A}$$

The static output voltage is

$$V_C = 12\ \text{V} - I_C \cdot (10\ \text{k}\Omega) = 7.18\ \text{V}$$

From the static I_E we can find

$$r_e = \frac{26\ \text{mV}}{0.485\ \text{mA}} = 54\ \Omega$$

Plugging r_e into the above gain equation, the result is a gain of –10.3.

Continuing with the amplifier design equations,

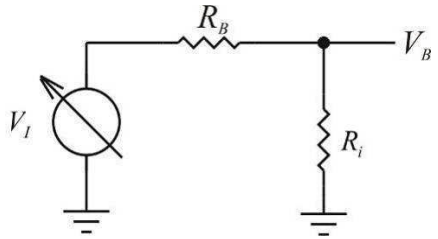
$$A_v = -\alpha \cdot \frac{R_L}{r_e + R_E} \Rightarrow R_E = \frac{\alpha \cdot R_L}{A_v} - r_e$$

$$R_{E1} = \left(\frac{V_{EE}}{V_{TH}}\right) \cdot R_E \;;\; R_{E2} = \frac{1}{\left(\frac{1}{R_E} - \frac{1}{R_{E1}}\right)}$$

Select R_L for r_o and set $V_C \approx (V_{CC} - V_B)/2$ for maximum linear output voltage range. Then $I_C = (V_{CC} - V_C)/R_L$.

How can the static R_i be measured? Form a voltage divider, as shown below. R_B is a known, measured resistor. Start with a measured value of around 100 kΩ. Set V_I = 1 V and measure V_B. Then using the divider formula,

$$V_B = \frac{R_i}{R_B + R_i} \cdot V_I \Rightarrow R_i = \frac{V_I}{V_B - V_I} \cdot R_B$$



To find the incremental r_i, measure v_B for v_I = 0 V. Change v_I to 100 mV and measure v_B. Then using the divider formula,

$$r_i = \frac{\Delta v_I}{\Delta v_B - \Delta v_I} \cdot R_B$$

where Δv_I = 100 mV – 0 V = 100 mV and Δv_B is the change in v_B corresponding to Δv_I. Is $R_i = r_i$? Only if the port circuit loop is linear.

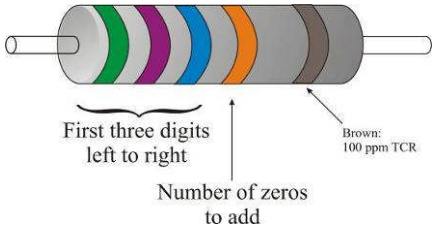
The theoretical value of $r_i = r_B = (\beta + 1)\cdot(r_e + R_E)$ = 145 kΩ. For more accuracy, measure β on a DMM and use its value.

Temperature Effects

The most commonly-used resistors are carbon-film or metal-film 5 % and 1 % tolerance resistors. They are marked with color bands, from the left edge to the right, as shown for a 1 % resistor.

The 5 % resistors have three bands plus an optional fourth band near the right edge. The 1 % resistors have four bands plus the

optional right-edge band which indicates the *temperature coefficient* (TC). The first 2 bands on 5 % and three bands on 1 % are the significant figures, given as a color code. The next band to the right is the number of zeros to add. A 576 kΩ, 1 % resistor is shown below.



Two additional colors are included for the rightmost band: gold is 0.1 and silver is 0.01. These are used for low-value resistors.

Color (abbreviation)	Numeral
Black (BLK)	0
Brown (BRN)	1
Red (RED)	2
Orange (ORG)	3
Yellow (YLW)	4
Green (GRN)	5
Blue (BLU)	6
Violet (VIO)	7
Gray (GRY)	8
White (WHT)	9

For the above resistor, a color sequence of green, violet, blue, orange → 5, 7, 6, 3 or 576000 = 576 kΩ, 1 %. A sequence of brown, green, black → 1, 5, 0 or 15 Ω, 5 %. The final zero adds zero zeros, which is none. For a 1 % resistor with yellow, white, white, red → 4, 9, 9, 2 is 49.9 kΩ. Orange, black, brown, yellow → 3, 0, 1, 4 or 3010000 = 3.01 MΩ, 1 %. A final band indicates TC. Brown is T1 = 100 ppm/°C. The colors are in the order of the rainbow.

BJT β also varies somewhat with temperature, about +1 %/°C. This causes α to vary by $1/(\beta + 1)$, about 100 parts per million (100 ppm). BJTs with β values that are not too close to one will have negligible α dependence on temperature. BJTs usually have β over 20 and typically over 100.

Stable gain results from making R_E dominate over r_e. The tradeoff for stable gain through R_E domination is reduced gain. More stages

are required but gain is more stable. r_e not only varies with I_E but also with temperature. The more complete formula for calculating r_e is

$$r_e = \frac{V_T}{|I_E|}\ ,\ V_T = \frac{k_B \cdot T}{q_e} \approx 25.85\,\text{mV}\ ,\ 300\,\text{K}$$

where V_T is the *thermal voltage* of p-n junctions; k_B (Boltzmann's constant) and q_e (electron charge) are physical constants. T is the *absolute temperature*, the temperature as measured in Kelvins (K). Absolute zero Kelvin is –273.16 °C. Ambient (air) temperature in an enclosure is about 300 K (27 °C ≈ 80 °F) and many electronic component characteristics are based on this temperature.

Resistors also have temperature coefficients which are the measure of how much the resistance will change with temperature. It is usually given in units of ppm/°C. Typical metal-film resistors are 100 ppm/°C. That is, the *fractional* (%) *TC* is

$$TC\%(R) = \frac{\Delta R / R}{\Delta T}$$

$\Delta R/R$ is the *fractional* change in R caused by ΔT. If a 10.0 kΩ resistor has a TC = 100 ppm/°C (which is also 100 ppm/K; Δ1 °C = Δ1 K), then for a resistor specified over a temperature range of 25 °C ± 15 °C, the maximum fractional deviation in its resistance from its 25 °C value will be (±100 ppm/°C)·(15 °C) = ±1500 ppm, or ±0.15 %. The actual change in the resistance is

$$(10.0\ \text{k}\Omega)\cdot(\pm 1500\cdot 10^{-6}) = (10.0\ \text{k}\Omega)\cdot(\pm 0.0015) = \pm 15\ \Omega$$

where ± indicates the *tolerance range*. A 10 kΩ, ±5 % resistor has a tolerance range from (10 kΩ)·(1 – 5 %) = (10 kΩ)·(0.95) = 9.5 kΩ to (10 kΩ)·(1.05) = 10.5 kΩ.

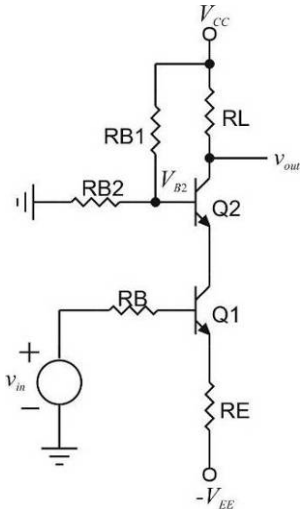
The fractional TC of β is

$$TC\%(\beta) = 1\,\%/^\circ\text{C} = \frac{d\beta / \beta}{dT} = \frac{1}{\beta}\cdot\frac{d\beta}{dT} \approx \frac{1}{\beta}\cdot\frac{\Delta\beta}{\Delta T}$$

$\Delta\beta/\beta$ is the fraction of β that changes for a temperature change of ΔT. The fraction is unitless. Hence a fractional TC will have units of 1/°C = °C^{-1} or 1/K = K^{-1}.

Cascode Amplifier

Not only can single-BJT configurations be cascaded, some combinations of them perform better than single-BJT amplifiers. A CE stage followed by a CB stage is one of them. It is called a *cascode* amplifier, as shown below. Q1 is the CE stage and Q2 is the CB stage. Beginning with r_o, it is that of the CB output stage, or $r_o = R_L$. The CE input resistance is the cascode r_i and is (using the β transform)



$$r_i = R_B + (\beta_1 + 1)\cdot(r_{e1} + R_E)$$

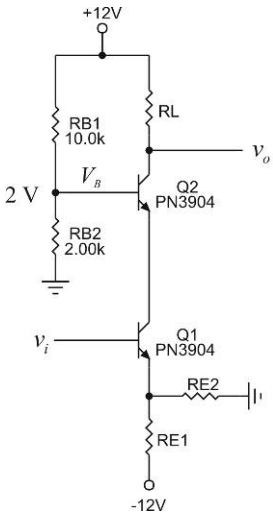
The voltage gain (using the transresistance method) is

$$A_v = -\alpha_1 \cdot \alpha_2 \cdot \frac{R_L}{r_{e1} + R_E + R_B/(\beta_1 + 1)}$$

Both r_M of Q2 and R_L of Q1 are infinite and cancel when A_v(CE) is multiplied by A_v(CB). The Q2 emitter port resistance is

$$r_{E2} = r_{e2} + (R_{B1} \parallel R_{B2})/(\beta_2 + 1)$$

It does not affect the voltage gain because the output voltage develops across R_L instead. There will be a (small) voltage drop across r_{E2} that can affect dynamic behavior when circuit reactance is included. However, we are only concerned with quasistatic behavior for now.



The r_M of the cascode is that of the CE stage; the CB stage does not determine the common i_{e1} which, after α_1 becomes $i_{c1} = i_{e2}$. Then another α loss occurs in Q2 and both $\alpha_1 \cdot \alpha_2$ appear as factors in the gain expression.

Lab Experiment: Cascode Amplifier

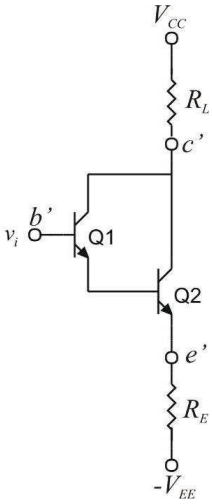
For the cascode amplifier shown, choose $V_C = V_O$ to maximize the linear range at the output, from V_B to V_{CC} = 12 V. That is, set

$$V_C = \frac{12\text{ V} - V_B}{2} + V_B = 7\text{ V}$$

Then the range is $\Delta V_O = v_o = v_c = \pm 5$ V for $\Delta v_I = v_i = \pm 1$ V. Set $I_E = 485$ μA, then calculate values for R_{E1}, R_{E2}, R_L, and A_v. $V_B = 2$ V might not be the optimal value for maximum output range, depending on the choice of A_v. Later, we will derive a formula for V_B that maximizes the range of v_O.

To keep the Q2 base voltage constant (as a voltage source), connect a 0.1 μF to 1 μF ceramic capacitor from it to ground so that varying base current will not cause v_{B2} to vary.

Darlington Circuit



Another combination of 2 BJTs is called a *Darlington* circuit (after the inventor). The two BJTs combine to form the equivalent of a high-β BJT, shown with primed terminal designations. The collector currents combine at *c'*, the equivalent emitter current is that of Q2, and the *b'* current is the base current of Q1.

The Darlington circuit can be used as an amplifier as shown. If the output is taken from across R_L then Q2 is a CE stage; if from R_E then Q2 is a CC stage. In either case, $i_{e1} = i_{b2}$.

Applying the β-transform to both Q1 and Q2 of this combination,

$$r_i = (\beta_1 + 1) \cdot (r_{e1} + (\beta_2 + 1) \cdot (r_{e2} + R_E))$$

The gain to the collector is the sum of the gains of Q1 and Q2 because their collector currents add;

$$A_v = A_{v1} + A_{v2} = -\alpha_1 \cdot \frac{R_L}{r_{e1} + (\beta_2 + 1) \cdot (r_{e2} + R_E)} - \alpha_2 \cdot \frac{R_L}{r_{e2} + R_E + r_{e1}/(\beta_2 + 1)}$$

Had there been R_B in the base of Q1, it would have appeared referred to the emitter of Q1 as $R_B/(\beta_1 + 1)$ in the first gain term and then that resistance is referred again by Q2 to its emitter as $(R_B/(\beta_1 + 1))/(\beta_2 + 1)$ or about $R_B/(\beta + 1)^2$, a very reduced resistance.

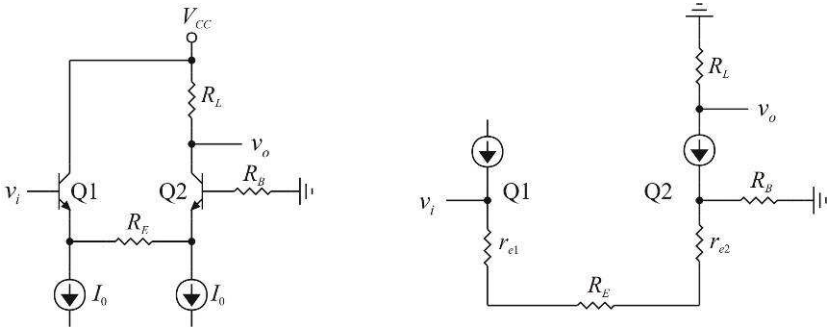
The Darlington amplifier gain can be simplified by approximating it as

$$A_v \approx -\alpha_2 \cdot \frac{R_L}{r_{e2} + R_E}\ ,\ \beta_1, \beta_2 >> 1$$

The emitter resistance of Q1 is large, causing its gain contribution to A_v to be small. In effect, Q1 is a CC stage that provides large r_i while Q2 as a CE stage provides the voltage gain.

Emitter-Coupled Diff-Amp

Another quite useful and common 2-BJT combination is the *emitter-coupled differential amplifier* pair of BJTs (or 2-BJT *diff-amp* for short), shown below with its incremental equivalent circuit.



The emitter current sources, I_0, do not appear in the incremental circuit model because they are static (constant-current) sources (as is V_{CC}, a static voltage source). Constant-current sources are opened and constant-voltage sources are shorted in the incremental model. A current source has no change in current for a change in voltage across it. Then $\Delta v/\Delta i \rightarrow \infty\ \Omega$, an open circuit. Likewise, a change in current from a voltage source does not change its voltage: $\Delta v/\Delta i = 0\ \Omega$, a short circuit. Thus for incremental models, independent current sources are opened and independent voltage sources are shorted.

A *parameter* is a mathematical concept of a constant, such as a component value, that can be changed for repeated circuit analysis. The parameter is changed, then held constant while the analysis is performed and the effect on the circuit observed. Design specifications are often given as a list of circuit parameters with their allowable ranges. The resistance parameters of interest for the emitter-coupled diff-amp are

$$r_o = R_L\ ;\ r_i = (\beta_1 + 1) \cdot (r_{e1} + R_E + r_{e2} + R_B\ /(\beta_2 + 1))$$

The β transform is used at the base of Q1 to refer emitter resistance to the base with (β_1 + 1). Then in the emitter circuit, R_B is referred from the Q2 base to the emitter circuit using the β transform a second time, by dividing R_B by (β_2 + 1). The voltage gain is

$$A_v = +\alpha_2 \cdot \frac{R_L}{r_{e1} + R_E + r_{e2} + R_B/(\beta_2 + 1)} = +\alpha_2 \cdot \frac{R_L}{r_{e1} + R_E + r_{E2}}$$

Q1 functions as a CC stage and Q2 as a CB stage. Neither inverts the amplified waveform. Only the CE configuration inverts.

The 2-BJT amplifier has an input that is *differential*; the difference voltage between the two inputs is amplified: $v_i = v_{i+}$ and the grounded base resistor of Q2 is $v_{i-} = 0$ V. Then the amplifier input voltage is the differential voltage,

$$v_i = v_{i+} - v_{i-}$$

If the Q2 base resistor were not grounded but driven with another input voltage, it would amplify with the same gain magnitude as v_{i+} except that it would be inverted; the sign of its gain would be negative.

With two inputs and one output, the gain of the amplifier is

$$A_v = \frac{v_o}{v_i} = \frac{v_o}{v_{i+} - v_{i-}}$$

The two paths to the output add by superposition:

$$v_o = A_{v+} \cdot v_{i+} - A_{v-} \cdot v_{i-}$$

The amplifier is made symmetrical by adding $R_{B1} = R_B = R_{B2}$. For a symmetric circuit, gains for the two paths are the same;

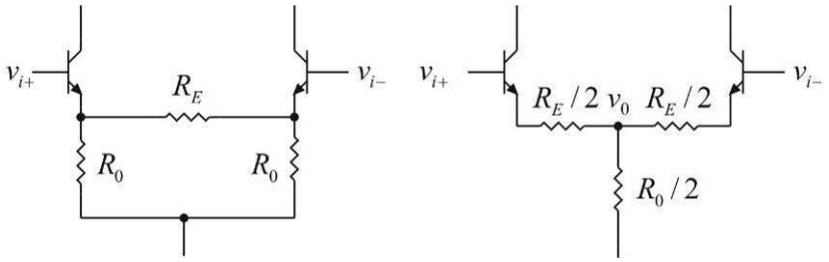
$$A_v = A_{v+} = A_{v-}$$

Substituting into the v_o equation, $v_o = A_v \cdot (v_{i+} - v_{i-}) = A_v \cdot v_i$.

The I_0 current sources are static and are often implemented by a resistor, R_0, to the negative supply. R_0 and the –12 V supply provide a "long-tailed" emitter current source of I_0. A *long-tail* is a current source approximated by a large voltage in series with a large resistor. If R_0 is much larger than the other resistances in the emitter circuit, such as R_E, r_{E1} or r_{E2}, it approximates a current source.

An alternative circuit variation is to have one current source with a value of $2 \cdot I_E$. It can also be approximated by long-tail sources. The

two diff-amp emitter configurations are shown below. The left circuit has a Π (pi) resistor network; the right circuit has a T network. The emitter-to-emitter resistance of the input loop of both is R_E.



When R_0 is included for greater accuracy in the gain calculation of the T configuration,

$$r_{M1} = (r_{E1} + \frac{R_E}{2}) + \frac{R_0}{2} \| (r_{E2} + \frac{R_E}{2})$$

If the circuit is symmetrical, $r_{E1} = r_{E2}$ and $r_{M2} = r_{M1}$.

$R_0/2$ forms a divider in the T network with $r_{E1} + R_E/2$ so that the input voltage to the Q2 CB stage is attenuated at $R_0/2$ to v_0. The transfer function is

$$\frac{v_0}{v_{i+}} = \left(\frac{R_0/2}{(r_{E1} + R_E/2) + R_0/2} \right)$$

Then $v_0 < v_{i+}$ is the input to the CB stage and the diff-amp circuit is no longer symmetrical.

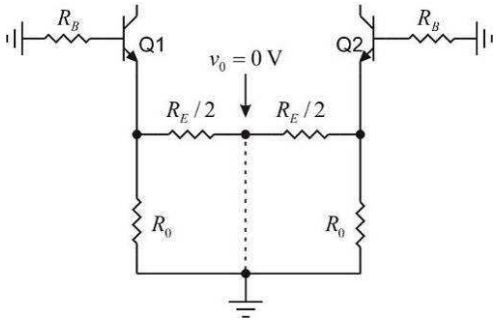
Another way to view the asymmetry is that $R_0/2$ forms a current divider with $r_{E2} + R_E/2$ and diverts some of i_{e1} away from the Q2 emitter. The current divider, like the voltage divider, is an intermediate attenuation;

$$\frac{i_{e2}}{i_{e1}} = \left(\frac{R_0/2}{(r_{E2} + R_E/2) + R_0/2} \right)$$

Either voltage or current divider transfer function is included in the gain formula for the diff-amp, and $A_{v+} < |A_{v-}|$. They are equal for circuit symmetry whenever $r_{E1} = r_{E2}$.

A fully-differential amplifier has both differential inputs and outputs. If both BJTs have equal collector resistors, then the output is

$$v_o = v_{o+} - v_{o-} = v_{C2} - v_{C1}$$

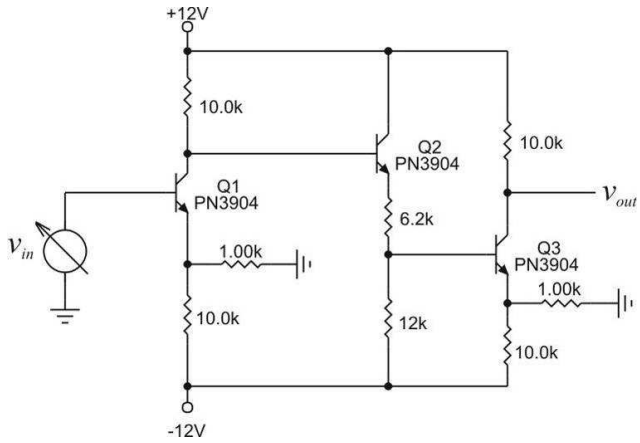


The gain magnitudes from either input to either output are the same for a symmetrical diff-amp because the divider R_0 forms the same attenuation for both sides of the amplifier. As shown, a symmetrical fully-differential amplifier with Π configuration and balanced inputs ($v_{i+} = -v_{i-}$) has a center node voltage, v_0, that is an incremental 0 V or *virtual ground*. By setting v_0 to 0 V, r_M for either side can be found by splitting the circuit down the middle, as shown. Equivalently, $R_0/2$ for the T configuration can be split into two parallel resistors of value R_0. Each half is moved to a side of the diff-amp with $R_E/2$. Then the r_M of either side is

$$r_M = R_0 \parallel \frac{R_E}{2} + r_E$$

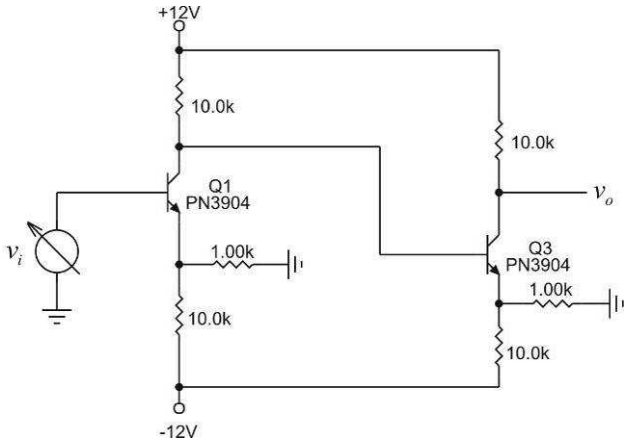
and the differential gain is the gain of either side.

Cascade CE Stages



Cascade stages were shown (page 22) that can interact to form dividers. The cascade amplifier with three stages demonstrates the divider effect. The first and third stages of this amplifier are CE stages with a CC stage between them for buffering. The buffering essentially eliminates the divider effect by isolating the CE stages from each other.

To show the divider effect, removal of the middle stage results in the simpler amplifier shown below.



The second-stage r_i forms a divider with the collector resistor of the first stage, or *loads* the first-stage output. Because the first stage does not have zero r_o, it forms a divider with r_i of the second stage. Two methods of analysis can be used for loaded stages:

The *thevenizing method*: Calculate the stage-1 gain with the second stage disconnected. Then v_{o1} is the open-circuit (Thevenin) voltage. Find r_o, the resistance in series with this voltage. The two form a Thevenin equivalent output circuit. Then connect r_i of the second stage and calculate the stage-2 gain. The Thevenin output resistance of the first stage becomes a base resistance for the second stage.

The gains are as follows.

$$A_{v1} = \frac{v_{c1}}{v_i} = -\alpha_1 \cdot \frac{R_{L1}}{r_{e1} + R_{E1}} = -(0.993) \cdot \frac{10.0\,\text{k}\Omega}{54\,\Omega + 909\,\Omega} = -10.31$$

$$A_{v3} = -\alpha_3 \cdot \frac{R_{L3}}{r_{e3} + R_{E3} + R_{L1}/(\beta_3 + 1)}$$
$$= -(0.993) \cdot \frac{10.0\,\text{k}\Omega}{54\,\Omega + 909\,\Omega + 10.0\,\text{k}\Omega/151} = -9.65$$

Then the total gain is $A_v = A_{v1} \cdot A_{v3} = 99.52$.

The *loaded-divider method*: Calculate stage-1 gain while loaded by stage 2:

$$A_{v1} = \frac{v_{b3}}{v_i} = -\alpha_1 \cdot \frac{R_{L1} \parallel r_{i3}}{r_{e1} + R_{E1}}$$
$$= -(0.993) \cdot \frac{10.0\,\text{k}\Omega \parallel [151 \cdot (54\,\Omega + 909\,\Omega)]}{54\,\Omega + 909\,\Omega} = -9.65$$

The loading of the second stage causes A_{v1} to be less than the open-circuit value of the thevenizing method. The loaded second-stage gain is the gain from the actual node voltage, v_{b3} (instead of open-circuit Thevenin voltage v_{c1}) to v_o;

$$A_{v3} = -\alpha_3 \cdot \frac{R_{L3}}{r_{e3} + R_{E3}} = \frac{v_o}{v_{b3}}$$
$$= -(0.993) \cdot \frac{10.0\,\text{k}\Omega}{963\,\Omega} = -10.31$$

A_{v3} has no base resistance in r_M because it is calculated from the actual stage-3 input voltage. In the loaded-divider method, the loading is taken into account in calculation of the stage-1 gain while in the Thevenizing method, it is taken into account in calculation of the stage-3 gain. In either case, the loading is included.

Combining the stage gains, $A_v = A_{v1} \cdot A_{v3} = 99.52$, the same as for the thevenizing method.

When building the above circuit, keep in mind error tolerances on parts and in calculations. Expect agreement to within a few percent. What is the purpose of the stage-2 CC stage? Stage-3 $r_{i3} = 140.5$ kΩ and stage-1 $r_{o1} = 10.0$ kΩ. The stage-2 input and output resistances are $r_{i2} = 2.606$ MΩ and $r_{o2} = 4.12$ kΩ. Then $r_{o1} << r_{i2}$ and $r_{i3} >> r_{o2}$. Stage 2 reduces the effects of loading.

Feedback Amplifiers

Four Amplifier Types

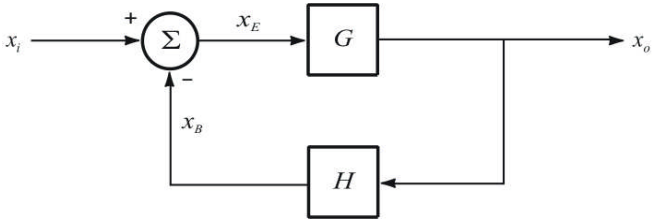
On page 4, two types of amplifiers are shown: voltage and current. As two-port devices, amplifiers can have either a voltage or current input or output. The four combinations are given in the table.

Amplifier Type	Input	Output	Gain
Voltage	v_i	v_o	$A_v = v_o/v_i$
Transresistance	i_i	v_o	$R_m = v_o/i_i$
Transconductance	v_i	i_o	$G_m = i_o/v_i$
Current	i_i	i_o	$A_i = i_o/i_i$

The two new types are *transconductance* and *transresistance* amplifiers. These remaining types combine current and voltage at input and output. As a result, they are not unitless as are voltage and current amplifiers but have units of resistance or 1/resistance (conductance). A transresistance amplifier outputs a voltage with a current input and its transfer function is of the form $v_o/i_i = R_m$. The resistance is a *trans*resistance, meaning that it transfers across from input to output port. A similar meaning is attached to transconductance.

Feedback

Cascaded amplifier stages have gains that are only as accurate and stable as each of the stages. As more stages are added, accuracy and changes with time and temperature, or *drift*, increases. A kind of circuit that reduces drift is the *feedback amplifier*. The general block diagram for feedback is shown below.



A *block diagram* is a *functional* description of a circuit whereas a circuit diagram tells us its *structure. Design* is the activity of starting with a given function, usually as a *specification*, and ending with a structure that behaves as specified.

Block diagrams are a graphic representation of algebraic equations. They show the flow of cause and effect. Written inside the block is the *transfer function* - one of the four types of amplifier gain - and is the ratio of the output of a block over the input quantity. The round symbol with Σ inside is a summing block, or *summer*, that adds or subtracts depending on the sign by the input arrows.

In the feedback diagram, the forward path through the amplifier is *G* and the feedback path is *H*. In general, *x* can represent either a voltage or current. x_B is the fed-back quantity that returns some of the output to the input. Usually *H* is a resistive voltage divider made of resistors and can be made stable and accurate. Then x_B is an attenuated x_o. It is compared with the input by subtracting the two to result in the error quantity,

$$x_E = x_i - x_B = x_i - H \cdot x_o$$

Then the forward path amplifies the error as the output quantity:

$$x_o = G \cdot x_E$$

The block diagram is a graphic representation of the algebra of feedback. Substituting x_E into the above equation,

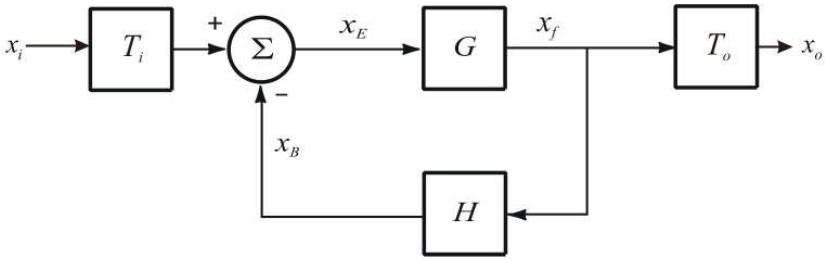
$$x_o = G \cdot (x_i - H \cdot x_o) = G \cdot x_i - G \cdot H \cdot x_o$$

Solving for the transfer function, the *closed-loop gain* is

$$A(\text{cl}) = \frac{x_o}{x_i} = \frac{G}{1 + G \cdot H}$$

This is the closed-loop *feedback formula.*

The challenge in going between block diagram and circuit is to identify the quantities of the block diagram in the circuit. It is often the case that adjoining circuits that are not part of the feedback loop itself will be intertwined with loop circuitry, making it harder to separate them. For the general case, it is best to include two more blocks on either side of the loop, as shown below.



These casacaded blocks (which are circuits performing some *function*) are multiplied to the loop feedback formula;

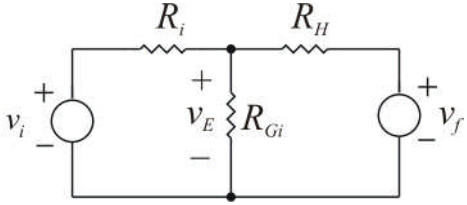
$$\text{closed-loop } A = \frac{x_o}{x_i} = T_i \cdot \left(\frac{G}{1 + G \cdot H} \right) \cdot T_o$$

The feedback loop gain is the factor between T_i and T_o. The feedback loop output is x_f, the feedback quantity that is the input to H. Most of the decision-making in analyzing feedback circuits is to choose x_E and x_f, the error and feedback quantities.

Just as there are four kinds of amplifiers, there are also four kinds of feedback circuits. The summing block adds either circuit voltages or currents. Only quantities of the same kind - with the same units - can be added or subtracted. If $T_i v_i$ is a voltage, then x_B must be a voltage, v_B. The four feedback possibilities are tabulated below.

Amplifier Type	$T_i \cdot x_i$	Σ	x_E	x_f	G	H
voltage	v	v	v_E	v_f	A_v	A_v
transresistance	i	i	i_E	v_f	R_m	G_m
transconductance	v	v	v_E	i_f	G_m	R_m
current	i	i	i_E	i_f	A_i	A_i

In circuits, Σ can be found as voltages summed around loops (by Kirchhoff's voltage law, KVL), currents can be summed at nodes (by KCL) or either can be summed by superposition, as demonstrated by the following circuit.



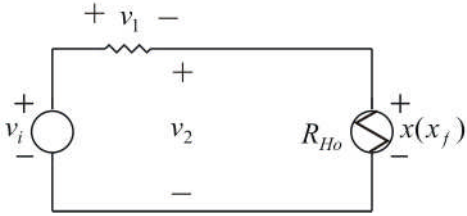
R_{Gi} is the input resistance of the G block and the voltage across it is v_E, the error voltage. The two loops of the circuit add voltages to v_E from the two sources: the input, v_i and the feedback block input, v_f. The divider from v_i to v_E is T_i and the divider from v_f to v_E is $-H$. The negative sign is included because in the block diagram, $x_E = -x_B$.

Applying the voltage-divider formula and superposition (whereby the contributions of the sources are added one at a time),

$$v_E = \left(\frac{R_{Gi}}{R_{Gi} + R_i}\right) \cdot v_i + \left(\frac{R_{Gi}}{R_{Gi} + R_H}\right) \cdot v_f = T_i \cdot v_i + (-H) \cdot v_f$$

Similarly, for currents, superposition can be applied at an error summing node instead.

The input error quantity is *not* generally determined by the circuit. The choice of error quantity x_E does affect the choice of input topology. This can be seen from the following input circuit.



The output of the H block is represented by a generalized source (either Thevenin or Norton equivalent) consisting of transmittance $x(x_f)$ and source resistance R_{Ho}. The feedback-circuit input is a voltage source in series with an input resistance across which is voltage v_1. Across the H source is a node resistance and across it is v_2.

If v_1 is chosen as v_E, the H-path port is made a Thevenin circuit and the input forms a loop around which voltage summation occurs. If v_2 is chosen for v_E instead, then voltage divider superposition results in a summer for the error voltage.

These are not the only choices of x_E. Converting the input and feedback ports to Norton equivalent circuits results in a common node with voltage v_E across its node resistance. KCL can be applied at the node to result in an error current, i_E. The feedback-circuit input summer is generally determined by choice of error quantity and not by the circuit itself. The same kind of argument applies to feedback quantities. Once x_f is chosen, then either a loop (for i_f) or node (for v_f) as the equivalent pickoff circuit results. Often, either choice can lead to a successful analysis.

Op-Amps

What happens when the forward-path gain, G, becomes very large? In mathematical notation, $G \to \infty$ and $1/G \to 0$. The feedback-loop formula is rewritten by dividing numerator and denominator by G;

$$\frac{x_f}{T_i \cdot x_i} = \frac{G}{1+G \cdot H} = \frac{1}{1/G+H}$$

Then letting $G \to \infty$, this reduces to

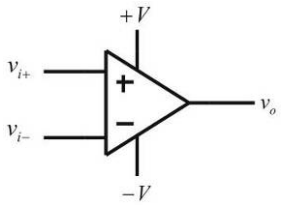
$$\frac{x_f}{T_i \cdot x_i} = \frac{1}{H}$$

G can have inaccurate and unstable gain but if the gain is large, it is sufficient to make the closed-loop gain depend dominantly on H. If H is a passive divider made of stable and accurate resistors, then this kind of feedback amplifier, known as an *op-amp*, is also accurate and stable. Its high-gain forward path need only have high gain to achieve high accuracy and stability.

An *op-amp* is an ideal differential-input voltage amplifier with high gain. It has

- a differential input: $v_i = v_{i+} - v_{i-}$
- high voltage gain: $K \to \infty$
- $r_i \to \infty\ \Omega$ (open-circuit input)
- $r_o \to 0\ \Omega$ (voltage-source out)

The circuit symbol for an amplifier in general is a triangle, and this includes op-amps.



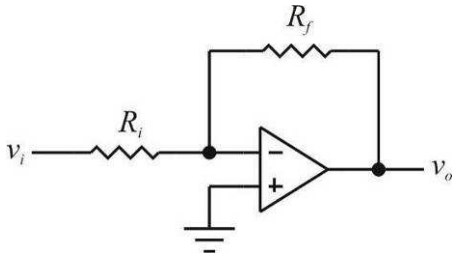
The transfer function, or gain, of the op-amp is a voltage gain and is

$$\frac{v_o}{v_i} = \frac{v_o}{v_{i+} - v_{i-}} = K >> 1$$

Without feedback, a very large K is not useful and op-amps are not used without the addition of a feedback circuit.

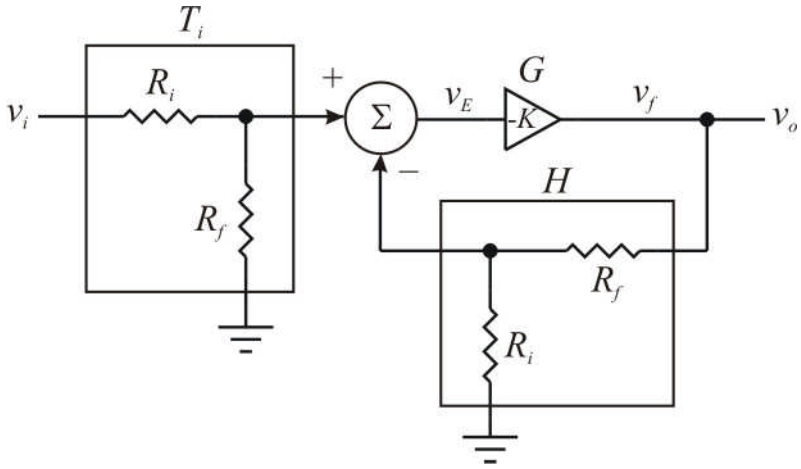
Inverting Op-Amp

Op-amps have two configurations, inverting and non-inverting. The inverting configuration is shown below.



This circuit can be made more like the block diagram by redrawing it as shown below. The output quantity is v_o, also the obvious choice for feedback quantity: $v_o = v_f$. There is no T_o block; that is, $T_o = 1$.

The error quantity, v_E is the voltage at the inverting (–) input of the op-amp to the noninverting (+) input at ground. Then the error is a voltage which is either summed as voltages in series around a loop or by superposition. Here we use superposition.



The boxes are drawn to envision two-port circuits with ground as the common negative terminal of input and output ports. T_i is the attenuation (gain < 1) of the voltage divider formed by R_i and R_f between v_i and the op-amp inverting input;

$$T_i = \frac{R_f}{R_i + R_f}$$

The same resistors in the opposite direction form a reverse divider between the output voltage, v_o, and the error-summing port, the input to the op-amp inverting terminal and ground;

$$H = -\frac{R_i}{R_i + R_f}$$

The divider does not invert and thus lacks the – sign of the summing block. Consequently, H must be given a negative sign. It cancels the negative sign of the summing block in the feedback diagram on which the closed-loop gain formula is based.

The forward-path gain is that of the op-amp from its inverting input to output. The G transfer function follows from the op-amp transfer function:

$$v_o = K \cdot (v_{i+} - v_{i-}) = K \cdot (0\ \text{V} - v_{i-}) = -K \cdot v_E$$

Then $G = v_o/v_E = -K$. Substituting these transfer functions into the feedback formula, the closed-loop voltage gain is

$$\text{closed-loop}\ \ A_v = \frac{v_o}{v_i} = T_i \cdot \left(\frac{G}{1 + G \cdot H} \right) \cdot T_o$$

or

$$A_v = \left(\frac{R_f}{R_i + R_f} \right) \cdot \frac{-K}{1 + (-K) \cdot \left(-\frac{R_i}{R_i + R_f} \right)} \cdot 1$$

Simplifying the algebra, this becomes

$$A_v = -\frac{K \cdot R_f}{R_f + (1 + K) \cdot R_i}$$

Now divide numerator and denominator by $1/K$ and let K go to infinity, keeping in mind that as K gets larger, $1/K$ goes to zero *in the limit*. Mathematically, this is written as

$$\lim_{K\to\infty}\left(\frac{1}{K}\right)=0$$

Then the ideal closed-loop op-amp gain is

$$\lim_{K\to\infty}A_v=\lim_{K\to\infty}\left(-\frac{R_f}{\frac{R_f}{K}+\frac{(1+K)}{K}\cdot R_i}\right)=-\frac{R_f}{0+1\cdot R_i}\Rightarrow$$

$$A_v=-\frac{R_f}{R_i}\ ,\ K\to\infty$$

This simple formula for the gain of the op-amp inverting configuration makes it easy to design amplifiers with an accurate gain by choosing two resistors to set the gain. The resistors can be made as accurate as required, though K must be large enough to achieve the accuracy. For example, let R_f = 10.0 kΩ, R_i = 1.00 kΩ; then A_v = –10.

The error at the input side of G (at the summing block) caused by finite gain is v_E and as a fraction of v_i it is derived from the basic feedback block diagram equation for v_E:

$$v_E=v_i-H\cdot v_o=v_i-H\cdot(G\cdot v_E)$$

Solve for v_E and divide by v_i;

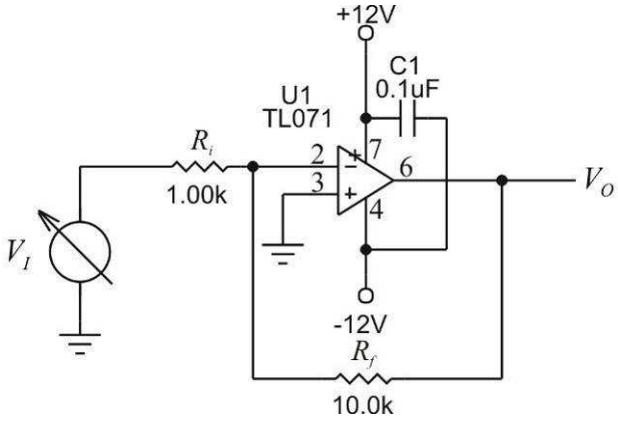
$$\frac{v_E}{v_i}=\frac{1}{1+G\cdot H}$$

The $1 + G{\cdot}H$ also appears in the closed-loop gain formula and is the factor by which feedback improves amplifier performance: the *feedback factor* (or *return ratio*). It is one plus the *loop gain*, $G{\cdot}H$. When v_E/v_i is multiplied by G, the closed-loop gain results:

$$A_v=\frac{v_o}{v_i}=\frac{v_o}{v_E}\cdot\frac{v_E}{v_i}=G\cdot\frac{1}{1+G\cdot H}$$

If the loop gain of a feedback amplifier is 99, then the gain accuracy will be no better than 1/(1 + 99) = 0.01 = 1 %. Integrated circuit (IC) op-amps have typical gains in the range of 50 k to 200 k.

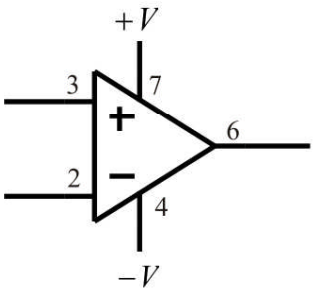
Lab Experiment: Inverting Op-Amp

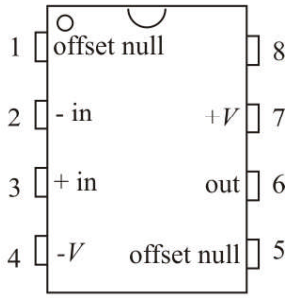


The circuit diagram exemplifies a common practice on circuit (or *schematic*) diagrams to leave off units of component values: Ω, μF. The component *designator* is its number, such as C1, or U1 for the IC. By circuit-diagram shorthand, the other terminals of the supplies are not shown but are connected to ground. C1 bypasses varying supply currents around the op-amp so that the ±12 V supply nodes remain constant in voltage.

The TL071 op-amp is a single op-amp in an 8-pin dual in-line package made by Texas Instruments. The TL071 has industry-standard pin assignments (or *pinout*) as shown for a single op-amp in an 8-pin dual in-line plastic (DIP) or small-outline IC (SOIC) package. Pin numbers begin with 1 at the mark (a dot) and proceed counter-clockwise *around* the package as shown in a top view. DIP packages are circuit-board *through-hole* technology in that the pins of the IC go through holes in the *etched circuit-board* (ECB).

8-Pin Single Op-Amp





SOIC packages have pins that are bent to contact exposed metal pads on the surface of the ECB and are *surface-mount* technology (SMT). SMT parts are 4 or more times smaller than through-hole parts of the same type, but through-hole parts are easier to hand-wire and they are recommended for lab experiments for that reason.

SOIC packages have half the *pin pitch* of DIP packages, the distance between the IC pins in a row. For DIPs it is 0.1 inch and for SOICs it is half that, or 50 mils. (A *mil* is a milli-inch.) IC packages with smaller pin pitches are common but harder to solder using a soldering iron. SOICs can be soldered (by most persons) with an Antex iron and standard iron tip. Smaller tips are available for smaller pitches. The pinouts are the same between DIP and SOIC parts.

A variable voltage supply can be used (as before) for V_I and a DMM to measure gain:

$$A_v = \frac{\Delta v_O}{\Delta v_I}$$

Compared to the calculated gain based on op-amp feedback theory, the agreement between calculated and measured values should be very good. Discrepancy indicates the tolerance of the resistors.

Noninverting Op-Amp

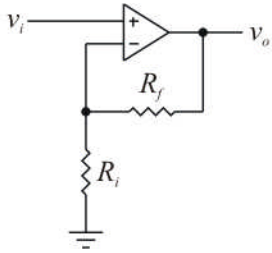
The other op-amp configuration is *noninverting*. It is shown below along with a feedback model using a 2-port block. Unlike the inverting configuration, there is no T_i block.

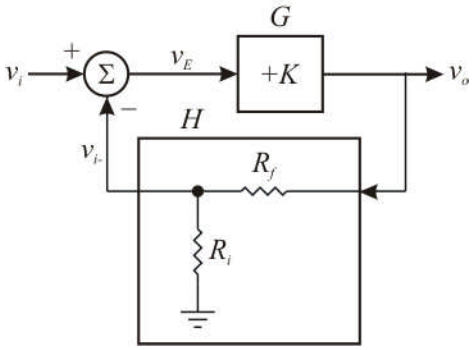
To perform feedback analysis, the first task is to choose which variables will be v_f and v_E. As for the inverting op-amp, $v_f = v_o$ and $T_o = 1$. The variable that will be chosen for v_E is v_{i+} of the op-amp (to ground). Then v_i feeds directly into the error summer ($T_i = 1$) and the feedback to the summer,

$$v_B = v_{i-} = \frac{R_i}{R_f + R_i} = H$$

The summer in the circuit is the input stage of the op-amp. The inverting input of the op-amp supplies the negative sign of the summer. The summer and forward path are the op-amp;

$$v_o = K \cdot (v_{i+} - v_{i-}) = G \cdot (v_i - H \cdot v_o)$$





This equation can be solved for the closed-loop gain;

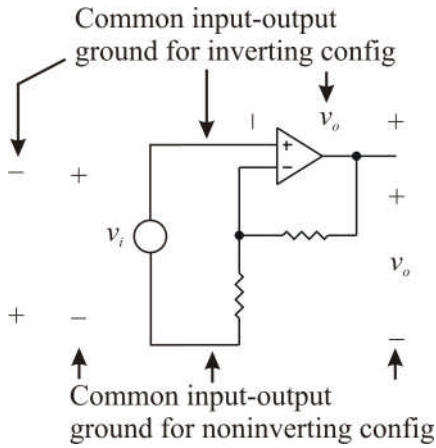
$$A_v = \frac{v_o}{v_i} = \frac{G}{1 + G \cdot H} = \frac{K}{1 + K \cdot \left(\dfrac{R_i}{R_f + R_i} \right)}$$

Then if K is allowed to go to infinity,

$$\lim_{K \to \infty} A_v = \lim_{K \to \infty} \frac{1}{\dfrac{1}{K} + \left(\dfrac{R_i}{R_f + R_i} \right)} = \frac{R_f + R_i}{R_i} = \frac{R_f}{R_i} + 1$$

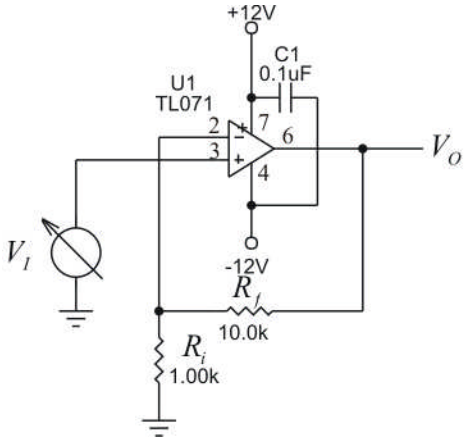
Using the same values of resistors as for the inverting op-amp example, the closed-loop gain is $A_v = +11$.

The noninverting and inverting op-amp configurations have the same interconnection of components - the same *topology*. What is different is where ground is connected, as shown below.



Ground is the common node shared between the two-port input and output negative terminals of v_i and v_o. If R_i is grounded and v_{i+} made the + terminal of v_i, then the op-amp is noninverting; if v_i and v_o are flipped so that their terminals are interchanged, then the op-amp is inverting. In the noninverting configuration, v_i adds to the divider voltage and this results in the "+ 1" in the noninverting op-amp gain.

Lab Experiment: Noninverting Op-Amp

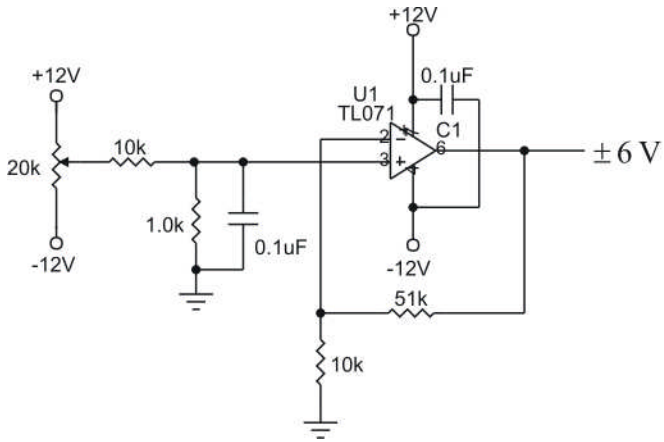


The noninverting op-amp circuit requires only a minor rewiring of the inverting op-amp, as shown. Repeat the same procedure as for the inverting op-amp and see for yourself the difference in gain, both in polarity (noninverting) and in magnitude (×11, not ×10).

Op-amps are generally very linear (closed-loop) except when the limits of their range are approached. Gain variation as static voltage is

varied finds these limits. Nonlinearity over the op-amp range can be detected by measuring the gain at different static input voltages. For instance, measure Δv_O for Δv_I from 0.3 V to 0.7 V, a change of 0.4 V around V_I = 0.5 V. Then compare Δv_O for the same Δv_I around a static V_I = –0.5 V.

A variable-voltage bench power supply can be made from the ±12 V fixed supplies using the following circuit. The 20 kΩ part on the left is a *potentiometer*, a variable resistive voltage divider. Use one that has a shaft or knob that can be turned by hand. The variable voltage source makes use of the noninverting op-amp circuit with the gain changed for ±6 V output. For more voltage resolution and a ±1 V range, make the 51 kΩ feedback resistor 10 kΩ.



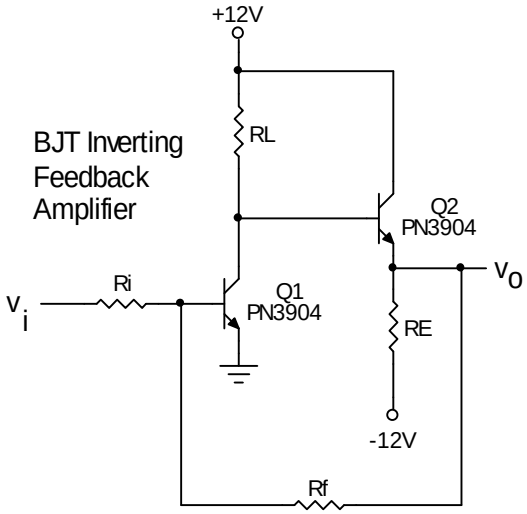
If you have two variable voltage sources (and you can by building two of the above circuits), then disconnect the op-amp + input from ground and connect it to a second variable source. Observe what effect it has on v_O. Derive a formula for the noninverting op-amp gain with both sources and compare it to the inverting and noninverting op-amp gain formulas.

BJT Inverting Feedback Amplifier

Op-amps make feedback circuit design simpler because the amplifier has already been designed. The high gain of op-amps, however, makes them behave slowly, and lower-gain *discrete* (not integrated) BJT amplifiers are sometimes more optimal. Between full amplifier integration and discrete-component BJTs are transistor arrays: several BJTs on a common silicon substrate that are made

with the same process and have matching parameters such as β and especially V_{BE} at a given I_C and substrate temperature. We will make use of BJT arrays in lab experiments to benefit from the matching on which some circuit concepts are based.

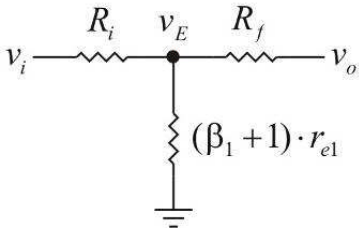
One benefit of matching V_{BE}s of BJTs is in the emitter-coupled diff-amp. If the V_{BE}s do not match they cause an *input offset voltage* error that is amplified as a static input voltage. Op-amps also amplify static voltages with high gain and this error at the output can be significant.



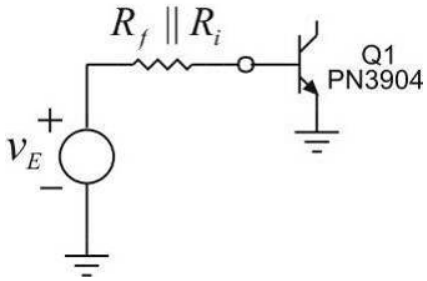
A 2-BJT inverting feedback amplifier is shown above. Feedback analysis is no different for discrete circuits such as this than for op-amps. Choose $x_f = v_o$. Then $T_o = 1$. This is similar to the inverting op-amp except that the input resistance of Q1, or

$$r_{B1} = (\beta_1 + 1) \cdot r_{e1}$$

is in parallel with and loads the feedback divider resistors, as shown below. (Op-amp inputs are ideally open circuits.) The circuit forms two merged voltage dividers. Superposition can be applied for error-summing.



The addition of r_{B1} complicates analysis by giving us a choice of how to proceed. For the first choice, let $x_E = v_{b1}$ with r_{B1} disconnected. Thevenize the resulting divider and reconnect, as shown below.



Then v_E is the Thevenin open-circuit error voltage and $R_f \| R_i$ is the Thevenin series resistance. The expression for v_E is the superposition of the sources v_i and v_o as applied to the feedback divider:

$$v_E = \left(\frac{R_f}{R_f + R_i}\right)\cdot v_i + \left(\frac{R_i}{R_f + R_i}\right)\cdot v_o = T_i \cdot v_i - (-H)\cdot v_o$$

This one equation has within it both $T_i = v_E/v_i$ and the transfer function from $v_f = v_o$ to v_E, or $H = -v_E/v_f$;

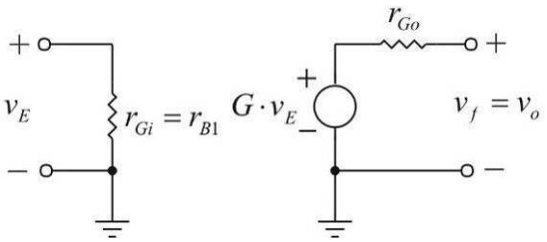
$$T_i = \left(\frac{R_f}{R_f + R_i}\right),\ H = -\left(\frac{R_i}{R_f + R_i}\right)$$

Next, in calculation of the forward-path gain, G, $R_f \| R_i$ is included as base resistance and is β-transformed to the emitter of Q1 to become part of r_M. The stage gains, with stage-1 open-circuit output (Q2 base disconnected) are

$$G_{v1} = -\alpha_1 \cdot \frac{R_L}{r_{e1} + \dfrac{R_f \parallel R_i}{\beta_1 + 1}},\ G_{v2} = \frac{R_E}{r_{e2} + R_E + \dfrac{R_L}{\beta_2 + 1}}$$

Then $G = G_{v1} \cdot G_{v2}$. The first stage is a CE and the second is a CC.

The model of G as a two-port block is shown below.



The r_o of G is

$$r_{Go} = \left(r_{e2} + \frac{R_L}{\beta_2 + 1} \right) \| R_E \approx r_{e2} + \frac{R_L}{\beta_2 + 1} \text{ , } R_E >> r_{e2} + \frac{R_L}{\beta_2 + 1}$$

r_o is small enough to be neglected and the G amplifier considered a voltage source. If it were included, it would have formed a voltage divider with R_f; a more rigorous analysis would include it. R_f is usually much larger than R_E but its effect on r_{Go} could be included as a shunt (incremental) resistance across R_E. Even so, if R_E is reduced slightly in the expression for G_{v2}, it will change G_{v2} very little and can usually be neglected, as we have.

The other way to handle the loading of the Q1 base resistance is to choose $v_E = v_{b1}$ with loading included, so that v_E is not an open-circuit voltage but is the actual (loaded) base node voltage. Then we have a superposition involving two overlapping dividers and the expression for error voltage is more complicated;

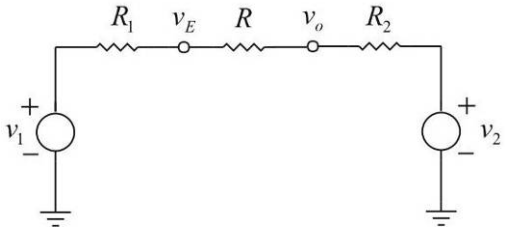
$$v_E = \left(\frac{R_f \| r_{B1}}{R_f \| r_{b1} + R_i} \right) \cdot v_i + \left(\frac{R_i \| r_{B1}}{R_i \| r_{B1} + R_f} \right) \cdot v_o$$

The parallel loading of r_{B1} appears in the voltage-divider formulas. From the expression for v_E we can extract

$$T_i = \left(\frac{R_f \| r_{B1}}{R_f \| r_{b1} + R_i} \right) \text{ , } H = -\left(\frac{R_i \| r_{B1}}{R_i \| r_{B1} + R_f} \right)$$

Two-Port Loading

In calculating r_{Go} of the amplifier of the previous section, R_f would need to be included as shunting R_E at the output but not as $R_f + R_i \| r_{B1}$. This is because v_E is assumed to be the actual voltage at the base node, calculated to include the loading on the v_E node by R_f. The *Two-Port Loading Theorem* lets us calculate v_E or v_o of a loaded divider (as shown below in general) by assuming that the other voltage is already calculated. To calculate v_E, assume that v_o is already calculated to include the loading on the v_o node by $R + R_1$. Then when calculating v_o, we can also assume that v_E has already been calculated.



To show that this produces correct results, we will derive v_E and v_o for the above general circuit using superposition:

$$v_E = \frac{R + R_2}{R + R_1 + R_2} \cdot v_1 + \frac{R_1}{R + R_1 + R_2} \cdot v_2$$

$$v_o = \frac{R_2}{R + R_1 + R_2} \cdot v_1 + \frac{R + R_1}{R + R_1 + R_2} \cdot v_2$$

where the resistance ratios result from applying the voltage-divider formula. Solving algebraically to eliminate v_2 in v_E and v_1 in v_o,

$$v_E = \frac{R}{R_1 + R} \cdot (v_1 - v_o) + v_o$$

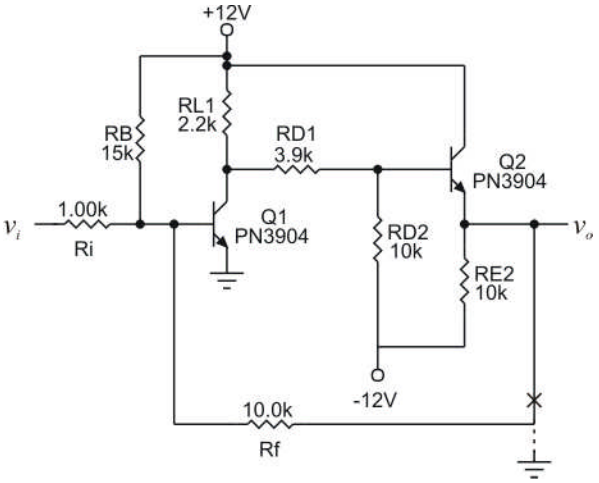
$$v_o = \frac{R}{R_2 + R} \cdot (v_2 - v_E) + v_E$$

These two formulas are the voltage-divider formulas for a divider with voltages at each end (v_1 and v_o for the first equation). The v_E equation assumes v_o is given, as though it were a voltage source, and

the v_o equation assumes v_E is given. That is, the assumed v_E and v_o are the loaded (actual) node voltages, not the open-circuit voltages.

Lab Experiment: Inverting Amplifier

The following two-BJT inverting feedback amplifier (referred to as amplifier 10) is a refinement of the previous circuit, to bias the output at zero volts. R_B is added to provide Q1 base current (so that it need not all be supplied through R_f), and the R_{D1}, R_{D2} divider reduces the voltage from the Q1 collector so that the output is at 0 V.



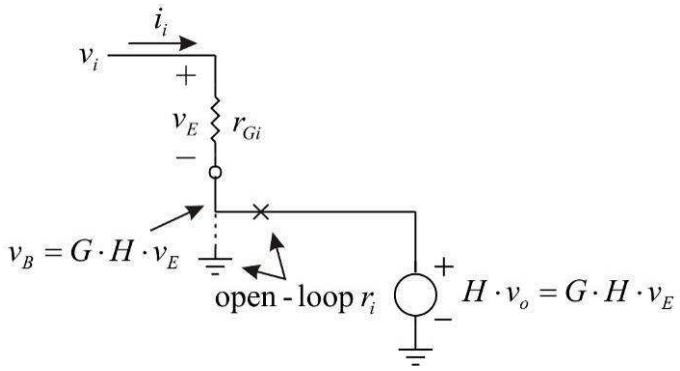
As usual, we are interested in finding the closed-loop voltage gain. To measure T_i and G, disconnect R_f from v_o and ground it, as shown. This opens the feedback loop so that the gains of its blocks can be measured. Measure $-H$ by disconnecting R_f from ground and driving it with a variable voltage source while measuring v_E with v_i grounded.

Reconnect R_f to the output and measure the closed-loop r_o as follows. Fix v_I at some voltage (0.1 V will do) and measure v_O with the output node open. Then attach a 1.0 kΩ resistor from v_O to ground and measure the loaded v_O. Calculate r_o using the voltage-divider formula;

$$r_o = \left(\frac{v_{Go}}{v_o} - 1 \right) \cdot (1.0\,\text{k}\Omega) = \left(\frac{\Delta v_{GO}}{\Delta v_O} - 1 \right) \cdot (1.0\ \text{k}\Omega)$$

The Effect of Feedback on r_i and r_o

The summing loop of a feedback circuit with an error voltage, v_E, is shown below.



The error voltage is the input of G which has input resistance r_{Gi}. The open-loop r_i can be found by opening the feedback loop and grounding r_{Gi}, as shown in the above diagram. Then

$$r_i = \frac{v_i}{i_i} = \frac{v_E}{i_i} = r_{Gi}$$

r_{Gi} is "bootstrapped" by the feedback voltage, v_B, in that the lower end of r_{Gi} is made to follow the top end in voltage with only the error voltage as the difference.

From the basic feedback equations,

$$v_E = v_i - v_B = v_i - G \cdot H \cdot v_E$$

or

$$v_E = \frac{v_i}{1 + G \cdot H}$$

When the loop is closed, then the input current,

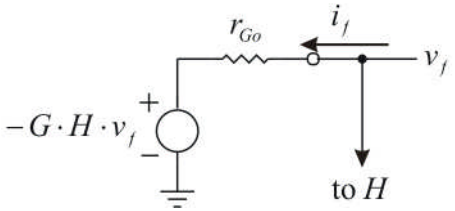
$$i_i = \frac{v_E}{r_{Gi}} = \frac{v_i / (1 + G \cdot H)}{r_{Gi}}$$

With feedback,

$$r_i = \frac{v_i}{i_i} = \frac{v_i}{\dfrac{v_i/(1+G\cdot H)}{r_{Gi}}} = (1+G\cdot H)\cdot r_{Gi}$$

The effect of feedback is to increase the input resistance across v_E by the feedback factor, $1 + G{\cdot}H$.

At the output, with $x_f = v_f$, the output resistance can be found for the closed-loop amplifier. The output circuit is shown below. The node that applies is the output of the loop itself, v_f, not $v_o = T_o{\cdot}v_f$. The G output is modeled as an output port with a Thevenin equivalent circuit. The open-loop output resistance of G is r_{Go}.



Without feedback, the voltage source is independent of the output and can be grounded. Then only r_{Go} is across the output port and the

$$\text{open-loop } r_o = \frac{v_f}{i_f} = r_{Go}$$

With feedback, the output voltage source is dependent on v_f, the node it is supplying. Apply KVL to the output voltage loop:

$$v_f = -G\cdot H\cdot v_f + i_f\cdot r_{Go} \Rightarrow v_f = r_{Go}\cdot\frac{i_f}{1+G\cdot H}$$

Substitute for v_f in

$$r_o = \frac{v_f}{i_f} = \frac{\left(r_{Go}\cdot\dfrac{i_f}{1+G\cdot H}\right)}{i_f} = \frac{r_{Go}}{1+G\cdot H}$$

Voltage feedback reduces open-loop output resistance r_{Go} by the feedback factor $1 + G{\cdot}H$.

In both cases of input and output resistance, the feedback loop for a voltage amplifier was changed toward the ideal: infinite r_i and zero r_o. For $x_E = i_E$, a similar analysis shows that feedback reduces r_i by the

feedback factor, $1 + G{\cdot}H$. And for $x_f = i_f$, r_o is increased by $1 + G{\cdot}H$. In all cases, the feedback amplifier port resistances become closer to their ideal values by the feedback factor. All four types of amplifiers are benefited by feedback by the feedback factor.

Feedback Improves Linearity

Feedback has additional advantages beyond idealizing the three basic amplifier parameters. It also improves amplifier linearity and reduces distortion. This is important in audio, video, and instrumentation applications. In the three BJT configurations, the dominance of R_E over r_e reduces temperature and gain drift, and also distortion, because r_e changes with I_E. As R_E increases, i_e decreases for a given input voltage until the gain is zero, thus defeating the purpose for the amplifier. Feedback has the same tradeoff between gain and linearity but with high loop gain, open-loop linearity can be significantly improved.

We can bring nonlinearity into feedback analysis as ε, the nonlinear error in G, and let the linear part be K;

$$G = K + \varepsilon$$

Then substitute G into the feedback gain formula:

$$A_v = \frac{G}{1 + G \cdot H} = \frac{K + \varepsilon}{1 + (K + \varepsilon) \cdot H}$$

For small nonlinearity (low distortion), $K >> \varepsilon$ and we can approximate the denominator as

$$1 + (K + \varepsilon) \cdot H \approx 1 + K \cdot H$$

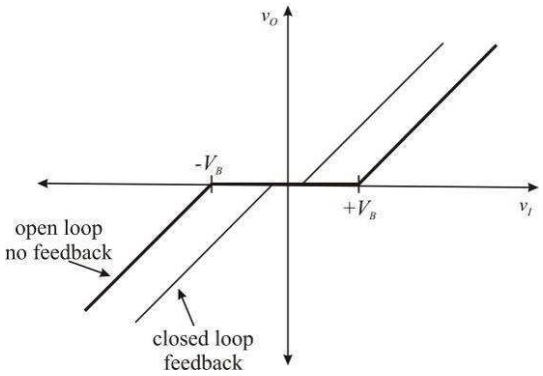
Then substituting into A_v,

$$A_v \approx \frac{K}{1 + K \cdot H} + \frac{\varepsilon}{1 + K \cdot H} \quad , K >> \varepsilon$$

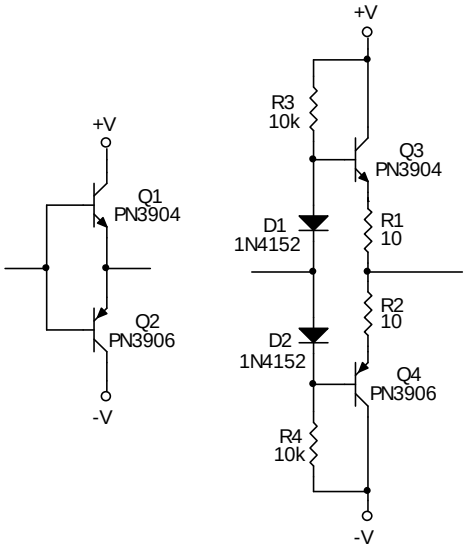
The first term in A_v is the ordinary linear feedback formula. The second term accounts for nonlinearity. It has been reduced from its open-loop value of ε by the feedback factor, $1 + G{\cdot}H$. An amplifier with 1 % total harmonic distortion (THD) and a loop gain of 99 will have about 1 %/100 = 0.01 % distortion based on feedback analysis. Other factors can contribute to nonlinearity that are not removed by

feedback such as nonlinearity in H (resistors change value slightly with voltage or current changes) and in T_i and T_o.

A *deadband* is nonlinearity in an amplifier where around zero volts output, there is no voltage, as shown below.



With feedback, the deadband width, $2 \cdot V_B$, is reduced. Deadband is characteristic of the output stages of *bipolar* (+ and –) output amplifiers such as the one shown below on the left.



It is a NPN CC stage in parallel with a PNP CC. When $v_I > V_{BE}$, Q1 conducts to the output load to ground. When $v_I < -V_{BE}$, Q2 conducts and v_o becomes negative. Within $\pm V_{BE}$, neither BJT conducts and $v_o = 0$ V. The deadband voltage, $V_B = V_{BE}$. The nonlinearity is caused

by inadequate biasing of the BJTs. The range of v_I over which they conduct must overlap so that one or the other is always conducting.

The modified circuit on the right adds biasing components. R3 and R4 supply current through diodes D1 and D2. The diode drop, V_D, of each is comparable to V_{BE} and both transistors are then biased for conduction. To control the bias current that flows through both of them, emitter resistors R1, R2 are added. They have a small value of resistance because the voltage across them is $V_D - V_{BE}$ which is also small. Static circuit design for this circuit depends on knowing the $V_{BE}(I_E)$ of the BJTs and $V_D(I_D)$ of the diodes. Both are usually specified over a wide range and precise design is not possible. However, the amount of bias current is not critical either, and this bias scheme is commonly used.

This output stage is a *complementary CC* stage that can be used to provide more output current from op-amps by adding it as the output stage after the op-amp. As an emitter-follower stage, it has a voltage gain of nearly one. In the *crossover region* around 0 V where both BJTs are conducting, emitter current is minimum and r_e is largest, reducing gain. However, both BJTs drive the output and the CCs are in parallel, causing the r_o of each to be in parallel and halved.

The Miller Effect

Miller's Theorem is an application of feedback theory that can be used to quickly find r_i of inverting amplifiers with a feedback resistance, as shown below. Resistance R_f is the feedback element. The amplifier output is

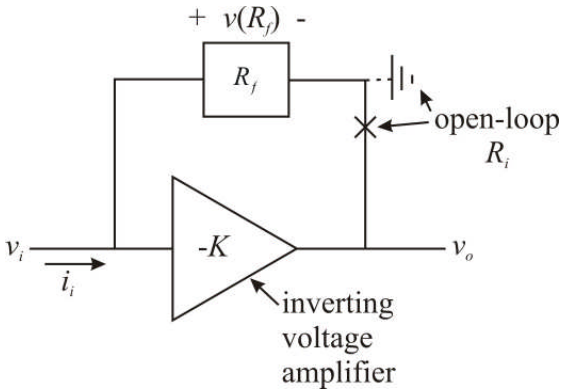
$$v_o = -K \cdot v_i$$

The voltage across R_f is

$$v_i - v_o = v_i - (-K) \cdot v_o = (1+K) \cdot v_i$$

The open-loop r_i can be found by disconnecting R_f from the output and grounding it, as shown. Then $r_i = v_i/i_i = R_f$. For the closed-loop amplifier,

$$\text{closed-loop } r_i = \frac{v_i}{i_i} = \frac{v_i}{\left(\frac{(1+K) \cdot v_i}{R_f} \right)} = \frac{R_f}{1+K}$$

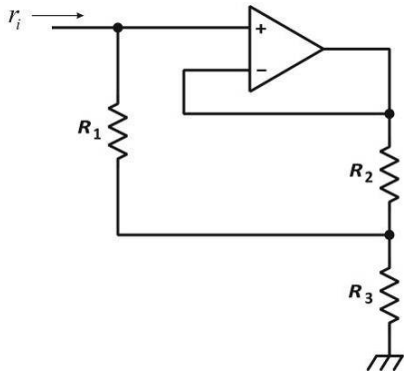


As we have seen previously, feedback reduces r_i by the feedback factor - in this case, $1 + K$. This is the *Miller effect*. If R_f were instead a capacitor, it would appear larger as $(1 + K)\cdot C_f$ at the input and could reduce amplifier speed.

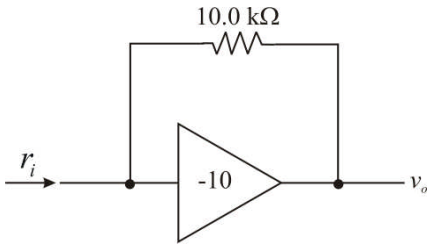
A more deliberate and precise resistance multiplier, shown below, uses an op-amp buffer with $A_v = +1$. Solving the circuit for r_i results in an equation that shows the Miller effect;

$$r_i = R_1 \cdot \left(1 + \frac{R_3}{R_2}\right) + R_3$$

where $K = R_3/R_2$. R_1 is said to be "bootstrapped" by the Miller effect to appear as a larger resistance from the input than it is.



As an example, an inverting amplifier, shown below, has a gain of –10 and $R_f = 10$ kΩ. What are the Miller equivalent resistances?



The input resistance because of R_f is

$$r_i = \frac{10.0\ \text{k}\Omega}{1+10} = 909\ \Omega$$

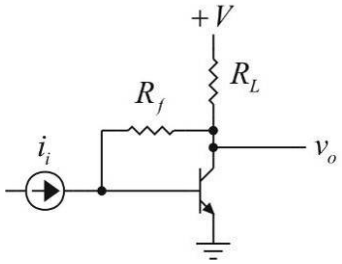
The inverting op-amp inverting (–) input is like the above circuit. The v_{i-} node is driven by feedback to be low in resistance and is called a *virtual ground* because (except for v_E which is small if loop gain is large) it is held near ground by feedback. (Another name is *quasistatic ground*.)

At the output, the ideal inverting voltage amplifier has zero resistance. However, the shunt path of R_f will otherwise appear as

$$r_o = \left(\frac{K}{1+K}\right) \cdot R_f = \frac{10}{1+10} \cdot (10.0\,\text{k}\Omega) = 9.09\,\text{k}\Omega$$

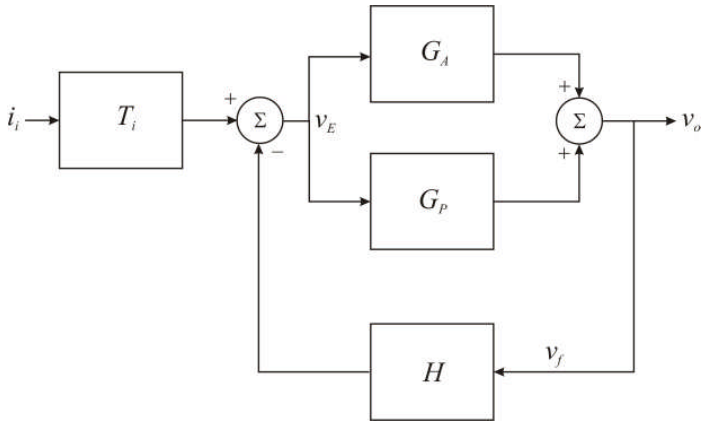
Shunt-Feedback Amplifier

The *shunt-feedback amplifier* (SFA) is another simple one-BJT stage, shown below.



Instead of v_i, this amplifier has a current input. The other terminal of i_i could go to ground, but as a current source, it is somewhat arbitrary. In this inverting amplifier, R_f provides a feedback path. With no R_E, the BJT gain is relatively large and, like op-amps, will depend on feedback to keep the closed-loop gain stable.

The circuit also has two forward paths. It is not uncommon to have an active path, through the collector current source of the BJT, and also a passive path in reverse through the feedback path. This is the case here. R_f and R_L form a divider with v_b as input and v_o as output. The block diagram is shown below.



For feedback analysis, let $v_E = v_b$ and $v_f = v_o$. Then, following the block diagram and using the two-port loading theorem,

$$T_i = \frac{v_E}{i_i} = \frac{v_b}{i_i} = r_\pi \parallel R_f$$

The input current develops v_b across the resistance at the base node. It is r_π to ground in parallel with R_f to the collector node. Applying the two-port loading theorem, the collector is regarded as already loaded by R_f and its voltage is assumed to have been found and appears as though it were a voltage source, v_o.

The feedback path is

$$H = \frac{-v_E}{v_o} = \frac{-v_b}{v_o} = -\frac{r_\pi}{r_\pi + R_f}$$

The voltage-divider formula is applied, and with no inversion by the divider the negative sign is added to reflect the negated Σ input on the block diagram. The active and passive forward paths are

$$G_A = -\frac{R_f \parallel R_L}{r_m} \; ; \; G_P = \frac{R_L}{R_L + R_f}$$

The total forward-path gain is

$$G = G_A + G_P = \frac{R_L}{R_L + R_f} \cdot \left(1 - \frac{R_f}{r_m}\right)$$

The closed-loop transresistance is

$$R_m = \frac{v_o}{i_i} = T_i \cdot \frac{G_A + G_P}{1 + (G_A + G_P) \cdot H}$$

Substituting from the gains of the blocks,

$$R_m = (r_\pi \parallel R_f) \cdot \frac{\frac{R_L}{R_L + R_f} \cdot \left(1 - \frac{R_f}{r_m}\right)}{1 + \left[\frac{R_L}{R_L + R_f} \cdot \left(1 - \frac{R_f}{r_m}\right)\right] \cdot \left(-\frac{r_\pi}{r_\pi + R_f}\right)}$$

Applying $r_m = r_e/\alpha$, this reduces to

$$R_m = (-\alpha \cdot R_f + r_e) \cdot \frac{R_L}{\frac{R_f}{\beta + 1} + r_e + R_L}$$

If R_L is made a current source, then $R_L \to \infty$, the rational factor approaches one, and

$$R_m = -\alpha \cdot R_f + r_e$$

BJTs usually have $\beta >> 1$. For $\beta \to \infty$,

$$R_m = (-R_f + r_e) \cdot \left(\frac{R_L}{R_L + r_e}\right)$$

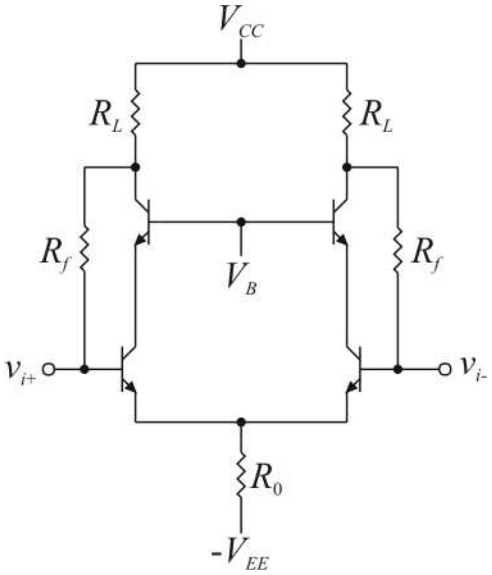
And if $R_L >> r_e$, then let $r_e = 0\ \Omega$ and

$$R_m = -R_f$$

which is the same result as though the BJT were an inverting op-amp.

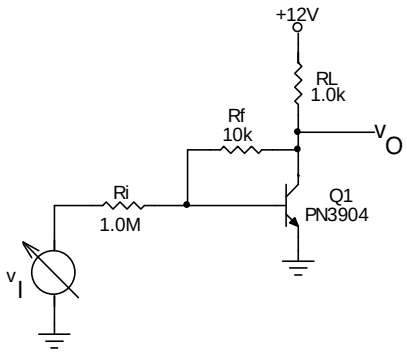
Shunt-feedback amplifiers have multiple variations. Not only might R_L be replaced by a current source or the BJT replaced by an inverting op-amp, the BJT could be replaced by a Darlington configuration for higher β or a cascode stage for faster response. Emitter-coupled BJT pairs can be made into shunt-feedback

amplifiers (shown below) by adding R_f between base and collector of each of them, and additionally, those stages could be cascode stages for a fast differential amplifier with local (intrastage) feedback.



Lab Experiment: Shunt-Feedback Amplifier

The input current source for a BJT shunt-feedback amplifier is approximated as a long-tail source: a large resistance in series with a large voltage. As $R_i \to \infty$, the input source approaches a current source. Thus R_i has been made large relative to R_f. For $V_I = 10$ V, $I_I = 10\,\mu$A. Let $i_i = \Delta i_I = 10\,\mu\text{A} - 5\,\mu\text{A} = 5\,\mu$A. Then calculate and measure what $v_o = \Delta v_O$ should be. Compare and reconcile results.



The design formula used to calculate the value of load resistance for biasing the BJT is

$$R_L = \frac{R_f}{\beta} \cdot \frac{V_{CC} - V_O}{V_O - V_{BE}}$$

Solving,

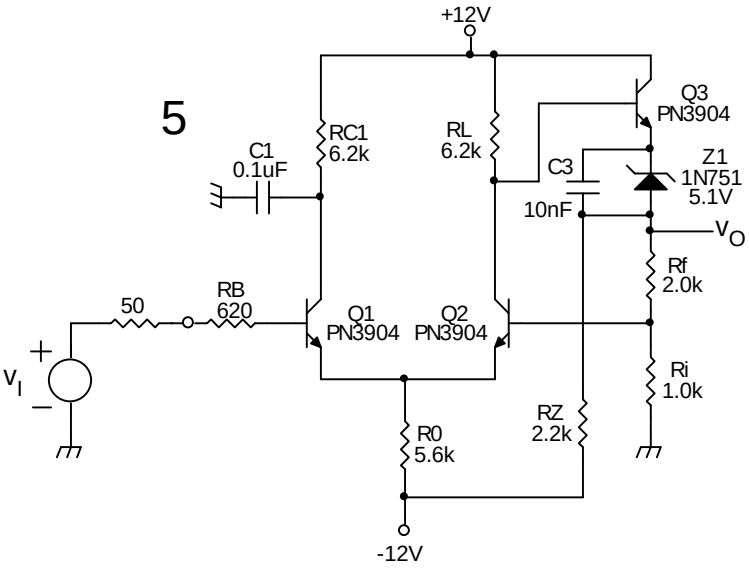
$$V_O = \left(R_L \parallel \frac{R_f}{\beta} \right) \cdot \left(\frac{V_{CC}}{R_L} + \frac{V_{BE}}{R_f / \beta} \right)$$

$$= (1.0\ \text{k}\Omega \parallel (10\ \text{k}\Omega/150)) \cdot (12\ \text{mA} + (0.65\ \text{V}/667\ \Omega)) = 5.19\ \text{V}$$

where $10\ \text{k}\Omega/150 \approx 667\ \Omega$. The left factor is then about $400\ \Omega$, $(0.65\ \text{V}/667\ \Omega) \approx 975\ \mu\text{A}$ and the right factor is about $12.98\ \text{mA}$.

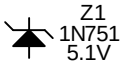
3-NPN Feedback Amplifier

The 2-BJT diff-amp is a good choice for the differential input stage of an op-amp. The goal is to design a simple 3-BJT op-amp. The circuit is shown (amplifier 5), copied from a computer-aided design (CAD) circuit drawing program (CircuitMaker).



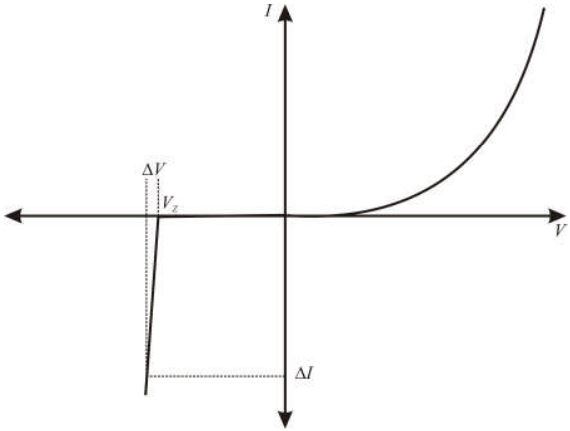
The feedback divider R_f, R_i is included to make this a complete amplifier. The + input is at v_i and the – input at the base of Q2.

The new circuit symbol is shown below:



This is a *zener* or *avalanche diode*. It is designed to break down when the reverse-biasing voltage polarity is applied to it. At the breakdown voltage of 5.1 V, the electric field across the diode p-n junction is strong enough to cause electrons to be freed from atomic orbits and accelerated, where they collide with other atoms with sufficient energy to free yet more electrons. The resulting *avalanche* effect produces a very stable voltage across the diode at its rated voltage. By "stable" is meant that both the voltage TC and the incremental resistance are low ($r_Z \approx 0\ \Omega$). A change in current through the diode causes negligible voltage change across it and it behaves like a static voltage source.

The $V(I)$ function of an avalanche diode is shown below.



As shown on the $V(I)$ plot,

$$r_Z \approx \frac{\Delta V}{\Delta I} \approx 5\,\Omega$$

where ΔV and ΔI are both taken as positive changes. Avalanche diodes are used as *voltage references* when a stable voltage is required. The rated voltage is given at a rated current that is chosen to minimize TC(V_Z). In the reverse direction of forward biasing, an avalanche diode behaves as an ordinary forward-conducting diode with about a 0.65 V drop at 1 mA.

The Zener effect begins to occur for breakdown voltages below about 5 V and is caused by quantum electron tunneling. This strange electronic behavior also occurs in tunnel diodes, which are designed to have breakdown voltages near 0 V! Tunnel diodes also have a negative-resistance region of operation through which they switch extremely quickly. (They consequently were used as trigger generator components in oscilloscopes in the 1960s.) For reference-voltage generation, they are suboptimal (though available) in that their incremental resistance is too high and varies with current. Thus only avalanche diodes are of use as voltage references, though the "Zener" name is attached to either in practice.

In the BJT amplifier, an avalanche diode is used as a voltage "shifter" or *voltage translator*: a device or circuit that changes the static voltage from one node to another without changing the incremental voltage. Because the voltage across the diode is constant with current (when reverse-biased), it is relevant to static circuit design. The change in V_{E3} appears at the anode (lower terminal) of the diode but is 5.1 V lower in static voltage. The capacitor shunting the diode reduces its electrical noise (from all those electron-atom collisions) and stabilizes it. A 5.1 V diode is near-optimal in voltage value (6.2 V is optimal) because around this voltage, $\mathrm{TC}(V_Z) \approx 0$; the voltage remains constant with temperature change.

The first step in circuit analysis is the static (large-signal) analysis, to determine BJT bias. From it, the incremental parameters, such as incremental emitter resistance, r_e, are found. We again apply the BJT T-model with parameters β (constant over I_E) and r_e. The simple T-model is actually capable (with correct bias values) of producing accuracy comparable to computer circuit simulators (like the one in CircuitMaker) with a few manual calculations.

Beginning with the static analysis and assuming 1 mA of Q1, Q2 emitter currents, refer R_B to the emitter with the β transform. At 1 mA, PN3904 V_{BE} = 0.65 V, and for Q1, Q2,

$$I_0 = 2 \cdot I_E = \frac{12\ \mathrm{V} - 0.65\ \mathrm{V}}{5.6\ \mathrm{k\Omega} + 670\ \Omega/151} = 2.025\ \mathrm{mA} = 2 \cdot (1.013\ \mathrm{mA})$$

The solution for I_0 is *iterative* in that the value from the above equation can be used to determine from the *v-i* junction relationship a new V_{BE}, and the I_0 equation is invoked again. The 1 mA assumption is valid and

$$V_{C1} = 12\ \mathrm{V} - (0.993) \cdot (1.013\ \mathrm{mA}) \cdot (6.2\ \mathrm{k\Omega}) = 5.76\ \mathrm{V}$$

C_1 bypasses the collector of Q1 to ground and keeps varying voltage from occurring at the collector node. Although C_1 eliminates the Miller effect at the Q1 base (from collector-base capacitance, C_{bc}), the variation in i_{c1} caused by v_i is lost and does not contribute to the gain at the output, v_o. (This will be improved in the next feedback amplifier.) The same static current through R_L drops the same open-circuit voltage - that is, the voltage without the loading of Q3. With the Q3 base disconnected, $V_{C2} = 5.76$ V. This voltage is midway between the +12 V supply and the input ground near the base voltage, and operates the BJT near its maximum-power bias point.

Similarly, the choice of $R_{C1} = R_L$ for both collectors results in about half of the supply voltage at the collectors for 0 V in and out. This is intentional because at $V_{CC}/2$, *thermal distortion* is minimized. More exactly, when the static $V_{CE} = (V_{CC} + V_{EE})/2$, the BJT operates at maximum power and minimum Δp_Q. The emitter static voltage is usually near the base voltage (of 0 V in this circuit) and "V_{EE}" is a Thevenin equivalent voltage that is small compared to V_{CC}. In this case, $V_C \approx V_{CC}/2$ is a reasonable thermal criterion. Thermal distortion in the amplified waveform is caused by a change in r_e from a change in power dissipation of Q1 and Q2 as v_I changes. The BJT power dissipation is

$$p_Q = v_{CE} \cdot i_C$$

As it changes with the waveform, the total (static + incremental) p_Q changes and the temperature of the BJT itself changes by

$$\Delta T = R_\theta \cdot \Delta p_Q$$

where R_θ is the *thermal resistance*, with units of °C/W or K/W, from the BJT junction to the ambient temperature. As the BJT temperature increases, r_e increases at about 0.33 %/K and V_{BE} decreases by approximately –2 mV/°C.

Δp_Q is minimum when the operating-point for the collector voltage is $V_{CC}/2$. At this bias voltage, $p_Q(v_{CE})$ is maximum and at the vertex of an inverted parabolic function. Any change of v_{CE} away from the vertex at $V_C = V_{CC}/2$ is minimum compared to any other point on the parabola. Thus, we consider it the voltage to choose for V_{CE} in static design. It is not quite optimal for maximum range, and the tradeoff between them is your decision as designer. If one or the other is a specified requirement, the bias can be set to satisfy the requirement while allowing the other to be as best as it can be under

the constraint of the requirement. A final fact on thermal distortion: by keeping the static BJT power low, the change in power is that much lower.

At the Q2 collector node, disconnect the base of Q3 and find the Thevenin open-circuit voltage. Then "looking into" the Q3 base, we "see" a resistance of

$$r_{i3} = (\beta_3 + 1)\cdot(r_{e3} + r_Z + R_Z||(R_f + R_i))$$

without an output load. When $V_O = 0$ V, $V_I = 0$ V, by specification. The path for Q3 current is through emitter biasing resistor R_Z. At $V_O = 0$ V, it conducts $I_{E3} = I_Z = 12$ V/2.2 kΩ = 5.45 mA and from this we can find $r_{e3} = 26$ mV/5.45 mA ≈ 4.8 Ω. In series with it is

$$(R_f + R_i)||R_Z = (3\ \text{k}\Omega)||(2.2\ \text{k}\Omega) = 1.27\ \text{k}\Omega$$

The external emitter resistance dominates $r_{e3} + r_Z$ as is desired for good circuit design and

$$r_{i3} = (151)\cdot[4.8\,\Omega + 5\,\Omega + 1.27\,\text{k}\Omega] = 193.2\,\text{k}\Omega$$

Then R_L forms a divider with r_{i3} of

$$\frac{r_{i3}}{r_{i3} + R_L} = \frac{193.2\,\text{k}\Omega}{193.2\,\text{k}\Omega + 6.2\,\text{k}\Omega} \approx 0.969$$

or almost one. The loading effect of r_{i3} on Q2 collector node resistance is small, and when an external resistance is placed across V_O, it will not reduce the above divider attenuation by much (unless it is nearly a short circuit).

We can now determine the static Q2 collector voltage as the open-circuit voltage (5.76 V) times the divider loading of it to result in an actual (loaded) $V_{C3} = (5.76\ \text{V})\cdot(0.969) = 5.58$ V. Continuing the static analysis toward the output, at 5.45 mA, $V_{BE3} = 0.694$ V and

$$V_Z \approx 5.1\,\text{V} - (5\,\Omega)\cdot(20\,\text{mA} - 5.45\,\text{mA}) = 5.03\,\text{V}$$

Adding V_{BE3} to $V_Z = 5.0$V, the total voltage drop from V_{C3} to V_O is 5.72 V. Then

$$V_O = V_{C3} - 5.72\ \text{V} = -0.14\ \text{V}$$

Ideally, this voltage should be 0 V. Using 5 % resistor values, it is as close as is practical. This output voltage *offset error* will be fed back to the amplifier input and reduced further by feedback.

The function of the Q3 stage is as a *buffer*; it presents a relatively high resistance to the previous stage (at the Q2 collector) and a low resistance to drive the output, as the ideal voltage amplifier.

For incremental feedback analysis, first identify the input and output quantities as v_i and v_o. Choose v_o as the feedback quantity, and thus $x_f = v_o$ and $T_o = 1$. Multiple choices exist for the error quantity and the art of feedback analysis includes the skill in choosing x_f and x_E so that the analysis is made as simple and perspicuous as possible. One possibility is

$$v_E = v_i - v_{b2}$$

where $v_{b2} = v_b$(Q2), the Q2 base voltage. The G input (v_E) port has for its negative terminal the actual v_{b2} as loaded by the Q2 base circuit of H. The loading effect on H is included in H and the diff-amp r_{M1} excludes $R_{B2} = R_f \,||\, R_i$. The effect of loading by the Q1 base on the H divider in this choice of v_E is in H because v_E is chosen to be the already-loaded Q2 base voltage. For this choice of v_E,

$$r_{M1} = [r_{e1} + R_B\,/(\beta+1)] + r_{e2} = r_{E1} + r_{e2} \approx 52\,\Omega + 670\,\Omega/151 = 58\,\Omega$$

with R_0 neglected because it is much greater than the other resistances shunting it in the emitter circuit.

Q1, Q2 form a BJT diff-amp, and the voltage sources driving the input loops of Q1, Q2 might be a better choice as a differential v_E;

$$v_E = v_i - v_B$$

This Thevenized alternative for error will be chosen instead, where v_B is the thevenized H source. This choice is preferred because H is kept simple; it is a resistive voltage divider. And also, the diff-amp remains symmetrical because the Thevenin resistance of H, or $(R_f||R_i)$, is included in the base circuit of Q2 for calculation of r_{M1} of the first stage of G and can balance with R_B of Q1 to keep the diff-amp stage symmetrical. Let the input source of Q2 be the Thevenin equivalent circuit of the feedback divider so that $R_{B2} = R_i||R_f \approx 667\ \Omega$. This is close to $R_B = 670\ \Omega$. In effect, we are placing the loading on the feedback divider of the Q2 base (or r_{B2}) into G so that

$$H = \frac{v_{B\,2}(\text{oc})}{v_o} = \frac{R_i}{R_f + R_i} = \frac{1}{3} \approx 0.333$$

Then the output of H is the unloaded Thevenin (open-circuit) voltage of the divider.

For both choices of error quantity, the input quantity v_i appears directly ($T_i = 1$) in v_E. Only the feedback loop itself remains in the block diagram, like a noninverting op-amp.

As statically determined, the diff-amp-stage emitter bias current of about 2 mA is split equally between Q1 and Q2. We can calculate

$$r_{e1} = r_{e2} = r_e \approx 26 \text{ mV}/(I_0/2) \approx 26 \ \Omega$$

Furthermore, for a balanced diff-amp stage,

$$R_{B1} = 670 \ \Omega \approx R_f \| R_i = (2.0 \text{ k}\Omega) \| (1.0 \text{ k}\Omega) \approx 667 \ \Omega$$

Then the transresistance of the first stage - the resistance across which the (differential) input voltage develops the emitter current (that becomes output current after an α loss) is

$$r_M = \frac{R_{B1}}{\beta_1 + 1} + r_{e1} + r_{e2} + \frac{R_{B2}}{\beta_2 + 1} = 2 \cdot (r_e + R_B/(\beta + 1)) = 2 \cdot r_E$$

This assumes that $R_0 >> r_{M1}$ so that it appears as a current source; it is 92 times larger than r_{M1}. The second expression for r_M applies equally to r_{M1} or r_{M2} because it assumes a balanced or symmetrical input circuit: the parameters of the Q1 side equal those of the Q2 side.

Having settled upon the error quantity and error-voltage loop, we can proceed to find the forward path voltage gain. The output is a voltage, v_o, and the error quantity, which is the input of G, is voltage v_E. Therefore, G is a voltage gain.

Continuing the analysis of the gain, G, consider the cascaded second stage of Q3. The interstage loading can be accounted for by including the shunt input resistance of the Q3 stage in the calculation that loads the collector resistance. The Q2 collector node resistance is

$$R_{c2} = R_L \| (\beta + 1) \cdot (r_{e3} + r_Z + R_Z \| (R_f + R_i))$$

Alternatively, if the input to Q3 were opened and the open-circuit gain calculated, then the loading appears in the base resistance of Q3 as $R_I/(\beta + 1)$. Either method of calculating interstage loading will give correct results when the gains are combined. For the CC stage,

$$r_{E3} = r_{e3} + r_Z + R_Z \| (R_f + R_i) = 4.8 \ \Omega + 5 \ \Omega + 1.27 \text{ k}\Omega = 1.28 \text{ k}\Omega$$

Having now obtained r_{E3}, the output CC stage voltage gain is

$$G_2 = \frac{r_Z + (R_f + R_i) \parallel R_Z}{r_{E3} + R_L/(\beta_3 + 1) + [r_Z + (R_f + R_i) \parallel R_Z]} = \frac{1.37\,\text{k}\Omega}{1.41\,\text{k}\Omega} = 0.971$$

The static emitter currents of Q1, Q2 let us determine

$$r_{e1} = r_{e2} = 26\text{ mV}/(1.013\text{ mA}) = 25.5\,\Omega$$

Then the first stage diff-amp gain is

$$G_1 = -\alpha_2 \cdot \frac{R_L}{R_B/(\beta_1 + 1) + r_{e1} + r_{e2} + (R_f \parallel R_i)/(\beta_2 + 1)}$$

Because of circuit symmetry, this can be approximated as

$$G_1 = \alpha_2 \cdot \frac{R_L}{2 \cdot [R_B/(\beta + 1) + r_e]} \approx (0.993) \cdot \frac{6.2\,\text{k}\Omega}{60\,\Omega} \approx 102.6$$

Then the forward path voltage-amplifier gain is

$$G = G_1 \cdot G_2 = (102.6) \cdot (0.971) = 99.6 \approx 100$$

To be more precise, a tiny amount of v_i passes to the output through the emitter of Q2 to its base, as i_{b2}, causing v_{b2} that is attenuated (in the reverse direction) through H and r_{Go} to the output. The gain of this nefarious path backwards through H is negligible.

To find H, recount that the error voltage sums across the diff-amp input as

$$v_E = v_i - v_B$$

where $v_B = [R_i/(R_f + R_i)] \cdot v_o = (1/H) \cdot v_o$. Then

$$H = \frac{v_B}{v_o} = \frac{R_i}{R_f + R_i}$$

The closed-loop gain is then found from the feedback formula;

$$A_v = T_i \cdot \frac{G}{1 + G \cdot H} \cdot T_o = 1 \cdot \frac{100}{1 + (100) \cdot (0.333)} \cdot 1 = 2.92$$

The relatively low forward gain of 100 causes A_v to be less than $1/H$ by 2.8 %. The closed-loop output error should be about

$$-0.14\text{ V}/(33.3) = -4.2\text{ mV}$$

an amount quite acceptable for many applications. This amplifier thus depends on feedback for gain and offset precision.

Matched-Transistor Circuits

Matched-BJT Circuits

An entire category of new BJT amplifier circuits was discovered at Tektronix in the 1960s, mostly by Barrie Gilbert, George Wilson, and Art Metz. These *translinear circuits* are based on the close matching of BJTs that is possible in integrated or *monolithic* (*mono* = one, *lith* = rock) circuits. By placing BJTs close together on a single silicon wafer, not only do their β values match well, their I_S values match even better. I_S is the *saturation current* of the BJT junctions which determines V_{BE} at a given I_E. The equation for the forward-biased (conducting) voltage of a p-n junction is

$$V_F = V_T \cdot \ln\left(\frac{I}{I_S} - 1\right) \approx V_T \cdot \ln(I / I_S) \ , I >> I_S$$

where V_T is the usual 26 mV at room temperature and ln is the natural (base *e*) logarithm. For a PN3904 BJT, $I_S \approx 10$ fA or $10 \cdot 10^{-15}$ A. Then for $I = 1$ mA,

$$V_{BE} = (26\,\text{mV}) \cdot \ln(10^{-3}\ \text{A}/10^{-14}\ \text{A}) = (26\,\text{mV}) \cdot \ln(10^{11})$$
$$\approx (26\,\text{mV}) \cdot (25.33) \approx 0.65\,\text{V}$$

which is the value we have been using all along. Diodes and *b-c* junctions are made with a higher breakdown voltage than BJT *b-e* junctions by reducing the *doping*, the concentration of atoms that contribute electrons (*n*) or holes (*p*) in the silicon lattice to make *n* or *p* material. Lower doping concentration, *N*, increases the avalanche breakdown voltage, *BV*, in inverse proportion. (The distribution of doping concentration across a junction also affects breakdown voltage.) With lower concentration, I_S is also less. At the same current there is a higher forward-biased voltage across more lightly-doped junctions for equal junction areas. I_S is proportional to junction area,

A; a larger junction has lower junction voltage at the same current. To summarize: $N\uparrow \Rightarrow BV\downarrow$, $I_S\uparrow$, $V_F\downarrow$; $A\uparrow \Rightarrow I_S\uparrow$, $V_F\downarrow$.

This theory might seem inconsistent with actual part data, such as the common diodes listed in the table below. The 1N4148 V_F should be higher than that of the lower-breakdown 1N4152, but it is not. The areas appear to be the same with the same rated current. And though it has a larger current rating and hence junction area, for its *BV* a decade higher, the 1N647 V_F should be higher than that of the 1N4152.

Diode Type	*BV*, V	V_F(1 mA), V	max I_D, mA
1N4152	40	0.63	200
1N4148	75	0.57	200
1N647	400	0.62	400

These discrepancies are caused in part by wide tolerances on V_F, series resistance, and the difficulty of determining an accurate I_S. They are also affected by whether the junction n and p doping is abrupt on each side (step junction) or *graded* (doping concentration linear with distance) for higher breakdown voltage. These variations cause high breakdown-voltage junctions to have incremental resistances of close to $r_d = 2\cdot V_T/|I_D|$. Thus, for discrete diodes,

$$r_d \approx \frac{52\,\text{mV}}{|I_D|}$$

They also illustrate the difficulty of matching discrete p-n junctions, especially of different design and manufacture. For BJTs and diodes made at the same time on the same substrate (or "chip"), these discrepancies largely disappear; monolithic junctions match well.

For discrete design, BJT matching is also more difficult than for monolithic BJTs. Transistors taken from the same manufacturing batch are likely to be better matched than not. Adjacent transistors from the same manufactured roll are about the best for matching short of manual matching on a *curve tracer*, an instrument for displaying $i_C(V_{CE}, I_B)$ and $i_C(V_{BE}, I_B)$.

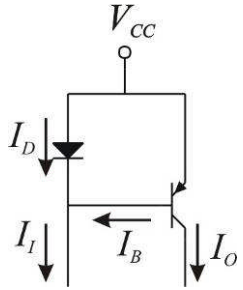
For *semi-discrete* design, BJT arrays such as the 3-NPN, 2-PNP CA3096 are available. The NPNs are matched and the PNPs, made of a different construction than the NPNs, are matched.

Current Mirrors

A *current mirror* is an inverting current amplifier constructed (or *implemented*) with matched BJTs. A current mirror is shown below. If the diode junction is matched to the PNP *b-e* junction, then

$$I_I = I_D + I_B$$

The junctions have the same voltage across them and if they are matched, then they have the same current through them, or $I_E = I_D$.



$$I_I = I_E + I_B = [(\beta + 1) + 1] \cdot I_B$$

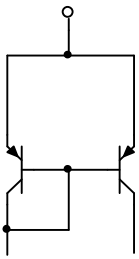
and

$$I_O = I_C = \beta \cdot I_B$$

The static current gain of the mirror is positive, given the directions of positive current in the diagram, and is

$$A_I = \frac{I_O}{I_I} = \frac{\beta}{\beta + 2} \approx 1 \, , \beta >> 1$$

Discrete diodes do not match well with BJT *b-e* junctions because the *b-e* junctions are designed for high emitter injection efficiency of its minority carriers into the base and this results in a low breakdown voltage of about 5 to 7 V. Discrete diodes usually have higher breakdown voltages. For the 1N4152, it is 40 V. The junctions are different and so is their $V(I)$.



To achieve a closer match, another BJT *b-e* junction could be used in place of the diode, as shown. Discrete BJTs taken from the same manufacturing roll will have a match of around 10 to 25 mV. Integrated

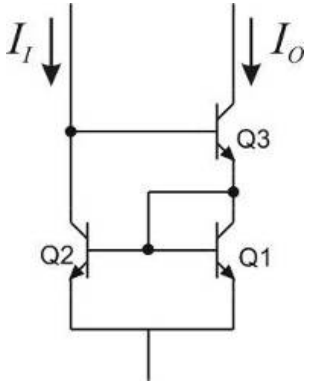
BJTs on the same substrate match to a few mV or less.

A better current mirror that requires 3 BJTs is the *Wilson current mirror*, shown below, invented by George Wilson of Tektronix. The BJTs are matched by $V(I)$ and β. Some of the input current, I_I, supplies I_{B3}. Then Q3 emitter current flows through the diode, Q1, and causes Q2 to conduct the same amount of current. Input current

$$I_I = I_{B3} + I_{C2}$$

Because of matching, $I_{E2} = I_{E1}$. By KCL,

$$I_{E3} = I_{E1} + I_{B2}$$



$$I_O = I_{C3} = \alpha \cdot I_{E3} = \alpha \cdot (I_{E2} + I_{B2}) = (\beta + 2) \cdot I_{B2} = \frac{\beta + 2}{\beta} \cdot I_{C2}$$

Then substituting for I_{C2} from I_I,

$$I_O = \alpha \cdot I_{E3} = \alpha \cdot \frac{\beta + 2}{\beta} \cdot (I_I - I_{B3}) = \frac{\beta}{\beta + 1} \cdot \frac{\beta + 2}{\beta} \cdot \left(I_I - \frac{I_O}{\beta} \right)$$

Solving for the current gain,

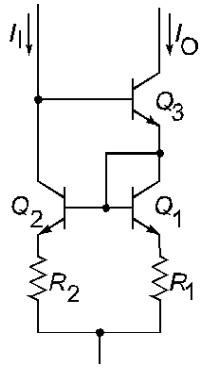
$$A_I = \frac{I_O}{I_I} = \frac{\left(\frac{\beta + 2}{\beta + 1} \right)}{1 + \left(\frac{\beta + 2}{\beta + 1} \right) \cdot \frac{1}{\beta}} = \frac{\beta \cdot (\beta + 2)}{\beta \cdot (\beta + 1) + (\beta + 2)} = \frac{(\beta + 1)^2 - 1}{(\beta + 1)^2 + 1}$$

(Note how the first expression on the right has the form of the feedback formula, suggesting a feedback interpretation of the circuit.)

Some values are plotted for the Wilson current mirror in comparison with the 2-BJT mirror.

β	Wilson A_I	2-BJT A_I
10	0.984	0.833
100	0.9998	0.980
150	0.99991	0.987
200	0.99995	0.990

For β = 200, the Wilson-mirror A_i is in error (from an ideal gain of one) by only 50 ppm. It is superior for precision current amplification to the 2-BJT mirror by at least two decades.



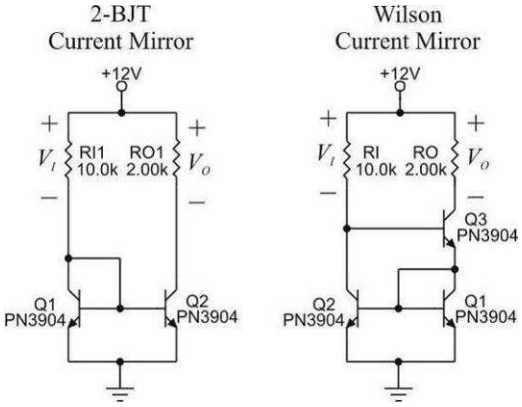
Wilson current mirror

By adding emitter resistors to the Wilson current mirror, as shown, the current gain becomes

$$A_i = \frac{i_O}{i_I} = \frac{(\beta+1)^2 - 1}{(\beta+1)^2 + 1} \cdot \frac{R_2}{R_1}$$

for matched-β BJTs. The β-dependent factor, like α, is even less sensitive in value to β variation than is α because of the β^2 terms. The external resistors usually dominate in determining the current gain, though in ICs, matched-β BJTs can also achieve fairly accurate A_i without them.

Lab Experiment: Current Mirrors



The current gain of the two current mirrors shown above are measured and compared with their theoretical gain formulas by also measuring on a DMM the values of β of the BJTs. Input and output currents are measured by either using a current-measuring DMM in series with input or output, or by measuring the voltage across R_I and R_O and using Ohm's Law.

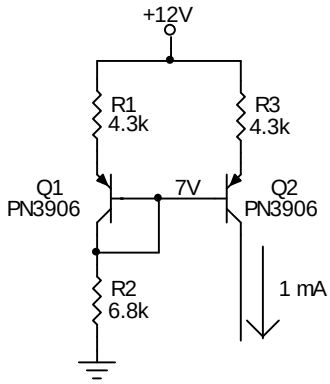
By using current mirrors, a given current gain can be achieved using transistors having a wide range of β values. These techniques overcome the inherent limitations in the BJT as a component, and achieve higher performance with limited parts.

The 2-BJT and Wilson mirrors were previously introduced and the Wilson mirror transfer function derived. The simpler 2-BJT current mirror uses two transistors and optionally two resistors. It has been used in previous amplifiers. By matching Q1 to Q2 by making it another BJT of the Q2 type (and preferably from the same manufactured batch of BJTs) bias currents can be established as shown below for PNP BJTs.

The circuit will maintain its current gain over a wide range of currents. No selection of BJTs by β is required nor is the gain significantly affected by $\beta(T)$. The 2-BJT current-mirror gain is dependent on matching of the two R_E resistors for $V(R_E) >> V_{BE}$.

When there is no R_E ($R_E = 0\ \Omega$), then for matched βs and *b-e* junctions,

$$i_I = i_{C1} + i_{B2} = i_{E1} + \frac{i_{C2}}{\beta}$$



Solving for i_{E1},

$$i_{E1} = i_I - \frac{i_{C2}}{\beta} = i_I - \frac{i_O}{\beta}$$

The *b-e* junctions are in parallel and have the same voltage across them. Then for matched junctions, $i_{E1} = i_{E2} = i_{C2}/\alpha$, and

$$i_O = i_{C2} = \alpha \cdot i_{E1} = \alpha \cdot \left(i_I - \frac{i_O}{\beta} \right)$$

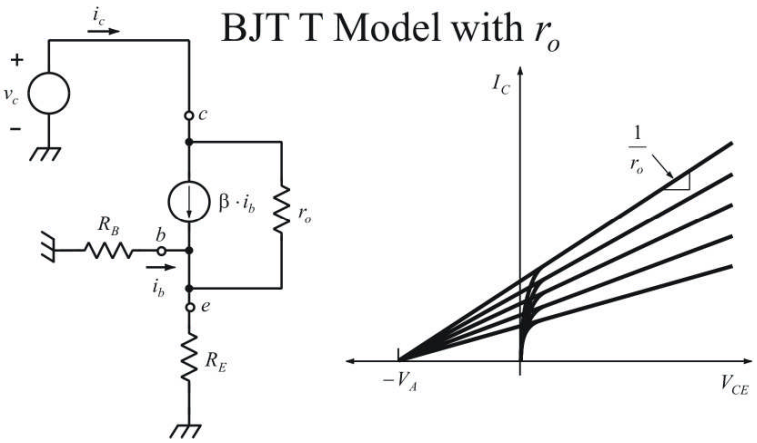
With a little rearranging, the current gain is

$$A_i = \frac{i_O}{i_I} = \frac{\beta}{\beta + 2}$$

BJT Model Parameter r_o

In the BJT T model i_C is independent of v_{CE}. Plots of $i_C(v_{CE})$ with I_B as a parameter are displayed on *transistor curve tracer* instruments. They are theoretically flat curves for the T model, but real transistors have a slope and if projected to the left, they intersect the horizontal axis (where $i_C = 0$ mA) at what is called the *Early voltage*, V_A, (after J. M. Early) which is typically about 100 V. It corresponds to an equivalent resistance between collector and emitter called r_o:

$$r_o \equiv \frac{|V_{CE}| + V_A}{|I_C|}, \; |I_C| > 0 \text{ mA}$$



When r_o is included in circuit analysis, it becomes more complicated. (See *Designing Amplifier Circuits*, D. Feucht, for a more detailed development.) By inspecting the circuit diagram above, some of the effects of r_o can be identified. First, r_o provides an additional path from the input loop (emitter) to the output. It is a passive path, in the form of a voltage divider that contributes the forward-path gain,

$$F = \frac{R_L}{R_L + r_o}$$

Second, r_o shunts R_L. The gain formula without r_o can be modified to include it. A CE approximation in somewhat the form of a transresistance gain is

$$\text{CE } A_v \approx -\frac{R_L \parallel r_o}{R_B / \beta_0 + r_m + (R_E \parallel r_o) \cdot (1 - F)} + F$$

This reduces to

$$\text{CE } A_v \approx -\frac{R_L}{\frac{R_B / \beta_0 + r_m}{(1 - F)} + R_E \parallel r_o} + F$$

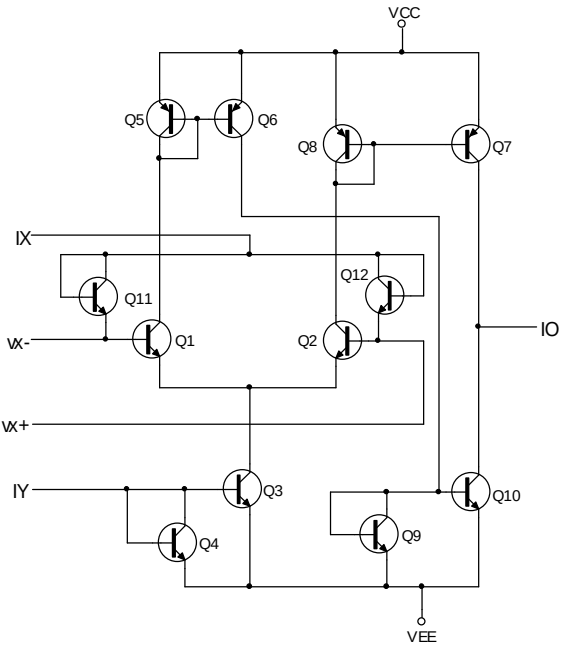
For large r_o, $1 - F$ is slightly less than one, just as α_0 is. If $1 - F \approx \alpha_0$, then CE voltage gain,

$$\text{CE } A_v \approx A_v\big|_{r_o \to \infty} + F$$

where the first term, which is negative for the CE stage, is the voltage gain calculated disregarding r_o. F reduces gain magnitude. The effect of r_o is usually, though not always, insignificant.

Translinear Circuits

The 2-BJT diff-amp can be driven by a pair of matched junctions to form a current amplifier as shown in the simplified circuit diagram of a repeated half of the LM13700.



Diodes Q11, Q12 match Q1, Q2 and form an example of a *translinear circuit*. (BJTs are used as diodes for better junction matching even for monolithic circuits.) It has the property that the ratios of currents in the diodes equal the ratio of BJT currents. This is evident when Kirchhoff's Voltage Law (KVL) is applied to the diode loop and the BJT input loop. Let $i_C(\text{Q11}) = i_{D-}$ and $i_C(\text{Q12}) = i_{D+}$. Then

$$i_{D-} + i_{D+} = I_X$$

For a perfect current mirror, $i_C(\text{Q3}) = I_Y = i_C(\text{Q1}) + i_C(\text{Q2})$, for $\alpha = 1$ ($\beta >> 1$). The diff-amp output current is differential and is

$$i_O = i_{O+} - i_{O-}$$

With this nomenclature, applying KVL to the diode loop,

$$v_x = v_{x+} - v_{x-} = -V_T \ln\left(\frac{i_{D+}}{I_s}\right) + V_T \ln\left(\frac{i_{D-}}{I_s}\right) = -V_T \ln\left(\frac{i_{D+}}{i_{D-}}\right)$$

For matched *b-e* junctions, I_s values match and cancel. Applying KVL to the BJT loop,

$$v_x = v_{x+} - v_{x-} = V_T \ln\left(\frac{i_{O+}}{I_s}\right) - V_T \ln\left(\frac{i_{O-}}{I_s}\right) = V_T \ln\left(\frac{i_{O+}}{i_{O-}}\right)$$

Equating and solving for the current ratios,

$$\frac{i_{O+}}{i_{O-}} = \frac{i_{D-}}{i_{D+}}$$

It can be shown algebraically that if

$$\frac{a}{b} = \frac{c}{d}$$

then

$$\frac{a-b}{a+b} = \frac{c-d}{c+d}$$

Applying this algebraic identity to the above circuit equations,

$$\frac{i_O}{I_Y} = \frac{i_{O+} - i_{O-}}{i_{O+} + i_{O-}} = -\frac{i_X}{I_X} = -\frac{i_{X+} - i_{X-}}{i_{X+} + i_{X-}}$$

where input $i_X = i_{D+} - i_{D-}$. This can also be expressed as a current-gain transfer function;

$$\frac{i_O}{i_X} = -\frac{I_Y}{I_X}$$

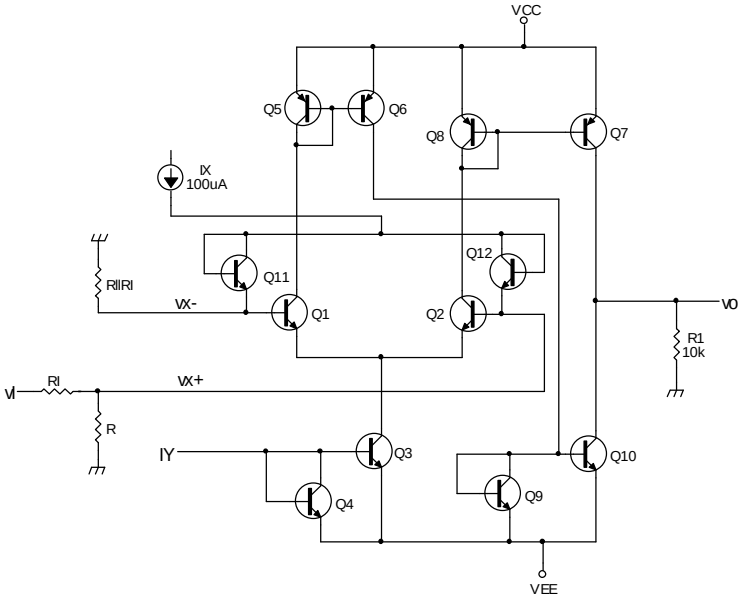
The translinear cell is a linear differential-input, differential-output current amplifier. Because the static current ratios I_X and I_Y set the current gain, by varying either of them, the gain is varied. The diff-amp stage output current, through the mirrors with current gain of one, is the amplifier output current, and the amplifier gain is given as the diff-amp stage gain above. It is inverting because the diodes are

common-anode. An alternative common-cathode connection, with anodes connected to the diff-amp bases, has positive gain. The LM13700 can be used as a two-quadrant multiplier or VGA.

The input circuit is shown below. This amplifier inputs a unipolar I_Y and a unipolar (positive) v_I. When speed is a consideration, it is best to make v_I the gain control and I_Y (which becomes i_Y) the faster waveform.

When $v_I = 0$ V, then the two sides of the circuit are symmetrical and $i_{D-} = i_{D+} = I_X/2$. Let i_X be the differential current,

$$i_X = i_{D+} - i_{D-}$$



Translinear circuits are easier to analyze using the variable, x, to represent the fraction of current that is conducted by one side. The range of x is ±1. Let

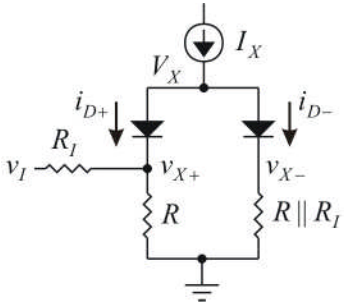
$$i_{D+} = x \cdot I_X$$

By Kirchhoff's Current Law (KCL),

$$i_{D+} + i_{D-} = I_X$$

and thus

$$i_{D-} = (1 - x) \cdot I_X$$



To switch all the current from one diode to the other, an infinite differential voltage is required. Because x is a hyperbolic tangent function of v_X ($x = \tanh(v_X/2 \cdot V_T)$), the current only asymptotically approaches a complete switchover. Huge voltages are theoretically required to switch decades of currents at near-infinitesimal values. In practice, ±200 mV will switch about 99.9 % of the current of a differential BJT or diode pair; $x \approx \pm 0.999$.

Translinear circuit design requires that a decision be made about the full-scale (fs) value of x. The zero-scale (zs) value is $x = 0.5$, where $i_{D+} = i_{D-}$. For a choice of $x(\text{fs}) = 0.75$, the ratio of diode currents is

$$\frac{x}{1-x} = \frac{0.75}{0.25} = 3$$

and $\ln(3) \approx 1.01 \approx 1$. The output fraction i_O / I_Y at full-scale is also 0.5, and the full-scale multiplier gain is 0.5. Therefore, I_Y must be twice as large as the desired i_O.

Substituting into the previously derived equation for v_I and applying a KVL equation involving the resistors,

$$v_X = -V_T \cdot \ln\left(\frac{x}{1-x}\right) = (i_{D+} \cdot (R \parallel R_I) + v_I) - i_{D-} \cdot (R \parallel R_I)$$

$$= i_X \cdot (R \parallel R_I) - \frac{(R \parallel R_I)}{R_I} \cdot v_I = \left[(2 \cdot x - 1) \cdot I_X - \frac{v_I}{R_I}\right] \cdot (R \parallel R_I)$$

Applying the transfer function, the fractional output current becomes

$$\frac{i_O}{I_Y} = -\frac{i_X}{I_X} = -\frac{v_I / R_I}{I_X} + \frac{V_T \cdot \ln\left(\frac{x}{1-x}\right)}{I_X \cdot (R \parallel R_I)}$$

This equation can be expressed in v_I as

$$\frac{v_I}{R_I} = (2 \cdot x - 1) \cdot I_X + \frac{V_T \cdot \ln\left(\frac{x}{1-x}\right)}{R \parallel R_I}$$

To determine the required range for $\pm v_I$, at full-scale let $x = 0.75$, $I_X = 100\ \mu A$, and $R \| R_I = 49.9\ k\Omega$. Then

$$\frac{v_I}{R_I} = \frac{100\,\mu A}{2} - \frac{(25.8\,mV) \cdot \ln(3)}{49.9\,k\Omega} \cong 50\,\mu A - 0.568\,\mu A = 49.43\,\mu A$$

The second term in the above equation accounts for the nonlinearity of the diodes in series with resistors. If $R \| R_I$ is made relatively large, the first term dominates and the output-current fraction becomes approximately a linear function of the input voltage. If x is restrained to be only slightly larger than 0.5, the currents do not deviate much from being balanced and linearity is maintained.

R_I can be chosen for a given v_I(fs) and R found from the parallel combination. For v_I(fs) = 4 V, then

$$R_I = \frac{4\,V}{49.43\,\mu A} = 80.92\,k\Omega \rightarrow 80.6\,k\Omega, \pm 1\%$$

and

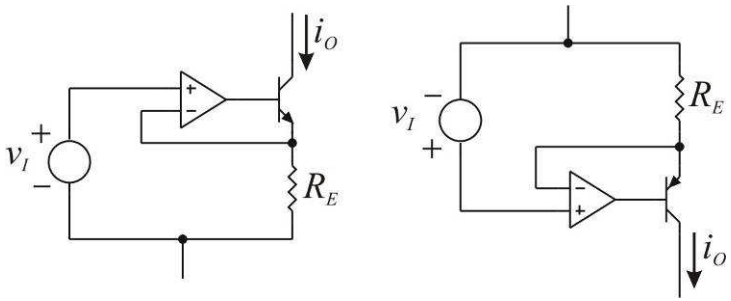
$$R = \frac{1}{\frac{1}{R_I} - \frac{1}{R \parallel R_I}} = 131\,k\Omega \rightarrow 130\,k\Omega, \pm 1\%$$

The full-scale gain from the full-scale Y input of $I_Y = 100\ \mu A$ to the output is $I_Y/2 = 50\ \mu A$ because $(2 \cdot x - 1)$ at the full-scale x of x(fs) = 0.75 is 0.5.

For greater accuracy and linearity, v_I, R_I must be a voltage-dependent current source. One well-used implementation of these circuits is the voltage-to-current (V-to-I) converter.

V-to-I Converters

The V-to-I converter is a transconductance amplifier, usually implemented with an op-amp for precision. One of each polarity is shown below.



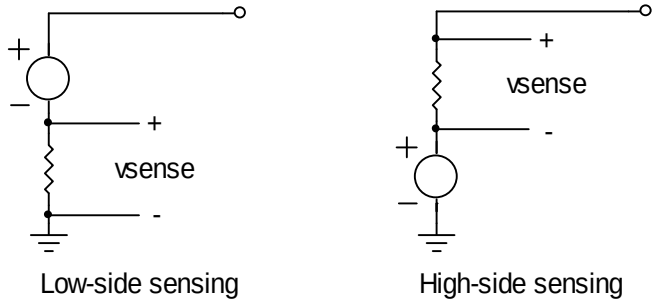
The op-amp input voltages are the same and v_I is applied to R_E. Then the output current is $I_O = \alpha \cdot (v_I / R_E)$. If a Darlington is substituted for the BJT, the α factor is essentially one. For a FET, it is also one.

The V_{BE} offset of the BJT does not affect the accuracy of $G_M = 1/R_E$ because the *b-e* junction is within the op-amp feedback loop. However, the op-amp input offset voltage, V_{IOS}, adds or subtracts from v_I to cause error. V_{IOS} is found in op-amp parts data and for low-cost op-amps is typically a few millivolts. A TL071 will have a typical value of about 5 mV.

The input voltage sources are usually precision resistors driven by a ground-based current source which could be another V/I converter.

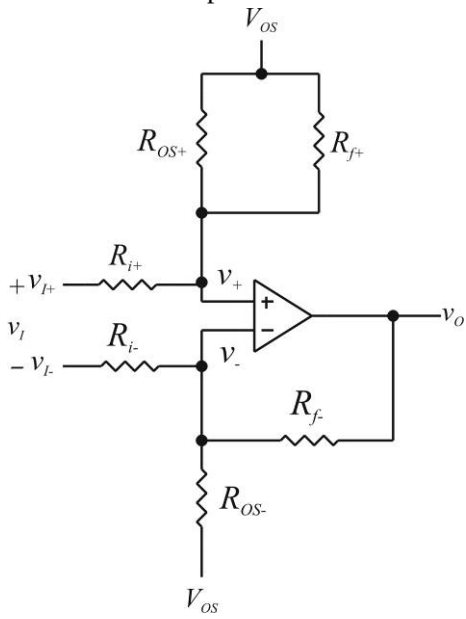
High-Side Current Sensing

Of the two most basic quantities of electronics, voltage and current, voltage is the easier to measure because it is an "across" quantity, a quantity that has a value with respect to some 0 V reference node which is *ground.* Current is not as easy to measure because it is a "through" quantity. To measure current, the branch in which it flows must be opened and a current sensor inserted. This has the disadvantage of intrusion into the circuit structure.



The current from a voltage source can be sensed using a resistor in series with the source. The resistor can be placed in series with either terminal, as shown.

If the source is *ground-referenced* - that is, one of its terminals is grounded - and the sense resistance is small, then low-side sensing can be feasible, though the voltage source is no longer truly ground-referenced. Low-side sensing has the advantage of a ground-referenced or "non-floating" sensed voltage. However, the low-side scheme can cause problems in some circuit designs.



High-side sensing requires that a differential, non-ground-referenced, or *floating* voltage be sensed because neither terminal of the sense resistor is at ground. At first, this appears to be an easy problem to solve. Amplifiers with differential inputs can be applied.

A straightforward scheme is to use a *differential-input op-amp* with a voltage gain of A_v, as shown. It uses both inverting and noninverting inputs. With current-sense resistor, R_s (not shown), the output scaling factor is $A_v \cdot R_s$ in V/A or transresistance units.

R_s develops a voltage, v_I, which is amplified by the *diff-amp*. To be differential, the gain magnitudes of the two inputs must be the same so that a change in the high-side voltage affecting both inputs equally does not affect the amplifier output. This is achieved when the resistors on each side are equal: $R_{x+} = R_{x-} = R_x$. Then for the ideal diff-amp, the voltage gain is

$$A_v = \frac{v_O}{v_I} = \frac{v_O}{v_{I+} - v_{I-}} = \frac{R_f}{R_i}$$

The offset voltage, V_{OS} is actually another input and it has a gain of

$$\frac{v_O}{V_{OS}} = \frac{R_f}{R_i \parallel R_{OS}} + 1$$

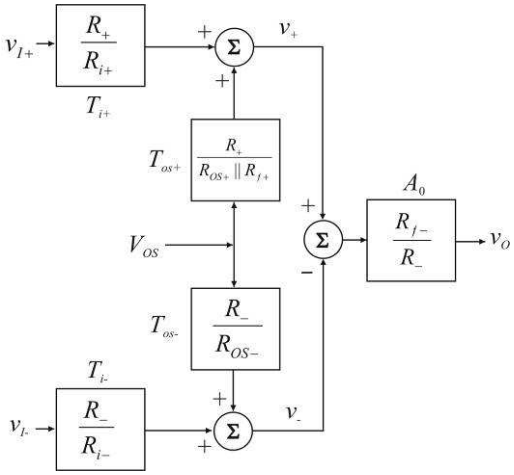
The two differential input voltages can be represented as a vector quantity with two components: a *differential-mode voltage*,

$$v_{DM} = v_{I+} - v_{I-}$$

and a *common-mode* (CM) *voltage,*

$$v_{CM} = \frac{v_{I+} + v_{I-}}{2}$$

which is the average of the input voltages. For high V_{CM}, V_{OS} subtracts from it so that $V_O = 0$ V or some lower voltage. The op-amp input offset voltage error is also multiplied by the common-mode gain and places a limit on the common-mode range.



The differential amplifier above - a *one-op-amp diff-amp* - can be more easily analyzed by representing it as the given block diagram where v_+ and v_- are the op-amp input voltages.

The noninverting op-amp configuration does not have a T_i but it is necessary for differential inputs to add a voltage divider, T_{i+}, so that the noninverting gain has the same magnitude as the inverting gain. In the block diagram, R_+ and R_- are the node resistances at the op-amp inputs. By using

$$R_+ = R_{i+} \parallel R_{OS+} \parallel R_{f+} \; ; \; R_- = R_{i-} \parallel R_{OS-} \parallel R_{f-}$$

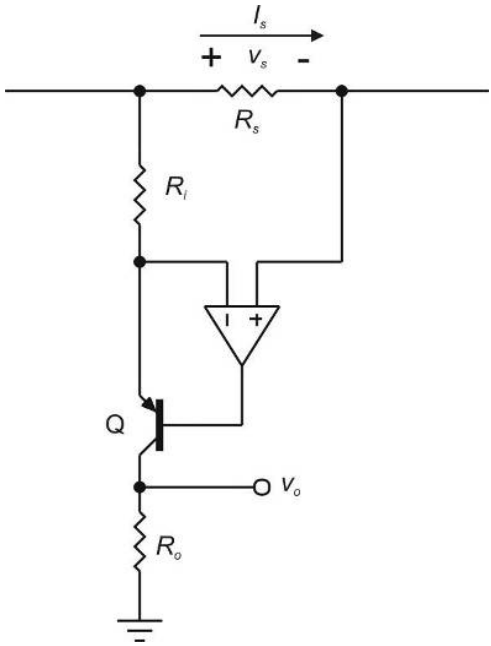
the resistive divider formulas are more simply represented. The block diagram shows the general case where the corresponding matched resistances of the two sides are not necessary made equal. For a differential input, they are made equal and, for instance, $R_{f+} = R_{f-}$.

Another somewhat more elegant scheme for high-side current sensing is shown. The voltage v_s, dropped across sense resistor R_s is replicated across R_i by the op-amp. The resulting current is v_s/R_i and it flows through PNP BJT Q to develop an output voltage, v_o, across R_o.

The transfer function of the circuit is

$$\frac{v_o}{i_s} = \alpha \cdot \frac{R_o}{R_i} \cdot R_s$$

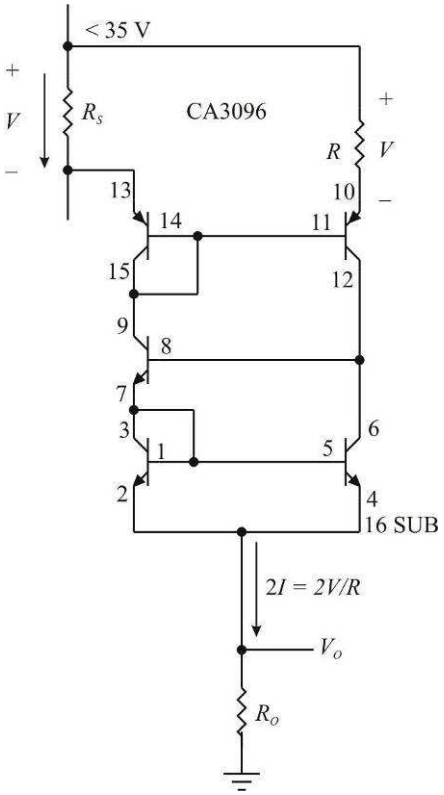
If the BJT has a high β, then $\alpha \approx 1$. This scheme is used in such commercial ICs as the Linear Technology LT6106.



What complicates both this scheme and differential amplifiers is common-mode voltage range. The supply voltages of the op-amps must exceed (or at least equal) the voltage range of both terminals of R_s. If the CM range is large, high-voltage amplifier circuits are required. The one-op-amp diff-amp gets around this, but at the expense of CM gain error. Extended CM voltage input range causes circuit imbalance from the tolerances of gain-setting resistors. For op-amp circuit resistors with a tolerance of $\pm \varepsilon$, the CM gain error tolerance is $\pm 3 \cdot \varepsilon$. It would be advantageous, therefore, to find a circuit that avoids these limitations and is limited in CM voltage only by BJT

c-e breakdown voltages. Such circuits would need to operate from current derived from the sensed circuit, without separate supplies.

High-side current-sense circuits implemented with an op-amp or other circuitry require a separate supply to power them. The positive supply must exceed or at least equal the high-side voltage and the amplifier often must either be floated or designed to operate from a relatively high voltage for a high common-mode range. This impediment can be eliminated by using circuits that are powered entirely from the sensed current loop.



The versatile CA3096 2 PNP, 3 NPN BJT array is used in the circuit shown above. The matched BJTs are important for accuracy. The lower three NPNs form a Wilson current mirror. The current into the pin 6 collector is replicated into the pin 9 collector. The *b-e* junction voltages of matched PNPs cancel with the same currents through them. Summing the voltages around the input loop, sensed voltage V is replicated across R, causing $I = V/R$ to flow through the

pin 12 PNP and into the mirror. With the same current through the other PNP, connected as a diode, its *b*-*e* voltage is the same. For a CA3096, the mismatch of PNPs (and also of NPNs) can be up to 7 mV. The PNPs have a typical β of 47 and minimum of 20. The NPNs have a typical β of 290. The mirror thus has a current gain of 0.99998. The bottom of the mirror circuit outputs $I_O = 2 \cdot I$ which develops output voltage V_O across R_O.

This circuit satisfies another requirement of high-side sense circuits, that of having high *common-mode* voltage *rejection* (CMR). If the CM voltage at the sense resistor varies while the current remains constant, the varying voltage across the circuit should not cause a change in output current. In other words, the output resistance of the circuit (looking up from R_O) should be that of a current source: very high. Tracing from the bottom of the circuit upward through diodes, eventually a collector is encountered. The output has the resistance of r_o of the PNP with collector pin 12. This r_o is the *c*-*e* resistance parameter of an extended BJT T model. It varies inversely with I_C in the model. In circuit operation, it accounts for the fact that as V_{CE} is increased, I_C will also increase slightly. CA3096 PNPs have a breakdown voltage limited to 35 V and so is the maximum CM voltage.

Viewed as a transresistance amplifier,

$$V_O = I_O \cdot R_O = 2 \cdot I \cdot R_O = 2 \cdot \frac{V}{R} \cdot R_O = 2 \cdot \frac{I_S \cdot R_S}{R} \cdot R_O$$

The gain is then

$$R_m = \frac{V_O}{I_S} = 2 \cdot R_O \cdot \left(\frac{R_S}{R} \right)$$

More generally, the voltage gain (whatever the sense-resistor value) is simply

$$A_v = \frac{V_O}{V} = 2 \cdot \frac{R_O}{R}$$

For a full-scale sensed voltage, V, of 0.25 V (2.5 A fs through $R_s = 0.1\ \Omega$, for instance, or 25 A through 10 mΩ), a value of $R = 100\ \Omega$ and $R_O = 1.00\ \text{k}\Omega$, then $A_v = 20$ and $V_O(\text{fs}) = 5$ V. Also, $I_O(\text{fs}) = 2 \cdot I(\text{fs}) = 5$ mA.

Lab Experiment: High-Side Current-Sense Circuit

A prototype of the above CA3096 high-side current-sense circuit was built with 1 % resistors: R = 100. Ω; R_O = 1.00 kΩ. The CMR of the circuit was tested with an Innovatia Floating Differential Source and the following data were obtained.

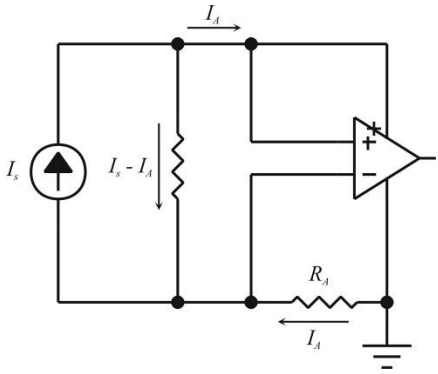
At full-scale current (lower 3 rows), r_o of the BJTs is lower and this is seen in the lower r_{out} values than at mid-scale current. The CMR also is not as good (is lower) at lower CM voltages. For better design, a lower fs current is recommended than 5 mA. A decade lower would be better; it increases r_{out} by about a decade.

V_- (– side of R_S), V	V, mV	V_O, V	r_{out}, kΩ over 10 V	r_{out}, kΩ over 20 V
10	128	2.52	38.5	125
20	128	2.68		
30	128	2.78	100	
10	250	4.74	18.0	60.6
20	250	4.92		
30	250	5.07	66.7	

The circuit can also be implemented with discrete BJTs if the parts are taken from the same manufacturing lot (same roll of transistors for through-hole parts), though monolithic BJTs such as the CA3096 are more assuring. For applications for which the above circuit is adequate, its advantage of simplicity and low parts count and cost is realized.

The amplifier current, $2 \cdot I = I_A$ taken from the sensed current loop, must be returned to the isolated loop. A path is shown as resistance R_A in the following generalized scheme for a high-side circuit. If I_A is small and R_A correspondingly large, then loop isolation is minimally impacted. R_A must be large enough to drop sufficient voltage for the amplifier to operate linearly.

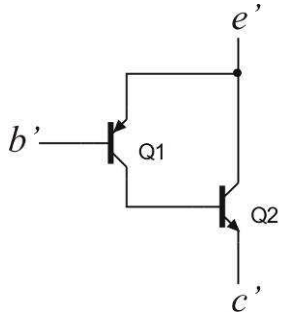
At zero-scale input current, this voltage can be zero, leaving the ground-referenced amplifier with zero volts across it. If the amplifier inputs are always conducting some current, then this minimum current must sustain the minimum operating voltage across R_A. To insure minimum operating voltage at zero-scale input current, R_A can be placed in series with an avalanche diode or a stack of diodes.



Alternatively, instead of returning R_A to ground, it can be returned to a large enough supply voltage to keep the amplifier linear with the negative supply terminal connected to the current-sense circuit output ground. Then even zero return current would not cause the amplifier to have insufficient voltage across it. The supply ground must be connected to the isolated sense loop only through R_A (or an alternative circuit for this return path branch). Then only the ground node of the supply is common to the current-loop circuit, through R_A, and no unintended current is shared by them.

Complementary BJT Configuration

A commonly-used circuit in the output stages of power amplifiers is the complementary two-BJT configuration, shown below. NPN BJTs are preferred power devices (just as n-channel MOSFETs are). The complementary PNP configuration is, like the Darlington configuration, a two-transistor combination that has three terminals like a single BJT. Consequently, the complementary PNP behaves like a PNP while using a NPN to implement the output.



The base of the equivalent PNP is the base of Q1, the Q1 emitter and Q2 collector node is the equivalent emitter, and the Q2 emitter is the equivalent collector. Apply KCL at the equivalent emitter to express the currents in terms of I_{E2};

$$I_E' = I_{E1} + I_{C2} = I_{E2} \cdot \left(\frac{1}{\alpha_1 \cdot (\beta_2 + 1)} + \alpha_2 \right)$$

The first term,

$$I_{E2} \cdot \frac{1}{\alpha_1 \cdot (\beta_2 + 1)}$$

is I_{E1}; $I_{C1} = I_{B2} = I_{E2}/(\beta_2 + 1)$. Then

$$I_{E1} = \frac{I_{C1}}{\alpha_1} = \frac{I_{E2}}{\alpha_1 \cdot (\beta_2 + 1)}$$

This can be rearranged to become

$$I_E' = I_{E1} + I_{C2} = I_{E2} \cdot \left(\frac{1}{\alpha_1 \cdot (\beta_2 + 1)} + \alpha_2 \right) = I_{E2} \cdot \left(1 + \frac{1}{\beta_1 \cdot (\beta_2 + 1)} \right)$$

where the second term,

$$I_{E2} \cdot \frac{1}{\beta_1 \cdot (\beta_2 + 1)} = I_{B1}$$

An equivalent β, or β' is

$$\beta' = \frac{I_C'}{I_B'} = \frac{I_{E2}}{I_{B1}} = \frac{I_{E2}}{I_{E2} \cdot \left(\frac{1}{\beta_1 \cdot (\beta_2 + 1)} \right)} = \beta_1 \cdot (\beta_2 + 1)$$

An equivalent α can also be solved for;

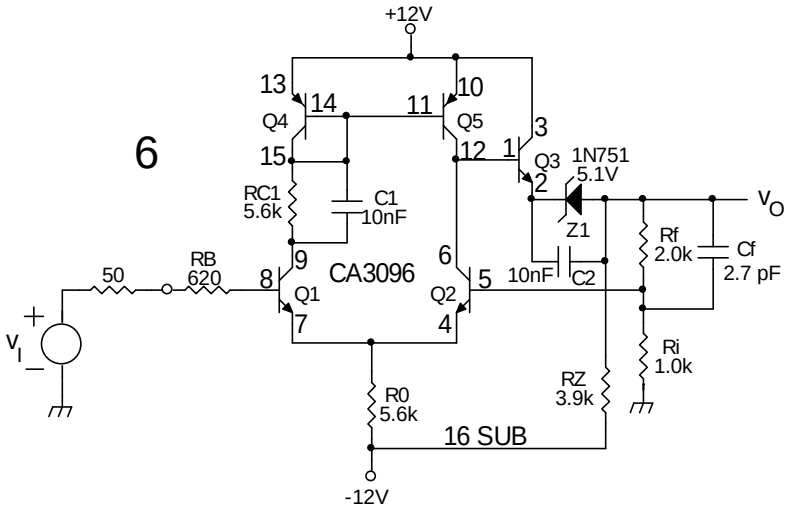
$$\alpha' = \frac{I_C'}{I_E'} = \frac{I_{E2}}{I_{E2} \cdot \left(1 + \frac{1}{\beta_1 \cdot (\beta_2 + 1)} \right)} = \frac{\beta_1 \cdot (\beta_2 + 1)}{\beta_1 \cdot (\beta_2 + 1) + 1} = \frac{\beta'}{\beta' + 1}$$

Thus, the formulas for β and α are consistent with each other relative to a single BJT. We consequently have some assurance in regarding the complementary PNP as having at least quasistatic functional

equivalency to a single PNP with the designated terminals corresponding to a single PNP.

Dynamically, the circuit is somewhat more complicated in that the two BJTs form a tight feedback loop with a loop gain dependent upon external circuit elements. If the PNP is a smaller, faster transistor than the more powerful NPN, then the NPN dominates dynamic response in the loop and this usually stabilizes the pair.

Noninverting BJT-Array Amplifier



With knowledge of current mirror circuits, the 3-BJT discrete feedback amplifier 5 can be improved by increasing its gain. The 3-NPN amplifier is embellished by using a *high-side* (connected to V_{CC}) current mirror in place of R_L, thus not wasting i_{c1}. Two discrete PNPs could be used to form a simple mirror, though monolithic arrays of BJTs are readily available. We will use the CA3096 BJT array for implementation. It has two matched PNPs and three matched NPNs. The improved amplifier (renamed amplifier 6) is shown, drawn in the CircuitMaker schematic editor program.

The difference from amplifier 5 is mainly in the collector circuits, where a 2-BJT mirror has replaced the load resistor. The gain of G nearly doubles when both collector currents contribute to the output. What was R_L is now a PNP collector current source and gain increases to the maximum allowed by r_{b3}.

The other change is that R_Z is increased in value so that I_{E3} and the avalanche-diode current is about 3 mA. Avalanche-diode voltages are specified at a given current and this should be considered when biasing them. The lower current will cause the avalanche voltage to be reduced from the previous feedback amplifier, to reduce the static output voltage offset. In the previous amplifier this offset error at the output is fed back to the input and amplified to reduce it by the feedback factor.

The CA3096 can be replaced by an Intersil HFA3096 with faster BJTs for dynamic speed improvement. For now, the CA3096 is sufficient for quasistatic (low-frequency) design. Typical βs, V_{BE} (1 mA), and I_S of the CA3096 BJTs are given by Intersil:

typical NPN $\beta = \beta_{NPN} = 390$; V_{BE} (1 mA) = 0.69 V ; $I_S \approx 2.5$ fA
typical PNP $\beta = \beta_{PNP} = 47$; V_{BE} (1 mA) = 0.66 V ; $I_S \approx 8$ fA

The static design of the amplifier proceeds very much like the 3-BJT design. R_{C1} is retained for minimization of the effects of thermal distortion (or *thermals*, as they are called) and is bypassed with capacitor C1 to reduce v_{c1} to near-zero.

Again, a goal of the static design is to make $V_O = 0$ V when $V_I = 0$ V. The current mirror makes the static design more challenging because it does not provide a fixed resistance (as was R_L) across which to develop a voltage to set V_{C2}. The voltage now depends on r_{i3} which varies with output load. The static design leaves the output open so that only the feedback divider and bias resistor R_Z load it. The much higher voltage gain, however, decreases the output offset error more than in the previous amplifier and reduction of it for the open-loop forward path is not as critical.

To derive the incremental design formulas, we will use the same

$$v_E = v_i - v_{Ho}$$

where $v_{Ho} = v_B$ is the open-circuit (Thevenized) output voltage of the feedback divider. With an open-circuit *H* source, *G* retains the base resistance of Q2 in

$$r_{M1} = 2 \cdot (r_e + R_B / (\beta_{NPN} + 1)) = 55.5\,\Omega$$

The current mirror has a current gain of

$$A_i = \frac{\beta_{PNP}}{\beta_{PNP} + 2} = 0.96$$

Its deviation from one merely reduces the gain increase from i_{c1} by about 4 %. The mirror takes a differential output - the difference of $i_{c2} - i_{c1}$ - and by inverting i_{c1} makes them add algebraically so that i_{o1} is about $i_{c2} + 0.96 \cdot i_{c2} \approx 1.96 \cdot i_{c2}$.

The effective R_L is $r_{i3} = r_{b3}$ of Q3 or

$$r_{b3} = (\beta_{NPN} + 1) \cdot [r_e + r_z + (R_f + R_i) \| R_Z] \approx (391) \cdot (1.709\ \text{k}\Omega) = 668\ \text{k}\Omega$$

In the chapter, "Cascode Amplifier Designs", the BJT parameter r_o is applied. It reduces the resistance of the collector node. For this first pass, we will ignore r_o. If I_E is reduced, r_o increases inversely with it, and it can be made large. In this circuit, it is not large enough to be ignored, though ignoring it does not greatly alter the result of the analysis.

The Q2 collector voltage, $v_{b3} = r_{b3} \cdot i_{b3}$, where i_{b3} is the output of the mirror added algebraically to i_{c2}. The voltage gain, A_{v2}, of second-stage Q3 is v_o/v_{b3}, an emitter-side voltage divider with an input resistance of r_{e3}. An equivalent way of thinking about A_{v2} is to let Q3 be a current amplifier of gain $\beta_{NPN} + 1$ having a load resistance of $r_z + (R_f + R_i) \| R_Z$. This is equivalent to calculating v_{b3} and dividing from there, which results in

$$G_2 = \frac{v_o}{v_{b3}} = \left(\frac{r_z + (R_f + R_i) \| R_Z}{r_z + (R_f + R_i) \| R_Z + r_{e3}} \right) = \frac{1.70\ \text{k}\Omega}{1.70\ \text{k}\Omega + 13.5\ \Omega} = 0.992$$

$$G_1 = \frac{v_{b3}}{v_i} = \alpha_{NPN} \cdot (1 + A_i) \cdot \frac{(\beta_{NPN} + 1) \cdot (r_{e3} + [r_z + (R_f + R_i) \| R_Z]}{r_{M1}}$$

$$= \frac{390}{391} \cdot (1.96) \cdot \frac{668.3\ \text{k}\Omega}{20.5\ \Omega} = 63.7\ \text{k}$$

$$G_0 = G_1 \cdot G_2 = 63.2\ \text{k}$$

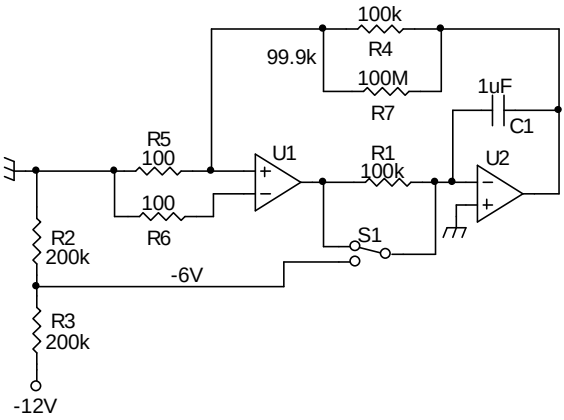
This circuit is a credible op-amp with this much forward-path gain.

High-Gain Amplifier Measurement

A prototype of the BJT-array amplifier 6 as shown above was measured for its open- and closed-loop gains. The pinout of the DIP package follows the same pattern as the TL071, with numbering proceeding counter-clockwise from the dot around the CA3096 package as viewed from the top. Because G of the 3-NPN amplifier is not huge, it can be measured using ordinary bench equipment. The

open-loop gain of this amplifier can be measured indirectly from the amount by which the closed-loop gain falls short of being 1/*H*. Of course, the accuracy of the open-loop gain requires that the resistors in *H* be accurately measured first. Another way is to measure error voltage, v_E, though it is small, and compute the loop gain from it using the feedback formulas.

Semiconductor manufacturers measure input offset and gain by adding an additional op-amp for testing as shown below. U1 is the amplifier under test and U2 is used to do the test. By the voltage divider of R4, R7 and R5, the output voltage of U2 is –1000 times the U1 input voltage. In the upper S1 position, U2 v_O is –1000 times the U1 input offset voltage. In the S1 lower position, the U1 output voltage is forced to change by feedback to +6 V and v_{O2} changes by Δv_{o2}.



Then the U1 open-loop gain is

$$K_1 = \frac{\Delta v_{O1}}{\Delta v_{I1}} = \frac{\Delta 6\text{ V}}{\left(\frac{-\Delta v_{O2}}{1000}\right)} = 1000 \cdot \frac{\Delta 6\text{ V}}{-\Delta v_{O2}}$$

BJT Input-Loop Linearity

A consideration of static circuit design is amplifier linearity. A BJT input loop includes what are essentially linear elements except the BJT *b-e* junction, a diode. Its voltage changes exponentially with emitter current so that changes away from the operating-point are exaggerated in the direction of the change; v_{BE} is superlinear

(increases more than the linear change) for a given positive i_e and decreases less than a linear change for decreases in i_E.

When linearity is a design criterion we need to be able to quantitatively approximate it. For v_i = 0 V (at the operating-point), let $i_E = I_0$, the operating-point current. Then for a change in input voltage, Δv_I, the (nonincremental) change in emitter current is

$$\Delta i_E = \frac{\Delta v_I - \Delta v_{BE}}{R_E}$$

The change in voltage around the input loop causing Δi_E is thus

$$\Delta v_I = R_E \cdot (\Delta i_E) + \Delta v_{BE}$$

The diode equation leads to

$$\Delta v_{BE} = V_T \cdot \ln\left(\frac{\Delta i_E}{I_S}\right) - V_T \cdot \ln\left(\frac{I_0}{I_S}\right) = V_T \cdot \ln\left(\frac{I_0 + \Delta i_E}{I_0}\right) = V_T \cdot \ln\left(1 + \frac{\Delta i_E}{I_0}\right)$$

The input-loop voltage change, Δv_I can be normalized to the linear change, $r_{M0} \cdot \Delta i_E$;

$$\frac{\Delta v_I}{r_{M0} \cdot \Delta i_E} = \frac{R_E \cdot \Delta i_E + V_T \cdot \ln(1 + \Delta i_E / I_0)}{\left(R_E + \frac{V_T}{I_0}\right) \cdot \Delta i_E}$$

This can be put in the form,

$$\frac{\Delta v_I}{r_{M0} \cdot \Delta i_E} = 1 + \left(\frac{V_T}{I_0 \cdot R_E + V_T}\right) \cdot \left(\frac{\ln(1 + \Delta i_E / I_0)}{\Delta i_E / I_0} - 1\right)$$

The deviation from linearity is the second term,

$$\varepsilon_v = \left(\frac{V_T}{I_0 \cdot R_E + V_T}\right) \cdot \left(\frac{\ln(1 + \Delta i_E / I_0)}{\Delta i_E / I_0} - 1\right)$$

The first factor shows the impact that the *b-e* junction nonlinearity, represented by the thermal voltage V_T, makes on the input loop and how a large R_E reduces nonlinearity. The second factor is the extent of the nonlinear exponential itself. It can be approximated as

$$\left(\frac{\ln(1 + \Delta i_E / I_0)}{\Delta i_E / I_0} - 1\right) \approx -\tfrac{1}{2} \cdot \left(\frac{\Delta i_E}{I_0}\right) + \tfrac{1}{3} \cdot \left(\frac{\Delta i_E}{I_0}\right)^2$$

The ratio $\Delta i_E/I_0$ can be expressed in Δv_I and V_{RE} quite simply as

$$\frac{\Delta i_E}{I_0} = \frac{\Delta i_E \cdot R_E}{I_0 \cdot R_E} = \frac{\Delta v_{RE}}{V_{RE}} \approx \frac{\Delta v_I}{V_{RE}} \ , \ \Delta v_{RE} = \Delta v(R_E) \ , \ \Delta v_I << V_{RE}$$

Nonlinearity can alternatively be analyzed as gain change caused by the change in r_M, the incremental input-loop transresistance, from its operating-point value at I_0;

$$\Delta r_M = \frac{V_T}{I_0 + i_e} - \frac{V_T}{I_0}$$

The fractional change in r_M is

$$\frac{\Delta r_M}{r_{M0}} = \frac{V_T \cdot \left(\frac{-\Delta i_E / I_0}{I_0 + \Delta i_E} \right)}{R_E + \frac{V_T}{I_0}} = \left(\frac{V_T}{R_E \cdot I_0 + V_T} \right) \cdot \left(\frac{-\Delta i_E / I_0}{\Delta i_E / I_0 + 1} \right)$$

Substituting for $\Delta i_E/I_0$,

$$\frac{\Delta r_M}{r_{M0}} = \left(\frac{V_T}{R_E \cdot I_0 + V_T} \right) \cdot \left(\frac{-\Delta v_{RE} / V_{RE}}{\Delta v_{RE} / V_{RE} + 1} \right)$$

Both fractional nonlinearity quantities consist of two factors. The first is the same in both and is a quasistatic voltage divider that scales the second factor, the nonlinearity error. As V_{RE} dominates V_T, the first factor becomes small whenever $R_E >> r_e$. As R_E dominates r_M the input loop becomes more linear.

Some values of the nonlinearity factors that depend only on $\Delta i_E/I_0$ (= $\Delta v_{RE}/V_{RE}$) are tabulated below. All changes are negative. For a +1 % (0.01) change in input-loop voltage, the voltage error - effectively the nonlinear component of v_I - is –0.4967 %.

$\frac{\Delta v_{RE}}{V_{RE}} \approx \frac{\Delta v_I}{V_{RE}}$	$\left(\frac{\Delta v_{RE} / V_{RE}}{\Delta v_{RE} / V_{RE} + 1} \right)$, %	$\left(1 - \frac{\ln\left(1 + \frac{\Delta v_{RE}}{V_{RE}}\right)}{\left(\frac{\Delta v_{RE}}{V_{RE}}\right)} \right)$, %
0	0	0
0.01	0.9909	0.4967
0.02	1.961	0.9869
0.05	4.762	2.420

0.10	9.091	4.690
0.20	16.67	8.839
0.50	33.33	18.91
1.00	50.00	30.69

As i_E deviates from I_0 and Δi_E increases, the nonlinearity error increases. For a 1 % deviation ($\Delta i_E/I_0 = \Delta v_{RE}/V_{RE} = 0.01$), then the linearity error is $-0.004967 = -1/200.5 = 7.65$ bits. The larger is $V_{RE} = V_{EE1} - V_{BE}$ (where V_{EE1} is the Thevenin voltage source of the emitter), the more linear the circuit, and applies to both factors of ε_v.

Thermal Distortion and Input Range

For the general case of a BJT input loop with a static base voltage of V_B, a supply voltage, $-V_{EE}$, driving a voltage divider at the emitter consisting of R_{E1} (to $-V_{EE}$) and R_{E2} (to ground), then the emitter voltage,

$$v_E = V_B - v_{BE}\ ,\ V_{EE1} = V_{EE} \cdot \frac{R_E}{R_{E1}}$$

In circuit design, the supply voltage, $-V_{EE}$, is usually given as is I_E for performance optimization. The Thevenin emitter resistance is

$$R_E = R_{E1} \parallel R_{E2} = \frac{V_{EE1} + V_E}{I_E} = R_{E1} \cdot \left(\frac{V_{EE1}}{V_{EE}} \right)$$

The design formulas for the emitter resistors are then

$$R_{E1} = \frac{V_{EE1} + V_E}{I_E} = \frac{V_{EE}}{V_{EE1}}\ ;\ R_{E2} = \frac{1}{\frac{1}{R_E} - \frac{1}{R_{E1}}}$$

To minimize thermal distortion, the change in BJT power dissipation is minimized by operating it at the vertex of a negative power parabola. The BJT stage can be modeled with the BJT as current source i dissipating $p_D = v \cdot i$ with v across the source (v_{CE}). The BJt is in series with a resistor R to voltage source V. Then

$$v = V - i \cdot R$$

and

$$p_D = V \cdot i - i^2 \cdot R$$

Maximum p_D occurs at the vertex of the downward-opening parabola where

$$i = \frac{1}{2} \cdot \frac{V}{R} \Rightarrow v = V/2$$

and $R = (V/2)/i$. By setting the BJT operating-point at max p_D, Δp_D is minimized, junction temperature change ΔT_J is also, and $\Delta v_{BE} = v_{be}$(thermal) is minimized, which is the goal of the derivation.

Thus for minimum thermal dynamic effects, set

$$V_{CE} = V_{CC}/2$$

In that case,

$$I_C = \frac{V_{CC}/2}{R_L}$$

Given that voltage gain for $R_E >> r_E = r_e + R_B/(\beta_0 + 1)$ is

$$-A_v \approx \alpha_0 \cdot \frac{R_L}{R_E} = \frac{V_{CC}/2}{V_{EE1}}$$

By increasing V_{EE1}, both range (or linearity) and I_E are increased. By increasing R_E, both A_v and I_E are decreased. In the BJT output loop, increasing V_{CC} increases range (or linearity), thermals, and V_O; increasing R_L increases A_v but decreases thermals, and V_O.

Voltage Reference

Another use of feedback amplifiers is in the design of *bandgap voltage reference* circuits. These circuits use temperature coefficients (TCs) of *b-e* junctions that are of opposite polarity and make them cancel to result in a stable output voltage. A bandgap voltage reference (circuit 12) is shown below based on a scheme somewhat like that invented by Paul Brokaw at Analog Devices Inc.

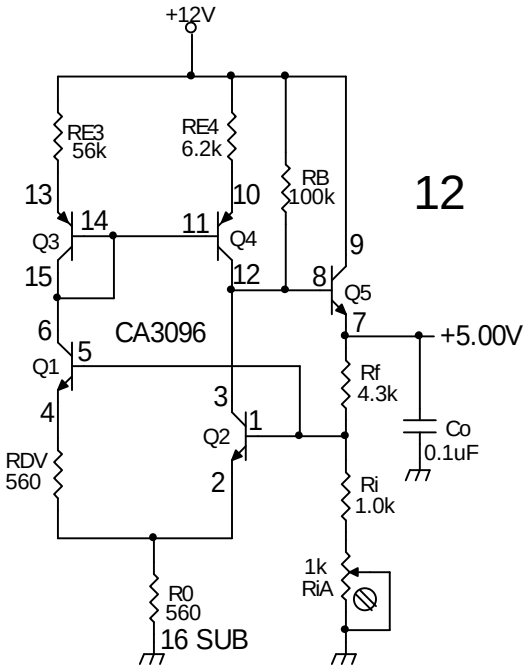
To explain TC cancellation, we backtrack to the *v-i* relationship for p-n junctions:

$$V = V_T \cdot \ln(I / I_S),\ I >> I_S$$

where the *thermal voltage*

$$V_T = \frac{k_B \cdot T}{q_e}$$

$$= (86.173\,\mu\text{V/K}) \cdot T \approx 25.85\,\text{mV} \cdot \left(\frac{T}{300\,\text{K}} \right)$$



V_T has a positive TC because the absolute temperature, T, (in kelvins, K) causes V_T to increase proportionally. In the diode equation, I_S is also temperature-dependent and has a negative TC. When the two effects are combined, $I_S(T)$ dominates, and the TC of a p-n junction or diode is about –2.0 mV/K, (or –2.0 mV/°C).

At a tropical room temperature of 300 K (26.84 °C ≈ 80.3 °F), the TC of V_T is the change in V_T per change in T, or

$$TC(V_T) = \frac{dV_T}{dT} \approx \frac{\Delta V_T}{\Delta T} = \frac{(86.173\,\mu\text{V/K}) \cdot \Delta T}{\Delta T} = 86.173\,\mu\text{V/K}$$

By multiplying V_T by a scaling factor and adding it to V, a voltage should result that has a zero TC.

To generate a scaled V_T in a circuit, two matched BJTs conduct currents with a ratio of (in the case of the above circuit) about

$$I_{C4}/I_{C3} = R_{E3}/R_{E4} = (56\ \text{k}\Omega)/(6.2\ \text{k}\Omega) = 9.0$$

If the current mirror gain of the circuit is worked out in detail, it is

$$A_i = \frac{i_O}{i_I} = \frac{\beta}{1+(\beta+1)\cdot\left(\dfrac{R_{E4}}{R_{E3}}\right)} - \frac{V_{BE4} - V_{BE3}}{i_I \cdot R_{E3}}$$

The difference in *b*-*e* voltages of mirror BJTs Q3, Q4 is

$$V_{BE4} - V_{BE3} = (60\ \text{mV/dec}) \cdot \log_{10}(I_{C4}/I_{C3}) \approx 57.3\ \text{mV}$$

using the estimate of 9 for the current ratio. For $\beta >> 1$, multiply numerator and denominator of the first term by $1/\beta$ and take the limit as $\beta \to \infty$. The mirror gain simplifies to

$$A_i = \frac{i_O}{i_I} \approx \frac{I_{C2}}{I_{C1}} \approx \frac{R_{E3}}{R_{E4}}$$

Using the more exact formula with $\beta = 390$ and $\Delta V_{BE} = 57.3$ mV, then the ΔV_{BE} term subtracts about 10 mV for $i_I = 100\ \mu\text{A}$ and $A_i = 9.05$. Then mirror output current is $I_{C2} = (9.05)\cdot(100\ \mu\text{A}) = 0.905$ mA. At the output, $V_O = 5$ V and $I_{E5} \approx 5\ \text{V}/5.75\ \text{k}\Omega \approx 0.87$ mA and I_{B5} is about 2.23 μA. $I_{C2} = I_{C4} - I_{B5} = 0.903$ mA and $I_{C2}/I_{C1} = 9.03$.

The difference in junction voltages between Q1 and Q2 while conducting different currents is what generates the ΔV_{BE} from which V_T and its positive TC is derived. The V_{BE} value for each BJT is calculated and the values subtracted;

$$V_{BE2} - V_{BE1} = \Delta V_{BE} = V_T \cdot \ln\left(\frac{I_{C2}}{I_S}\right) - V_T \cdot \ln\left(\frac{I_{C1}}{I_S}\right) = V_T \cdot \ln\left(\frac{I_{C2}}{I_{C1}}\right)$$
$$\approx V_T \cdot \ln(A_i) \approx V_T \cdot \ln(9.03) \approx 57\ \text{mV}$$

The I_S factors are equal for matched BJTs and cancel, leaving ΔV_{BE} temperature-dependent only on V_T for a constant current ratio. ΔV_{BE} is applied across R_{DV}. We have $I_{E1} = I_{C1}/\alpha_{NPN} = 100\ \mu\text{A}/0.997$ or about 100.3 μA. Then the design calculation of

$$R_{DV} = \Delta V_{BE}/\ I_{E1} = 57\ \text{mV}/0.1\ \text{mA} = 570\ \Omega \to 560\ \Omega,\ 5\ \%$$

The base voltage, V_B, of Q1, Q2 is

$$V_B = V_T \cdot \ln(I_{C2} / I_S) + R_0 \cdot (I_{E1} + I_{E2})$$
$$= V_T \cdot \ln(I_{C2} / I_S) + R_0 \cdot (I_{C2} / I_{C1} + 1) \cdot I_{E1}$$

Substituting $I_{E1} = \Delta V_{BE}/R_{DV}$,

$$V_B = V_T \cdot \ln(I_{C2} / I_S) + \frac{R_0}{R_{DV}} \cdot (I_{C2} / I_{C1} + 1) \cdot \Delta V_{BE} = V_{BE2} + V_0$$

The current through R_0 is $(I_{C2}/I_{C1} + 1)$ times the current through R_{DV}. Consequently V_0 changes with ΔV_{BE} and V_T. V_0 is proportional to absolute temperature and can be used as a temperature-sensor output. In the V_B equation, we have a *b-e* junction with its –2 mV/K TC in series with the positive-TC V_0. We want to design the reference so that V_O has a zero TC and that requires that $TC(V_B) = 0$. This can be expressed as

$$TC(V_B) = TC(V_{BE2}) + g \cdot TC(V_0) = 0$$

where g is the gain by which $TC(V_0)$ must be multiplied to cancel $TC(V_{BE2})$. From the V_B equation,

$$g = \frac{R_0}{R_{DV}} \cdot (I_{C2} / I_{C1} + 1)$$

Comparing terms to the V_B equation,

$$TC(V_{BE2}) = TC(V_T \cdot \ln(I_{C2} / I_S)) = -2.0\,\text{mV/K}$$
$$TC(V_0) = TC(\Delta V_{BE}) = TC(V_T) \cdot \ln(I_{C2} / I_{C1})$$

Then for $TC(V_B) = 0$,

$$TC(V_{BE2}) + g \cdot TC(\Delta V_{BE}) = 0 \Rightarrow g = \frac{-TC(V_{BE2})}{TC(\Delta V_{BE})}$$

Substituting for the TCs,

$$g = \frac{R_0}{R_{DV}} \cdot (I_{C2} / I_{C1} + 1) = \frac{-(-2.0\,\text{mV/K})}{(86.173\,\mu\text{V/K}) \cdot \ln(9.03)} = 10.55$$

Given $I_{C2}/I_{C1} = 9.03$ and solving for R_0 in

$$g = \frac{R_0}{R_{DV}} \cdot (I_{C2} / I_{C1} + 1) \Rightarrow$$

$$R_0 = \frac{g \cdot R_{DV}}{(I_{C2}/I_{C1}+1)} = \frac{(10.55)\cdot(560\,\Omega)}{10.03} = 589\,\Omega \rightarrow 560\,\Omega, 5\%$$

and

$$V_0 = \frac{R_0}{R_{DV}} \cdot (I_{C2}/I_{C1}+1) \cdot \Delta V_{BE} = (10.55)\cdot(57\,\text{mV}) = 601\,\text{mV}$$

With V_0 determined,

$$\begin{aligned} V_B &= V_{BE2} + V_0 = V_T \cdot \ln(0.903\,\text{mA}/2.5\,\text{fA}) + 0.601\,\text{V} \\ &= 0.688\,\text{V} + 0.601\,\text{V} = 1.29\,\text{V} \end{aligned}$$

The loop forward-path amplifier resembles the noninverting BJT-array amplifier, though with an asymmetrical diff-amp circuit. Because of the high load resistance at the base of Q5, the voltage gain is high enough that the feedback divider will determine the closed-loop gain. Then

$$V_O = V_B \cdot \left(\frac{R_f}{R_i + R_{iA}} + 1 \right)$$

where $R_i' = R_i + R_{iA}$ is the adjusted R_i for 5 V out. Doing divider math, $R_f/R_i = (5\,\text{V} - 1.29\,\text{V})/(1.29\,\text{V}) = 2.988$. Let $R_i' = 1.5\,\text{k}\Omega$ (pot wiper centered), and $R_f = 4.45\,\text{k}\Omega$. The closest 5 % value is 4.3 kΩ. The divider resistors are better chosen to be metal-film (instead of carbon film) ±1 % resistors with TCs of ±100 ppm. These resistors are low in cost and are replacing 5 % resistors for many applications.

From a feedback standpoint, there is no explicit v_I for this feedback circuit, though it is derived from the thermal properties of Q1, Q2. The static voltage V_B is itself the input quantity and what the feedback loop amplifies that is of interest is only this static quantity.

FET-Input Amplifier

The prior emphasis on BJTs does not diminish the importance of FETs from design consideration. FET gate inputs are nearly open circuits and this characteristic can be put to good use in amplifiers requiring high input resistance. FETs are also used in both discrete and integrated circuit design.

A FET has $\alpha = 1$ because it has no appreciable gate current; $i_D = i_S$, and the quasistatic transresistance of the JFET (corresponding to r_m of a BJT) is

$$r_{m0} = -\frac{V_P}{2 \cdot I_{DSS}}$$

where V_P is the JFET *pinchoff voltage* (corresponding to *threshold voltage*, V_{TH}, for MOSFETs) and I_{DSS} is the drain (or source) current at $V_{GS} = 0$ V. JFETs are *depletion-mode* devices in that at $V_{GS} = 0$ V, $I_D \neq 0$ mA. Reverse voltage must be applied gate to source to turn them off. MOSFETs are typically *enhancement-mode* devices in that they are off unless turned on by applying a gate-source voltage greater than V_{TH}. A 2N4416 n-channel JFET has typical parameters: $V_P = -4.5$ V, $I_{DSS} = 10$ mA, ±5 mA, and nominal $r_{m0} = 175\ \Omega$. For n-channel FETs, V_{GS} and V_P are negative and I_D and I_{DSS} are positive. For p-channel FETs, the polarities for both are opposite.

The quasistatic T model also applies to JFETs and MOSFETs, though for static design, the *v-i* relationships are needed. The total-variable relationship of drain current to v_{GS} in the linear region of operation is

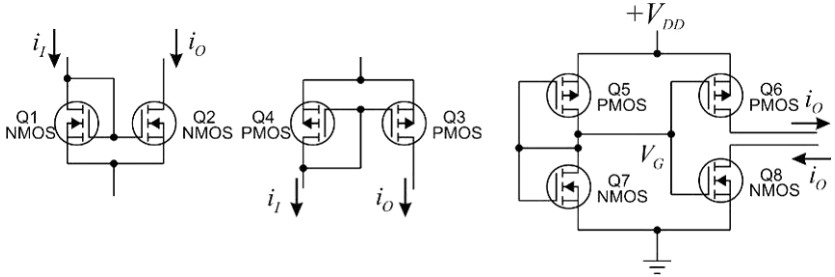
$$i_D = I_{DSS} \cdot \left(\frac{v_{GS}}{V_P} - 1 \right)^2, v_{DS} > v_{GS} - V_P$$

BJT r_e depends on I_E and for FETs, r_m depends on V_{GS} or I_D;

$$r_m = \frac{v_{gs}}{i_d} = \frac{r_{m0}}{1 - \frac{V_{GS}}{V_P}} = \frac{r_{m0}}{\sqrt{\frac{I_D}{I_{DSS}}}}$$

FET transconductance is derived by taking the derivative of $i_D(v_{GS})$ and evaluating it at $v_{GS} = V_{GS}$.

The T model for FETs is essentially the same as for BJTs with the following conversions: the terminals *b, e, c* → *g, s, d*; $r_e \rightarrow r_m$, and the dependent current source, $\beta \cdot i_b \rightarrow v_{gs}/r_m$. Then the circuit theorems that apply to BJTs apply to FETs. The transresistance method of finding gain applies.



N Mirror P Mirror

Matched-BJT circuits, to an extent, carry over to FETs, such as current mirrors but not bandgap references. Both polarities of mirror are shown for MOSFETs.

With gate and drain connected, the n-channel mirror Q1 $V_{GS1} = V_{DS1}$. The drain voltage on Q1 will be whatever is required to conduct i_I. Then Q2 has the same V_{GS} and for matched MOSFETs, $i_O = i_I$. This also applies to the p-channel mirror. The circuit on the right outputs equal currents of opposite polarity. V_G adjusts so that $I_5 = I_7 = I_D$. Then for matched n and p MOSFETs, $I_6 = I_8 = I_D$. I_D can be adjusted by adding a resistor in series with the source of Q5.

JFETs are readily available as discrete devices, and MOSFET arrays are available in the CMOS 4000-series logic family as the 4007UB, with some of the connections of the complementary current mirror shown above. A similar part is the CA3600.

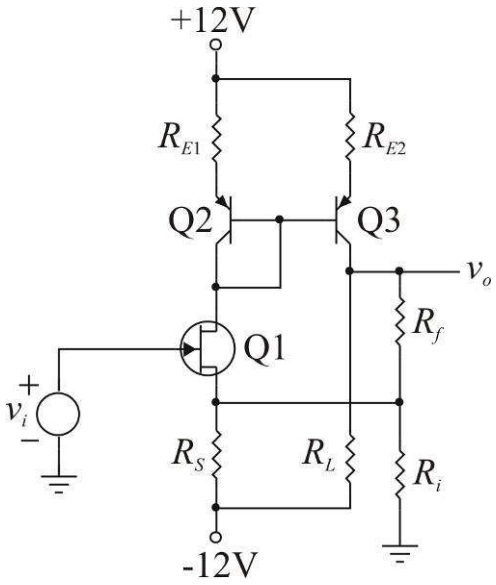
JFET-input amplifiers are used wherever a high input resistance is required. One use of JFETs is shown in the following circuit, a buffer amplifier sometimes found after the V/div attenuator in vertical amplifiers of oscilloscopes.

The n-channel JFET, Q1, operating-point is set by choosing a drain current, I_D, and then calculating (using V_P and I_{DSS}) the value of V_{GS}. An optimum I_D is the current at which the TC of the JFET V_{GS} is zero, or V_{GSZ}. From JFET physics,

$$V_{GSZ} = V_P + 0.8\ \text{V}$$

That is, irrespective of the JFET polarity, if V_{GS} is made to be 0.8 V on the conducting side from pinchoff, it will change least with temperature change. The transconductance at the zero-TC point is

$$r_{mz} = \frac{r_{m0}}{1 - \dfrac{V_P + 0.8\ \text{V}}{V_P}} = \frac{V_P}{0.8\ \text{V}} \cdot r_{m0}$$



R_S is calculated from the chosen (V_{GS}, I_D) operating-point for the JFET. The transistor characteristic curves are found in the JFET data that show the terminal *v-i* relationships. For this circuit,

$$R_S = \frac{12\ \text{V} - V_{GS}}{I_D}$$

The current mirror inputs I_D through Q2 and outputs I_{C3} at the mirror gain. For a static 0 V out, $R_L = 12\ \text{V}/A_i \cdot I_D$ where $A_i \approx R_{E1}/R_{E2}$. The R_E are chosen so that the Q2 base voltage is about half the supply voltage, or 6 V, to minimize thermal effects.

The amplifier is complete in itself without feedback, but a feedback divider is included. As with any feedback analysis, we first choose x_f and x_E. As usual, $x_f = v_o$ is an easy choice. For x_E, choose as before for similar feedback amplifiers:

$$v_E = v_i - H \cdot v_o$$

where H is the divider attenuation. For biasing, R_i and R_S form a divider that can be thevenized to include the −12 V supply. R_S is chosen to set I_D. The feedback error loop becomes JFET r_m in series with $r_{Ho} = R_f \| R_i'$ where $R_i' = R_i \| R_S$. Applying the transresistance method,

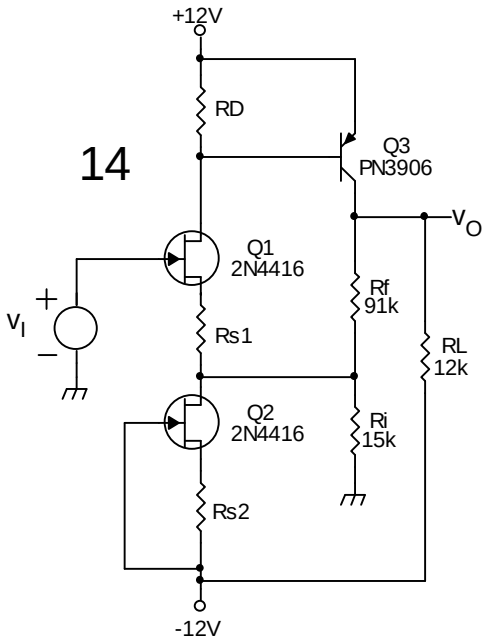
$$G = -\frac{(R_{E1} / R_{E2}) \cdot R_L}{r_m + r_{Ho}}$$

and

$$H = -\frac{R_i'}{R_f + R_i'}$$

The forward gain, G, is affected by the values of the feedback-divider resistors. If $R_f + R_i'$ is not much larger than R_L, then in G, R_L should be replaced by the parallel combination $R_L \parallel (R_f + R_i')$.

`



A simplification of this circuit with larger G sets A_i near maximum by setting $R_{E2} = 0\ \Omega$, removing Q2, and setting the value of R_{E1} for biasing to V_{BE3}/I_{D1}. Then A_i is larger and less stable.

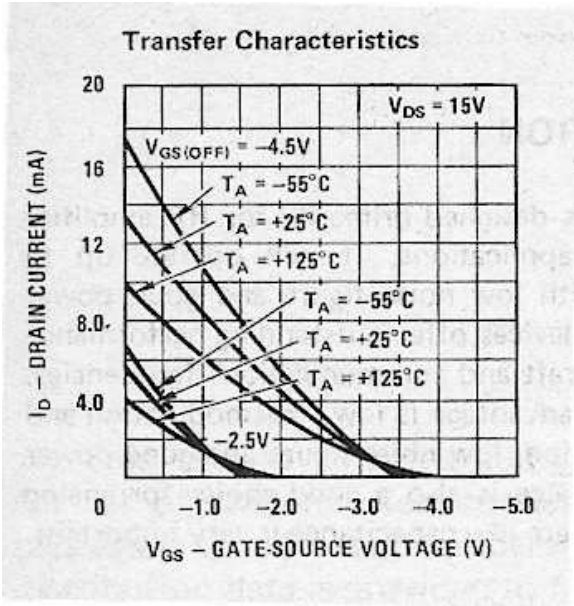
The result is JFET-input amplifier (14), shown using matched JFETs. This circuit is more likely to be found in higher performance applications such as test equipment. For R_{S1}, R_{S2} set to 0 Ω, then Q2 is biased at $V_{GS} = 0$ V and $I_{D2} = I_{DSS}$. As Q1 conducts the same current, V_{GS1} is also zero (because of matching with Q2) and $V_{S1} = 0$ V. With no voltage across the feedback divider, r_{Ho}, it does not divert any of I_{D2} from Q1. The biasing of Q3 at 12 V/R_L = 1 mA

sets its $V_{BE2} = 0.66$ V and $R_D = 0.66\ \text{V}/I_{D1}$. For the nominal 2N4416 I_{DSS} of 10 mA, then $R_D = 66\ \Omega \rightarrow 68\ \Omega$, 5 %. The Miller C_c of Q1 is small with essentially no voltage gain to the drain. The static voltage is one *b-e* junction drop down from +12 V and Q1 will be subject to thermals. By adding a shunt RC in series with the drain, the static V_{DS} can be set near $V_{DD}/2 = 6$ V. The shunt C avoids the Miller effect.

An alternative design using R_{S1}, R_{S2} sets the biasing of Q1, Q2 so that they operate around the zero-TC point of a 2N4416 JFET, at a V_{GS} of

$$V_{GSZ} = -4.5\ \text{V} + 0.8\ \text{V} = -3.7\ \text{V}$$

At this value of V_{GS}, I_D is found from the 2N4416 data. The curves of $I_D(V_{GS})$ are plotted as shown, taken from the National Semiconductor *FET Databook* (1977).



Two sets of plots are shown at three different temperatures for $V_P = -2.5$ V and -4.5 V. In both cases, the curves at different temperatures intersect at a value of V_{GS} that is 0.8 V less in magnitude than V_P. Using the curve for $V_P = -4.5$ V and an ambient (air) temperature of $T_A = 25$ °C, $I_{DZ} = I_D(V_{GSZ}) = 0.8$ mA. At this low value of I_D,

$$r_m = \frac{V_P}{0.8\,\text{V}} \cdot \frac{-V_P}{2 \cdot I_{DSS}} = (5.625) \cdot \frac{4.5\,\text{V}}{2 \cdot (13\,\text{mA})} = 974\,\Omega$$

This is a rather large r_m compared to $r_{m0} = 173\ \Omega$ and can reduce gain appreciably in a follower circuit. The zero-TC scheme is better applied to single-JFET circuits for which static and quasistatic stability (or *precision*) and not speed is important.

The static zero-TC design requires that

$$R_{S2} = \frac{3.7\,\text{V}}{0.8\,\text{mA}} = 4.63\,\text{k}\Omega \rightarrow 4.7\,\text{k}\Omega$$

and $R_D = 825\ \Omega \rightarrow 820\ \Omega$. For 0 V across the feedback divider output resistance, R_{S1} must equal R_{S2}. The biasing requirement of the R_{S1}, R_{S2} values cause the first-stage gain, with its large r_M, to be low, and the gain of G depends mainly on G_2. For audio and other amplifiers not needing flat gain down to 0 Hz, R_{S1} can be shunted with a capacitor large enough that the time constant is below the lowest frequency of interest. Or the time constant can be chosen for emitter peaking, to cancel the pole caused by C_{c3}.

Circuit Dynamics

Reactance

The three basic passive circuit *elements* are resistance, capacitance, and inductance. The last two are *reactances.* They do not dissipate power by converting electric power to heat as do resistances. They store it instead. Capacitors store *charge, q* (which is what electricity *is*). Inductors store *flux*, λ. The defining equations for *C* and *L* are

$$q = C \cdot v \; ; \lambda = L \cdot i$$

Reactances differ from resistance in that the relationship between voltage and current for them involves time, making them *dynamic* relationships. Both charge and flux relate to current and voltage through time, and they are defined as

$$q = i \cdot t \; ; \; \lambda = v \cdot t$$

Current, *i*, is the rate of charge; $i = q/t$. Voltage is the flux rate, $v = \lambda/t$. The unit of *q* is the coulomb (C) which is A·s, and for flux is the volt-second (V·s). These are total-variable equations. For a small change in charge or flux,

$$i = \frac{dq}{dt} \; ; \; v = \frac{d\lambda}{dt}$$

The incremental form of the defining equations for *C* and *L* are

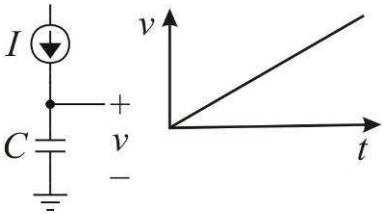
$$dq = C \cdot dv \; ; \; d\lambda = L \cdot di$$

Dividing each side by $dt \approx \Delta t$,

$$\frac{dq}{dt} = i = C \cdot \frac{dv}{dt} \; ; \; \frac{d\lambda}{dt} = v = L \cdot \frac{di}{dt}$$

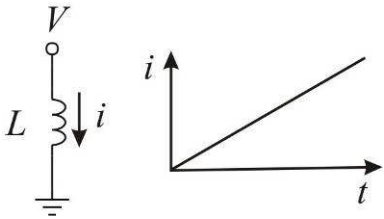
These are the basic *v-i* relationships for *C* and *L*. The current through a capacitor is proportional to the rate of change of voltage across it, $dv/dt \approx \Delta v/\Delta t$. This is the slope of the plotted waveform. For a constant voltage, there is no change in charge and the current is zero.

A current source driving a capacitor, as shown, will generate a constantly increasing voltage, or a voltage *ramp*. Similarly a constant voltage applied across an inductor will cause a current ramp in it with slope *di*/*dt*. Both the capacitor voltage and inductor current will increase indefinitely, for a capacitor over infinite time is an open circuit and an inductor is a short circuit.



Reactance is a kind of "resistance" that we can derive from the defining equations for *C* and *L* and from the *v*-*i* equations. By Ohm's Law, incremental resistance is

$$r = \frac{v}{i}$$



The *v*-*i* equations for *C* and *L* can also be put in this form, though not directly as they are, for what would *dv*/*i* = *dt*/*C* mean? The *v*-*i* relationship for capacitance when solved for *dv* is

$$dv = \frac{1}{C} \cdot i \cdot dt$$

We are already familiar with the concept from calculus of a *differential*, *dx*. In the above equation, *v* is a function of *t*, or *v*(*t*). Time, *t*, is the *independent variable* because it can be chosen arbitrarily. Once it is chosen, *v* depends on the value of *t* and is the *dependent* variable.

Both *v* and *i* are functions of time, and we can set the independent variable $dt = \Delta t$. Then

$$dv(t) = \frac{1}{C} \cdot i(t) \cdot \Delta t$$

Within the time interval Δt, *i*(*t*) can change. If Δt is a small interval, then presumably *i*(*t*) will not change much in it and *dv*(*t*) can be approximated as

$$dv(t) \approx \Delta v(t) = \frac{1}{C} \cdot \bar{i}(t) \cdot \Delta t$$

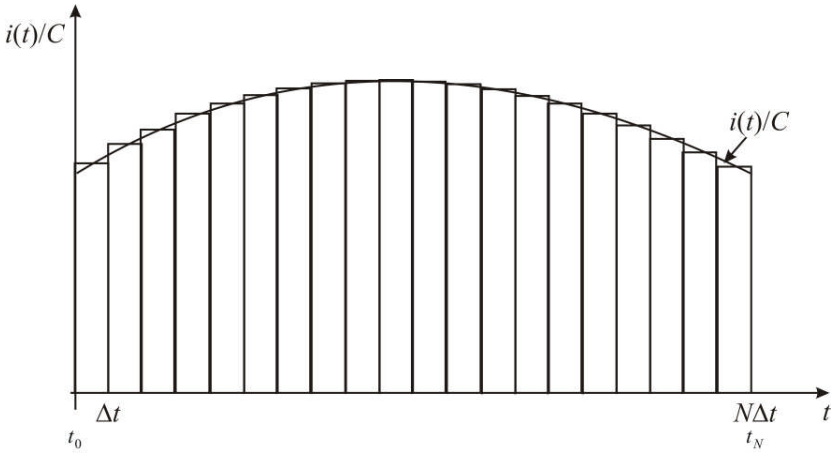
where the bar over i indicates its average over Δt. Now let t be a succession of these time intervals. The end of each interval is

$$t_1 = \Delta t \ , \ t_2 = 2 \cdot \Delta t \ , \ \ldots , \ t_n = n \cdot \Delta t \ \text{ where } n = 1, 2, \ldots$$

In general, the end of the nth time interval is $n \cdot \Delta t$, where n is a positive integer. Then for each n there is a corresponding $\bar{i}(t_n)$ and $\Delta v(t_n)$;

$$dv(t_n) \approx \Delta v(t_n) = \frac{1}{C} \cdot \bar{i}(t_n) \cdot \Delta t$$

For each time interval ending in $t = t_n$, we approximate $i(t) \approx \bar{i}(t_n)$. A plot of $i(t)/C$ is shown below. The stepped approximation to it using interval averages of $i(t_n)$ is $\bar{i}(t_n)/C$. The current in each interval is constant over the interval at the average value.



The geometric interpretation of the equation for $\Delta v(t)$ is that for each Δt interval, the area of the rectangle of width Δt and height $\bar{i}/C$ is Δv for that interval. As the independent variable, Δt, decreases to zero, the dependent variable, Δv approaches dv. In other words, as the time interval is made smaller, Δv becomes a more accurate approximation of dv.

To obtain $v(t)$, add all the dv intervals over all the Δt intervals of time, from $t = t_0$ to t_N. Let this time interval have N intervals. Then we

are interested in knowing what Δv is from $t_0 = 0$ s to $t_N = N \cdot \Delta t$. This overall change in v, or Δv, is the sum of all the Δt interval changes:

$$\Delta v = \Delta v(t_1) + \Delta v(t_2) + \Delta v(t_3) + \cdots + \Delta v(t_N)$$

This can be rewritten using a more compact mathematical sigma (Σ) notation;

$$\Delta v = \sum_{n=1}^{N} \frac{\bar{i}(n \cdot \Delta t)}{C} \cdot \Delta t \quad \text{from } t = 0 \text{ s to } t_N$$

As $\Delta t \to 0$ in the limit, then $\Delta v(t_n) \to dv$. As Δt is reduced, there are more Δt intervals between t_0 and t_N. In the limit, N increases to infinity. When the summation is taken to the limit, the result is the calculus operation of the *integral*.

$$\Delta v = v(t_N) - v(t_0) = v\Big|_{t_0}^{t_N} = \lim_{\Delta t \to 0} \sum_{n=1}^{(t_N - t_0)/\Delta t} \frac{\bar{i}(n \cdot \Delta t)}{C} \cdot \Delta t = \int_{t_0}^{t_N} \frac{i(t)}{C} \cdot dt$$

The integral sign looks like an S for summation - summation in the limit, where there are an infinite number of area intervals of 0+ width. Instead of an average i over a time interval, the value of $\bar{i}(t)$ for the infinitesimally thin interval is $i(t)$ where t is the point in time of that interval.

Derivatives and Integrals

Now we have the two basic operations of calculus. (An *operation* is a function of a function.) Both are based on limits and shrinking Δt to zero. The *derivative* is

$$y(t) = \frac{dx(t)}{dt} = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t}$$

and the *definite integral* is

$$\Delta y = y\Big|_{t_a}^{t_b} = \lim_{\Delta t \to 0} \sum_{n=1}^{(t_b - t_a)/\Delta t} x(n \cdot \Delta t) \cdot \Delta t = \int_{t_a}^{t_b} x(t) \cdot dt$$

What is "definite" are the endpoints of its interval, $[t_a, t_b]$ of Δy. These endpoints of the integration are also called *limits*, not to be confused with the operation of "taking the limit" (lim).

If the limits of integration are removed, then the *indefinite integral* results, the inverse operation of the derivative. It is written

$$y(t) = \int x(t) \cdot dt$$

As the inverse operation,

$$\frac{dy(t)}{dt} = \frac{d}{dt}\int x(t) \cdot dt = \int \frac{dx(t)}{dt} \cdot dt = \int dx(t) = x(t)$$

The integral "undoes" what the derivative does. The derivative of $y(t)$ on a graph is the slope of $y(t)$ at any point in time, t. The integral of $x(t)$ is the area under $x(t)$. The steeper the slope of the function, the more area it accumulates per time interval. Then the rate of change of area is the function itself. In the approximation, the height of the rectangles is the average value of the function being integrated, and the rate at which the heights change is its derivative. Because Δt is constant from interval to interval, the change in height not only changes the area but the rate of change is also the original function.

Using calculus notation, we can express

$$v = \frac{1}{C} \cdot \int i(t) \cdot dt \ , \ i = \frac{1}{L} \cdot \int v(t) \cdot dt$$

where

$$q(t) = \int i \cdot dt \ , \ \lambda(t) = \int v \cdot dt$$

For $i(t)$, the total charge that accumulates over time is the sum of all the instantaneous currents times dt. And similarly, flux is the accumulated voltage over time. For constant v and i, the above integrals simplify to

$$Q = I \cdot t \ , \ \lambda = V \cdot t$$

Sine-Waves

We can now return to the quest to find v/i for reactances. Reactance takes on meaning in association with a particular kind of waveform, the *sinusoid* or "sine-wave". From trigonometry, a sine waveform is

$$x(t) = X \cdot \sin[\theta(t)] = X \cdot \sin(\omega \cdot t)$$

The angle, θ, in electronics is the *phase angle* (or *phase*) of the sine-wave and X is its *amplitude.* These two parameters define the sine waveform. The phase is expressed in *frequency*, ω, as

$$\theta = \omega \cdot t$$

The measurement of angles, like those of other quantities, is arbitrary. The ancient world had a 360-day year so we have 360 degrees in one *revolution* (abbreviated *rev*) of the circle. In trigonometry, angles are defined by unitless ratios of lengths (circular arc length divided by radius) and thus "angle units" are pseudo-units. The "natural angle unit", from trig, is the arc length of one rev of the unit circle - its circumference - which is $2 \cdot \pi$. To distinguish this number from degrees, it is completed by adding the pseudo-unit of *radians*, which tells us what the number means.

Frequency is the rate of change of phase with time and is $\omega = \theta/t$ when it is constant. More generally, as a function of time it is

$$\omega = \frac{d\theta}{dt}$$

It is sometimes expressed in "radians/s" or "rad/s" but because the radian is not really a unit, for consistency (as required in math programs such as MathCAD and by good engineers), the unit of ω is the inverse or reciprocal of seconds, or s^{-1}. Then a ks^{-1} is understood to mean $1000\ s^{-1} \neq 1\ (ks)^{-1} = 1/1000 \cdot s$.

From trigonometry, the relationship between sinusoidal and exponential functions involves complex numbers. *Euler's equation* relates trigonometric functions sine and cosine to complex numbers;

$$e^{j\theta} = \cos\theta + j \cdot \sin\theta$$

where instead of i (which is well-established as the variable denoting current), j is used in engineering instead; by definition,

$$j \equiv \sqrt{-1}$$

Then sinusoids can be expressed as complex numbers, as $e^{j \cdot \omega t}$.

Complex numbers have two forms, *rectangular* and *polar*. In that order, complex number x is

$$x = \mathrm{Re}\{x\} + j \cdot \mathrm{Im}\{x\} = \| x \| \cdot e^{j\theta}$$

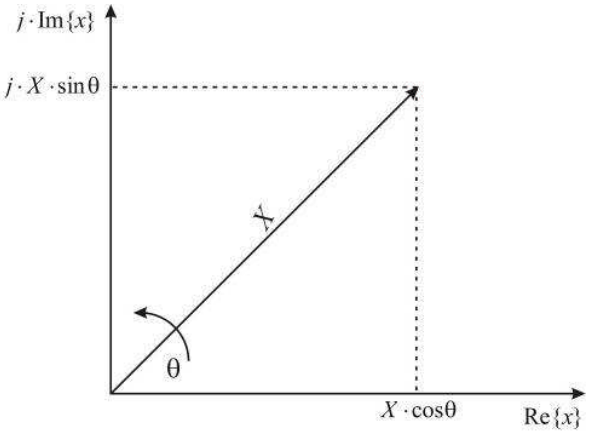
In rectangular form, Re{x} is the real component of x and Im{x} is the imaginary component. The real component of $e^{j\omega \cdot t}$ is Re{$e^{j\cdot\omega t}$} = cos($\omega \cdot t$), where $\theta = \omega \cdot t$, and the imaginary component is Im{$e^{j\cdot\omega t}$} = sin($\omega \cdot t$). In polar form, ||x|| is the magnitude and θ is the phase angle. The forms are equivalent, and the relationship between them is

$$X = \| x \| = \sqrt{[\mathrm{Re}\{x\}]^2 + [\mathrm{Im}\{x\}]^2} \ , \ \theta = \tan^{-1}\left(\frac{\mathrm{Im}\{x\}}{\mathrm{Re}\{x\}}\right)$$

Expressions for sinθ and cosθ can be derived from Euler's equation by adding and subtracting $e^{j\cdot\theta}$ and $e^{-j\cdot\theta}$;

$$\cos\theta = \frac{e^{j\cdot\theta} + e^{-j\cdot\theta}}{2} \ ; \ \sin\theta = \frac{e^{j\cdot\theta} - e^{-j\cdot\theta}}{2\cdot j}$$

Sine and cosine differ only by phase; $\sin\theta = \cos(\theta - \pi/2)$. This difference is expressed in the more general complex-exponential form of sinusoids.



Sinusoids can be represented as vectors with two components, amplitude and phase. When in vector form, as shown, the amplitude is referred to as the *magnitude* of the vector.

By definition, the positive direction of θ is CCW as shown - the direction the vector rotates in time at the frequency of ω.

Sometimes it is convenient to write complex quantities in polar form as a magnitude followed by $\angle\theta$ to denote the phase. For example, 120 V $\angle$45° is such a designation.

Reactance Magnitude and Phase

This brings us to the final step in preparing to define reactance. In general, from calculus,

$$\frac{d}{dt}e^{a\cdot t} = a \cdot e^{a\cdot t}$$

where a is a constant. This derivative will be used without proof. For $a = 1$, the rate of change of the exponential function is the value of the function itself. Its slope increases at the same rate as the function.

Applying the derivative of the exponential function, the rate of change (or d/dt) of $\exp(j\cdot\omega\cdot t)$ is

$$\frac{d}{dt}e^{j\cdot\omega\cdot t} = j\cdot\omega\cdot e^{j\cdot\omega\cdot t} = j\cdot\omega\angle\omega\cdot t$$

With this derivative, the *v*-*i* relationships for C and L can now be expressed when the voltages and currents are sinusoids, written in polar form. For a voltage of

$$v(t) = V\cdot e^{j\cdot\omega\cdot t}$$

(where the magnitude, V, is constant) across a capacitance, C, then the current through it is

$$i = C\cdot\frac{dv(t)}{dt} = C\cdot\frac{d}{dt}\left(V\cdot e^{j\cdot\omega\cdot t}\right) = j\cdot\omega\cdot C\cdot(V\cdot e^{j\cdot\omega\cdot t})$$

Current i is also a sinusoid and is of exponential form. However, the voltage and current have a phase difference as expressed in $i(t)$;

$$i(t) = I\cdot e^{j\cdot(\omega t+\phi)} = I\cdot e^{j\cdot\omega\cdot t}\cdot e^{+j\phi}$$

where ϕ is the phase by which the current leads the voltage. Then

$$i(t) = I\cdot e^{j\cdot\omega\cdot t}\cdot e^{+j\phi} = j\cdot\omega\cdot C\cdot V\cdot e^{j\cdot\omega\cdot t}$$

The current and voltage are of the same frequency and the exponentials of $j\cdot\omega\cdot t$ cancel from each side of the equation leaving

$$I\cdot e^{+j\phi} = (j\cdot\omega)\cdot C\cdot V = j\cdot(\omega\cdot C\cdot V)$$

Noting that

$$\frac{1}{j} = \frac{1}{j} \cdot \frac{j}{j} = \frac{1}{\sqrt{-1}} \cdot \frac{\sqrt{-1}}{\sqrt{-1}} = \frac{j}{-1} = -j$$

and rewriting,

$$\frac{V}{I} \cdot e^{-j \cdot \phi} = \frac{1}{j \cdot \omega \cdot C} = -j \cdot \left(\frac{1}{\omega \cdot C} \right) = j \cdot X_C$$

where the *capacitive reactance*, $X_C = -1/\omega \cdot C$. Use Euler's equation and equate the real and imaginary components:

$$\frac{V}{I} \cdot (\cos\phi - j \cdot \sin\phi) = j \cdot X_C \Rightarrow \frac{V}{I} \cdot \cos\phi = 0 \ , \ \frac{V}{I} \cdot \sin\phi = -X_C$$

Equating and solving for ϕ in the first equation (assuming $V/I \neq 0$),

$$\phi = \cos^{-1}(0) = \pm\frac{\pi}{2} = \pm 90°$$

Substituting for ϕ, $\sin\phi = \pm 1$ and the second equation becomes

$$\frac{V}{I} = \mp X_C$$

Magnitudes are always non-negative. Thus $V/I = -X_C \geq 0\ \Omega$ for $\phi = +90°$. The current leads the voltage in a capacitance by 90°, or voltage lags in phase behind current by 90°, making the reactance phase –90°:

$$Z_C = \frac{v_C(t)}{i_C(t)} = j \cdot X_C = |X_C| \angle -90°$$

A similar derivation for inductance is as follows.

$$v(t) = L \cdot \frac{di(t)}{dt} \Rightarrow V \cdot e^{j \cdot (\omega \cdot t + \phi)} = L \cdot \frac{d}{dt}(I \cdot e^{j \cdot \omega \cdot t}) = j \cdot \omega \cdot L \cdot I \cdot e^{j \cdot \omega \cdot t}$$

Here, ϕ is the angle by which the voltage leads the current. Then simplifying,

$$\frac{V}{I} \cdot e^{+j \cdot \phi} = j \cdot \omega \cdot L = j \cdot X_L$$

where $X_L = \omega \cdot L$ is the *inductive reactance*. Invoking Euler's theorem,

$$\frac{V}{I}\cdot(\cos\phi + j\sin\phi) = j\cdot\omega\cdot L \Rightarrow \cos\phi = 0\,,\ \frac{V}{I}\cdot\sin\phi = X_L$$

From the first equation, $\phi = \pm 90°$ and for the magnitude, $V/I \geq 0$;

$$Z_L = X_L \ \angle\phi = +90°$$

Euler's equation was used to find ϕ, though a graphic method can produce an intuitive answer without equations. Refer back to the vector diagram in the complex plane. For a phase angle of +90°, the vector is vertical, pointing upward, and has only an imaginary component. For reactance with magnitude X_L, then $\phi = +90°$ and

$$Z_L = \frac{v_L(t)}{i_L(t)} = X_L \angle +90°$$

Capacitive reactance is negative and $-j$ is a vertical vector pointing downward.

You might have noticed throughout this section that equations for C and L are the same except with v and i interchanged. C and L are *duals*. The concept of *duality* extends also to circuits. Two circuits are duals of each other whenever the correspondence between them has the following exchanges.

Circuit Duality	
C	*L*
G	*R*
I source	*V* source
open	short
loop	node
series	parallel

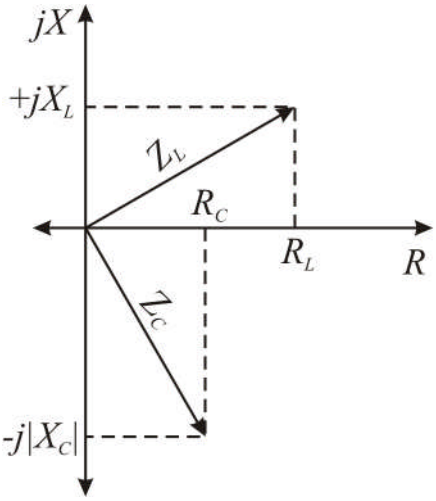
Impedance

When reactance is added to resistance, the result is the more general quantity, *impedance*:

$$Z = R + j\cdot X$$

The following vector plots show two impedances:

$$Z_L = R_L + j \cdot X_L \ , \ Z_C = R_C + j \cdot X_C = R_C - j \cdot |X_C|$$



The real component is resistive and the imaginary component is reactive. The circuit interpretation of Z_L is a resistance of R_L in series with an inductance with a positive reactance of X_L. For Z_C, it is a resistance, R_C, in series with negative capacitive reactance, X_C. We can see from the impedance vectors that resistance is a special case of impedance when there is no reactance. And reactance is the case of zero resistance - an "imaginary resistance".

Ohm's Law applies more generally to impedance:

$$\text{Ohm's Law: } v = Z \cdot i$$

keeping in mind that reactance, and hence impedance, is defined for complex exponential voltage and current waveforms. The addition of imaginary numbers for reactance literally adds another dimension to circuit analysis.

The Frequency Domain

Reactance is a function of ω, not time, and circuit analysis involving it is in the *frequency domain*. In this domain, the waveforms are of the form $e^{j \cdot \omega t}$, which is the complex sinusoid. How circuits affect waveform magnitude and phase is of interest. Amplifier gains in the frequency domain show the *frequency response* of the

amplifier. We now investigate what we can learn about circuit behavior in the frequency domain.

The steady-state frequency domain is the domain of the imaginary axis, $j\cdot\omega$, corresponding to $e^{j\cdot\omega\cdot t}$, the generalized sinusoid. This is a waveform that has constant amplitude. Sine-waves more generally have decreasing or increasing amplitudes which must be included in circuit theory. For more general frequency analysis and design, a real component, σ, is added to the imaginary $j\cdot\omega$ to result in the *complex frequency*

$$s = \sigma + j\cdot\omega$$

Then waveforms of circuits in the *complex-frequency domain*, or *s-domain*, are of the form

$$x(t) = X\cdot e^{s\cdot t} = X\cdot e^{(\sigma + j\cdot\omega)\cdot t} = e^{\sigma\cdot t}\cdot e^{j\cdot\omega\cdot t}$$

The $e^{j\cdot\omega\cdot t}$ factor is sinusoidal but $e^{\sigma\cdot t}$ is new. If σ, a real number, is negative, then the result is a decaying function of time that asymptotically approaches zero. In the language of limit theory,

$$\lim_{t\to\infty} e^{\sigma\cdot t} = 0$$

If $\sigma > 0$, then the function exponentially increases without bounds. Circuits with this response have practical limits to their ranges of v and i and when those limits are encountered, the circuit behaves nonlinearly and *saturates* by staying within the range limit.

The *s*-domain gives us new theoretical abilities that are highly useful in designing circuits. To develop them, we first rewrite the reactances of C and L in the *s*-domain by observing that, in general,

$$\frac{dx(t)}{dt} = \frac{d}{dt}X\cdot e^{s\cdot t} = s\cdot X\cdot e^{s\cdot t} = s\cdot x(t)$$

If the derivative of a waveform is *s* times the waveform, then *s* can be regarded as a differentiation operator on the condition that x is a complex exponential in *s*. For

$$v(t) = V\cdot e^{s\cdot t}$$

then the *v-i* relationship of capacitance can be expressed in the *s*-domain as

$$i = C \cdot \frac{dv}{dt} = C \cdot s \cdot v \Rightarrow Z_C = \frac{v}{i} = \frac{1}{s \cdot C}$$

Similarly, for inductance,

$$i(t) = I \cdot e^{s \cdot t}$$

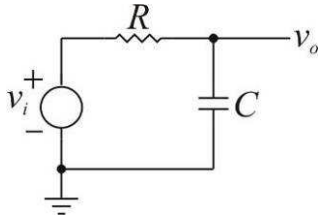
$$v = L \cdot \frac{di}{dt} = L \cdot s \cdot i \Rightarrow Z_L = \frac{v}{i} = s \cdot L$$

We can now write circuit equations in the *s*-domain directly, without using calculus, by using Z_C and Z_L. Both are functions of *s*, not *t*, and the constraint in their use is that the voltages and currents be complex exponential waveforms.

It can be shown in higher mathematics that waveforms of arbitrary periodic waveshape can be decomposed into a sum of complex exponentials (the *Fourier series*). Even aperiodic waveforms such as a step of voltage, can be applied. The complex-exponential waveform constraint will not be a hindrance in ordinary circuit design and the concept of impedance will be useful for arbitrary waveshapes.

The RC Integrator

To apply circuit analysis in the *s*-domain, we use it to analyze the following commonly-used circuit, the *RC integrator*.



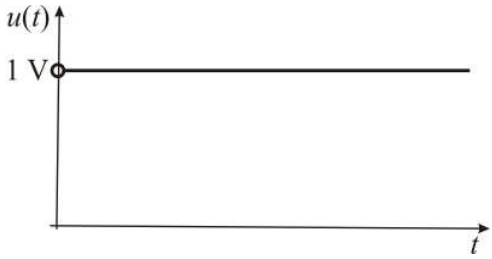
Apply the voltage-divider formula to find the transfer function:

$$T_{RCI}(s) = \frac{v_o(s)}{v_i(s)} = \frac{Z_C}{Z_C + R} = \frac{\frac{1}{s \cdot C}}{\frac{1}{s \cdot C} + R} = \frac{1}{s \cdot R \cdot C + 1}$$

The coefficient of *s* in the denominator is the *time constant*, τ, of the circuit:

$$\tau = R \cdot C$$

The units of the time constant show that it is a time; R has units of Ω, C has units of s/Ω, and their product is s (seconds).



Let v_i be a *unit step* waveform:

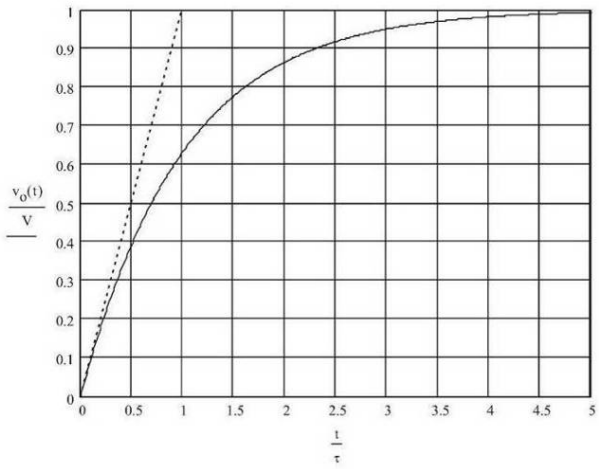
$$u(t) = \begin{cases} 0\ \text{V}, t \le 0 \\ 1\ \text{V}, t > 0 \end{cases}$$

This is not a periodic waveform, as the plot shows.

When the RC integrator is driven by $u(t)$, the response is shown by the plot below. The horizontal axis is in time constants and the vertical axis is *normalized:* scaled to a full-scale value of 1.

The response is exponential. After 5 time constants, the voltage has risen to within 0.674 % of the final (asymptotic) value of one. The time response is

$$v_o(t) = V \cdot (1 - e^{-t/\tau}) \ , \ t > 0 \text{ s}$$



As shown by the dotted line, the initial slope of the curve reaches the final value in one time constant. The function is about 0.632 or 63 % of its final value after one time constant.

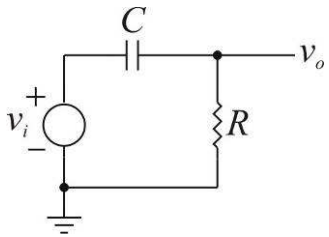
This response can be derived from the *s*-domain transfer function (using the inverse Laplace transform) but is beyond the scope of coverage here. Keeping in mind, however, that *s*-domain analysis assumes complex exponential waveforms, the response is related to $T_{RCI}(s)$. The value of s that makes the denominator zero,

$$s \cdot \tau + 1 = 0$$

is $s = p = -1/\tau$, the root of the polynomial in the denominator. It is called a *pole* of $T_{RCI}(s)$. (Think of a tent. The tent height "goes to infinity" at the poles as does $T_{RCI}(s)$.) The unit of p is s^{-1} which is frequency. In this case, p is a real number, not imaginary, and the response is of the form $e^{-\alpha \cdot t}$ where $\alpha = 1/\tau$. The meaning of this will become clearer after another example of a circuit in the *s*-domain.

The RC Differentiator

A circuit similar to the RC integrator is the *RC differentiator*, shown below.



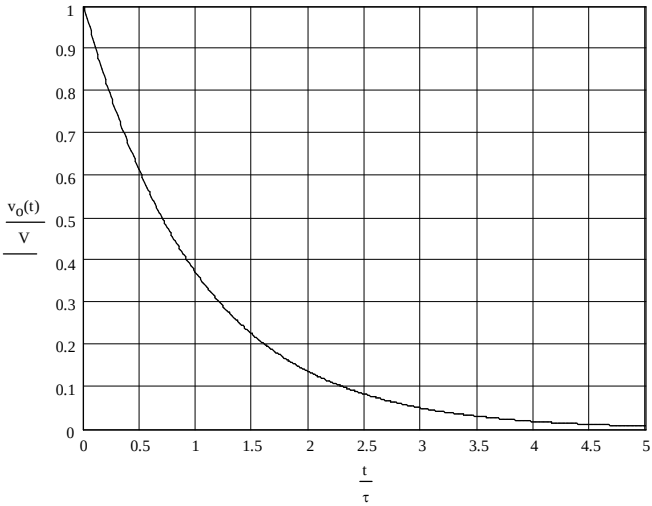
The transfer function is found by ordinary circuit analysis, except that impedances are in the *s*-domain;

$$T_{RCD}(s) = \frac{v_o(s)}{v_i(s)} = \frac{R}{R + 1/s \cdot C} = \frac{s \cdot R \cdot C}{s \cdot R \cdot C + 1}$$

The *s*-domain response differs from the RC integrator in the numerator. Its root is $s = z = 0$ and it makes $T_{RCD} = 0$. Thus z is called a *zero*. RC differentiator response to a unit step is shown below.
The response, like that of the RC integrator, is exponential;

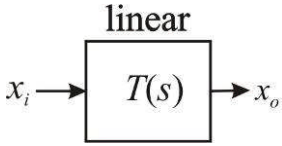
$$v_o(t) = V \cdot e^{-t/\tau}$$

and as before, $\tau = R \cdot C$. The pole of T_{RCD} is the same as that of T_{RDI}. What is different is the appearance of a zero. The *s* multiplied by the rest of the transfer function reminds us of *s* as a differentiation operator. The reciprocal of *s*, or $1/s$ is an integration operator, as can be seen in the expression for capacitive reactance, $1/s \cdot C$, wherein current is integrated to result in voltage.



Transfer Functions in *s*

From the previous two circuits in the *s*-domain it is apparent that complex exponential response in time is somehow related to the transfer function in *s*. The *transfer function*, *T*(*s*), of a circuit is the linear multiplier to the input quantity that gives the output quantity. *Transfer function* is synonymous with gain, attenuation, or *transmittance* - different words for the same two-port concept.



Transfer functions in the *s*-domain are in the form of a fraction of polynomials which is a *rational function*:

$$T(s) = \frac{x_o(s)}{x_i(s)} = \frac{N(s)}{D(s)}$$

where x_i and x_o are of the form $X \cdot e^{s \cdot t}$. *N* and *D* are polynomials in *s* of the form

$$a_n \cdot s^n x(s) + a_{n-1} \cdot s^{n-1} x(s) + \cdots + a_0$$

By substituting for x_i and x_o into the general transfer function,

$$T(s) = \frac{x_o}{x_i} = \frac{X_o(s) \cdot e^{s \cdot t}}{X_i(s) \cdot e^{s t}} = \frac{X_o(s)}{X_i(s)}$$

The exponential time functions cancel, leaving the transfer function in *s* only. This applies only when the circuit is linear.

T(*s*) can be rewritten as

$$D(s) \cdot x_o(s) = N(s) \cdot x_i(s)$$

From this we can see that *N* modifies the input, x_i, and characterizes the driven or *steady-state response*, $x_{ss}(s)$, while *D* characterizes the *natural* or *transient response*, $x_{tr}(s)$. The two responses add to give the total response:

$$x_o(s) = x_{ss}(s) + x_{tr}(s)$$

The steady-state response is caused by x_i which can last indefinitely. In limit notation, it can be expressed in the time domain as

$$x_{ss}(t) = \lim_{t \to \infty} x_o(t)$$

The transient response decays away with time;

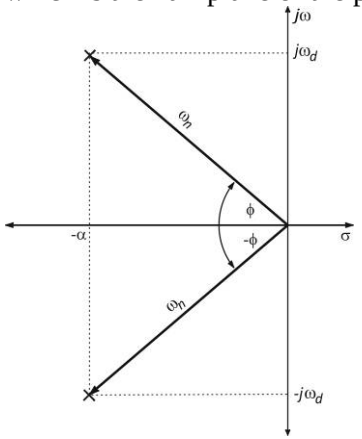
$$\lim_{t \to \infty} x_{tr}(t) = 0$$

As x_{tr} decays away, only the driven steady-state response from $x_i(t)$ is left. The transient response is also called the *natural* response because it is caused by initial non-zero voltages or currents of reactances in the circuit and is the response of the circuit itself without an input, x_i. The poles of the transfer function, in *D*(*s*), characterize the transient response while the zeros affect the steady-state response.

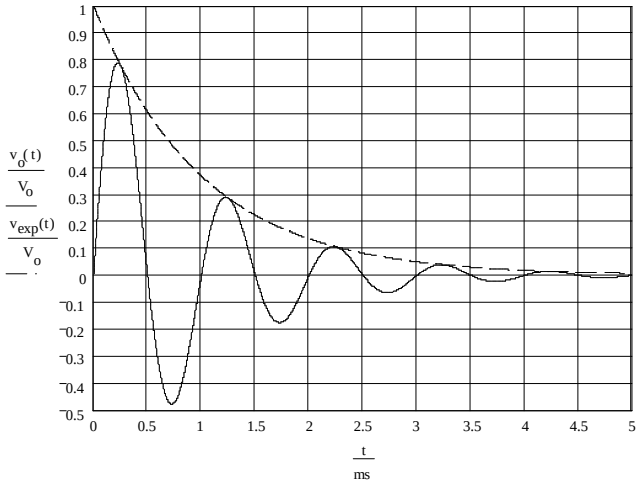
The real part of poles and zeros as complex frequencies produces the exponential responses of the RC integrator and differentiator which is the transient response. The imaginary component of poles and zeros produces the steady-state sinusoidal response. Together they constitute the total response. A complex pole will produce a sinusoidal response with exponentially decaying or growing amplitude. Poles or zeros that are complex numbers always appear in the transfer function in complex conjugate pairs. A pole or zero that is complex can be expressed as a pair at

$$s = -\alpha \pm j \cdot \omega_d$$

where $-\alpha$ is the real frequency corresponding to the exponential factor in the time response, and ω_d is the imaginary component of the sinusoidal factor. The negative sign before α lets the frequency, α, be positive in the left half of the *s*-plane (or the *left half-plane*, LHP), which is the half-plane of the poles of practical circuits.



A complex pole-pair (poles are shown by × and zeros by circles) is shown in the left half-plane of the complex-frequency plane. Each pole, shown as a vector, has a magnitude of ω_n, a real frequency of $-\alpha$ and an imaginary frequency of $\pm j \cdot \omega_d$. A negative imaginary frequency, $-j \cdot \omega_d$, inverts a sinusoid in phase (by π, or 180°). The *pole angle* is ϕ. As ϕ increases, the imaginary component increases over the real component and the sinusoid in the time response becomes greater. For $\phi = \pi/2$ (90°), the response is a steady-state (constant amplitude) sine-wave. For $\phi = 0$, there is no sinusoid in the response (no "ringing") and it is completely exponential. The following plot is of a complex pole-pair with $f_d = 1$ kHz and $\alpha = 1/(2 \cdot \tau) = 500$ Hz. The exponential decay of the amplitude is the dashed curve called the *envelope* of the waveform.



The N and D polynomials can be factored so that the roots are apparent in each factor.

$$T(s) = \frac{x_o(s)}{x_i(s)} = \frac{(s - z_1) \cdot (s - z_2) \cdots}{(s - p_1) \cdot (s - p_2) \cdots}$$

When the polynomials are *normalized* (so that $a_0 = 1$, by dividing through N and D by their a_0) then the constant factor is the gain. The transfer function is then in its generally more useful *normalized* form:

$$T(s) = T_0 \cdot \frac{\left(\frac{s}{z_0}\right) \cdots \left(\frac{s}{z_1} + 1\right) \cdot \left(\frac{s}{z_2} + 1\right) \cdots}{\left(\frac{s}{p_0}\right) \cdots \left(\frac{s}{p_1} + 1\right) \cdot \left(\frac{s}{p_2} + 1\right) \cdots}$$

where $-p_i$ are the pole frequencies and $-z_i$ the zero frequencies. Because useful poles are negative, the custom is to talk of p or z as the pole or zero frequency, though they are located at $-p$ or $-z$ in the s-domain. The zeros and poles at the origin are included in $T(s)$, and there can be multiple (s/ω) factors of each. Wherever s appears, it is the numerator of a ratio of frequencies of a zero or pole. As $t \to \infty$, $s \to 0$. This is expected because s is a frequency, and as $0\ s^{-1}$ is approached, more and more time goes by per cycle of sinusoid, approaching a constant value in the limit. When s is set to zero, the normalized $T(s) = T_0$, the quasistatic (frequency-independent) gain.

In normalized form, the transfer functions of the RC integrator and differentiator are

$$T_{RCI}(s) = \frac{1}{s/\omega_p + 1}\ ,\ \omega_p = 1/R{\cdot}C$$

$$T_{RCD}(s) = \frac{s/\omega_p}{s/\omega_p + 1}\ ,\ \omega_p = 1/R{\cdot}C$$

The differentiator has a zero at the origin ($0\ s^{-1}$) and a pole at $-\omega_p$. The significance of ω_p for the zero is that it is the frequency at which its gain magnitude is 1. For both circuits, the pole magnitude, $p = \omega_p$, is at the same frequency.

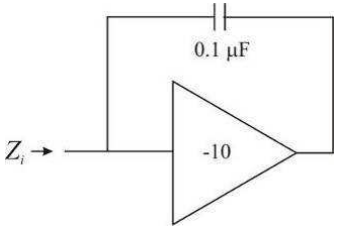
In much of the engineering literature, to distinguish the function in the *s*-domain from the corresponding function in the time domain, it is written with a capital letter:

$$x(t) \rightarrow X(s)$$

However, we already have a convention where upper- and lower-case letters signify static and incremental quantities. Consequently, we will retain the same function symbol and recognize which function it is by which domain (*t* or *s*) symbol appears in it.

Capacitance Multiplier

Before proceeding further into circuit dynamics, we take in another circuit in the *s*-domain, one that uses the Miller effect and is a *capacitance multiplier*.



Using Miller's Theorem in the *s*-domain,

$$Z_i = \frac{Z_f}{K+1} = \frac{1/s \cdot C}{10+1} = \frac{1}{s \cdot (11 \cdot C)}$$

The effective *C* is eleven times larger so that at the input node,

$$C_i = 11 \cdot C = 1.1 \text{ µF}$$

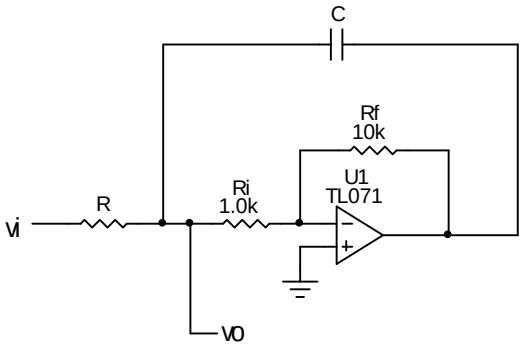
The C multiplier circuit is useful for designing into a circuit a capacitance larger than is practical from a capacitor. Some capacitors become very large or expensive above certain values. Not only is the C value amplified by the C multiplier circuit, the TC of the capacitor is also retained in the behavior.

The usefulness of this kind of circuit is not limited to multiplication of capacitance because the Miller effect on Z_f also applies to the other basic elements as the following table shows.

Z_f	Z_i
R	$R/(K+1)$
L	$L/(K+1)$
C	$(K+1)\cdot C$

Lab Experiment: RC Integrator with C Multiplier

The C multiplier is combined in this experiment with the RC integrator to result in an RC integrator with a lower pole frequency without using a larger capacitor.



The design replaces C in the RC integrator with the capacitance multiplier circuit. We can do this because the negative terminal of the C-multiplier capacitance is grounded, the same as the RC integrator. Select the values for R and C_i (and C) based on a specified time constant for the step response.

The TL071 op-amp is the $-K$ amplifier, with noninverting (+) input grounded and ±12 V supplies. Choose the closed-loop gain to be K of the C multiplier. Place C around the inverting op-amp as shown for $K = 10$. Drive the input of the circuit with a function generator set to square-wave output. Set the high and low levels on the square-wave to be about ±1 V. Square-waves approximate step inputs at a frequency low enough so that the transient response decays away before the next transition of the square-wave. Calculate the time constant from waveform exponential decay.

Frequency Response of Magnitude

The magnitude and phase of $T(s)$ over steady-state frequency, $s = j\cdot\omega$, is called the *frequency response.* It is the steady-state

response of the circuit to an input sinusoid in the form of an imaginary exponential. To find the frequency response of $T(s)$, substitute $s = j \cdot \omega$ and separate the magnitude and phase factors in the resulting $T(j \cdot \omega)$.

To show this by example, the RC-integrator response is

$$T_{RCI}(j \cdot \omega) = \frac{1}{\frac{j \cdot \omega}{\omega_p} + 1}, \ \omega_p = 1/R \cdot C$$

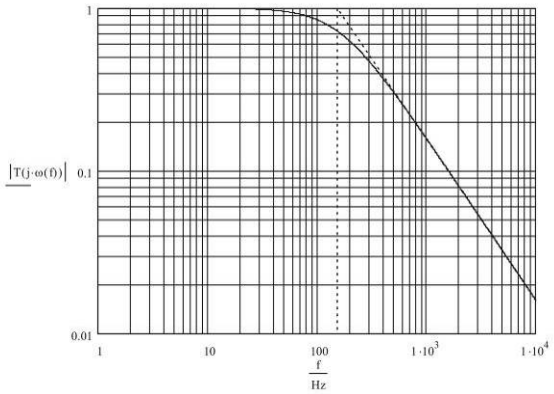
The formulas in the "Sine-Waves" section apply. The magnitude of the frequency response is

$$\| T_{RCI}(j \cdot \omega) \| = \frac{1}{\sqrt{\left(\frac{\omega}{\omega_p}\right)^2 + 1}}$$

The phase is

$$\angle T_{RCI}(j \cdot \omega) = -\tan^{-1}\left(\frac{\omega}{\omega_p}\right)$$

Magnitude and phase, when plotted, are called *frequency-response plots* or *Bode plots*. The RC-integrator frequency-response magnitude plot is shown below for $R = 1$ kΩ and $C = 1$ μF.



The time constant, $\tau = R \cdot C = (1\ \text{k}\Omega) \cdot (1\ \mu\text{F}) = 1$ ms; $\omega_p = 1/\tau = 1\ \text{ks}^{-1}$. The horizontal axis on frequency-response plots is the magnitude or absolute value of pole or zero frequency. On the plot, radian

frequency is not used on the frequency axis but the per-revolution (or *per-cycle*) frequency with units of Hz.

To keep the two angular scales of phase separate, different variable symbols are used for the two frequencies. They are related by

$$\omega = 2 \cdot \pi \cdot f$$

where f has scaling in hertz (Hz) and ω in s^{-1}. This departs from the usual scientific and engineering practice of representing quantities with symbols without regard to their units. Units are instead included in the values of the quantity:

$$\text{value} = \text{number} \cdot \text{unit (or scale)}$$

Because both frequency scales appear so often, different symbols for them are used for convenience; ω is frequency in s^{-1} and f is in Hz. The actual unit (as found in MathCAD) for Hz is s^{-1} because the $1/2 \cdot \pi$ factor of Hz is unitless. It is, however, a scaling factor for the number of the frequency. When frequencies appear in the form of ratios, as in normalized transfer functions, either ω or f can be used when each of the values in the ratio have the same scaling. For instance, (s/100 Hz) requires that s also be in Hz.

The above magnitude plot is of a single pole. It has a *break frequency* or *corner frequency* at the pole frequency,

$$\omega_p = p = 1 \text{ ks}^{-1} \approx 159 \text{ Hz}$$

where the magnitude of $T(j \cdot \omega)$ begins to decrease toward an eventual slope of –1 on the log-log plot. Far below ω_p, the plot is flat at a value of one. An asymptote can be drawn for the descending part of the graph above ω_p, shown as the dotted line. It intersects 1 at the pole frequency. This *asymptotic approximation* for the magnitude response of a real pole (on the σ-axis in the s-plane) is simply a flat response to ω_p, then a line on the log-log plot descending with a slope of –1.

To derive the –1 slope, take the (decimal or common) logarithm;

$$\log \| T_{RCI}(j \cdot \omega) \| = \log \frac{1}{\sqrt{\left(\frac{\omega}{\omega_p} \right)^2 + 1}} = -\tfrac{1}{2} \cdot \log \left(\left(\frac{\omega}{\omega_p} \right)^2 + 1 \right)$$

For $\omega >> \omega_p$, the plot approaches the asymptotic line and $\omega/\omega_p >> 1$. Therefore, we can drop the 1 term and the approximation becomes

$$\log \| T_{RCI}(j\cdot\omega) \| \approx -\tfrac{1}{2}\cdot\log\left(\left(\frac{\omega}{\omega_p}\right)^2\right) = -\tfrac{1}{2}\cdot 2\cdot\log(\omega/\omega_p)$$
$$= (-1)\cdot\log(\omega/\omega_p)$$

Then

$$\frac{\log \| T_{RCI}(j\cdot\omega) \|}{\log(\omega/\omega_p)} = -1$$

A similar approximation can be made for $\omega << \omega_p$. The asymptote is a horizontal line having zero slope. Where the two asymptotic lines intersect is the pole frequency.

At the pole frequency, the magnitude has decreased to

$$\| T_{RCI}(j\cdot\omega_p) \| = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2} \approx 0.7071$$

The frequency at which the magnitude has decreased to this value, for a single-pole circuit or any other, is called the *bandwidth*, f_{bw}, and is a measure of the quickness of response or "speed" of a circuit.

Step Response

In the time domain, a corresponding measure of circuit quickness to bandwidth is the *risetime* of its response to a unit step. By definition, risetime, t_r, is the time that the step response takes to go from 0.1 of its step height (or "amplitude", to generalize from sine-waves) to 0.9 of it. For the RC integrator, the response is a real exponential and the risetime for it can be found by solving for the times at 10 % and 90 %.

$$v_o(t) = V\cdot(1-e^{-t/\tau}) \Rightarrow t = -\tau\cdot\ln\left(1-\frac{v_o(t)}{V}\right)$$

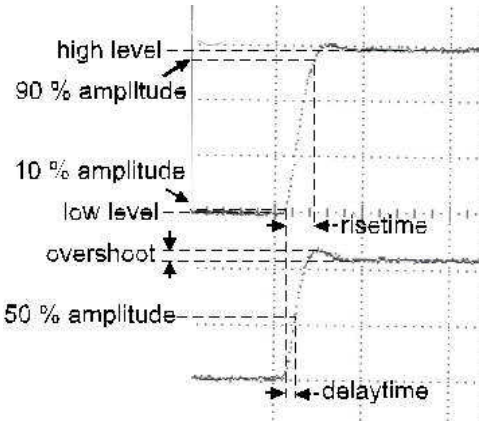
$$t_r = t_f - t_i = -\tau\cdot\ln\left(\frac{1-\frac{0.9\cdot V}{V}}{1-\frac{0.1\cdot V}{V}}\right) = -\tau\cdot\ln(0.1/0.9) = \tau\cdot\ln(9) \approx 2.2\cdot\tau$$

The risetime of the RC integrator example is about (2.2)·(1 ms) or about 2.2 ms.

The risetime for a circuit with (or approximated by) a single-pole bandwidth of $\omega_{bw} = 1/\tau_{bw}$ is

$$t_r \approx 2.2 \cdot \tau_{bw} = \frac{2.2}{\omega_{bw}} = \frac{2.2}{2 \cdot \pi \cdot f_{bw}} \approx \frac{0.35}{f_{bw}}$$

Another commonly-used measure of circuit speed from step or *pulse* response is *delay time*. It is shown along with risetime in the following oscilloscope screen picture (*oscillograph*).



The *delay time* is defined as the time the pulse response takes to reach 50 % of the total rise of the pulse. For the RC integrator (or any single-pole circuit), it is

$$t_d = t_f - t_i = -\tau \cdot \ln\left(\frac{1 - \frac{0.5 \cdot V}{V}}{1 - \frac{0.0 \cdot V}{V}}\right) = -\tau \cdot \ln(0.5) = \tau \cdot \ln(2) \approx 0.693 \cdot \tau \approx 0.7 \cdot \tau$$

The delay time of the RC integrator example is thus about 700 μs. Delay time is quite useful in high-speed digital logic applications because it is a measure of the propagation delay through gates and other logic circuits.

The step in the oscillograph exceeds (or *overshoots*) the final level before returning to it. This overshoot is an exponentially-decreasing, or *damped*, sinusoid, and the circuit has a complex pole-pair. In more advanced treatments of this subject, the peak value of the overshoot can be derived if the pole angle, ϕ, is known from the circuit transfer function. The fractional peak overshoot is

$$M_p = e^{-\pi/\tan\phi}$$

Some common values are listed in the following table.

ϕ, deg	M_p, %
0	0
30	0.433
45	4.321
60	16.30

The pole locations in s can be approximated by measuring the step overshoot and using the table or equation in reverse.

Two equal poles of time constant τ have a response to a voltage step of V of

$$v_{step}(t) = V \cdot (1 - (1 + \frac{t}{\tau})) \cdot e^{-t/\tau}$$

The risetime, t_r, is defined as the time interval between the $0.1 \cdot V$ and $0.9 \cdot V$ points of the response;

$$\frac{t_r}{\tau} = \frac{t(0.1 \cdot V) - t(0.9 \cdot V)}{\tau} = \ln\left(\frac{0.9}{1 + \frac{t(0.1 \cdot V)}{\tau}}\right) - \ln\left(\frac{0.1}{1 + \frac{t(0.9 \cdot V)}{\tau}}\right)$$

This iterates to a solution of $t_r \approx (3.8897 - 0.53183) \cdot \tau \approx 3.3579 \cdot \tau$.

A single pole has a risetime of $t_r \approx 2.1972 \cdot \tau \approx 2.2 \cdot \tau$. Then t_r for two poles is about 53 % longer than for a single pole. The approximation of the risetime for two poles is

$$t_r \approx \sqrt{(2.2 \cdot \tau)^2 + (2.2 \cdot \tau)^2} = \sqrt{2} \cdot (2.2 \cdot \tau) \approx 1.414 \cdot (2.2 \cdot \tau) \approx 3.107 \cdot \tau$$

This value is low by about 7.5 %. A better approximation formula for risetime adjusts for this difference;

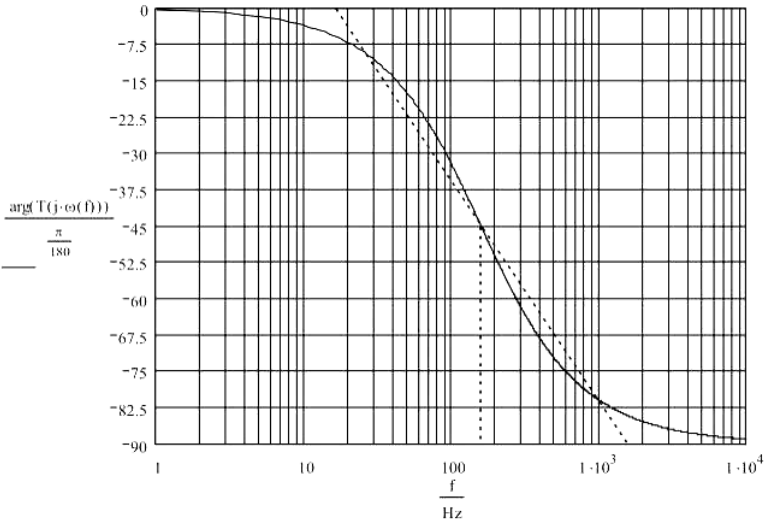
$$t_r \approx 1.075 \cdot \sqrt{\sum_i {t_{ri}}^2} \approx 1.1 \cdot \sqrt{\sum_i {t_{ri}}^2}$$

Frequency Response of Phase

The magnitude plot of $T(s)$ is only half the frequency-response information. The other plot is that of the phase. For the RC integrator,

$$\angle T_{RCI}(j \cdot \omega) = -\tan^{-1}\left(\frac{\omega}{\omega_p}\right)$$

where for our particular circuit, ω_p = 1 ks^{-1} ≈ 159 Hz. Phase is plotted on the semi-log plot below.



The vertical (phase) axis is linear and the frequency axis is logarithmic. The phase response for a single pole is 0° at low frequency. Then at about a decade from the pole, it begins to decrease more. At the pole it is at –45°. And it settles to –90° about a decade later. If a piecewise-linear approximation is made (dotted line on graph), the maximum error from the actual curve is about 6° - acceptable for most circuit design. Thus, a single pole has the effect of reducing the phase by 90° over two decades, one on each side of the pole frequency, and with a slope of 45°/dec.

RC Differentiator Frequency Response

The RC differentiator transfer function for steady-state response is

$$T_{RCD}(j\cdot\omega) = \frac{j\cdot\omega\cdot R\cdot C}{j\cdot\omega\cdot R\cdot C+1} = \frac{j\cdot\omega/\omega_p}{j\cdot\omega/\omega_p+1}\ ,\ \omega_p = 1/R\cdot C$$

To put a complex rational function into rectangular form, multiply numerator and denominator by the conjugate of the denominator:

$$T_{RCD}(j\cdot\omega)=\frac{j\cdot\omega/\omega_p}{j\cdot\omega/\omega_p+1}\cdot\frac{-j\cdot\omega/\omega_p+1}{-j\cdot\omega/\omega_p+1}=\frac{\left(\frac{\omega}{\omega_p}\right)^2+j\cdot\left(\frac{\omega}{\omega_p}\right)}{\left(\frac{\omega}{\omega_p}\right)^2+1}$$

The real and imaginary components are now separated and the magnitude is

$$\|\,T_{RCD}(j\cdot\omega)\,\|=\frac{\sqrt{\left[\left(\frac{\omega}{\omega_p}\right)^2\right]^2+\left(\frac{\omega}{\omega_p}\right)^2}}{\left(\frac{\omega}{\omega_p}\right)^2+1}=\frac{\left|\left(\frac{\omega}{\omega_p}\right)\right|\cdot\sqrt{\left(\frac{\omega}{\omega_p}\right)^2+1}}{\left(\frac{\omega}{\omega_p}\right)^2+1}$$

This then becomes

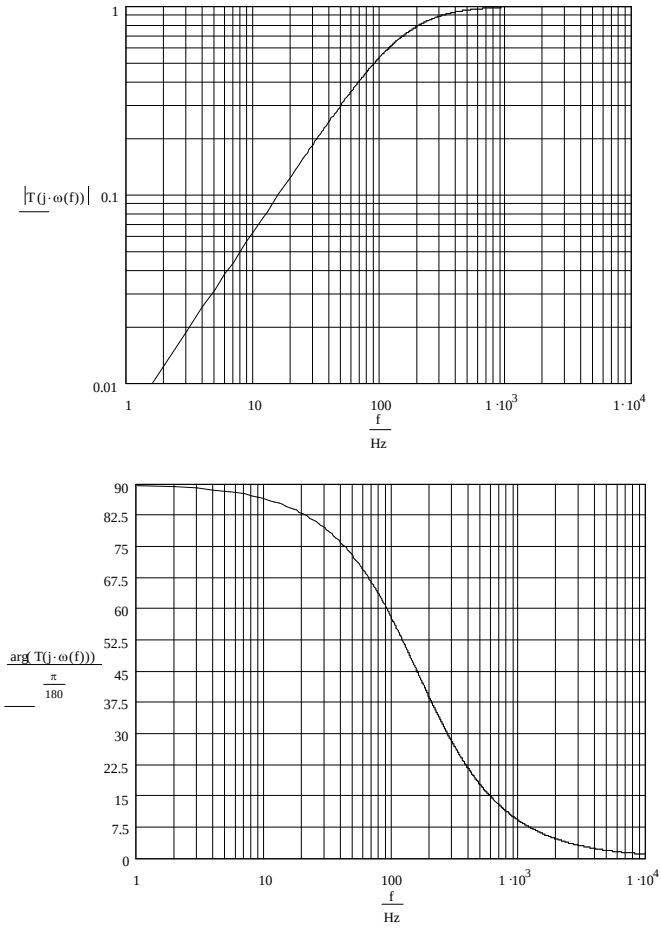
$$\|\,T_{RCD}(j\cdot\omega)\,\|=\frac{\left|\left(\frac{\omega}{\omega_p}\right)\right|}{\sqrt{\left(\frac{\omega}{\omega_p}\right)^2+1}}$$

The phase is

$$\angle T_{RCD}(j\cdot\omega)=\tan^{-1}\left(\frac{\omega/\omega_p}{(\omega/\omega_p)^2}\right)=\tan^{-1}(\omega_p/\omega)$$

On the frequency-response plots, the zero at the origin causes the magnitude to increase from the origin until ω_p where it becomes flat at one. The phase for $\omega << \omega_p$ is +90° (Atan(∞) = 90°). It then decreases over two decades to zero for $\omega >> \omega_p$ (Atan(0) = 0°), crossing through +45° at ω_p.

The general rules for Bode-plot construction are that poles cause a –1 slope per pole in magnitude and a change in phase of –90° over two decades around the pole frequency. Poles at the origin have a magnitude decrease from the origin with a –1 slope and have a constant –90° (–π/2) phase shift. Negative or LHP zeros cause a +1 slope per zero and a +90° change in phase over two decades around the zero frequency.

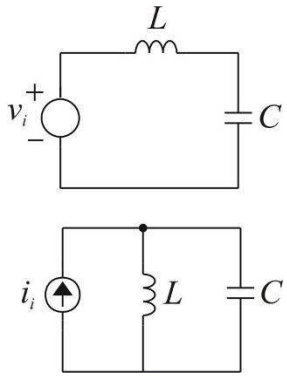


Zeros at the origin increase from the origin with a +1 slope and add a constant +90° ($\pi/2$) phase shift. The contributions of the poles and zeros are added on Bode plots for the combined response. By using asymptotic approximations, the piecewise-linear plots can be quickly sketched on graph paper.

Poles and zeros with positive real components are in the right half-plane (RHP) of *s* and have different, generally undesirable behavior. However, zeros in the RHP are often found in phase-shift networks in filters, oscillators, and impedance meters. RHP poles correspond to growing exponentials with time, are inherently unstable in actual circuits, and are to be avoided in design.

Resonance

A circuit with capacitive and inductive reactances *resonates* or oscillates at the damped frequency, ω_d, of the complex pole. If the circuit lacks resistance, the conjugate poles are *undamped* by resistance and the response is a steady-state sine-wave. Circuits that can resonate this way are shown below.



What is the impedance across the port at the source? The two reactances of the upper circuit are in series, forming a *series resonance*;

$$Z_s = s \cdot L + \frac{1}{s \cdot C} = \frac{s^2 \cdot L \cdot C + 1}{s \cdot C}$$

The denominator has a pole at the origin. The quadratic numerator has zeros at

$$s = \pm j \cdot \frac{1}{\sqrt{L \cdot C}} = \pm j \cdot \omega_n \Rightarrow \omega_n = \frac{1}{\sqrt{L \cdot C}}$$

where ω_n is the undamped *natural* or *resonant frequency*. The transfer function is not yet in the preferred form, with ratios of frequencies wherever *s* occurs, but it can be manipulated into that form as

$$Z_s = \frac{s^2 \cdot L \cdot C + 1}{s \cdot C} = \frac{\left(\frac{s}{\omega_n}\right)^2 + 1}{s \cdot \sqrt{\frac{C}{L}} \cdot (L \cdot C)} = Z_n \cdot \frac{\left(\frac{s}{\omega_n}\right)^2 + 1}{\left(\frac{s}{\omega_n}\right)}$$

where Z_n is the *resonant impedance* and is

$$Z_n = \sqrt{\frac{L}{C}}$$

Resonances are characterized by ω_n and Z_n, and can be located as a resonant point on a *reactance plot*, a plot of $\log||Z(\omega)||$ versus $\log \omega$. Substituting $s = j \cdot \omega$ to find the steady-state response,

$$Z_s(j \cdot \omega) = Z_n \cdot \frac{-\left(\frac{\omega}{\omega_n}\right)^2 + 1}{j \cdot \left(\frac{\omega}{\omega_n}\right)} = j \cdot Z_n \cdot \frac{\left(\frac{\omega}{\omega_n}\right)^2 - 1}{\left(\frac{\omega}{\omega_n}\right)}$$

The impedance of the series resonance at 0 Hz is infinite and at $\omega \to \infty$ is infinite - an open circuit on both sides of resonance. At the resonant frequency, $Z_s(j \cdot \omega_n) = 0\ \Omega$, a short circuit.

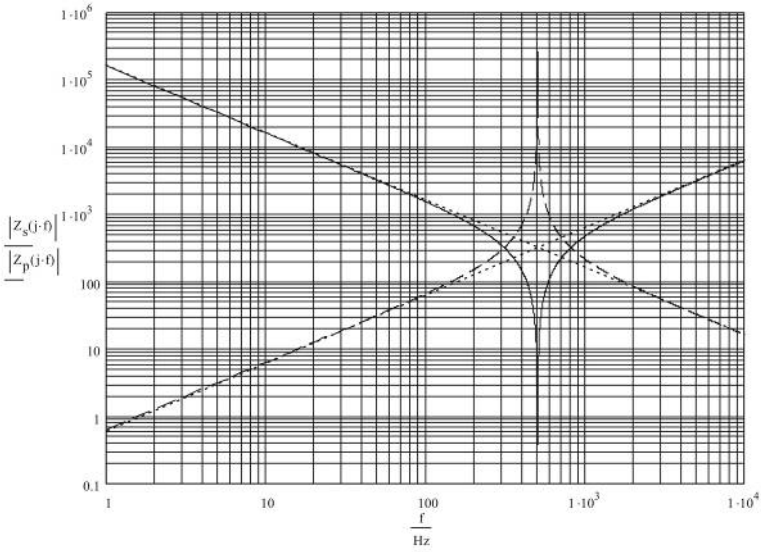
The parallel resonant circuit is the dual of the series resonant circuit and has an input impedance of

$$Z_p(s) = \frac{(s \cdot L) \cdot (1/s \cdot C)}{(s \cdot L) + (1/s \cdot C)} = \frac{s \cdot L}{s^2 \cdot L \cdot C + 1} = Z_n \cdot \frac{\left(\frac{s}{\omega_n}\right)}{\left(\frac{s}{\omega_n}\right)^2 + 1}$$

The parallel resonance is the series resonance with the rational factor inverted. Then for $s = j \cdot \omega$,

$$Z_p(j \cdot \omega) = Z_n \cdot \frac{\left(\frac{j \cdot \omega}{\omega_n}\right)}{\left(\frac{j \cdot \omega}{\omega_n}\right)^2 + 1} = j \cdot Z_n \cdot \frac{\left(\frac{\omega}{\omega_n}\right)}{1 - \left(\frac{\omega}{\omega_n}\right)^2}$$

Asymptotic approximations of series and parallel resonances are shown (dotted) with exact (solid) plots for $L = 100$ mH and $C = 1$ μF. Frequency on the horizontal axis is in Hz.



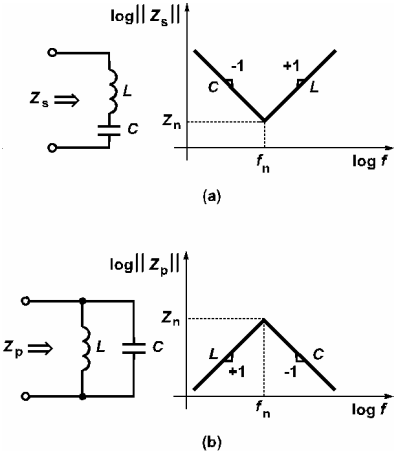
The resonant parameters are f_n = 503.3 Hz and Z_n = 316.2 Ω. The two dotted lines are the asymptotes with slopes of ±1. They intersect at the *resonant point* of (ω_n, Z_n). The solid plot is Z_s, pointing downward, and the dashed plot is Z_p, pointing upward at resonance. On a log-log plot, zero for Z_s at resonance goes to $-\infty$ while Z_p at resonance goes to $+\infty$. The asymptotic approximation for Z_s is V-shaped and for Z_p it is an inverted V. The asymptotic plots intersect at the resonant point.

At resonance, the reactance magnitudes are equal and are

$$X_L = \omega_n \cdot L = \frac{L}{\sqrt{L \cdot C}} = \sqrt{\frac{L}{C}} = Z_n$$

$$-X_C = \frac{1}{\omega_n \cdot C} = \frac{\sqrt{LC}}{C} = \sqrt{\frac{L}{C}} = Z_n$$

The phase of the two reactances at resonance is not equal but is inverted, and for Z_s they subtract resulting in a short circuit. For Z_p, they also cancel each other resulting in an open circuit. Asymptotic magnitude plots of resonances are shown below for series (a) and parallel (b) resonances.



Resonances are damped by resistance in series with Z_s or in parallel with Z_p. By including R, the response is no longer a constant-amplitude sine-wave but is damped, causing it to decay. R causes the imaginary pole-pair of the undamped resonance to move off the $j\omega$-axis and into the LHP so that $\alpha > 0$. The pole angle is no longer 90° as it is for an undamped resonance, and $\phi < 90°$.

Damping causes the undamped imaginary pole- (or zero-) pair to be a complex conjugate pair that is found by solving a quadratic factor in the transfer function of the form

$$\left(\frac{s}{\omega_n}\right)^2 + 2\cdot\zeta\cdot\left(\frac{s}{\omega_n}\right) + 1$$

ζ is the *damping* and has a simple relationship to the pole angle, ϕ;

$$\zeta = \cos\phi$$

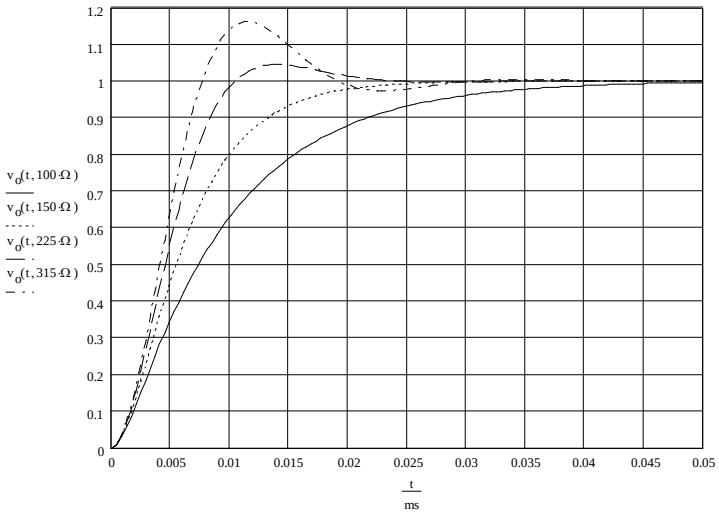
Then

$$\alpha = \omega_n \cdot \cos\phi = \zeta \cdot \omega_n$$

As ζ is made larger, the real frequency α becomes larger relative to ω_d, the imaginary *damped frequency*, which is somewhat less than the undamped resonant frequency, ω_n. The relationship between them is

$$\omega_d = \omega_n \cdot \sqrt{1-\zeta^2} = \omega_n \cdot \sin\phi$$

Some values of ζ and ϕ, and the corresponding kinds of step response, are given in the table below. The accompanying plots closely approximate these responses and are keyed in the table.



ϕ, deg	ζ	Plot Key	Response Type
0	1.58	solid	overdamped ($\zeta > 1$)
0	1	dotted	critically damped ($\zeta = 1$)
30	0.866	(not shown)	underdamped, MFED ($\zeta < 1$)
45	0.707	dashed	underdamped, MFA ($\zeta < 1$)
60	0.500	dot-dash	underdamped ($\zeta < 1$)
90	0	(not shown)	undamped ($\zeta = 0$)

At $\phi = 30°$, the *group* or *envelope delay*, $-d\phi/d\omega$, is "maximally flat". This is negative the slope of the phase plot. If the slope remains constant the group delay is "flat" and the phase decreases linearly with frequency. The waveform is distorted least because all frequency components of the waveform are delayed equally.

At 45°, the MFA or "maximally-flat amplitude" response has the "flattest" (constant) frequency (not step) response without any resonant peaking. MFED has very little overshoot in the step response (0.43 %) and is considered optimal for many amplifier designs requiring high accuracy or low distortion over their bandwidth.

The overdamped response is undesirable if amplifier speed has any importance. The critically-damped response is at the boundary of overshoot and has two equal poles on the real axis. Both the MFED and MFA responses have less risetime and are quicker than the critically damped response. By tolerating a small amount of overshoot, risetime and delay time can be significantly improved.

For some applications, pulse waveshape accuracy or a flat frequency response are not important and a larger pole angle is tolerable. The final response, with $\phi = 60°$, clearly shows "ringing" - the sine component of the response. Its risetime is lowest of all though it takes the most time to settle to the final step value.

Amplifier dynamic design is a tradeoff between time-domain overshoot or frequency-domain peaking, risetime, delay time, and distortion. Oscilloscope vertical amplifiers - the amplifiers that drive the vertical axis of the graph displayed by the 'scope - are one of the most critical applications involving such tradeoffs. Vertical amplifiers are usually designed to have a pole angle around 30° for the best compromise between an accurate pulse waveshape and a flat frequency response.

The simple relationship between damping and resonant impedance can be used to select a value of resistance to damp a resonance:

$$\zeta_s = \frac{R}{2 \cdot Z_n} \ , \ \zeta_p = \frac{Z_n / 2}{R}$$

For critical damping $\zeta = 1$ and R equals the combined magnitude values of the reactances. At resonance, $-X_C = X_L = Z_n$. For series resonance, $2 \cdot Z_n$ are in series with R and for critical damping, $R_s = 2 \cdot Z_n$. For parallel resonance, the elements are in parallel and have a combined $Z_n/2$ that must equal R_p for critical damping.

A final passing note about terminology: ζ in control theory is called the *damping ratio* and α is the *damping factor*. Our interest in quadratic response frequently encounters ζ in relation with damping. It is the central damping parameter from which α is easily derived. To simplify the repeated referrals to ζ, I have chosen to simply call it *damping*.

Transfer Function Poles from Circuits

To extend circuit analysis and design to include reactive elements, the main challenge is to find the transfer function and input and output port impedances, the three central design parameters of amplifiers. They can be analyzed by writing basic circuit equations and solving them, though for circuits of any complexity, the algebra quickly becomes unwieldy. We are faced with two difficult problems:

1. Find the coefficients of $A(s)$ expressed in circuit elements;
2. Factor the pole and zero polynomials.

Solving for the polynomial roots requires factorization. The usual simplification of this problem is to truncate the polynomials after the quadratic (s^2) term on the basis that the higher-degree coefficients will be small and have a small effect on circuit behavior. This is often valid because for good circuit design, stages are isolated from each other by buffering and reactances are separated from each other. This approximation by either a linear (s) or quadratic (s^2) polynomial effectively solves the factoring problem.

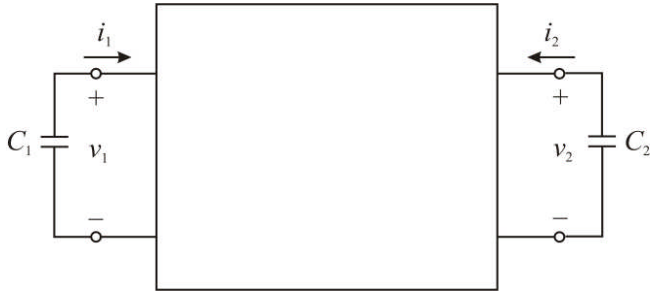
For the first problem, polynomial coefficients take on a general form that can be related to circuit elements. Whenever the circuit equations - using the basic laws (ΩL, KVL, KCL) - are written, each capacitance adds $1/s{\cdot}C$ to the equations and each inductance, $s{\cdot}L$. For n (irreducible) reactances, somewhere in the algebra is an n-degree polynomial, and it will be $D(s)$ and possibly $N(s)$. (Reactances in series or parallel reduce to a single equivalent.) The term of highest degree, $a_n{\cdot}s^n$, will have n reactances (C or L) as products in a_n. As a frequency, s has units of inverse seconds, s^{-1}, and each term (such as $a_0 = 1$) is unitless. Thus a_m, the coefficient of s^m, must have units of s^m to cancel the s^{-m} unit of s^m.

In particular, a_1 will have terms in it that are time constants, with units of seconds, with one term for each reactance. Each reactance can be viewed as attached to a port into the rest of the circuit. The corresponding circuit resistance of the time constant for a given reactance is the port resistance with the other reactive elements set to zero (Cs open, Ls shorted). We will consider only capacitances. With all capacitors removed (open-circuited ports; capacitances are zero), the open-circuit resistance of port k of capacitance C_k can be found.

To show how this works, consider the following two-port circuit. The ports are not amplifier input and output but are at the terminals of two capacitors. The following development is general but it can pertain to C_c and C_e of the single-stage BJT amplifier. The port equations relating port currents i_1 and i_2 to port voltages v_1 and v_2 require four resistance parameters:

$$v_1 = (R_{11} \| 1/s \cdot C_1) \cdot i_1 + R_{12} \cdot i_2$$

$$v_2 = R_{21} \cdot i_1 + (R_{22} \| 1/s \cdot C_2) \cdot i_2$$



The impedance of the capacitors is $1/s{\cdot}C_k = -v_k/i_k$ where voltage and current are of the same port, indexed by k. The negative sign refers the port impedance to the other side of the port, where C_k is ("looking into C_k") because the direction of current by port convention is reversed by it.

Hence the $1/s{\cdot}C_k$ are paralleled with their open-port resistances, R_{kk}. Solving for port 1, its open-port resistance, R_{11}, is

$$R_{11} = \frac{v_1}{i_1},\ i_2 = 0\ (C_2 = 0),\ C_1 = 0$$

Current of the other port, i_2, is set to zero by opening the port ($C_2 = 0$) and C_1 is removed to isolate R_{11} in the first equation. For port 2,

$$R_{22} = \frac{v_2}{i_2},\ i_1 = 0\ (C_1 = 0),\ C_2 = 0$$

In general, for n C ports, all are open-circuited while R_{kk} is found. Thus we have a method for finding the *open-circuit time constants* (*OCTC*s) of the C_k of a circuit. For Ls, the dual condition applies and inductance ports are shorted instead of opened. For mixed Cs and Ls, open Cs and short Ls.

For a circuit with all real poles, $D(s)$ is the product of the pole factors

$$(s\cdot\tau_1 + 1)\cdot(s\cdot\tau_2 + 1)\cdot\cdots\cdot(s\cdot\tau_n + 1) =$$
$$s^n\cdot(\tau_1\cdot\tau_2\cdot\cdots\cdot\tau_n) + \cdots + s^2(\tau_1\cdot\tau_2 + \tau_1\cdot\tau_3 + \cdots) + s\cdot(\tau_1 + \tau_2 + \cdots + \tau_n) + 1$$

Instead of using pole frequencies, the corresponding time constants, τ_i are used. The terms of the coefficient of s (or a_1) are the OCTCs. After the OCTCs are found, they can be used to find amplifier bandwidth.

Bandwidth from OCTCs of Real Poles

OCTCs are useful for approximating the most important dynamic parameter of an amplifier, its *bandwidth* (or its corresponding time-domain parameter, *risetime*). A circuit with n (irreducible) reactances has a transfer function with n poles. Assume that the poles are real and are $-p_1$, $-p_2$, … Then the normalized (with quasistatic gain of one) transfer function is of the form

$$T(s) = \frac{1}{\left(\frac{s}{p_1}+1\right)\cdot\left(\frac{s}{p_2}+1\right)\cdots\left(\frac{s}{p_n}+1\right)}$$

$$= \frac{1}{\frac{s^n}{p_1 \cdot p_2 \cdot \cdots \cdot p_n} + \cdots + s\cdot\left(\sum_{j=1}^{n}\frac{1}{p_i}\right)+1}$$

Bandwidth, f_{bw}, is defined as

$$\left\|T(j\,\omega_{bw})\right\|^2 = \tfrac{1}{2}$$

where T is (unitless) voltage or current gain. The magnitude of T is

$$\left\|T(j\cdot\omega)\right\| = \left\|\frac{1}{\left(\frac{j\cdot\omega}{p_i}+1\right)\cdot\left(\frac{j\cdot\omega}{p_2}+1\right)\cdots\left(\frac{j\cdot\omega}{p_n}+1\right)}\right\|$$

$$= \frac{1}{\left\|\frac{j\cdot\omega}{p_i}+1\right\|\cdot\left\|\frac{j\cdot\omega}{p_2}+1\right\|\cdot\cdots\cdot\left\|\frac{j\cdot\omega}{p_n}+1\right\|}$$

Each factor in the denominator can be reduced to

$$\left\|\frac{j\cdot\omega}{p}+1\right\| = \sqrt{\left(\frac{\omega}{p}\right)^2+1}$$

Then squaring to get rid of the square-roots,

$$\left\|T(j\cdot\omega)\right\|^2 = \frac{1}{\left(\left(\frac{\omega}{p_1}\right)^2+1\right)\cdot\left(\left(\frac{\omega}{p_2}\right)^2+1\right)\cdots\left(\left(\frac{\omega}{p_n}\right)^2+1\right)}$$

$$= \frac{1}{\left(\frac{\omega^{2\cdot n}}{p_1\cdot p_2\cdot\cdots\cdot p_n}\right)+\cdots+\omega^2\cdot\sum_{k=1}^{n}\frac{1}{p_k^2}+1}$$

A two-pole approximation drops higher-degree terms and bandwidth is approximated by the formula

$$\left\|T(j\cdot\omega_{bw})\right\|^2 \approx \frac{1}{\omega_{bw}^2\cdot\sum_{k=1}^{n}\frac{1}{p_k^2}+1} = \frac{1}{2}$$

Solve by equating denominators. The *OCTC bandwidth* is

$$\omega_{bw} \approx \frac{1}{\sqrt{\sum_{k=1}^{n}\frac{1}{p_k^2}}}$$

Expressed as a time constant,

$$\tau_{bw} = \frac{1}{\omega_{bw}} \approx \sqrt{\sum_{k=1}^{n}\tau_k^2}$$

where τ_k are the OCTCs. The circuit dynamics are reduced to having a single equivalent pole at ω_{bw}. This approximation always results in a bandwidth that is less than the actual bandwidth and is a conservative performance estimate. It is accurate to the extent that the poles are separated in value and the lowest-frequency pole dominates. The worst case is a circuit that has n equal poles that roll off from ω_{bw} with a slope of $-n$. Then the higher-degree coefficients are significant as interactions of the time constants.

For the single-stage BJT amplifier, the goal is to find the OCTCs of C_c and C_e. The capacitance values are given in or calculated from transistor specifications. That leaves the open-circuit resistances at each of the capacitance ports. These can be found by circuit analysis of a generalized single-stage BJT. Once found, the resulting formulas can be used as *templates* for finding the OCTCs of single-BJT amplifier stages generally.

Quadratic Pole-Pair Bandwidth

A different approximation of bandwidth from that of the previous section can be made by including the quadratic (s^2) term and ignoring higher-degree terms in the transfer function polynomial. This approximation differs from the OCTC approximation in that it does not assume real poles, only two poles that need not be real. Consider such an amplifier transfer function with only a pole-pair, in normalized form. Its bandwidth is found from

$$|M(j\cdot\omega_{bw})|^2=\frac{1}{|a\cdot(j\cdot\omega_{bw})^2+b\cdot(j\cdot\omega_{bw})+1|^2}=\frac{1}{2}$$

This reduces algebraically as follows:

$$|(1-a\cdot\omega_{bw}^2)+j\cdot b\cdot\omega_{bw}|^2=2$$

$$(1-a\cdot\omega_{bw}^2)^2+(b\cdot\omega_{bw})^2=2$$

$$1-2\cdot a\cdot\omega_{bw}^2+a^2\cdot\omega_{bw}^4+b^2\cdot\omega_{bw}^2=2$$

This reduces to a quadratic equation in ${\omega_{bw}}^2$:

$$-a^2\cdot\omega_{bw}^4-(b^2-2\cdot a)\cdot\omega_{bw}^2+1=0$$

Solving for ${\omega_{bw}}^2$,

$$\omega_{bw}^2=\frac{2\cdot a-b^2}{2\cdot a^2}\pm\sqrt{\left(\frac{2\cdot a-b^2}{2\cdot a^2}\right)^2+\frac{1}{a^2}}=\left(\frac{1}{a}-2\cdot\left(\frac{b}{2\cdot a}\right)^2\right)\pm\sqrt{\frac{2}{a^2}=\frac{b^2}{a^3}+4\cdot\left(\frac{b}{2\cdot a}\right)^2}$$

It is often more useful to express the pole-pair as in

$$\left|\frac{1}{(s/\omega_n)^2+2\cdot\zeta\cdot(s/\omega_n)+1}\right|^2=\frac{1}{2}$$

where a and b in ${\omega_{bw}}^2$ are expressed in pole magnitude, ω_n, and damping, ζ,

$$a=1/\omega_n^2\ ;\ b=2\cdot\zeta/\omega_n$$

Substituting and simplifying yields the *quadratic bandwidth formula*,

$$\frac{\omega_{bw}}{\omega_n} = \frac{f_{bw}}{f_n} = \sqrt{1 - 2 \cdot \zeta^2 + \sqrt{4 \cdot \zeta^4 - 4 \cdot \zeta^2 + 2}}$$

This is the exact quadratic-pole bandwidth formula as a function of ω_n and ζ. Some values are tabulated below.

ζ	f_{bw}/f_n	ζ	f_{bw}/f_n
5.0	0.1010	0.9	0.7461
2.0	0.2666	0.8	0.8709
1.5	0.3742	$\frac{\sqrt{2}}{2} \approx 0.7071$	1.0000
1.2	0.4994	0.6	1.1482
1.1	0.5628	0.5	1.2720
1.0	0.6436	0.4	1.3745

Approximate Quadratic Bandwidth

An approximation made by Paul E. Gray and Campbell Searle sets the term of higher degree than two equal to zero;

$$1 - 2 \cdot a \cdot \omega_{bw}^2 + b^2 \cdot \omega_{bw}^2 \approx 2$$

Solving for bandwidth, the *approximate quadratic bandwidth formula* results:

$$\omega_{bw} \approx \frac{1}{\sqrt{b^2 - 2 \cdot a}} = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2 \cdot \zeta^2 - 1}} \ , \ \zeta > \frac{\sqrt{2}}{2}$$

This approximation is applicable only for responses more damped than MFA (with pole angle $\leq$ 45°), which is usually the case for wideband amplifiers. As ζ decreases, the approximation becomes less accurate so that for $\zeta = 1$, the above frequency ratio is 0.7071 instead of 0.6436, about 10 % high. Thus it is still useful for real poles.

To compare the OCTC and approximate quadratic bandwidths, let two real poles have time constants τ_1 and τ_2. Then the pole polynomial coefficients are

$$a = \tau_n^2 \ ; \ b = \tau_1 + \tau_2$$

By the quadratic approximation,

$$\frac{1}{\omega_{bw}} = \tau_{bw} = \sqrt{b^2 - 2\cdot a} = \sqrt{(\tau_1+\tau_2)^2 - 2\cdot(\tau_n^2)} = \sqrt{\tau_1^2 + 2\cdot(\tau_1\cdot\tau_2 - \tau_n^2) + \tau_2^2}$$

For quadratic real poles, $\tau_n{}^2 = \tau_1\tau_2$. Then the middle term becomes zero and

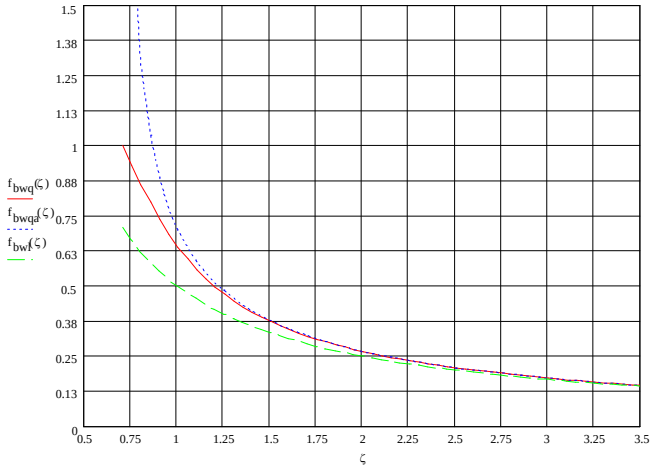
$$\tau_{bw} = \sqrt{\tau_1^2 + \tau_2^2}$$

This is the same as the OCTC formula for bandwidth.

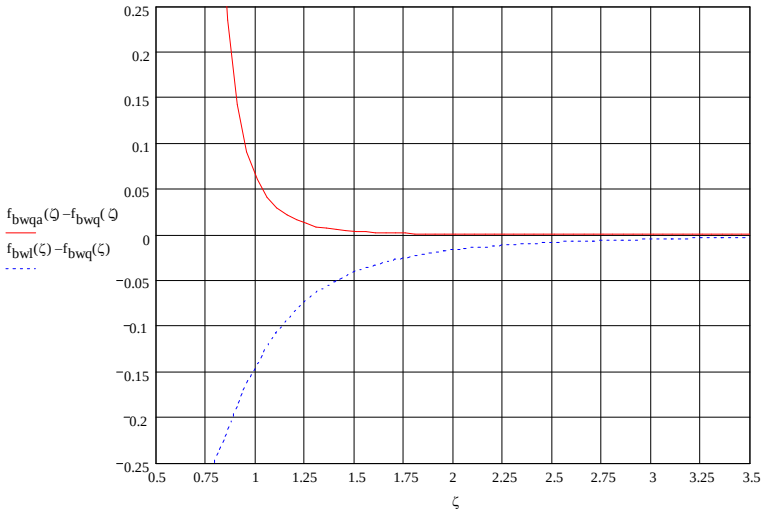
If the poles are widely separated, then $a \approx 0$ and a *dominant-pole bandwidth* approximation is

$$\omega_{bw} \approx \frac{1}{b} = \omega_n \cdot \frac{1}{2\cdot\zeta} \;;\; \tau_{bw} \approx \sum_{j=1}^{2} \tau_j = b$$

The resulting τ_{bw} is a simple summation of OCTCs, or b, whereas in the OCTC bandwidth, based on the linear (s) term in $D(s)$, it was the square-root of the sum of the squares of possibly more than two OCTCs, a value that is less than the above τ_{bw} and somewhat more accurate. The three quadratic bandwidth formulas are plotted below.



The OCTC bandwidth formula reduces n-degree pole polynomials to the equivalent of a real pole at the OCTC bandwidth. For complex poles with a low pole angle, f_{bw} is not altered much, though as ζ decreases (and pole angle increases), error increases superlinearly.



Plots of the error as functions of ζ of the approximate quadratic bandwidth (f_{bwqa}) and approximate linear bandwidth (f_{bwl}) relative to the exact quadratic bandwidth show that the approximate quadratic bandwidth and by equivalence, the OCTC bandwidth, will be high. The linear sum-of-TCs method will be low and have greater error. As damping increases, poles separate, one becomes dominant, and the bandwidth approximations become more accurate.

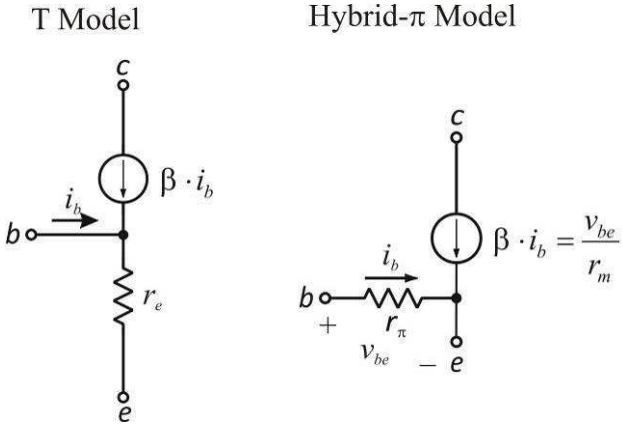
At critical damping ($\zeta = 1$), the approximate quadratic bandwidth is in error by about 7 % whereas the linear bandwidth approximation is low at –15 % with about twice the error. At about a decade of pole separation ($\zeta \approx 1.74$), the linear bandwidth has about –2.5 % error and the approximate quadratic bandwidth error is well within 1 %. At a pole separation ratio of about 34, $\zeta = 3$ and both are within ±1 % error. For underdamped ($\zeta < 1$) pole-pairs, the error of both approximations increases quickly and they become inapplicable.

The OCTC and approximate quadratic bandwidths are based on no more than first or second-degree polynomial approximations, yet can be generally useful, especially in feedback amplifiers because more than two close poles cause a phase delay of greater than 180° which tends to make feedback loops unstable. Consequently, a typical feedback amplifier with single- to double-pole rolloff has no more than two close lower-frequency poles, though others will be at higher frequencies. Therefore, these methods are not impractical for many transistor amplifiers.

Transistor Dynamics

Dynamic BJT Model

The T model of the BJT served us well in quasistatic analysis and can be extended to include dynamic behavior. In many electronics textbooks, the T model is used in a slightly different form, the *hybrid-π* model, converted by applying the β transform to r_e as shown.



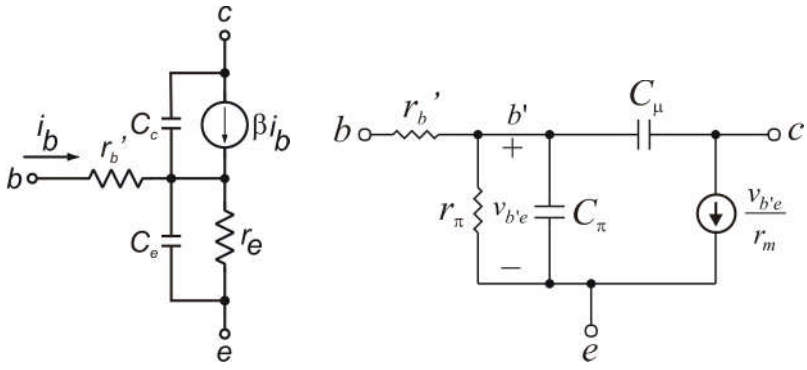
The difference of the hybrid-π model is that r_e is referred to the base by the β transform as

$$r_\pi = (\beta + 1) \cdot r_e$$

Also, the current source is written in equivalent form;

$$\beta \cdot i_b = \frac{v_{be}}{r_m}$$

Capacitances are now included in the models, as shown below. Added to the quasistatic hybrid-π and T models are *b-e* capacitance, C_π, and *b-c* capacitance C_μ. It is not uncommon to see C_π referred to as C_{be} or C_e and C_μ as C_{bc} or C_c instead. The T model (left) has the same capacitances.



An additional *base spreading resistance*, r_b' has been added from the internal base node to the base terminal. This resistance of 10 to 100 Ω exists because base current travels some distance through resistive base material. It figures prominently in high-speed BJT modeling. C_μ or C_c is typically less than 1 pF in fast BJTs to 10 pF or more in larger transistors.

C_π or C_e is important because as frequency increases, it shunts away current that would flow through r_e - the current that produces collector current - and β decreases. Another way to look at it is that the reactance of C_e reduces

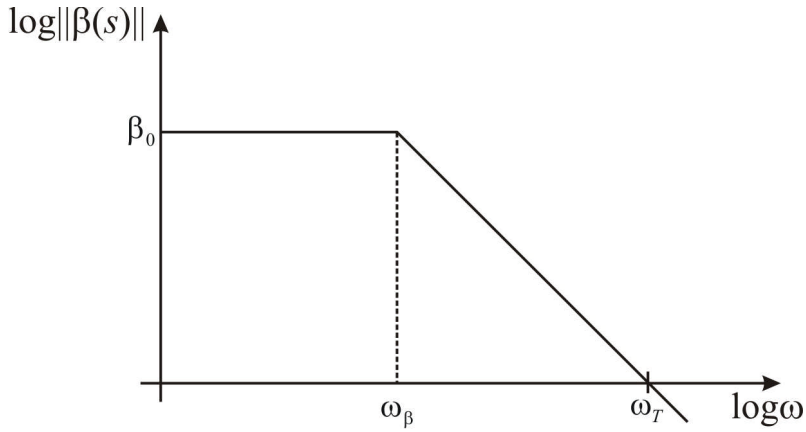
$$Z_\pi = r_\pi \,||\, 1/s{\cdot}C_\pi = \; r_\pi \,||\, 1/\,s{\cdot}C_e$$

and base current develops a decreasing v_{be} across it with frequency. For the same base current, the collector current decreases with frequency and β is no longer constant but is a function of frequency.

To derive an equation for $\beta(s)$, start with the parallel RC in the base circuit of the hybrid-π model;

$$Z_\pi = \frac{r_\pi \cdot (1/\,s \cdot C_e)}{r_\pi + (1/\,s \cdot C_e)} = \frac{r_\pi}{s \cdot r_\pi \cdot C_e + 1} = r_\pi \cdot \frac{1}{\dfrac{s}{\omega_\beta} + 1} = \frac{r_\pi}{s \cdot \tau_\beta + 1}$$

The time constant, $\tau_\beta = r_\pi{\cdot}C_e = 1/\omega_\beta$, where ω_β is the pole frequency of $\beta(s)$. The quasistatic β is distinguished from $\beta(s)$ as β_0. The fraction of i_b through r_π is multiplied by β_0 to produce the collector current. $\beta(s)$ is plotted below.



Using the current-divider formula,

$$\frac{i(r_\pi)}{i_b} = \frac{1/s \cdot C_e}{1/s \cdot C_e + r_\pi} = \frac{1}{s \cdot r_\pi \cdot C_e + 1} = \frac{1}{s \cdot \tau_\beta + 1}$$

Then

$$\beta(s) = \frac{i_c(s)}{i_b(s)} = \frac{i_c}{i(r_\pi)} \cdot \frac{i(r_\pi)}{i_b} = \beta_0 \cdot \frac{1}{s \cdot \tau_\beta + 1} = \frac{\beta_0}{s \cdot \tau_\beta + 1}$$

This also can be derived from a voltage standpoint as

$$\beta(s) = \frac{i_c}{i_b} = \frac{v_{be}/r_m}{i_b} = \frac{Z_\pi \cdot i_b / r_m}{i_b} = \frac{Z_\pi}{r_m} = \frac{r_\pi}{r_m} \cdot \frac{1}{\frac{s}{\omega_\beta} + 1}$$

The BJT transresistance is

$$r_m = \frac{v_{be}}{i_c} = \frac{v_{be}}{\alpha_0 \cdot i_e} = \frac{r_e}{\alpha_0} = \frac{(\beta_0 + 1) \cdot r_e}{\beta_0} = \frac{r_\pi}{\beta_0}$$

Then $r_\pi / r_m = \beta_0$ and the result for $\beta(s)$ is the same.

At ω_β, $\beta(s)$ begins to decrease asymptotically (or "roll off") with a –1 slope until it crosses 1 at

$$\omega_T = \beta_0 \cdot \omega_\beta \Rightarrow f_T = \beta_0 \cdot f_\beta$$

The bandwidth of β is ω_β and the *unity-gain* (gain-of-one) *frequency* is ω_T, the *gain-bandwidth product* of β.

In gain equations, α also appears. The frequency-dependent α is

$$\alpha(s) = \frac{\beta(s)}{\beta(s)+1} = \frac{\dfrac{\beta_0}{s/\omega_\beta + 1}}{\dfrac{\beta_0}{s/\omega_\beta + 1} + 1} = \frac{\beta_0}{\beta_0 + 1} \cdot \frac{1}{\dfrac{s}{(\beta_0 + 1)\cdot\omega_\beta} + 1}$$

This reduces to

$$\alpha(s) = \alpha_0 \cdot \frac{1}{\dfrac{s}{\omega_T / \alpha_0} + 1} = \alpha_0 \cdot \frac{1}{s(\alpha_0 \cdot \tau_T) + 1}$$

where α_0 is the quasistatic α. The pole frequency of $\alpha(s)$ is slightly greater than ω_T and for $\alpha_0 \approx 1$, it is approximately ω_T or f_T. BJT f_T (as given in BJT data) is a practical upper limit on frequency for amplifiers in that current is attenuated above f_T. The typical value of f_T for a PN3904 or PN3906 is 300 MHz at 10 mA of static collector current. Integrated BJTs made with high-speed processes exceed f_T of 10 GHz and discrete BJTs are commercially available above 1 GHz.

The expression for $\beta(s)$ can be simplified by constraining its validity to the *high-frequency* (hf) *region* between f_β and f_T. The result is a *BJT hf model*, valid only in the hf region. Ordinarily, $\beta_0 >> 1$ and if it is approximated as infinite ($\beta_0 \to \infty$), then the hf β is

$$\beta_{hf} = \lim_{\beta_0 \to \infty} \frac{\beta_0}{s/\omega_\beta + 1} = \lim_{\beta_0 \to \infty} \frac{1}{s/\omega_T + 1/\beta_0} = \frac{1}{s/\omega_T} = \frac{1}{s \cdot \tau_T}$$

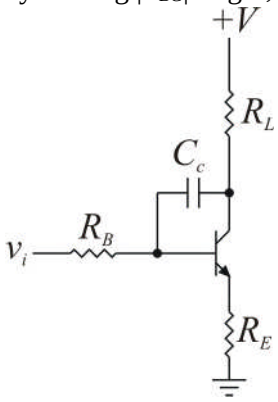
where $\tau_T = 1/\omega_T$. β_{hf} has a pole at the origin, and at ω_T the magnitude of β is one. β_{hf} removes the break in $\beta(s)$ at ω_β and the –1 slope of the hf region extends leftward and upward into the *low-frequency* (lf) *region* below f_β. In the hf region,

$$\beta_{hf} + 1 = \frac{s \cdot \tau_T + 1}{s \cdot \tau_T} \quad \text{and} \quad \alpha_{hf} = \frac{1}{s \cdot \tau_T + 1}$$

OCTC of C_c

Before deriving the open-circuit resistance, R_{bc}, across C_c, some general observations can be made. First, C_c is the capacitance of the base-collector junction. As reverse-bias voltage across this junction

increases, the junction widens. The junction-capacitance dielectric thickens causing C_c to decrease. BJT amplifier speed can be increased by making $|V_{BC}|$ larger, though BJT breakdown voltage constrains it.



Second, the CE configuration can have significant voltage gain from base to collector, and it is inverted. The circuit diagram shows C_c "simulated" as external to a quasistatic BJT. C_c is across the *b*-*c* nodes with an inverting (negative) voltage gain from *b* to *c*. The Miller effect occurs, making C_c effectively larger at the base by the base-to-collector gain, K_v, plus one, and should be expected to appear in the derivation of R_{bc}.

Applying Miller's Theorem, the portion of base capacitance, C_b, contributed by C_c is

$$C_b(C_c) = (1 + K_v)\cdot C_c$$

where K_v is the magnitude of the inverting voltage gain from base to collector. For small base-node resistance (to ground), R_b, it is close to

$$K_v \approx |A_v| \text{ , } R_b \text{ small}$$

where

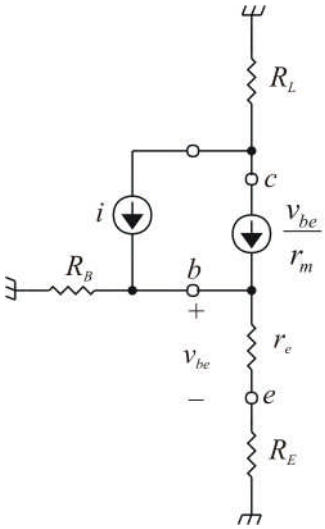
$$R_b = R_B \,\|\, [(\beta_0 + 1)\cdot(r_e + R_E)] = R_B \,\|\, r_B$$

The Miller effect also causes R_L to be effectively $R_L/(1 + K_v)$. The time constant at the base node is that of R_b with base capacitance, C_b to ground. Assuming base input resistance, r_B, is large, then $R_b \approx R_B$ and the base time constant is

$$\begin{aligned}\tau_b &= [R_B + R_L/(1 + K_v)]\cdot[(1 + K_v)\cdot C_c] \\ &= R_B\cdot[(1 + K_v)\cdot C_c] + R_L\cdot C_c\end{aligned}$$

This time constant is essentially the C_c OCTC.

We now proceed to find the open-circuit *b*-*c* resistance, R_{bc}, which forms a time constant with C_c. The OCTC of C_c, unlike C_e, involves both input and output loops of the BJT stage. The general one-stage circuit model for R_{bc} analysis is shown below.



The current source, *i*, between collector and base is applied to the *b*-*c* port in place of C_{bc} to find R_{bc}. KCL (as the node-voltage method) is applied at the base node;

$$\frac{v_b}{R_B} - \frac{v_{be}}{r_m} - i + \frac{v_b}{r_e + R_E} = 0$$

The emitter voltage divider formula gives v_{be};

$$v_{be} = v_b \cdot \frac{r_e}{r_e + R_E}$$

Noting that $r_e = \alpha_0 \cdot r_m$, the current source value is

$$\frac{v_{be}}{r_m} = \alpha_0 \cdot \frac{v_b}{r_e + R_E}$$

Substituting into the KCL equation and solving for *i*,

$$i = \frac{v_b}{R_B} + \frac{v_b}{(\beta_0 + 1) \cdot (r_e + R_E)} = \frac{v_b}{R_B \parallel (\beta_0 + 1) \cdot (r_e + R_E)} = \frac{v_b}{R_b}$$

having applied the α-β relationship,

$$1 - \alpha_0 = \frac{1}{\beta_0 + 1}$$

Solving for v_b, $v_b = R_b \cdot i$ where

$$R_b = R_B \parallel (\beta_0 + 1) \cdot (r_e + R_E) = R_B \parallel r_B$$

At the collector node,

$$v_c = -R_L \cdot \left(\frac{v_{be}}{r_m} + i \right) = -R_L \cdot \left(\alpha_0 \cdot \frac{v_{be}}{r_e} + i \right) = -R_L \cdot \left(\alpha_0 \cdot \frac{R_b}{r_e + R_E} + 1 \right) \cdot i$$

Finally,

$$R_{bc} = \frac{v_{bc}}{i} = \frac{v_b - v_c}{i}$$
$$= R_b + R_L \cdot \left(1 + \alpha_0 \cdot \frac{R_b}{r_e + R_E} \right) = R_L + R_b \cdot \left(1 + \alpha_0 \cdot \frac{R_L}{r_e + R_E} \right)$$

R_{bc} includes resistances in both the input loop (R_b, r_e, R_E) and output loop (R_L). R_{bc} is the resistance of a single time constant across the *b*-*c* nodes. It is composed of two time-constant terms, one associated with the collector node,

$$\tau_{cc} = R_L \cdot C_c$$

and the other with the base node,

$$\tau_{cb} = R_b \cdot [C_c \cdot (1 + K_v)] \ , \ K_v = \alpha_0 \cdot \frac{R_L}{r_e + R_E}$$

where the Miller effect from the base-to-collector voltage gain effectively increases C_c at the base node by $(1 + K_v)$ as it appears in R_{bc}. (In τ_{cc}, we see that no Miller effect occurs at the collector node.) These are *not* two separable time constants but are components of the one time constant,

$$\tau_c = \tau_{cc} + \tau_{cb}$$

They can be separated in design by adding a CB stage after a CE stage to make a cascode amplifier. For a CB stage for which $R_B = 0\ \Omega$, $\tau_c = \tau_{cc}$ and $R_{bc} = R_L$. For $R_B = 0\ \Omega$, R_{bc} is entirely in the collector circuit as R_L, separated from interaction with the base.

R_{bc} is maximum and bandwidth minimum when $R_L \to \infty$, but not for $R_B \to \infty$ for which $R_{bc} = (\beta_0 + 1) \cdot (r_e + R_E + R_L)$. For $R_E \to \infty$, then $R_{bc} = R_L + R_B$. Maximum bandwidth is achieved where a minimum occurs at

$$\min \ R_{bc} = R_L \ , \ R_B = 0\ \Omega$$

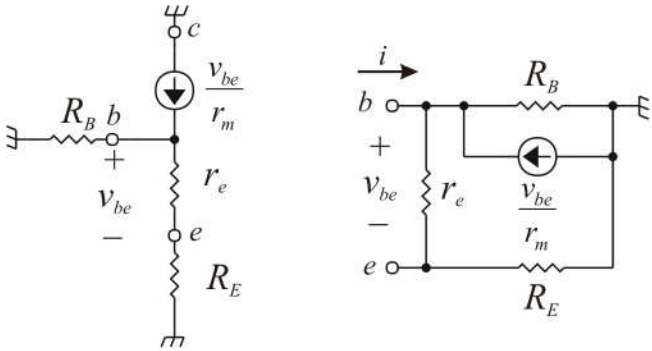
In this case (that of a CB), R_{bc} is isolated to the output circuit and the time constant of C_c is entirely at the collector node. For a finite R_B,

the minimum $R_{bc}(R_E) \rightarrow R_L + R_B$ as $R_E \rightarrow \infty$. R_{bc} is independent of R_E whenever $R_B = \beta_0 \cdot R_L$, for then $R_{bc} = R_L + R_B = (\beta_0 + 1) \cdot R_L$.

OCTC of C_e

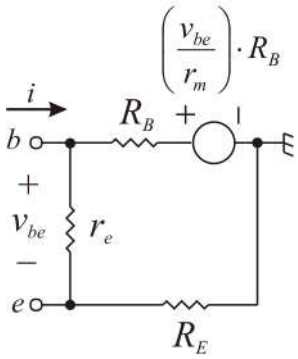
Having found R_{bc} and the time constant of C_c, this leaves the C_e time constant. The complication of C_c was that it involved both input and output circuits. The complication of C_e is not that - R_{be} is constrained to the input loop - but it affects $\alpha(s)$ and thereby affects the output differently than C_c.

From the following circuit with C_e and C_c removed, we can derive the open-circuit R_{be}. The circuit is redrawn (right) to make the *b*-*e* port more obvious.



The port resistance is r_e shunted by a circuit that is reduced by thevenizing the current source and R_B, as shown below. Then r_e is in parallel with a branch having a dependent source;

$$i = \frac{v_{be}}{r_e} + \frac{v_{be} - \left(\frac{v_{be}}{r_m}\right) \cdot R_B}{R_B + R_E} = v_{be} \cdot \left(\frac{1}{r_e} + \frac{1 - \frac{R_B}{r_m}}{R_B + R_E} \right)$$



The ground is included only for circuit reference and does not affect the *b*-*e* port resistance. Solving,

$$R_{be} = \frac{v_{be}}{i} = r_e \parallel \frac{R_B + R_E}{1 - \dfrac{R_B}{r_m}}$$

When the parallel formula is invoked,

$$R_{be} = \frac{r_e \cdot (R_B + R_E)}{r_e - (r_e / r_m) \cdot R_B + R_B + R_E} = \frac{r_e}{r_M} \cdot (R_B + R_E)$$

where the emitter-referred input-loop transresistance again appears:

$$r_M = r_e + R_E + R_B / (\beta_0 + 1)$$

This can be rewritten in a slightly different form that has a circuit interpretation:

$$R_{be} = r_e \cdot \left\{ (\beta_0 + 1) \cdot \frac{R_B / (\beta_0 + 1)}{r_M} + \frac{R_E}{r_M} \right\} = r_\pi \cdot \frac{R_B / (\beta_0 + 1)}{r_M} + r_e \cdot \frac{R_E}{r_M}$$

The left term has a fraction that is a voltage divider. The resistances of the divider are all referred to the emitter circuit. It is the fraction of the emitter input-loop voltage that is developed across the base resistance of the circuit, referred to the emitter. The $\beta_0 + 1$ factor refers r_e to the base as r_π by this fraction. The right term is a voltage divider that is the fraction of emitter-referred input-loop voltage across R_E. R_{be} is thus a combination of base-referred and emitter-referred r_e, and the voltage dividers give the fractions of each.

From the BJT dynamic model, $r_e \cdot C_e = \alpha_0 \cdot \tau_T$. Then the general formula for $\tau_{be} = R_{be} \cdot C_e$ is

$$\tau_{be} = \frac{\tau_\beta \cdot \dfrac{R_B}{\beta_0 + 1} + (\alpha_0 \cdot \tau_T) \cdot R_E}{r_M} = (\alpha_0 \cdot \tau_T) \cdot \frac{R_B + R_E}{r_M}$$

Now that we have a general formula for τ_{be}, we can see what effect R_B and R_E have on the time constant. As R_E is made large,

$$\lim_{R_E \to \infty} R_{be} = r_e$$

Then the *b-e* time constant, $\tau_{be} \to \alpha_0 \tau_T$. Whenever $R_E >> r_e$, we can largely disregard the effect of τ_{be} on frequency response because it is near τ_T, the high end of the useable frequency range of the BJT. (However, it *does* affect base input impedance which can affect dynamic response.) For $R_E = 0\ \Omega$, $R_{be} = r_\pi \| R_B$. For $R_B >> r_e$ or $R_B \to \infty$, $R_{be} \to r_\pi$ and the slowest time constant results: $\tau_{be} = \tau_\beta$. The value of a large R_E or small R_B is apparent; either reduces τ_{be} and results in a faster amplifier.

OCTC of C_L

A third and final capacitance of interest to include in a single-stage BJT dynamic model is the collector capacitance to ground, C_L. This is usually the loading of the amplifier output by the next stage. For the CB, C_c is isolated from the input circuit and the output pole is simply

$$\text{CB } \tau_L = R_L \cdot (C_c + C_L)$$

where R_L is the collector node resistance.

We found for the CE that the collector time constant component of C_c, $\tau_{cc} = R_L \cdot C_c$, appears in the transfer function without a Miller-effect multiplier. For a general voltage amplifier, the capacitance of C_f from input to output nodes appears at the output node as a Miller-transformed capacitance to ground, and is found by invoking the output form of Miller's Theorem:

$$z_o = \left(\frac{K_v}{1 + K_v} \right) \cdot Z_f = \left(1 + \frac{1}{K_v} \right) \cdot Z_f$$

Miller's Theorem extends to impedance and can be derived from the same circuit used to derive the formula for r_i. Unless a voltage

amplifier has output resistance, the shunt contribution of Z_f will not matter because it is shorted by the amplifier voltage-source output.

A BJT collector (or FET drain) is instead a current output. By Miller's Theorem, C_c will appear to be effectively $(1 + 1/K_v)\cdot C_c$. R_L is the resistance of τ_L and it also affects K_v. As R_L varies, so does K_v. The time constant is

$$\tau_L = R_L \cdot \left[\left(1 + \frac{1}{K_v} \right) \cdot C_c \right] = R_L \cdot \left(1 + \frac{r_m}{R_L} \right) \cdot C_c = (R_L + r_m) \cdot C_c$$

For large R_L, K_v is large and $R_L >> r_m$. Then τ_L approaches $R_L \cdot C_c$, as τ_{cc}. As R_L and K_v together approach zero, the resulting pole factor is $(s \cdot r_m \cdot C_c + 1)$. It forms a constant-magnitude or *all-pass* network with the right half-plane zero of the passive path through C_c and adds only phase shift without affecting amplitude or bandwidth.

In cascode amplifiers, the CE-stage $K_v << 1$ and $(1 + 1/K_v)$ becomes large as the equivalent R_L and K_v decrease to nearly zero. The Miller multiplier of C_c consequently makes the Miller output capacitance appear very large. However, because this capacitance is used to calculate the collector time constant, R_L is correspondingly small, negating the output Miller effect. The time constant that results has a capacitance of C_c. Therefore, C_c at the collector is used at its actual value whenever R_L affects both the Miller gain and the time-constant resistance. The Miller effect is only applied in dynamics calculations to the base node. This also applies to FETs; only the gate capacitance is multiplied by $(1 + K_v)$.

The C_c capacitance is added to C_L to form the collector-node time constant,

$$\text{CE } \tau_L = R_L \cdot (C_L + C_c) \text{ , } R_b \approx 0\ \Omega$$

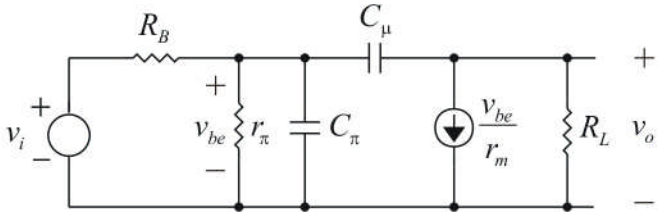
when base-node resistance (to ground) is insignificant and C_c is effectively across the collector node. This time constant is the third pole of the BJT stage. Otherwise, C_c has its own OCTC (τ_c) and

$$\text{CE } \tau_L = R_L \cdot C_L$$

For the CC configuration, if the collector is quasistatically grounded, there is no τ_L. This is achieved by capacitively bypassing any collector series resistance.

Textbook CE Stage

A classic derivation from basic circuit equations of the CE stage that is found in most circuits textbooks uses the hybrid-pi BJT incremental model without R_E. The circuit is shown below and KCL is applied to base and collector nodes.



After some algebraic reconfiguring,

$$\text{KCL @ } c\text{: } v_{be} = -v_o \cdot \frac{r_m \parallel (-1/s \cdot C_\mu)}{R_L \parallel (1/s \cdot C_\mu)}$$

$$\text{KCL @ } b\text{: } \frac{v_i}{R_B} = \frac{v_{be}}{R_b \parallel (1/s \cdot C_\mu) \parallel (1/s \cdot C_\pi)} - \frac{v_o}{1/s \cdot C_\mu} \text{ , } R_b = R_B \parallel r_\pi$$

When v_{be} of KCL at the collector is substituted into the base-node equation and simplified,

$$\frac{v_i}{R_B} = \frac{v_{be}}{R_b \parallel (1/s \cdot C_\mu) \parallel (1/s \cdot C_\pi)} - \frac{v_o}{1/s \cdot C_\mu}$$

Noting that $r_\pi \| r_m = r_e$, then after more algebra,

$$A_v = \frac{v_o}{v_i} = -\alpha_0 \cdot \frac{R_L}{r_e + R_B/(\beta_0 + 1)} \cdot \frac{-s \cdot r_m \cdot C_\mu + 1}{s^2 \cdot [(R_b \cdot C_\pi) \cdot (R_L \cdot C_\mu)] + s \cdot [R_b \cdot C_\pi + (R_L + R_b \cdot (1 + \alpha_0 \cdot \frac{R_L}{r_e})) \cdot C_\mu] + 1}$$

In the coefficient of s in the denominator are the familiar results of the OCTC resistance derivations for the single-BJT stage. Both time-constants that are products in the s^2 coefficient, a, are terms in the s coefficient, b, along with an additional Miller-capacitance term for C_μ. If it were zero, then the quadratic polynomial would factor into two real poles at the two time constants in a. Because b is larger, the two poles are always real and overdamped.

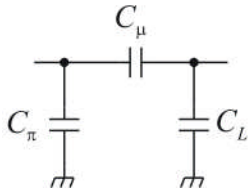
The zero is new and it has a negative linear term which results in a positive root - a zero in the right half-plane (RHP) of the *s* plane. RHP zeros have the same magnitude response as LHP zeros but their phase response is that of a LHP pole; magnitude increases while phase decreases. (They are undesirable for feedback loops because they destabilize them.)

We will need to check on the frequency (or time constant) of the RHP zero when finding bandwidth. If it is much higher in frequency than the OCTC pole bandwidth (as is usual), then it has negligible effect. However, if it is near the OCTC bandwidth magnitude, it can extend it until a higher-frequency pole rolls it off.

Single-Stage BJT Bandwidth

We have identified the three BJT-stage capacitances and their OCTCs, yet earlier (in regard to bandwidth estimation) noted that for *n* capacitances (and no inductances) the transfer function pole polynomial should be of degree *n*, with a s^n term. Each reactive element (capacitance or inductance) in a circuit contributes an *s* to the circuit equations, either as $s \cdot L$ or $1/s \cdot C$. Yet with three capacitances, the CE BJT-stage has only as much as an s^2 term. Why?

In loop or node equations, some reactances can combine into one effective reactance because they are in series or parallel. A less obvious dependence is capacitors that form a loop, as is the case in the general BJT stage, with the loop shown below.



Although (besides ground) two separate nodes are involved, capacitance cannot be reduced to a simple two-terminal equivalent. Three capacitances form equivalent capacitances at three ports. The resulting BJT node capacitances are

$$C_{be} = C_\pi + C_\mu \,||\, C_L \;;\; C_{bc} = C_\mu + C_\pi \,||\, C_L \;;\; C_o = C_L + C_\pi \,||\, C_\mu$$

where || is the parallel math operation, not a circuit description. The three capacitances form a dependence in these nodal equivalents as all

three are involved in all nodal capacitance expressions and consequently are not independent of each other.

The pole polynomial can be expressed in the single-stage BJT time constants as

$$D(s) = s^2 \cdot (\tau_e \cdot \tau_{cc}) + s \cdot (\tau_e + \tau_{cc} + \tau_{cb}) + 1$$

The Miller effect at the base contributes only to the linear coefficient, in τ_{cb}, and as it increases, ζ increases without any effect on pole magnitude, which is set by the two quadratic time constants as

$$\tau_n = \sqrt{\tau_e \cdot \tau_{cc}}$$

For $\tau_{cb} = 0$ s, the polynomial has LHP real roots at τ_e and τ_{cc}, both of which are often much less than τ_{cb}. Then τ_n is small relative to the linear coefficient, the pole-pair is overdamped and the poles are real and separated. With wide pole separation ($\zeta >> 1$), τ_n of the quadratic term of the polynomial is much less than the linear term and can be discarded because the lower-frequency pole will dominate.

The roots of $D(s)$ can be found by solving for them from OCTCs. Expressed in quadratic parameters, the poles are

$$p_{1,2} = \frac{1}{\tau_{1,2}} = -\frac{\zeta}{\tau_n} \cdot \left(1 \pm \frac{\sqrt{\zeta^2 - 1}}{\zeta} \right)$$

$$\tau_n = \sqrt{\tau_e \cdot \tau_{cc}} \;\; ; \;\; \zeta = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_e}{\tau_{cc}}} + \sqrt{\frac{\tau_{cc}}{\tau_e}} + \frac{\tau_{cb}}{\tau_n} \right)$$

This solution for the single-stage poles includes the interactions of C_e and C_c. The poles are real; $\zeta \geq 1$, and ζ is minimum whenever $\tau_e = \tau_{cc}$:

$$\zeta_{\min} = 1 + \left(\frac{\tau_{cb}}{2 \cdot \tau_n} \right)$$

For widely separated real poles and a dominant τ_{cb}, the TC values of the poles can easily be approximated. The low-frequency TC $\approx b \approx \tau_{cb}$. Having one pole value, and knowing that the quadratic coefficient, a, is the product of the real poles, then the high-frequency TC $\approx a/b \approx \tau_n^2/\tau_{cb}$. In general, for TCs τ_1 and τ_2 of real poles,

$$\frac{a}{b} = \frac{\tau_1 \cdot \tau_2}{\tau_1 + \tau_2} \approx \tau_2 \text{ , } \tau_2 << \tau_1$$

The OCTCs of the general single-stage BJT amplifier, now expressed in circuit elements, can be used to find the single-stage bandwidth. The OCTCs, τ_e and τ_c, and the *time constant components* of τ_c are

$$\tau_e = \tau_{be} = R_{be} \cdot C_e$$
$$\tau_c = R_L \cdot C_c + R_b \cdot [(1 + K_v) \cdot C_c] = \tau_{cc} + \tau_{cb}$$

Stage bandwidth can be found from the OCTC bandwidth formula as

$$\tau_{bw} = \sqrt{\tau_e^2 + \tau_c^2}$$

With the BJT-stage template for the OCTCs already found, this is an easy method to use. For cases where a circuit pole polynomial is given but not the circuit elements, then an equivalent to the OCTC bandwidth formula is the approximate quadratic formula. The most accurate, of course, is the exact quadratic bandwidth formula.

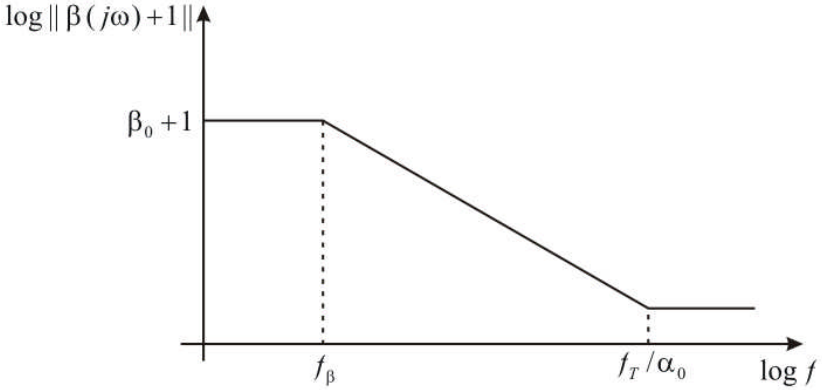
Complex pole-pairs can exist when additional reactances are introduced into a stage. Without additional reactances, feedback loops can also cause resonances that appear in the closed-loop transfer functions. In the high-frequency region of transistor operation, impedance gyration can cause resonances with additional circuit reactances. The exact or approximate quadratic bandwidth formulas give the bandwidth for complex pole-pairs from parameters f_n and ζ.

High-Frequency β Transform

In the high-frequency region ($f_\beta \leq f \leq f_T$) of BJT operation, device impedances are transformed, or *gyrated*, to become other kinds of circuit elements. For instance, BJT base resistance in the hf region is gyrated to appear as inductance at the emitter. This inductance can (and often does) resonate with emitter capacitance if the resonant frequency is within the hf region. These impedance transformations become apparent when we begin using $\beta(s)$ in the β transform. Then the β-transform factor becomes

$$\beta(s)+1=\frac{\beta_0}{s\cdot\tau_\beta+1}+1=(\beta_0+1)\cdot\frac{\frac{s\cdot\tau_\beta}{\beta_0+1}+1}{s\cdot\tau_\beta+1}=(\beta_0+1)\cdot\frac{s\cdot\alpha_0\cdot\tau_T+1}{s\cdot\tau_\beta+1}$$

where the last step applies $\tau_\beta = \beta_0 \cdot \tau_T$ and the quasistatic α_0.



A magnitude plot of $\beta(j\cdot\omega) + 1$ shows that between BJT β bandwidth, f_β, and f_T/α_0 ($\approx f_T$ for $\alpha_0 \approx 1$), β decreases in proportion to frequency, having a log-log-plot slope of –1. This causes the BJT collector current to decrease with increasing frequency and also rotates the impedance vector of components in the input loop of the BJT, an effect called *impedance gyration.*

Applying the frequency-dependent β transform to the base port,

$$Z_b = Z_\pi + (\beta(s)+1)\cdot Z_E$$

where Z_E is the impedance external to the emitter. External impedance across the base port, Z_B, refers to the emitter port as

$$Z_e = \frac{Z_\pi + Z_B}{\beta(s)+1}$$

These are the frequency-dependent expressions of the $\beta(s)$ transform over the entire BJT frequency range of interest.

The $\beta(s)$ transform can be simplified to apply only in the hf region as $\beta_{hf}(s)$, and has as its transform factor

$$\beta_{hf}+1=\frac{1}{s/\omega_T}+1=\frac{s/\omega_T+1}{s/\omega_T}=\frac{s\cdot\tau_T+1}{s\cdot\tau_T}$$

This is a pole at the origin with a *transverse* (crossing at one) *frequency* of ω_T and also a zero at ω_T. The break in the curve at ω_β is removed and the –1 slope extended leftward toward the origin. The break at ω_T remains and above ω_T, $||\beta(j\cdot\omega) + 1|| = 1$.

To apply the β_{hf} transform to circuits, start with emitter resistance R_E, referred to the base. The simplified β_{hf} transform equation for Z_b with emitter impedance of Z_E is

$$Z_b(\text{hf}) = \left(\frac{1}{s\tau_T} + 1\right) \cdot Z_E$$

Z_π is omitted because when its pole is removed at f_β,

$$Z_\pi(\text{hf}) = \lim_{\beta_0 \to \infty} \frac{r_\pi}{s \cdot \tau_\beta + 1} = \lim_{\beta_0 \to \infty} \frac{r_\pi / \beta_0}{s \cdot \tau_T + 1/\beta_0} = \frac{r_m}{s \cdot \tau_T} \approx 0\,\Omega$$

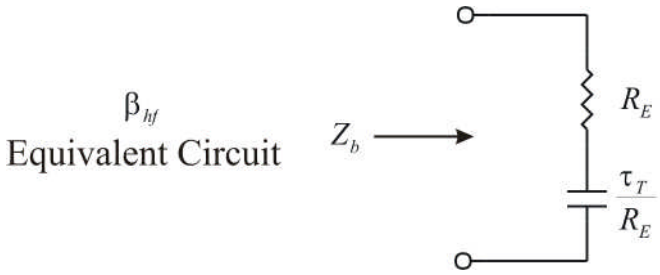
for $R_E >> r_e$. At the base port, with $Z_E = R_E$ in the emitter,

$$Z_b = (\beta_{hf} + 1) \cdot R_E = R_E + \frac{R_E}{s \cdot \tau_T}$$

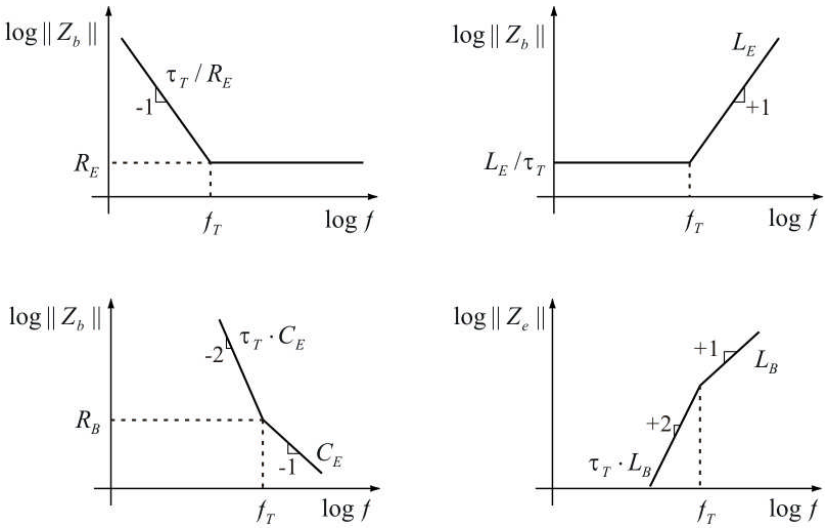
The second term can be rewritten as

$$\frac{R_E}{s \cdot \tau_T} = \frac{1}{s \cdot \left(\frac{\tau_T}{R_E}\right)}$$

and has the form of capacitive reactance where the gyrated resistance is τ_T/R_E in series with R_E. The base equivalent circuit is shown below.



Z_b for the three basic kinds of Z_E (R, L, and C) is shown below on reactance plots, for Z_E of R_E, L_E, and C_E; and (lower-right) Z_e for L_B. The bottom two plots are for the two less-familiar cases of when impedance is gyrated to the $-R$ axis. In these cases, resonances can be formed whenever one of these plots intersects another from the rest of



the circuit with a slope change of ±2. In ordinary series and parallel resonances, a +1 to –1 (parallel) or –1 to +1 (series) change in slope will occur at the resonant point. In the case of hf resonances, a 0 to +2 or –2 to 0 slope change also indicates a resonance. We will return to these more difficult cases in more detail later.

The use of β_{hf} simplifies hf analysis but what are we missing in using it? The *total-frequency* (lf and hf) base impedance, with $\beta(s) + 1$, is valid from 0 Hz to f_T, and is

$$Z_b = Z_\pi + [\beta(s) + 1] \cdot Z_E = Z_\pi + \left(\frac{\beta_o}{s \cdot \tau_\beta + 1} + 1 \right) \cdot Z_E$$

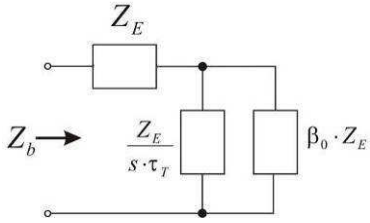
Much above f_β, Z_π becomes negligible when $X_C(C_e)$ substantially shunts r_π. Z_π is already half of r_π at f_β. By omitting Z_π,

$$Z_b \cong Z_E + \frac{1}{s \cdot \tau_T / Z_E + 1 / Z_E \cdot \beta_o}$$

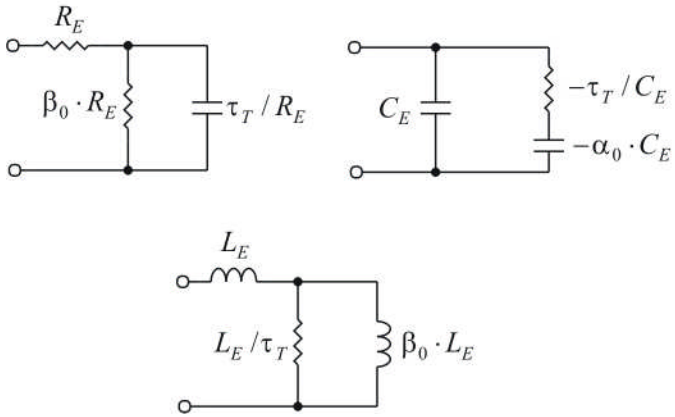
This can be rewritten as a *continued fraction* to express the input impedance in parallel-element form as

$$Z_b \cong Z_E + \frac{1}{\frac{1}{1/s \cdot (\tau_T / Z_E)} + \frac{1}{\beta_o \cdot Z_E}} = Z_E + \frac{1}{s \cdot (\tau_T / Z_E)} \| \beta_o \cdot Z_E$$

Written this way, the equivalent circuit can be drawn directly from the equation; Z_E is in series with two parallel elements: a reactance of value τ_T/Z_E in parallel with $\beta_0 \cdot Z_E$. The general equivalent base impedance is shown below.



By substituting R, $s \cdot L$, and $1/s \cdot C$ for Z_E, the equivalent circuits for the three elements can be determined. Then for any arbitrary linear network of R, L, and C, the transformed impedances can be substituted and combined to result in the total β-transformed impedance. Substituting the three basic circuit elements into Z_E, the transformed impedances are shown for R_E, C_E and L_E.



The difference between the total-frequency and the hf model can be seen for R_E as the addition of a shunt resistance across τ_T/R_E of $\beta_0 \cdot R_E$. At low frequency, the capacitance approaches an open circuit, leaving the two resistances in series. Then $Z_b = (\beta_0 + 1) \cdot R_E$, the quasistatic β-transformed base input resistance. For the hf-only base impedance, Z_b(hf), let $\beta_0 \to \infty$ ($\alpha_0 = 1$) and modify the above equivalent circuits accordingly.

Dividing Z_E by s in the general model of Z_b *gyrates* the impedance vector of Z_E by rotating its phase –90° in the complex impedance plane so that

$$R \to C$$
$$C \to -R$$
$$L \to R$$

The gyration of $C \to -R$ is interesting. These elements can form resonant circuits and oscillate. The reactance plots for the two cases, of C_E and L_B, were plotted previously. The impedance equations for them reduce to

$$Z_b(C_E) = \frac{s \cdot \tau_T + 1}{s^2 \cdot \tau_T \cdot C_E} \; ; \; Z_e(L_B) = \frac{s^2 \cdot \tau_T \cdot L_B}{s \cdot \tau_T + 1}$$

These resonances are duals and are damped by adding an element that "chops off" the corner with a slope change of ±2.

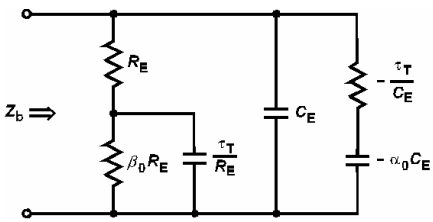
The equivalent circuits with negative elements for the two cases can be derived using some constructive algebra that is most easily demonstrated for Z_e;

$$Z_e(L_B) = \frac{s^2 \cdot \tau_T \cdot L_B}{s \cdot \tau_T + 1} = \frac{s^2 \cdot \tau_T \cdot L_B + (s \cdot L_B - s \cdot L_B)}{s \cdot \tau_T + 1}$$

$$= \frac{(s^2 \cdot \tau_T \cdot L_B + s \cdot L_B) - s \cdot L_B}{s \cdot \tau_T + 1} = s \cdot L_B - \frac{s \cdot L_B}{s \cdot \tau_T + 1}$$

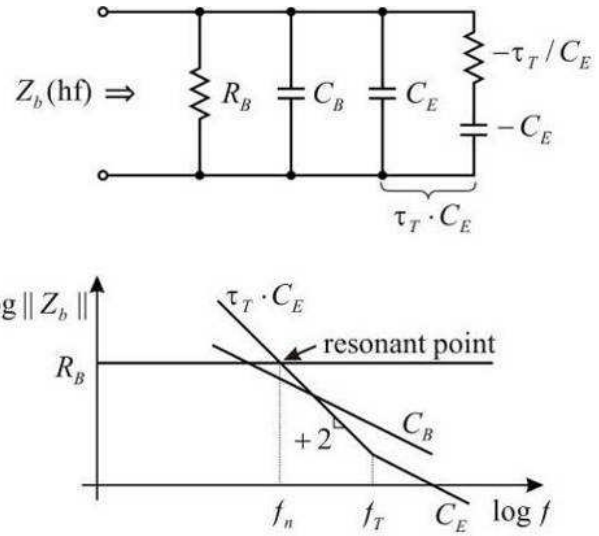
$$= s \cdot L_B + \frac{1}{\dfrac{1}{\left(-\dfrac{L_B}{\tau_T}\right)} + \dfrac{1}{s \cdot (-L_B)}}$$

An example is a shunt RC-loaded emitter-follower. A common-collector (CC) BJT amplifier has a shunt RC load for which $R_E = 470\,\Omega$ and $C_E = 10$ pF. The PN3904 BJT has $\beta_o = 150$ and $f_T = 300$ MHz at $I_E = 10$ mA. What is Z_b?



The combination of gyrated impedances for R_E and C_E are shown combined using the total-frequency model. To find circuit values, first calculate $\tau_T = 1/(2{\cdot}\pi{\cdot}f_T) = 531$ ps and $\alpha_o = 0.993 \cong 1$. Then $\tau_T/R_E = 1.13$ pF and $-\tau_T/C_E = -53.1\ \Omega$; $\beta_o{\cdot}R_E = 70.5$ kΩ and $-\alpha_o{\cdot}C_E \cong -10$ pF. A hf equivalent circuit omits $\beta_o{\cdot}R_E$ and $-\alpha_0{\cdot}C_E$ becomes $-C_E$.

Whether Z_π is negligible depends on the other elements in the circuit. If base reactance creates a resonance with the emitter impedance near f_β, then Z_π is probably significant. For this circuit, $r_e \cong 2.6\ \Omega$ and $C_\pi = \tau_T/r_e = 204$ pF.



One method of compensating the base input impedance is to add a series RC circuit from the base to ground that has positive values corresponding to those of the –R, –C branch. These branches cancel, leaving no –R to form a resonance. The compensated network is an all-pass filter with an impedance that is a resistance. An *all-pass circuit* has an impedance magnitude that is flat over frequency (resistive) but has phase delay.

What happens if the component tolerances of the compensating RC do not match –R or –C? The reactance plot of Z_b for the shunt R_E, C_E shows why the series RC damps the resonance. The equivalent circuit is shown first. On the reactance plot, C_B "bypasses" the resonant point and damps it, causing the change in slope to remain ±1 at each intersection of the asymptotic impedances. As C_B increases (and moves to the left on the plot), resonance damping increases. The

C_B of series-RC compensation has the same effect with a resulting impedance that shunts the rest of the equivalent circuit as a resistance. In CB stages with R_B, shunt base RC compensation is also applicable.

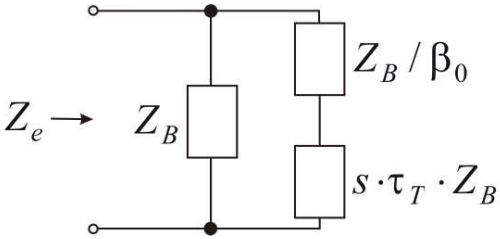
We have been examining hf behavior at the base port. A similar circuit derivation as that of Z_b applies to the emitter port and results in

$$Z_e = \frac{Z_\pi + Z_B}{1 + \beta_o / (s \cdot r_\pi \cdot C_\pi + 1)}$$

Approximating $Z_\pi \cong 0$ for $Z_B >> Z_\pi$, we obtain the continued fraction;

$$Z_e \cong \frac{1}{\dfrac{1}{Z_B} + \dfrac{1}{s \cdot \tau_T \cdot Z_B + Z_B / \beta_o}} = Z_B \parallel (s \cdot \tau_T \cdot Z_B + Z_B / \beta_o)$$

The equivalent circuit, shown below, is the dual of Z_b. The general form of Z_e has two elements in parallel instead of in series as in Z_b, its dual.



Below f_β, as s goes to zero, $s \cdot \tau_T \cdot Z_B$ approaches a short circuit and $Z_e(R_B)$ becomes the lf value of

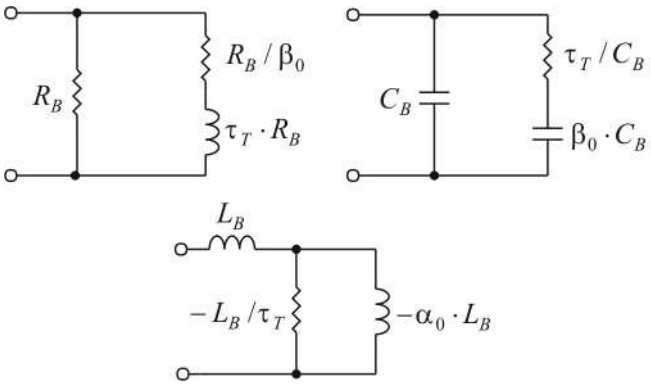
$$Z_e(\text{lf}) = \frac{1}{\dfrac{1}{R_B} + \dfrac{1}{R_B / \beta_0}} = \frac{R_B}{\beta_0 + 1}$$

The hf-only β_{hf} transform at the emitter is based on $\beta_{hf} = 1/s \cdot \tau_T$;

$$Z_e(\text{hf}) = \frac{Z_B}{1/s \cdot \tau_T + 1} = \frac{1}{\dfrac{1}{Z_B} + \dfrac{1}{s \cdot \tau_T \cdot Z_B}} = Z_B \parallel (s \cdot \tau_T \cdot Z_B)$$

and Z_B/β_0 is removed for $\beta_0 \rightarrow \infty$.

For the three basic elements, the resulting total-frequency impedances at the emitter port (looking into the emitter) are as shown below for $Z_B = R_B$, C_B, and L_B.



The hf contribution of $s \cdot \tau_T \cdot Z_B$ gyrates Z_e by +90° so that

$$R \rightarrow L$$
$$C \rightarrow R$$
$$L \rightarrow -R$$

Applying the hf approximations to τ_{be}, then

$$\tau_{be}(\text{hf}) = \tau_T \cdot \left(\frac{R_B + R_E}{R_E} \right) = \tau_T \cdot (K_i + 1) \ , K_i = R_B/R_E$$

K_i is the quasistatic current gain of a CE stage driven by an input current with R_B base input resistance shunting the base and $r_M \approx R_E$. By adding C_E across R_E, $Z_E = R_E \| (1/s \cdot C_E)$. Let $C_E = \tau_T / R_E$. Then

$$Z_E = \frac{R_E}{s \cdot \tau_T + 1}$$

Applying the β_{hf} transform, $Z_b(\text{hf}) = (\beta_{hf} + 1) \cdot Z_E$ and

$$Z_b(\text{hf}) = Z_E + \frac{Z_E}{s \cdot \tau_T} = \frac{R_E}{s \cdot \tau_T + 1} \cdot \left(1 + \frac{1}{s \cdot \tau_T} \right) = \frac{1}{s \cdot (\tau_T / R_E)}$$

The base impedance for the emitter circuit is reduced in the hf region to a capacitance, τ_T/R_E, and forms a pole with R_B. In this case in which Z_b is made to be capacitive, a time constant results at the base node of

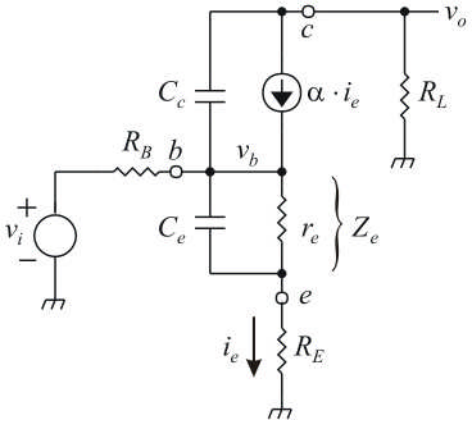
$$\tau_b = R_B \cdot (\tau_T / R_E) = K_i \cdot \tau_T$$

If one or more of the poles of an amplifier are in the lf region, then the bandwidth is set by the dominant lf pole. For fast amplifiers,

all of the poles affecting bandwidth are in the hf region and there are no dominant poles, making hf dynamic analysis unavoidable.

General Single-Stage CE Gain

The general single-stage CE voltage gain can be derived from basic circuit laws as the textbook CE with R_E added. It is a gargantuan grind but need only be done (right) once to produce a useful circuit template. The strategy is the same as for the textbook CE stage.



$$\text{KCL @ } b\text{: } \frac{v_i - v_b}{R_B} + \frac{v_o - v_b}{1/s \cdot C_c} + \alpha(s) \cdot i_e = i_e = \frac{v_b}{Z_e + R_E} \Rightarrow$$

$$\frac{v_i}{R_B} + s \cdot C_c \cdot v_o = \frac{v_b}{(1/s \cdot C_c) \parallel R_B} + (1 - \alpha) \cdot \frac{v_b}{Z_e + R_E}$$

$$\text{KCL @ } c\text{: } \frac{v_o}{R_L} + \frac{v_o - v_b}{1/s \cdot C_c} + \alpha(s) \cdot i_e = 0$$

$$\Rightarrow \frac{v_o}{(1/s \cdot C_c) \parallel R_L} = \frac{v_b}{\dfrac{1}{s \cdot C_c} \left\| \left(-\dfrac{Z_e + R_E}{\alpha(s)} \right) \right.}$$

where

$$Z_e = \frac{Z_\pi}{\beta_0 + 1} = \frac{1}{s \cdot C_e} \parallel r_e = \frac{r_e}{s \cdot r_e \cdot C_e + 1} = \frac{r_e}{s \cdot \alpha_0 \cdot \tau_T + 1}$$

C_e appears implicitly in the BJT time constant, $\tau_T = r_m \cdot C_e$ and $\alpha_0 \cdot \tau_T = r_e \cdot C_e$. Also, $\tau_\beta = \beta_0 \cdot \tau_T = r_\pi \cdot C_e = (\beta_0 + 1) \cdot r_e \cdot C_e$. Also, from the definition of $\beta(s)$,

$$\beta(s) = \frac{\beta_0}{s \cdot \tau_\beta + 1} \Rightarrow \beta(s) + 1 = (\beta_0 + 1) \cdot \frac{s \cdot \alpha_0 \cdot \tau_T + 1}{s \cdot \tau_\beta + 1}$$

Substituting v_b from the collector equation into the base equation and solving for the voltage gain,

$$A_v = \frac{v_o}{v_i} = \frac{1}{\left[(s \cdot R_B \cdot C_c + 1) + \frac{R_B}{(\beta(s)+1) \cdot (Z_e + R_E)} \right] \cdot \left[\frac{1}{s \cdot C_c} \middle\| \left(\frac{Z_e + R_E}{\alpha(s)} \right) \right]} - s \cdot R_B \cdot C_c$$
$$\frac{1}{s \cdot C_c} \middle\| R_L$$

Some of the sub-expressions within this imposing expression are worked out as follows.

$$Z_e + R_E = \frac{r_e + R_E \cdot (s \cdot \alpha_0 \cdot \tau_T + 1)}{s \cdot \alpha_0 \cdot \tau_T + 1} = (r_e + R_E) \cdot \frac{s \cdot \alpha_0 \cdot \tau_T \cdot \left(\frac{R_E}{r_e + R_E} \right) + 1}{s \cdot \alpha_0 \cdot \tau_T + 1}$$

$$(\beta(s) + 1) \cdot (Z_e + R_E) = (\beta_0 + 1) \cdot (r_e + R_E) \cdot \left(s \cdot \alpha_0 \cdot \tau_T \cdot \left(\frac{R_E}{r_e + R_E} \right) + 1 \right)$$

$$\frac{Z_e + R_E}{\alpha(s)} = (\beta_0 + 1) \cdot (r_e + R_E) \cdot \frac{s \cdot \alpha_0 \cdot \tau_T \cdot \left(\frac{R_E}{r_e + R_E} \right) + 1}{\beta(s) \cdot (s \cdot \tau_\beta + 1)}$$

$$\frac{1}{s \cdot C_c} \middle\| \left(- \frac{Z_e + R_E}{\alpha(s)} \right) = - \frac{r_e + R_E}{\alpha_0} \cdot \frac{s \cdot \alpha_0 \cdot \tau_T \cdot \left(\frac{R_E}{r_e + R_E} \right) + 1}{- s^2 \cdot \tau_T \cdot R_E \cdot C_c - s \cdot \frac{r_e + R_E}{\alpha_0} + 1}$$

$$\frac{1}{s \cdot C_c} \middle\| R_B = \frac{R_B}{s \cdot R_B \cdot C_c + 1} \; ; \; \frac{1}{s \cdot C_c} \middle\| R_L = \frac{R_L}{s \cdot R_L \cdot C_c + 1}$$

After some substitution and simplification,

$$A_v = \frac{R_L}{(s\cdot R_L\cdot C_c+1)\cdot\left((s\cdot R_B\cdot C_c+1)+\frac{R_B}{(\beta(s)+1)\cdot(Z_e+R_E)}\right)\cdot\left(-\frac{r_e+R_E}{\alpha_0}\cdot\frac{s\cdot\alpha_0\cdot\tau_T\cdot\left(\frac{R_E}{r_e+R_E}\right)+1}{-s^2\cdot\tau_T\cdot R_E\cdot C_c - s\cdot\frac{r_e+R_E}{\alpha_0}+1}\right)-s\cdot R_B\cdot C_c\cdot R_L}$$

In the reduction of the horrific denominator, cubic (s^3) terms appear but cancel. When flattened, the denominator of the denominator becomes the numerator and a final, more pleasing expression results;

$$A_v = -\alpha_0\cdot\frac{R_L}{r_M}\cdot\frac{-s^2\cdot\alpha_0\cdot\tau_T\cdot\frac{R_E}{\alpha_0}\cdot C_c - s\cdot(r_m+R_E/\alpha_0)\cdot C_c+1}{s^2\cdot[\tau_e\cdot(R_{c;e}\cdot C_c)]+s\cdot[\tau_e+\{R_L+R_b\cdot(1+\alpha_0\cdot\frac{R_L}{r_e+R_E})\}\cdot C_c]+1}$$

where $r_M = r_e + R_E + R_B/(\beta_0+1)$ and OCTC $\tau_e = \alpha_0\cdot\tau_T\cdot\frac{R_B+R_E}{r_M}$.

The OCTC of C_c, τ_c, is the second term in the linear coefficient. Also, in accord with circuit polynomials, the resistance of the SCTC of C_c with C_e shorted is

$$R_{c;e} = R_L + (R_B \parallel R_E)$$

The base node resistance is

$$R_b = R_B \parallel (\beta_0+1)\cdot(r_e+R_E)$$

For $r_m\cdot C_c >> \alpha_0\cdot\tau_T$, the RHP zero-pair has a dominant RHP zero at the time constant of the linear coefficient.

With R_E set to zero, A_v reduces to that of the textbook CE. What is different with $R_E \neq 0\ \Omega$ is not only the RHP zero-pair, which adds a zero term to the textbook CE of $(R_E/\alpha_0)\cdot C_c$. $\tau_{c;e}$ in a is not τ_{cc}, which is now a term of $\tau_{c;e.}$ τ_{cc} also appears in b but the second term of τ_c - the Miller term, τ_{cb} - is greater than the second term in $\tau_{c;e}$ which is $(R_B \| R_E)\cdot C_c$. Consequently, the poles are real. The general polynomial time constants are

$$\tau(a,b) = \tau_e$$
$$\tau(a) = R_{c;e}\cdot C_c = (R_L + R_B \parallel R_E)\cdot C_c = \tau_{cc} + (R_B \parallel R_E)\cdot C_c$$
$$\tau(b) = \tau_{cc} + R_b\cdot(1+K_v)\cdot C_c$$

CE Bandwidth

The OCTC linear pole coefficient for a BJT stage is

$$R_{be} \cdot C_e + (R_L + R_b \cdot (1 + K_v)) \cdot C_c$$

where R_b is the base node resistance. R_{bc} is composed of two terms, one associated with the collector node time constant component,

$$\tau_{cc} = R_L \cdot C_c$$

and the other with the base node,

$$\tau_{cb} = R_b \cdot [C_c \cdot (1 + K_v)] \ , \ K_v = \alpha_0 \cdot \frac{R_L}{r_e + R_E}$$

The Miller effect from the base-to-collector voltage gain effectively increases C_c at the base node by $(1 + K_v)$ as given in R_{bc}. The two terms of τ_c are not separable time constants but are components of the one time constant of C_c,

$$\tau_c = \tau_{cc} + \tau_{cb}$$

The other OCTC is

$$\tau_e = R_{be} \cdot C_e$$

The OCTCs of C_c and C_e can be rearranged as follows:

$$[R_{be} \cdot C_e + R_b \cdot C_c \cdot (1 + K_v)] + R_L \cdot C_c$$

Instead of being ordered by capacitance (as OCTCs), they are grouped by nodes *b* and *c*:

$$\tau_b = R_{be} \cdot C_e + R_b \cdot [C_c \cdot (1 + K_v)] = \tau_e + \tau_{cb}$$
$$\tau_{cc} = R_L \cdot C_c$$

The base-node resistance,

$$R_b = R_B \parallel [(\beta_0 + 1) \cdot (r_e + R_E)] = \frac{R_B \cdot (r_e + R_E)}{r_M}$$

is related to the R_{be} OCTC resistance by expressing

$$R_{be} = r_\pi \parallel \frac{R_B + R_E}{1 + \dfrac{R_E}{r_m}} = \frac{r_e \cdot (R_B + R_E)}{r_M}$$

Then combining equations,

$$\frac{1}{r_M} = \frac{R_b}{R_B \cdot (r_e + R_E)} = \frac{R_{be}}{(R_B + R_E) \cdot r_e} \Rightarrow R_b = \frac{R_B \cdot (r_e + R_E)}{(R_B + R_E) \cdot r_e} \cdot R_{be}$$

Multiplying the rational factor by R_E/R_E and applying the parallel-resistance formula, this reduces to

$$R_b = \left(\frac{R_B \parallel R_E}{r_e \parallel R_E} \right) \cdot R_{be} \Rightarrow R_{be} = \left(\frac{r_e \parallel R_E}{R_B \parallel R_E} \right) \cdot R_b$$

For the textbook CE case of $R_E = 0\ \Omega$, then $R_b = R_{be}$. For the more common case of $R_E >> r_e$ and large β_0, $R_b \approx R_B$ and

$$R_{be} \approx r_e\ ,\ R_E >> r_e, \beta_0 \to \infty$$

Then the OCTC of C_e is $\tau_e = R_{be} \cdot C_e \approx r_e \cdot C_e = \tau_T$. Thus τ_e in fast amplifiers, with $R_E >> r_e$ and small R_B, is near τ_T. Not uncommonly in fast amplifiers, $\tau_e < \tau_{cc}$, and to retain τ_e while omitting τ_{cc} results in a less accurate approximation.

By relating R_{be} to R_b, both terms in the base-node time constant τ_b can be expressed in R_b as

$$\tau_b = R_b \cdot \left[\left(\frac{r_e \parallel R_E}{R_B \parallel R_E} \right) \cdot C_e + (1 + K_v) \cdot C_c \right] = R_b \cdot C_b$$

The result is a single equivalent capacitance at the base node of

$$C_b = \left(\frac{r_e \parallel R_E}{R_B \parallel R_E} \right) \cdot C_e + (1 + K_v) \cdot C_c$$

τ_b is a time constant that applies to the base node as a port. When the base node loads a previous stage, C_b contributes to the previous stage as part of C_L.

By applying $\alpha_0 \cdot \tau_T = r_e \cdot C_e$, C_b becomes

$$C_b = \left(\frac{\alpha_0 \cdot \tau_T \cdot R_E}{(R_B \parallel R_E) \cdot (r_e + R_E)} \right) + (1 + K_v) \cdot C_c = \alpha_0 \cdot \left(\frac{\tau_T}{r_e + R_E} \right) \cdot \left(\frac{R_B + R_E}{R_B} \right) + (1 + K_v) \cdot C_c$$

The Miller term, $(1 + K_v)\cdot C_c$, from τ_{cb} dominates and the τ_e term only improves the approximation somewhat while τ_{cc} is absent. The single-pole approximation of bandwidth at the base node is valid for a dominant τ_{cb}, though a better single-pole approximation would include τ_{cc}, for it is not uncommonly about the same in value as τ_e, or greater for large R_L and large R_E (that makes $\tau_e \to \alpha_0 \cdot \tau_T$). The better single-pole approximation for the input-loop time constant is the OCTC bandwidth formula. Only when $\tau_{cc} << \tau_e$ and τ_{cb} does $R_b \cdot C_b$ offer an approximation of comparable accuracy.

The τ_b time constant is not an OCTC and instead mixes terms of different OCTCs. The difference in τ_{bw} that results from the OCTCs versus the *nodal TCs* when the root-sum-of-squares bandwidth is calculated is

OCTC

$$\tau_{bw} \approx \sqrt{\tau_e^2 + \tau_c^2} = \sqrt{\tau_{be}^2 + (\tau_{cb} + \tau_{cc})^2} = \sqrt{(\tau_{be}^2 + \tau_{cb}^2 + \tau_{cc}^2) + 2 \cdot \tau_{cb} \cdot \tau_{cc}}$$

Nodal TC

$$\tau_{bw} \approx \sqrt{\tau_b^2 + \tau_{cc}^2} = \sqrt{(\tau_e + \tau_{cb})^2 + \tau_{cc}^2} = \sqrt{(\tau_e^2 + \tau_{cb}^2 + \tau_{cc}^2) + 2 \cdot \tau_{cb} \cdot \tau_e}$$

The τ_{bw} differ in the last term under the radical in that $\tau_{cc} \neq \tau_e$. The nodal TC bandwidth is only accurate when $\tau_{cc} \approx \tau_e$. Equal bandwidth results only when

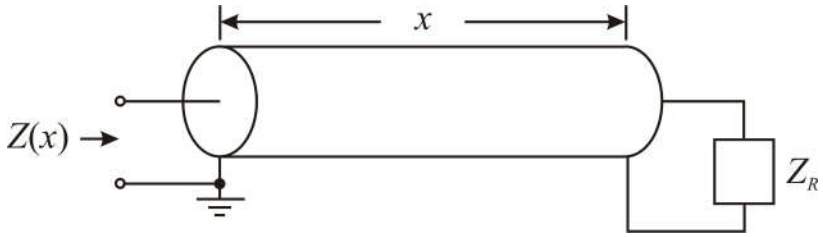
$$C_c = \frac{R_{be}}{R_L} \cdot C_e$$

The resulting C_c value for many BJT circuits happens to also be in the range of the right-side expression, which might hide the above discrepancy from being experimentally detected. While the nodal-TC bandwidth happens to be approximately correct much of the time, it is not theoretically correct and is thus not *necessarily* even approximately correct relative to the OCTC basis for bandwidth.

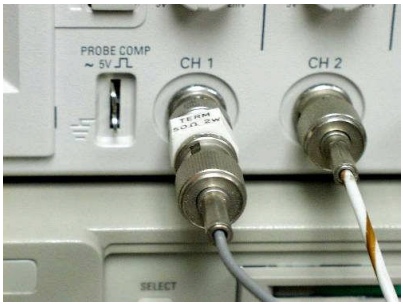
The OCTC bandwidth formula is based on real poles only. In the single-stage BJT amplifier, the poles are all real because poles of the τ_e and τ_c time constants are always real, as is the pole of τ_L. The quadratic coefficient, a, is always less than is required for $\zeta \geq 1$. What we have discovered in the textbook CE analysis is that zeros can possibly affect bandwidth. C_c bridges input and output loops and provides a path from input to output. It introduces a zero into the single-stage transfer function that sometimes should be included in dynamic analysis.

Transmission Lines

Wide-bandwidth or “high-speed” waveforms are delivered from a source to a separated load via *transmission lines*. These are a pair of conductors that along their length have distributed inductance in series with the conductors and distributed capacitance shunting them. The conductors also have series resistance, but if this is negligible, the line can be modeled as a *lossless* line, as shown, connected to a *terminating impedance* of Z_R at the load end of the line.



The symbol is that of a *coaxial* line where the inner conductor is surrounded by a cylindrical outer conductor that forms an electric shield for the inner wire.



Coaxial transmission lines are often terminated at both ends with quarter-turn twist-on and twist-off connectors called *BNC connectors* (as shown connected to an oscilloscope input), and the line is called a *coaxial* or “coax” (two syllables) *cable*. For Z_R, 50 Ω terminators have BNC connectors at each end (one male, one female), as shown.

The material between inner and outer conductors is usually a plastic in which the electric and magnetic fields of the cable are contained. The material has magnetic *permeability*, μ, of the distributed inductance per length,

$$\frac{L}{x} = \mu$$

and its distributed capacitance per length is the electric *permittivity* or *dielectric constant* of the material,

$$\frac{C}{x} = \varepsilon$$

Transmission lines have a *characteristic impedance*,

$$Z_n = \sqrt{\frac{\mu}{\varepsilon}} = \sqrt{\frac{L}{C}}$$

This is the same formula as for the resonant impedance of LC circuits. The other resonant-circuit parameter for transmission lines includes its spatial dimension as the *propagation speed*,

$$u_p = \frac{1}{\sqrt{\mu \cdot \varepsilon}} = \frac{x}{\sqrt{L \cdot C}}$$

It relates to wavelength and frequency by the wave propagation formula,

$$u_p = \lambda \cdot f = \frac{\lambda \cdot \omega}{2 \cdot \pi}$$

where λ is the *wavelength* of a propagating sine-wave and f is its frequency. Unlike discrete *lumped-parameter* circuit components that have impedances without regard to their spatial dimensions, a cable is a *distributed-parameter* impedance component for which the spatial dimension of distance, x, is significant. The *propagation time* of a sine-wave along the cable is

$$t_p = \frac{x}{u_p} = \left(\frac{2 \cdot \pi}{\lambda \cdot \omega}\right) \cdot x$$

For a free-space (no-material) medium, $u_p = u_c$, the speed of light, given as $u_c \approx 30$ cm/ns. This is a useful unit by which to remember u_c because it can be related directly at the electronics workbench to cables and circuits and their length dimensions.

Common coax cables have $Z_n = 50\ \Omega$. Those used in video applications have $Z_n = 75\ \Omega$. Other kinds of cable not designed for the transmission of high-speed waveforms also have transmission-line

parameters. For instance, SO4 cable is four-conductor power cable for three-phase electric-power transmission. SO4-16 (16 gage in the American Wire Gage, or AWG) has (as measured)

$$\mu \approx 0.75\ \mu\text{H/m}\ ;\ \varepsilon \approx 100\ \text{pF/m}$$

Then using the above formulas,

$$Z_n \approx 87\ \Omega\ ;\ u_p \approx 1.15 \cdot 10^8\ \text{m/s} = 11.5\ \text{cm/ns} = u_c/2.6$$

This value of u_p is lower than is typical of waveform propagation speed in fast amplifiers and coax cables, which is about $\frac{2}{3} \cdot u_c$ or $u_c/1.5$.

Coaxial (or "barrel") attenuators connect to 50 Ω terminators to provide 50 Ω input impedance. Commonly available attenuations are ×2, ×5, and ×10. The "×*N*" designation is used because on an oscilloscope screen, the waveform at the *input* to the attenuator is *N* times larger than the displayed waveform. The impedance (which is a resistance) is set to Z_n to prevent waveform reflections or sine standing waves on the transmission line. In the time domain, a step of voltage or current applied to the line will propagate to the other end, and if it is not terminated in Z_n, a reflection (of voltage or current) occurs that returns to the source end of the line.

As a circuit component, a transmission-line has an impedance that depends on its length, *x*;

$$Z(x) = Z_n \cdot \frac{Z_R + j \cdot Z_n \cdot \tan\left(\frac{2 \cdot \pi}{\lambda} \cdot x\right)}{Z_n + j \cdot Z_R \cdot \tan\left(\frac{2 \cdot \pi}{\lambda} \cdot x\right)} = Z_n \cdot \frac{Z_R + j \cdot Z_n \cdot \tan\left(\omega \cdot \frac{x}{u_p}\right)}{Z_n + j \cdot Z_R \cdot \tan\left(\omega \cdot \frac{x}{u_p}\right)}$$

A shorted line ($Z_R = 0\ \Omega$) reduces $Z(x)$ to an inductive reactance of

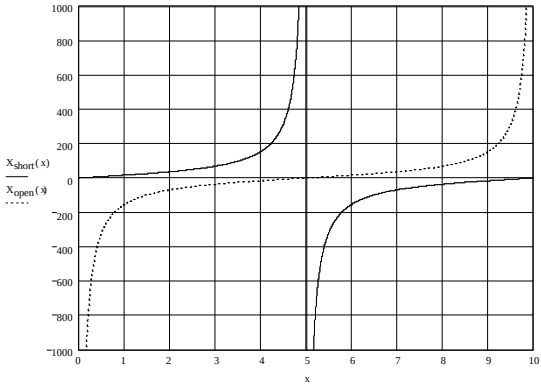
$$X_{short} = Z_n \cdot \tan\left(\frac{2 \cdot \pi}{\lambda} \cdot x\right) \approx Z_n \cdot \left(\frac{2 \cdot \pi}{\lambda}\right) \cdot x\ ,\ x << \lambda$$

An open line ($Z_R \to \infty$) has a reactance of

$$X_{open} = -\frac{Z_n}{\tan\left(\frac{2 \cdot \pi}{\lambda} \cdot x\right)} \approx -\frac{Z_n}{\left(\frac{2 \cdot \pi}{\lambda}\right) \cdot x},\ x << \lambda$$

These reactances are plotted below for $u_p = u_c/1.5$ and $f = 10$ MHz. Then $\lambda = u_p/f = 20$ m and x is plotted in meters.

At $\lambda/4$ (5 m), reactance is infinite - the equivalent of a parallel resonant circuit. At $\lambda/2$, the dual is that of a series resonant circuit and the impedance is zero. For an open line, series resonance occurs at $\lambda/4$ and parallel resonance at $\lambda/2$. If terminated in Z_n, it will be resistive of value Z_n independent of length.



Under a quarter-wavelength, an open line is capacitive (negative reactance, dotted curve). Combining u_p and Z_n, distributed (over x) inductance and capacitance can be expressed as

$$\frac{L}{x} = \mu = \frac{Z_n}{u_p} \; ; \; \frac{C}{x} = \varepsilon = \frac{1}{Z_n \cdot u_p}$$

For shorted or open line lengths much shorter than a wavelength,

$$L = \frac{X_L}{\omega} \approx Z_n \cdot \frac{x}{u_p} = Z_n \cdot \left(\frac{2 \cdot \pi}{\lambda \cdot \omega} \right) \cdot x \, , x << \lambda, \; Z_R = 0 \; \Omega$$

$$C = \frac{1}{-X_C \cdot \omega} \approx \frac{x / u_p}{Z_n} = \frac{\left(\frac{2 \cdot \pi}{\lambda \cdot \omega} \right) \cdot x}{Z_n} \, , x << \lambda, Z_R \to \infty \; \Omega$$

Transmission lines can be used to delay a waveform. Analog oscilloscope vertical amplifiers use a transmission line to delay the vertical waveform. This gives the horizontal-axis processing (the *time-base*) time to start the sweep before the leading edge of a step waveform arrives at the vertical deflection plates. The transmission line is used as a *delay line*.

A delay of $t_d = t_p$ appears in amplifier circuits and has an *s*-domain function of

$$e^{-s\cdot t_d} = \frac{1}{e^{s\cdot t_d}} \approx \frac{1}{s^2\cdot t_d^2 + s\cdot t_d + 1}$$

The exponential can be expanded into an infinite series of terms as a polynomial but the result is an infinite number of poles. A delay cannot be expressed exactly in the form of a rational function with a finite number of factors. It can only be approximated (as it is above) as a quadratic pole with pole locations in the *s*-domain at

$$p_{1,2} = \left(\frac{1}{t_d}\right)\cdot(-1 \pm j\cdot 1) = -\frac{1}{t_d}\cdot e^{\pm j\cdot\frac{\pi}{4}}$$

This complex pole-pair has a pole angle of 45° (MFA response) and a real frequency (the real component of *s* on the σ-axis) of $-1/t_d$. A better (*Padé*) approximation is

$$e^{-s\cdot t_d} = \frac{s^2\cdot\left(\frac{t_d}{2\cdot\sqrt{3}}\right)^2 - s\cdot(\frac{t_d}{2}) + 1}{s^2\cdot\left(\frac{t_d}{2\cdot\sqrt{3}}\right)^2 + s\cdot(\frac{t_d}{2}) + 1}$$

for which

$$p_{1,2} = \left(\frac{1}{t_d}\right)\cdot(-3 \pm j\cdot\sqrt{3})\ ;\ z_{1,2} = \left(\frac{1}{t_d}\right)\cdot(+3 \pm j\cdot\sqrt{3})$$

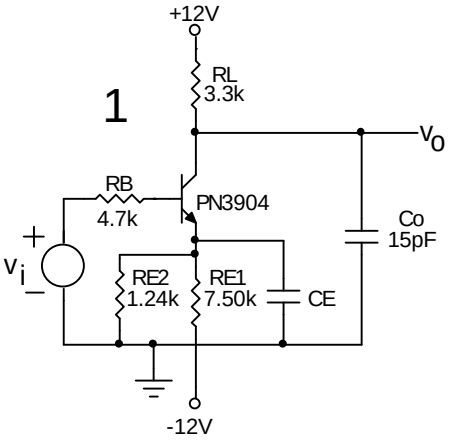
The quadratic zero-pair is in the RHP and forms an all-pass filter with the pole-pair. Both have pole angles of 30° (MFED response) and

$$\omega_{np} = \omega_{nz} = 2\cdot\sqrt{3}\cdot\left(\frac{1}{t_d}\right) \approx 3.46/t_d$$

Delay lines in amplifiers have a dynamic response that can only be approximated as a rational function in *s*. The exact delay-line transfer function is the *complementary error function* (erfc(*t*)). Its graph looks like an overdamped step function and can be approximately compensated with emitter peaking using more than one series RC compensator in parallel in the emitter circuit. In oscilloscopes, two or three adjustable time constants remove the erfc effects of the vertical-amplifier delay line on both step and sine-wave response.

CE Amplifier Design Refinement

Our emphasis now begins to transition from circuit analysis to design. Analysis is still dominant, but it is oriented toward finding parts values in circuits and determining circuit performance. The following CE amplifier (circuit number 1) has an output capacitance (C_o = 15 pF) of a typical ×10 oscilloscope probe for an oscilloscope of 100 MHz or less. Our goal for this circuit is to improve its bandwidth through a sequence of design refinements that demonstrate various aspects of amplifier design.



The starting-point is with the static design. The collector voltage range is maximum when the static collector voltage, V_C, is set so that $\Delta v_C = v_c$ has the same magnitude for both polarities. For range symmetry, $-v_b$ will change v_C by the same amount, opposite polarity, as $+v_b$, or

$$\Delta v_{C+} = -\Delta v_{C-}$$

$$|\Delta v_C| = K_v \cdot v_b = V_{CC} - V_C = V_C - v_b$$

where K_v is the |voltage gain| from base to collector. For $-v_b$, $\Delta v_C = V_{CC} - V_C$ and for $+v_b$, $\Delta v_C = -(V_C - v_b)$. The negative range lower limit is set by v_{CB} = 0 V. Solving for v_b,

$$v_b = \frac{V_C}{K_v + 1}$$

Substituting for v_b,

$$V_{CC} - V_C = V_C - \left(\frac{V_C}{1 + K_v} \right) = V_C \cdot \left(\frac{K_v}{1 + K_v} \right)$$

Solving for V_C,

$$V_C = \left(\frac{1 + K_v}{1 + 2 \cdot K_v} \right) \cdot V_{CC}$$

For large K_v, $V_C \approx V_{CC}/2$. For $K_v = 3$, $V_C = (4/7) \cdot V_{CC} = 6.86$ V so that $\Delta v_C = \pm 5.14$ V. Set $I_C = 1$ mA. Then $R_L = 5.14\ \text{k}\Omega \rightarrow 5.1\ \text{k}\Omega$, 5 %.

Having dynamic design in mind, the maximum-range R_L might not be the optimal choice because at the low end of the v_C range, where $v_{CB} = 0$ V, $C_c = 3.5$ pF. At 1.8 V it is down to 2.5 pF, enough of a reduction to motivate a trade of range for speed. Choose to set the minimum $v_{CB} = 1.8$ V. Then V_C is raised by 1.8 V above the max-range design to $V_C = 8.7$ V, $R_L = 3.3\ \text{k}\Omega$, and $\Delta v_C = \pm 3.3$ V. At this V_C, $C_c \approx 1.8$ pF. The collector bias current is 1 mA. The PN3904 has a typical $\beta_0 = 150$ in this current range. Then $\alpha_0 = 0.9934$, the emitter current is $I_E = 1.007$ mA, and $r_e = 26\ \text{mV}/I_E = 25.8\ \Omega$. The base-to-collector voltage gain magnitude is chosen to be

$$K_v = 3 = \alpha \cdot \frac{R_L}{r_e + R_E} = (0.9934) \cdot \frac{3.3\ \text{k}\Omega}{25.8\ \Omega + R_E}$$

which makes C_c appear as 7.2 pF at the base. The emitter-referred R_B is omitted from r_M because the gain sought is from the base, not from v_i. Solving for R_E, the Thevenin equivalent emitter resistance is $R_E = R_{E1} \| R_{E2} = 1.067\ \text{k}\Omega$. $V_{BE} \approx 0.65$ V and the Thevenin supply voltage must be set to 0.65 V + $I_E \cdot R_E = 1.72$ V.

$$1.72\ \text{V} = \frac{R_{E2}}{R_{E1} + R_{E2}} \cdot V_{EE} = \frac{R_E}{R_{E1}} \cdot V_{EE}$$

Substituting values and solving,

$$R_{E1} = \frac{V_{EE}}{1.72\ \text{V}} \cdot R_E = 7.425\ \text{k}\Omega \rightarrow 7.50\ \text{k}\Omega, 1\%$$

Then $R_{E2} = 1.255\ \text{k}\Omega \rightarrow 1.24\ \text{k}\Omega$, 1 %, and with these values, $R_E = 1064\ \Omega$. The quasistatic gain is

$$A_v = \frac{v_o}{v_i} = -\alpha_0 \cdot \frac{R_L}{r_E + R_E} = -\frac{3278\ \Omega}{1121\ \Omega} = -2.92$$

where

$$r_E = r_e + R_B/(\beta_0 + 1) = 25.8\,\Omega + 4.7\,\text{k}\Omega/151 = 56.93\,\Omega\,;\; r_M = 1121\ \Omega$$

CE Dynamics

Using the dynamic BJT model with $r_b' = 50\ \Omega$, the poles (inverted as OCTC time constants) are real ($C_E = 0$ pF) and, by component, are

$$\tau_e = R_{be} \cdot C_e = (\alpha_0 \cdot \tau_T) \cdot \left(\frac{R_B' + R_E}{r_M} \right)$$

$$= (0.993) \cdot (707\ \text{ps}) \cdot \left(\frac{5814\,\Omega}{1121\,\Omega} \right) = 3.64\,\text{ns}$$

$$\tau_{cb} = R_B' \cdot C_c \cdot (1 + K_v) = (4.75\,\text{k}\Omega) \cdot (1.8\,\text{pF}) \cdot (4) = 34.2\,\text{ns}$$

$$\tau_{cc} = R_L \cdot C_c = (3.3\,\text{k}\Omega) \cdot (1.8\,\text{pF}) = 5.94\,\text{ns}$$

where

$$R_B' = R_B + r_b'$$

The OCTCs are

$$\tau_e = \tau_{be} = 3.64\,\text{ns}$$

$$\tau_c = \tau_{cb} + \tau_{cc} = 40.14\,\text{ns}$$

Including C_o, the collector time constant is

$$\tau_L = R_L \cdot C_o = (3.3\ \text{k}\Omega) \cdot (15\ \text{pF}) = 49.5\ \text{ns} \Rightarrow 3.22\ \text{MHz}$$

In the base circuit, the source cable from the generator driving the prototype circuit input is about 1 m in length. The input resistance is high and can be approximated as open. From transmission-line theory, a length of less than a quarter-wavelength ($\lambda/4$) will be capacitive. The frequency up to which the 1 meter line is capacitive is

$$f_{\lambda/4} = \frac{u_p}{\lambda/4} = 4 \cdot \frac{u_c/1.5}{1\,\text{m}} = 4 \cdot \frac{2 \cdot 10^8\ \text{m/s}}{1\,\text{m}} = 800\ \text{MHz}$$

This exceeds f_T of the BJT and the cable can be regarded as a capacitance to ground. The cable capacitance was measured to be $C_l = 140$ pF, and connects to a generator of 50 Ω output resistance. The equivalent source input circuit is an RC integrator with a pole time constant of

$$\tau_i = R_G \cdot C_l = (50\ \Omega)\cdot(140\ \text{pF}) = 7.00\ \text{ns}$$

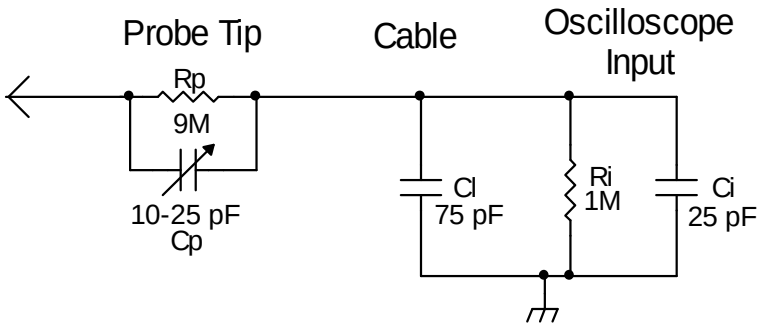
These capacitances result in a bandwidth time constant of

$$\tau_{bw} \approx \sqrt{\tau_e^2 + \tau_c^2 + \tau_L^2 + \tau_i^2} = 64.21\,\text{ns} \Rightarrow f_{bw} = 2.48\ \text{MHz}$$

The prototype circuit measured f_{bw} = 2.1 MHz. The calculated value is 15 % high. The prototype had an extra 2 pF of C_L that brings the theoretical calculation down to 2.29 MHz, or 9 % high.

Probe Loading

In reviewing the four time constants, it is obvious that τ_L has a major effect on bandwidth because of the probe loading at the output. A ×10 probe is preferred to a ×1 probe because its compensated voltage divider allows probe and oscilloscope to achieve a much higher bandwidth. The divider is shown below.



The probe tip contains a parallel RC with a time constant adjusted to be the same as the lower part of the divider: $R_p \cdot C_p = R_i \cdot (C_l + C_i)$. The time constant is typically

$$(1\ \text{M}\Omega)\cdot(75\ \text{pF} + 25\ \text{pF}) = 100\ \mu\text{s} \Rightarrow 1.59\ \text{kHz}$$

Then the adjusted capacitance of the probe is

$$C_p = \frac{100\,\mu\text{s}}{9\,\text{M}\Omega} = 11.1\,\text{pF}$$

C_p in series with $C_l + C_i$ is 10 pF. A Tektronix P6112 100 MHz ×10 probe connected to an oscilloscope with a 1 MΩ, 20 pF input specification measures about 15 pF on an RLC meter. This

demonstrates that instrumentation effects on measurement must always be kept in mind for circuit design in the laboratory.

Power Supply Bypassing

Not shown on the circuit diagram are 0.1 μF to 1 μF ceramic capacitors placed from supply terminals on the prototype board to ground. These *bypass capacitors* provide a dynamic return path through the supply for fast changes in current. Bypass capacitors should be added to all prototype circuits, as needed. They compensate for the imperfection of the power supplies as non-ideal voltage sources at the circuit. The leads that connect the circuit-board supply terminals to the power supply are inductive, and this inductance interferes with desired circuit operation by allowing different circuit loops powered from a common supply to interact through the common reactance of the supply.

Additionally, even if the supply leads were ideal (of zero impedance), practical voltage sources are usually regulated supplies with feedback loops in them of limited bandwidth. As load current varies with frequency, the supply cannot respond quickly enough to maintain a constant voltage output. Bypass capacitors are used as a passive remedy in supplies, but they too have limited decreasing reactance with frequency.

Capacitors have parasitic series inductance, mostly in their leads. Even chip capacitors have some series inductance. All capacitors also have some series resistance. These parasitic elements form a damped series-resonant circuit with the capacitance. The resonant frequency, f_r, determines the useful bandwidth of the capacitor as a capacitance. Above f_r, the capacitor appears inductive. Electrolytic capacitors have the highest charge density but also have the lowest f_r of typically from 100 kHz to 5 MHz. That is why ceramic capacitors are used for high-frequency bypassing. Their typical f_r range is 10 to 100 MHz. Plastic film capacitors resonate between 10 to 50 MHz and are larger and more expensive than ceramic capacitors but they are the best choice for precision applications such as op-amp integrators.

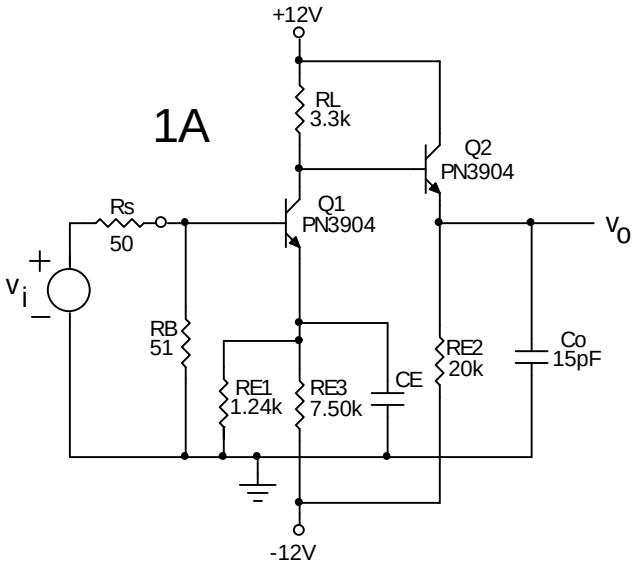
Most capacitors have enough series resistance to damp the internal resonance though they can still resonate with external circuitry. One scheme for wideband bypassing is to use an aluminum electrolytic capacitor of 10 to 47 μF in parallel with a ceramic

capacitor of 0.1 to 1 μF. If the combination resonates, some series resistance in the branch of the ceramic capacitor might be required.

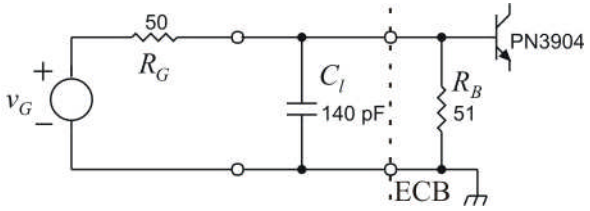
The circuit diagrams of amplifiers that appear in this book will not include the bypass capacitors to keep them less cluttered. However, they must be added for supply voltages to remain constant.

Buffered CE with Input Termination

To alleviate the C_o loading problem and at the same time eliminate τ_i, a revised circuit has a terminated transmission-line input and an emitter-follower (CC) stage to buffer the collector node from C_o. It is shown below as amplifier 1A.



In the bench setup, the prototype circuit is connected to a pulse generator via a 1 meter, 50 Ω cable as shown below.



The generator 50 Ω output is terminated on the *etched circuit-board* (ECB) end of the cable with R_B, a 51 Ω, 5 % resistor. Now that the cable is terminated in its characteristic impedance of 50 Ω, the 140 pF cable capacitance combines with distributed series cable inductance to cause the cable to appear at its characteristic impedance as a 50 Ω resistance at both (terminated) ends of the cable. The source consequently appears from the amplifier input as having a resistance of 50 Ω.

At the output, the CC output resistance, r_o, is much lower than R_L and the RC time constant formed with the scope probe is much smaller. It has an output time constant and bandwidth of about

$$r_o \cdot C_o = (3.3\ \text{k}\Omega/151 + 26\ \Omega)\cdot(15\ \text{pF}) = 0.718\ \text{ns} \Rightarrow 222\ \text{MHz}$$

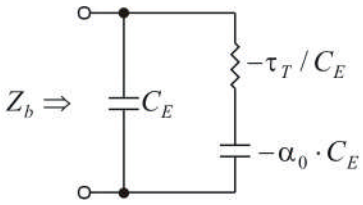
PN3904 f_T at 1 mA is about 225 MHz and τ_T = 707 ps. This pole is of little consequence; being near f_T, it will be much greater than f_{bw}.

At the buffered CE collector node, C_o is now replaced by the CC input impedance, Z_{b2}. Consider first the referral of R_{E2} to the CC base. In the lf region, it is (applying the β_0 transform) $(\beta_0 + 1)\cdot R_{E2}$ or about 3 MΩ and can be neglected. In the hf region, it is τ_T/R_{E2} in series with R_{E2}. The contribution to τ_L from R_{E2} is thus

$$(R_L + r_{b2}' + R_{E2})\cdot(\tau_T/R_{E2}) = (23350\ \Omega)\cdot(35.35\ \text{fF}) = 0.825\ \text{ns} \Rightarrow 193\ \text{MHz}$$

The value of base-referred hf C (0.035 pF) is tiny because R_{E2} is large, and can also be neglected. C_o refers to the base in the lf region as $C_o/(\beta_0 + 1) \approx 100$ fF - also a small capacitance for this circuit.

The total-frequency model of Z_{b2} is repeated below, where $C_E \approx C_o$, as derived previously. (See "High-Frequency β Transform".)



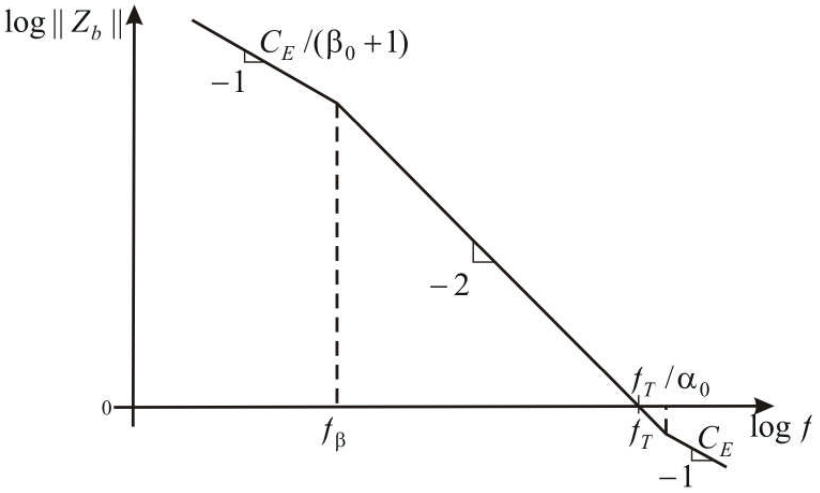
In the lf region, the series resistance, $-\tau_T/C_E$ is small (–47 Ω) and C_E is essentially in parallel with $-\alpha_0\cdot C_E$. Adding them,

$$C_E - \alpha_0 \cdot C_E = C_E \cdot (1 - \alpha_0) = \frac{C_E}{\beta_0 + 1}$$

This is C_E referred to the base in the lf region by $\beta_0 + 1$ (without gyration). The series –R, –C branch has a time constant of $\alpha_0 \cdot \tau_T$. Solving for the total-frequency base impedance of the above circuit,

$$Z_b(s) = \frac{s \cdot (\alpha_0 \cdot \tau_T) + 1}{\left[s \cdot \left(\frac{C_E}{\beta_0 + 1} \right) \right] \cdot (s \cdot \tau_\beta + 1)}$$

Its reactance plot is shown below; for amplifier 1A, $C_E = C_o$. The CC buffer isolates the base node from C_o by reducing it to $C_o/(\beta_0 + 1)$ in the lf region. This is the pole at the origin in Z_{b2}.



The β bandwidth of $f_\beta = 1.5$ MHz is less than the bandwidth of the amplifier and hf effects are significant. The time constant is

$$\tau_\beta = 1/\omega_\beta = 1/2 \cdot \pi \cdot f_\beta = 1/2 \cdot \pi \cdot (1.5\ \text{MHz}) = 106\ \text{ns}$$

A pole at f_β appears in Z_b. At the high end of the hf region, effective base capacitance is $C_E = C_o$ at f_T/α_0 and buffering is no longer effective in reducing the value of C_E.

Z_{b2} forms a divider with R_L. It can be either voltage- or current-driven; the result is the same:

$$\frac{Z_{b2}}{R_L + Z_{b2}} = \frac{s \cdot (\alpha_0 \cdot \tau_T) + 1}{s^2 \cdot [R_L \cdot \frac{C_E}{\beta_0 + 1} \cdot \tau_\beta] + s[R_L \cdot \frac{C_E}{\beta_0 + 1} + \alpha_0 \cdot \tau_T] + 1}$$

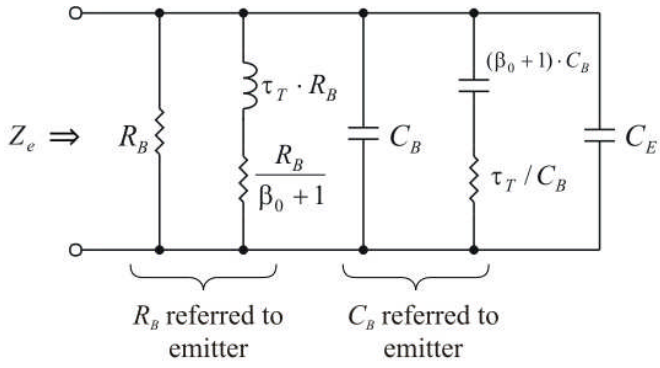
This reduces to

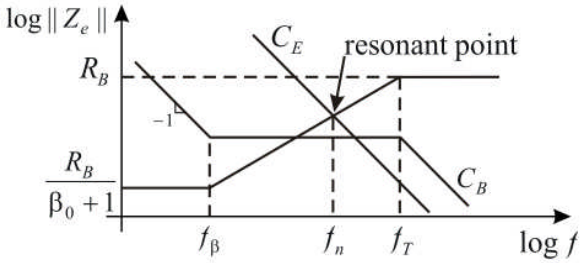
$$\frac{Z_{b2}}{R_L + Z_{b2}} = \frac{s \cdot (\alpha_0 \cdot \tau_T) + 1}{s^2 \cdot [(\alpha_0 \cdot \tau_T) \cdot R_L \cdot C_E] + s[\alpha_0 \cdot \tau_T + R_L \cdot \frac{C_E}{\beta_0 + 1}] + 1}$$

Both $\alpha_0 \cdot \tau_T$ and $R_L \cdot C_E$ appear as factors in a (the s^2 coefficient) and terms in b, but $R_L \cdot C_E$ is divided by $\beta_0 + 1$ in b. Were it not for $\beta_0 + 1$, both would be the time constants of real poles. Thus a is $\beta_0 + 1$ times larger than the product of the b terms and this results in complex poles. The quadratic pole-pair has

$$\tau_n = \sqrt{(\alpha_0 \cdot \tau_T) \cdot R_L \cdot C_E} \;;\; \zeta = \frac{\alpha_0 \cdot \tau_T + R_L \cdot \frac{C_E}{\beta_0 + 1}}{2 \cdot \tau_n} = \frac{\tau_\beta + R_L \cdot C_E}{2 \cdot (\beta_0 + 1) \cdot \tau_n}$$

The pole magnitude is the geometric mean of (or on log plots, the average distance between) $\alpha_0 \cdot \tau_T$ and $R_L \cdot C_E$, or in frequency, 226.5 MHz and 3.215 MHz, which is f_n = 27.0 MHz. The damping is ζ = (515.1 ps)/(5.898 ns) = 0.0873. The pole-pair is almost undamped because of C_E in the hf region.





However, at the Q1 collector is also $C_L \approx C_{c2} = C_c$(Q2) which is about 2 pF. Base capacitance tends to damp hf resonance caused by

C_E. The pole formed by R_L and C_{c2} is at 24.1 MHz. This is less than f_n = 27 MHz of the pole-pair and will damp it because the gyrated C_B is a resistance that is less than the resonant impedance on the reactance plot, as shown. (Compare to the plot on page 177.)

The hf damping, including $C_B = C_L$ in the hf approximation, is

$$\zeta_{hf} = \frac{\tau_B + \tau_T}{2 \cdot \sqrt{(R_L \cdot (C_E + C_B) \cdot \tau_T}} \; ; \; \tau_B = R_L \cdot C_B$$

Although C_{c1} is not part of the OCTC of $C_L = C_{c2}$, it is part of the collector node capacitance, the capacitance that damps the Q2 hf resonance. Thus Q2 C_B is $C_{c1} + C_{c2}$ = 4 pF. Substituting values of τ_B = 13.2 ns and τ_T = 707 ps, ζ_{hf} = 1.044. The resonance is gone leaving poles almost critically damped by the parasitic collector capacitance of the two BJTs. The two real poles are at

$$p_{1,2} = -\zeta \cdot f_n \cdot \left(1 \pm \frac{\sqrt{\zeta^2 - 1}}{\zeta}\right) = 21.33 \text{ MHz}, 36.29 \text{ MHz}$$

$$\Rightarrow 7.461 \text{ ns}, 4.386 \text{ ns}$$

The input is driven by a 50 Ω source, a typical resistance from function and pulse generator outputs. The amplifier input is terminated in 51 Ω, 5 %. This eliminates the input RC integrator and τ_i. The input resistance is about 25 Ω. With r_b' = 50 Ω for PN3904 BJTs, the base resistance is now

$$R_{B1}' = R_s \parallel R_{B1} + r_b' \approx 25\,\Omega + 50\,\Omega = 75\,\Omega$$

Reworking the equations for Q1 with R_B' = 75 Ω (C_E = 0 pF),

$$\tau_e = \tau_{be} = R_{be} \cdot C_e = (\alpha_0 \cdot \tau_T) \cdot \left(\frac{R_{B1}' + R_E}{r_M}\right) = (0.993) \cdot (707 \text{ ps}) \cdot \left(\frac{1149\,\Omega}{1101\,\Omega}\right) = 0.733 \text{ns} \quad \Rightarrow 217 \text{ MHz}$$

$$\tau_{cb} = R_{B1}' \cdot C_{c1} \cdot (1 + K_{v1}) = (75\,\Omega) \cdot (1.8 \text{ pF}) \cdot (4) = 0.54 \text{ ns} \;\Rightarrow 295 \text{ MHz}$$

Without the effect of Z_{b2},

$$\tau_{cc} = R_L \cdot C_{c2} = (3.3 \text{ k}\Omega) \cdot (1.8 \text{ pF}) = 5.94 \text{ ns} \;\Rightarrow 26.8 \text{ MHz}$$

An additional pole occurs at the collector, with Z_{b2} shunting C_{c2}, in addition to the single collector pole of τ_L.

The OCTCs of the first-stage BJT are

$$\tau_{e1} = \tau_{be1} = 0.733 \text{ ns} \Rightarrow 217 \text{ MHz}$$

$$\tau_{c1} = \tau_{cb1} + \tau_{cc1} = 6.48 \text{ ns} \Rightarrow 24.6 \text{ MHz}$$

The bandwidth of the Q1 BJT alone is the reciprocal of

$$\tau_{bw1} \approx \sqrt{\tau_{e1}^2 + \tau_{c1}^2} = 6.52 \text{ ns} \;\Rightarrow\; f_{bw} = 24.4 \text{ MHz}$$

If Q1 were left unbuffered by Q2 and connected to C_L = 15 pF at the Q1 collector, then

$$\tau_{L1} = R_L \cdot C_o = (3.3 \text{ k}\Omega)\cdot(15 \text{ pF}) = 49.5 \text{ ns} \;\Rightarrow 3.215 \text{ MHz}$$

The Q1 bandwidth in the unbuffered case is

$$\tau_{bw1} \approx \sqrt{\tau_e^2 + \tau_c^2 + \tau_L^2} = 49.93 \text{ ns} \;\Rightarrow f_{bw} = 3.2 \text{ MHz}$$

The circuit prototype of amplifier 1A measured f_{bw} = 3.2 MHz at the Q1 collector. With R_{B1} reduced from amplifier 1, collector dynamics associated with τ_L now dominates τ_{bw1}. With Q2 connected, collector quadratic poles are at 7.461 ns and 4.386 ns;

$$\tau_{bw1} \approx \sqrt{\tau_e^2 + \tau_c^2 + \tau_{L1}^2 + \tau_{L2}^2} = 10.84 \text{ ns} \Rightarrow f_{bw} = 14.7 \text{ MHz}$$

The prototype measured f_{bw} = 12.5 MHz.

CE with Emitter Peaking

The transresistance method extends to dynamic gain as the *transimpedance method*. Applied to the CE amplifier, the CE external emitter impedance is in series with the emitter port resistance,

$$r_E = R_B' / (\beta_0 + 1) + r_e = (75\,\Omega)/151 + 26\,\Omega = 26.5\,\Omega$$

The transimpedance with emitter peaking of C_E is

$$z_M = r_E + \frac{R_E}{s \cdot R_E \cdot C_E + 1} = (r_E + R_E) \cdot \frac{s \cdot (r_E \parallel R_E) \cdot C_E + 1}{s \cdot R_E \cdot C_E + 1}$$

The addition of C_E has added a pole and a zero to the transimpedance. The time constant of the zero is less than that of the pole because

($r_E \| R_E$) < R_E. For $R_E >> r_E$, the pole and zero are well-separated in frequency and compensation is effective.

At the collector of the CE amplifier, the impedance formed by R_L and C_L is

$$Z_L = \frac{R_L}{s \cdot R_L \cdot C_L + 1}$$

Combining Z_L with z_M, the amplifier lf gain is

$$A_v = -\alpha_0 \cdot \frac{Z_L}{z_M} = -\alpha_0 \cdot \frac{R_L}{r_E + R_E} \cdot \frac{s \cdot R_E \cdot C_E + 1}{[s \cdot (r_E \| R_E) \cdot C_E + 1] \cdot [s \cdot R_L \cdot C_L + 1]}$$

The quasistatic A_{v0} is independent of s and is the left factor. The transresistance pole and zero, being in the denominator of the gain expression, have become its zero and pole. The dominant collector-node pole of τ_{L1} = 7.46 ns can be cancelled (or *compensated*) by the emitter-circuit zero. Setting the two time constants equal,

$$R_E \cdot C_E = \tau_c = 7.46 \text{ ns} \Rightarrow 21.3 \text{ MHz}$$

The design value for C_E is

$$C_E = \frac{R_L}{R_E} \cdot C_L = \frac{\tau_L}{R_E} = \frac{7.46 \text{ ns}}{1.067 \text{ k}\Omega} = 7.0 \text{ pF} \rightarrow 10 \text{ pF, 5 \%}$$

The value is chosen to be somewhat larger to place the zero at a lower frequency. This partly compensates for the combined rolloff effect of both τ_L poles. The remaining higher-frequency pole of the emitter compensation is at a frequency with time constant

$$(r_E \| R_E) \cdot C_E = (26.5\ \Omega) \cdot (47 \text{ pF}) = 1.25 \text{ ns} \Rightarrow 128 \text{ MHz}$$

This uncancelled pole is within the hf region but is well over a decade above the uncompensated bandwidth and is not influential.

We already took into account the interstage loading poles in the hf region between f_β and f_T. As β varies with frequency, the Q1 base impedance gyrates by +90° so that R_{B1}' referred to the emitter appears inductive with a value of

$$R_{B1}' \cdot \tau_T = (75\ \Omega) \cdot (707 \text{ ps}) = 53.0 \text{ nH}$$

shunted by R_{B1}'. The Miller capacitance, from base node to ground through the collector, of $(1 + K_v) \cdot C_c$ = 7.2 pF gyrates at the emitter into a resistance of

$$\tau_T / (7.2 \text{ pF}) = 98.2\ \Omega$$

in parallel with 7.2 pF at the emitter.

If C_E were significantly increased, what would be the effect? Combining C_E = 47 pF with R_{B1}' = 75 Ω, the emitter-referred hf inductor resonates at

$$\tau_n = \sqrt{(\tau_T \cdot R_B') \cdot (C_E + (1+K_v) \cdot C_c)} = 1.693\,\text{ns} \Rightarrow 94\,\text{MHz}$$

This is within the BJT hf range. Its resonant impedance is

$$Z_n = \sqrt{\frac{\tau_T \cdot R_B'}{C_E + (1+K_v) \cdot C_c}} = \sqrt{\frac{53.0\,\text{nH}}{54.2\,\text{pF}}} = 31.3\ \Omega$$

As C_E increases, Z_n decreases and damping decreases. A shunt resistance of half this value, or 15.7 Ω, is required to critically damp the parallel resonance. The circuit shunt resistance is

$$R_E||\tau_T/(1 + K_v)\cdot C_c||R_{B1}' = 1064\ \Omega||98.2\ \Omega||75\ \Omega = 40.9\ \Omega$$

The CE BJT input loop is underdamped and

$$\zeta_p = \frac{1}{2} \cdot \frac{Z_n}{R} = \frac{(31.3\,\Omega)/2}{40.9\,\Omega} = 0.383 \Rightarrow \phi = 67.5°$$

with a pole angle of 67.5°. In the next chapter, the bandwidth of a quadratic pole-pair is derived as a function of f_n and ζ. With the quadratic bandwidth formula, we can calculate the extension of bandwidth caused by the underdamped, or peaked, response to be $1.39 \cdot f_n$ or $\tau_n/1.39$ = 1.218 ns ⇒ 131 MHz.

The step response of this pole-pair has about a 2-cycle ring as would be viewed on an oscilloscope - a rather underdamped circuit for a linear amplifier. The fractional peak overshoot is

$$M_p = e^{-\pi/\tan\phi} \cong 0.27 = 27\,\%$$

In the frequency domain, the frequency response magnitude peaks at

$$f_m = f_n \cdot \sqrt{1 - 2 \cdot \zeta^2} = 86.7\ \text{MHz}$$

with a peak amplitude at f_m of

$$M_m = \frac{1}{2 \cdot \zeta \cdot \sqrt{1-\zeta^2}} = 1.41$$

Without hf analysis, this resonance might not have been discovered, though it is at a frequency far above the dominant poles.

We now return to the circuit as it is, with C_E = 10 pF. We can expect it to not be as underdamped. Making the same calculations as before,

$$\tau_n = \sqrt{(\tau_T \cdot R_B') \cdot (C_E + (1+K_v) \cdot C_c)} = 0.955\,\text{ns} \;\Rightarrow 167\ \text{MHz}$$

$$Z_n = \sqrt{\frac{\tau_T \cdot R_B'}{C_E + (1+K_v) \cdot C_c}} = \sqrt{\frac{53.0\,\text{nH}}{17.2\,\text{pF}}} = 55.5\ \Omega$$

$$\zeta_p = \frac{1}{2} \cdot \frac{Z_n}{R} = \frac{(55.5\,\Omega)/2}{40.9\,\Omega} = 0.679 \;\Rightarrow \phi = 47.3°$$

$$M_p = e^{-\pi/\tan\phi} \cong 0.0549 = 5.5\,\%$$

$$f_m = f_n \cdot \sqrt{1 - 2 \cdot \zeta^2} = 46.9\ \text{MHz}$$

$$M_m = \frac{1}{2 \cdot \zeta \cdot \sqrt{1-\zeta^2}} = 1.003$$

The input loop is far better behaved, with nearly MFA response. For that, the bandwidth extension factor is $1.04 \cdot f_n$ = 174 MHz corresponding to a time constant for use in the OCTC calculation of 0.916 ns. This time constant is faster than τ_{e1}, and does not influence the bandwidth much.

An estimate of the buffered amplifier bandwidth can be assembled from τ_{c1} = 6.48 ns, the remaining uncompensated τ_{L2} = 4.386 ns, and if the faster two poles are included, the emitter-peaking pole, at 1.25 ns and the resonant pole-pair with bandwidth corresponding to 0.916 ns:

$$\tau_{bw} = \sqrt{(6.48\,\text{ns})^2 + (4.386\,\text{ns})^2 + (1.25\,\text{ns})^2 + (0.916\,\text{ns})^2} = 4.697\,\text{ns}$$
$$\Rightarrow 33.9\ \text{MHz}$$

With two dominant real poles, the response should have no measurable peaking because the underdamped input-loop pole-pair has its peaking frequency, f_m, above bandwidth.

The frequency-response measurements on the prototype circuit resulted in no appreciable magnitude peaking (M_m = 1) out to a bandwidth of 28 MHz, lower than the OCTC estimate of 34 MHz. Prototype measurements were taken with other values of C_E:

C_E, pF	M_m	f_m, MHz	f_{bw}, MHz
10	1.0	-	28
15	1.1	12	48
33	2.0	20	35
47	2.67	14.1	45

For $C_E = 47$ pF, the peaking was so great that Q1 was showing saturation at the bottom of the sine-wave for 300 mV input amplitude and 800 mV at the output.

In the time domain, the step response was observed to be slightly overdamped. With C_E = 47 pF, the bandwidth and risetime were again measured. A 600 mV output step had a risetime of 9 ns with a 4.2 ns input risetime. Compensating for the input risetime,

$$\text{output } t_r \approx \sqrt{(\text{measured } t_r)^2 - (\text{source } t_r)^2} = 8 \text{ ns}$$

with slight peaking on the leading edge followed by a roll-up of a few percent with a time constant of about 25 ns to the final step value. The single-pole bandwidth calculated from the risetime is

$$f_{bw} = 0.35/8 \text{ ns} = 43.8 \text{ MHz}$$

in good agreement with the step-response estimate.

The lf A_v poles were included from the previous BJT dynamic model but not included was the RHP zero. Its time constant and frequency are

$$\tau_{zRHP} = \frac{r_e + R_E}{\alpha_0} \cdot C_c = \frac{1074\,\Omega}{0.993} \cdot (1.8 \text{ pF}) = 1.945 \text{ ns} \Rightarrow 81.8 \text{ MHz}$$

The zero frequency is a half-decade higher than either the calculated or measured bandwidth, and has some slight effect on it.

If we use the quadratic coefficient, *a,* of the Q1 stage, quadratic bandwidth approximation can be applied;

$$a = (R_b \cdot C_e) \cdot (R_L \cdot C_c) = \tau_e \cdot \tau_{cc} = (0.733 \text{ ns}) \cdot (5.94 \text{ ns}) = 4.354 \text{ ns}^2$$

The linear coefficient is what is used for linear bandwidth estimation:

$$b = R_b \cdot C_e + [R_L + R_b \cdot (1 + K_v)] \cdot C_c = \tau_e + \tau_c = 0.733 \text{ ns} + 6.48 \text{ ns} = 7.213 \text{ ns}$$

The linear bandwidth estimate is $1/b$ = 22.1 MHz - low as expected. Invoking the quadratic bandwidth formula,

$$f_{bw} = \frac{1}{2 \cdot \pi \cdot \sqrt{b^2 - 2 \cdot a}} = 24.2 \text{ MHz}$$

This is closer than the 4-pole OCTC bandwidth and is somewhat low compared to the measured bandwidth of 28 MHz. These approximations ignore the effects of both resonances (Q1 input loop and Z_{b2}, Q1 collector node), yet give a reasonable idea of what to expect for bandwidth.

From a more general analysis involving C_e, C_c, and C_L, a more accurate form of a for the CE amplifier is

$$\begin{aligned} a &= (R_L \cdot C_L) \cdot (R_b \cdot (C_c + \tau_T / R_E)) + (R_L \cdot C_c) \cdot (R_b \cdot C_b) \\ &= \tau_L \cdot [(75\,\Omega) \cdot (1.8 \text{ pF} + 0.66 \text{ pF})] + (5.94 \text{ ns}) \cdot [(75\,\Omega) \cdot (7.86 \text{ pF})] \end{aligned}$$

This reduces to

$$a = 9.11\,\text{ns}^2 + 3.51\,\text{ns}^2 = 12.61\,\text{ns}^2$$

Substituting into the quadratic bandwidth formula,

$$f_{bw} = \frac{1}{2 \cdot \pi \cdot \sqrt{b^2 - 2 \cdot a}} = 30.7 \text{ MHz}$$

This is an acceptable estimate for a circuit of this complexity.

Dynamics Design Overview

We now stand back to review the overall dynamics of the buffered CE amplifier. We derived the lf gain with emitter compensation for τ_L and had derived the gain with only BJT capacitances and C_L. None of the poles or zeros were in the lf region (< 1.5 MHz). This is typical of fast amplifiers, except for compensation of nonlinear effects such as thermal effects. The BJT pole of τ_{c1} was calculated as a real OCTC, but interactions of C_{e1} and C_E, and C_L and C_o, produced quadratic pole-pairs and could not be treated as independent capacitances. The assumption that C_{c1} formed a real pole is based on the general single-stage BJT model with C_L at the collector node. Amplifier 1A, however, does not have a purely capacitive collector reactance because of Z_{b2}, and would require a more thorough analysis to develop a quadratic response including C_{c1}.

Dynamics analysis can easily become unmanageable for two kinds of reasons. We have already seen how approximate bandwidth

estimates are. First, parasitic capacitances are hard to estimate accurately; model parameters are inaccurate. Second, circuit dynamics is not easy to approximate accurately either for circuits of any complexity. Four capacitances of the Q1 stage of amplifier 1A form a series loop; they are not independent and the pole polynomial is thus a cubic. Yet we have sought separated real poles and quadratic pole-pairs.

To maintain manageable design, the general strategy for amplifier dynamics is to decompose the amplifier into separable *stages* which ideally are isolated from each other. In many cases, this is not possible, and the interstage interactions are controlled by designing each stage to have an input impedance that conforms to the model of the previous stage. Each of the various kinds of stages have design templates that give their dynamic response. The templates are worked out as we did for the general BJT stage, for instance, which has C_L loading at its output. The next stage would then be designed to have a capacitive input which (in part, at least) is this C_L. In the case of amplifier 1A, Q2 did not meet this design condition because Z_{b2} depended on the output load of the amplifier, which in various cases is ill-defined and thus variable.

What amplifier 1A has shown is that four interacting capacitances in the first stage push the given bandwidth estimation methods to their useful limit for design insight. Computer circuit simulation verifies a design by giving a more accurate bandwidth value. Keep in mind that our interest in using the non-computer approximate methods is not to determine accurate circuit behavior but to improve design insight for optimizing behavior. For this, approximate results are acceptable.

$\alpha(s)$ Compensation

We now take the CE amplifier through one last refinement that illustrates the benefit of the hf BJT concept. The use of C_E to compensate whatever is the dominant pole - often τ_L - has the side-effect of hf destabilization in the input loop. We found the input-loop resonance using hf analysis, and how it became increasingly underdamped with increasing C_E. This is not the best method of bandwidth extension because it can result in undesirable input-loop impedance, especially if there is a previous stage.

The hf base input impedance is

$$Z_b(\text{hf}) = (\beta_{hf} + 1) \cdot Z_E = \left(\frac{s \cdot \tau_T + 1}{s \cdot \tau_T} \right) \cdot \left(\frac{R_E}{s \cdot R_E \cdot C_E + 1} \right)$$

What is referred to as $\alpha(s)$ *compensation* sets

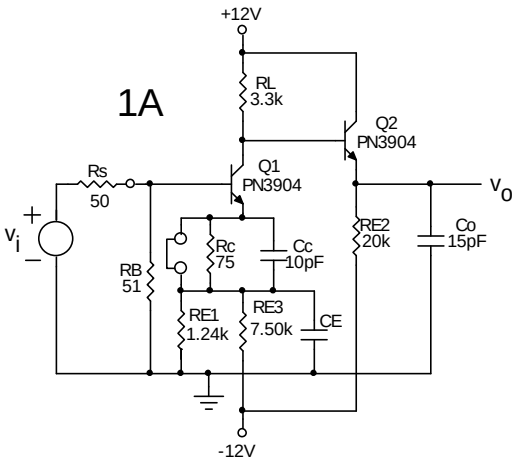
$$R_E \cdot C_E = \tau_T$$

so that the BJT has no hf resonance and the resulting

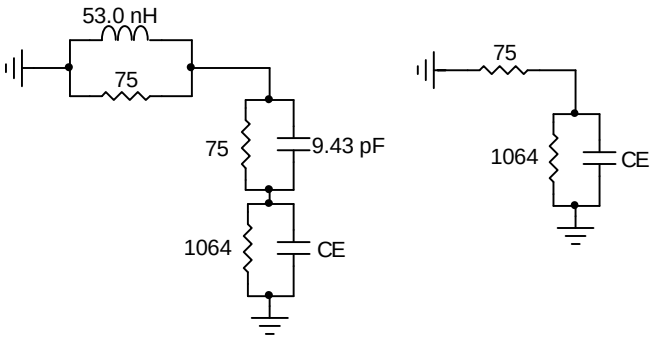
$$\text{compensated } Z_b(\text{hf}) = 1/s \cdot C_b \text{ , } C_b = \tau_T / R_E$$

The base-node capacitance is the hf base-referred emitter resistance.

The existing circuit can be modified to compensate for hf gyration of base resistance into inductance at the emitter by adding in series with R_E and C_E another shunt RC compensator, as shown below.



The hf model of the input loop is shown below for R_B' = 75 Ω. The base resistance is hf-referred to the emitter as a shunt RL.



This method for hf compensation of base resistance places a shunt RC in series with the emitter, where $C = \tau_T/R_B' = 9.43$ pF. (At 1 mA, a PN3904 has $f_T \approx 225$ MHz and $\tau_T \approx 707$ ps.) Then the shunt RL referred from the base and shunt RC form an all-pass network with a combined impedance that is the resistance R_B'. The resulting hf input loop is shown to the right. C_E compensates τ_L while in series above it the shunt RC compensates $\alpha(s)$ and the hf base impedance.

When implemented with a 10 pF capacitor instead of 9.43 pF, the hf compensation decreases the risetime. In the prototype circuit, a shorting wire across the upper RC was installed for comparison, and the difference was observed. The output with the RC in-circuit had a risetime of about 8 ns, significantly faster than the circuit risetime of 20 ns without it. The step was not underdamped.

For $C_E = 15$ pF, a shunt RC of $75\,\Omega$ in parallel with 10 pF resulted in $M_m = 1.08$, $f_m = 12$ MHz and $f_{bw} = 40$ MHz. This compares favorably with the time-domain equivalent of

$$0.35/t_r = 0.35/8 \text{ ns} = 44\text{MHz}$$

The quasistatic gain is, of course, affected by R_C. It adds to r_M and must be accounted for in quasistatic design. It also must be considered when making measurements with the RC shorting jumper because it changes the quasistatic reference amplitude for bandwidth measurements. In the above circuit, it decreases A_{v0} by about 7 %.

Amplifier Design

Slew Rate and Large-Signal Bandwidth

A *slew-rate-limited* response is a large-signal dynamic limitation in which sine-wave slopes become ramps. Slewing is caused by insufficient bias current to charge a capacitance at the maximum rate needed by a sine-wave of a given amplitude. The frequency at the onset of slewing for a full-scale (maximum-range) waveform is the *large-signal bandwidth*, f_{BW}.

With a 400 mV sine-wave amplitude at the output, *slewing* in the CE amplifier 1A prototype of the previous chapter was observed starting at about 15 MHz, and for a full-scale (±2.8 V) waveform, slewing began to be observed as low as 2 MHz. Sine-wave slope is maximum at zero voltage;

$$\text{max sine-wave slope} = \max \frac{d}{dt}(V \cdot \sin(\omega \cdot t)) = V \cdot \omega$$

The CE bias current at 1 mA is not sufficient to charge at a non-slewing rate above 15 MHz. The charge rate is limited to

$$\frac{I}{C} = \frac{dv(t)}{dt} = V \cdot \omega = (400\,\text{mV}) \cdot (2 \cdot \pi \cdot 15\,\text{MHz}) = 37.7\ \text{V/}\mu\text{s}$$

At I = 1 mA the onset of slewing occurs for C = 26.5 pF, a value in the range of the C_E circuit capacitances. The output range, as determined at the beginning of the experiment, was far greater than 0.4 V. Thus, f_{BW} is less than the observed 2 MHz.

With amplitude reduced to 200 mV and C_E = 47 pF, no slewing was observed up to the incremental bandwidth of f_{bw} = 34 MHz. With the upper shunt RC compensator added, the bandwidth increased to $f_{bw} \approx 100$ MHz. This frequency is near the emitter compensator pole, at 129 MHz and surpasses the RHP zero. At these frequencies, a faster oscilloscope than was used (100 MHz) is required for meaningfully accurate measurements.

For larger R_B' the hf-compensating C_E is reduced, and under 1 pF can pose an implementation problem. The parasitic capacitance of ¼ W axial-lead resistors is about ⅓ pF. The parasitic capacitance of

two in parallel (R_{E1} and R_{E2}) is thus about ⅔ pF. They can be used when C_E is applied to hf input compensation. Then τ_L is compensated by a more elaborate form of emitter circuit, a more elaborate and inclusive form of emitter peaking.

Multi-Stage Bandwidth

For amplifiers with real poles, maximum bandwidth results when stages have equal bandwidths. Single-pole stages have f_T values that limit gain for a given bandwidth. Stages with a dominant-pole frequency response trade off gain and bandwidth proportionally and have a constant *gain-bandwidth product*. If stage gain is made large, its bandwidth will be low. The root-sum-of-squares formula for time constants of real poles is minimized when no one time constant dominates. Hence the stages should have equal gains and thus equal bandwidths for maximum amplifier bandwidth. (The optimum number of stages and gain per stage for maximum bandwidth is derived in "Amplifier Stage Gain Optimization", *Designing High-Performance Amplifiers*, D. Feucht)

For n identical cascade stages with quadratic pole-pairs, each having MFA response and pole magnitude ω_n, the ratio of bandwidth, ω_{bw}, to ω_n is

$$\frac{\omega_{bw}}{\omega_n} = \sqrt{\sqrt{2^{1/n} - 1}}$$

(This is derived by substituting $\zeta = \sqrt{2}/2$ into the quadratic bandwidth formula presented in the section, "Quadratic Pole-Pair Bandwidth".) For $n = 2$, the amplifier will have a bandwidth reduction from the pole-pair magnitude, ω_n, of about 20 % or 0.8.

For n cascade stages of single-pole amplifiers with equal real poles of frequency p, the bandwidth is reduced by the fraction

$$\frac{\omega_{bw}}{p} = \sqrt{2^{1/n} - 1}$$

For n critically-damped stages with equal pole magnitudes,

$$\frac{\omega_{bw}}{\omega_n} = \sqrt{2^{1/2 \cdot n} - 1} \ , \ \zeta = 1$$

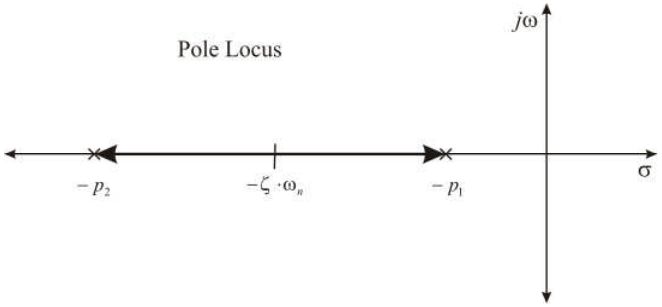
Each stage has two poles; hence the 2 in the denominator of the exponent.

Pole Separation

For two real poles, the quadratic polynomial describing them has $\zeta \geq 1$ and the pole values are found by solving the quadratic equation;

$$p_{1,2} = -\zeta \cdot \omega_n \cdot \left(1 \pm \frac{\sqrt{\zeta^2 - 1}}{\zeta} \right) = -\omega_n \cdot \left(\zeta \pm \sqrt{\zeta^2 - 1} \right)$$

They are separated from their center at $-\zeta \cdot \omega_n$ by $2 \cdot \omega_n \cdot \sqrt{\zeta^2 - 1}$.



For poles separated at $-p_1$ and $-p_2$, the damping factor is

$$\zeta = \frac{1}{2} \cdot \sqrt{\frac{(p_2 / p_1 + 1)^2}{p_2 / p_1}}$$

The pole separation ratio is

$$\frac{p_2}{p_1} = \frac{1 + \sqrt{\zeta^2 - 1} / \zeta}{1 - \sqrt{\zeta^2 - 1} / \zeta} = \frac{\zeta + \sqrt{\zeta^2 - 1}}{\zeta - \sqrt{\zeta^2 - 1}}$$

A table of some salient points of this function is given below along with the decade (log) separation. From the table it is apparent that ζ need not exceed 1 by much (1.74) to separate poles by a decade. By two decades, $\zeta \approx 5$ and their effect on phase is well-separated.

To be stable, a typical wideband feedback amplifier with single- to double-pole rolloff has no more than two close lower-frequency poles, though others will be at higher frequencies. Therefore,

bandwidth methods based on linear or quadratic approximations are not impractical for many transistor amplifiers.

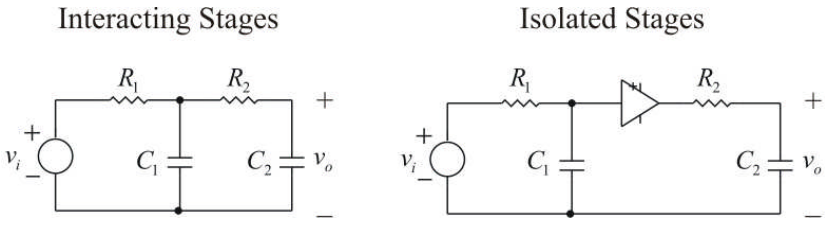
ζ	p_2/p_1	$\log(p_2/p_1)$
1	1	0
$\sqrt{2} \approx 1.414$	5.8284	0.7656
$\sqrt{3} \approx 1.732$	9.8990	0.9956
$\frac{11}{2 \cdot \sqrt{10}} \approx 1.7393$	10	1
2	13.928	1.1439
$e \approx 2.71828$	27.520	1.4396
3	33.971	1.5311
5.0500	100	2
10	398.00	2.60

Cascade Stage Interaction

The algebra describing the dynamic performance of BJT amplifier stages can quickly become so complicated that engineers will abandon algebraic analysis and let a computer circuit simulator solve a system of circuit equations *numerically* by iterating a solution as has been done manually for static circuit analysis. The big disadvantage of simulation is that it only outputs the behavior of a particular circuit under particular conditions and gives little design insight into how to choose parts values or optimize behavior. In contrast, the algebraic equations we have been deriving show how each component value affects the overall response. Equations put into forms that have memorable circuit interpretations are useful *formulas*.

The complications of dynamic analysis arise because C_c of a BJT amplifier stage dynamically couples it to adjacent stages. Although C_c is small - a few pF - it is comparable in value to capacitances in adjacent stages of the amplifier and cannot be ignored. The interstage interactions through this coupling lead to transfer functions that have poles (and zeros) that are hard to determine using loop or node analysis. We can make useful approximations by simplifying pole and zero determination based on open-circuit time constants. Beyond that, a powerful method of problem reduction - the extra-element theorem (EET) - allows us to isolate the effects of each reactance in a circuit, one reactance at a time.

To illustrate the stage interaction problem, consider two interacting stages of RC integrators, shown below.



To isolate the stages from each other, they can be separated by an ideal voltage buffer stage with no reverse transmittance through it. For the isolated stages, the transfer function (with a ×1 voltage amplifier as buffer) is

$$\frac{v_o}{v_i} = \frac{1}{(s \cdot R_1 \cdot C_1 + 1) \cdot (s \cdot R_2 \cdot C_2 + 1)}$$

The poles are independent of each other because the Rs and Cs that determine their respective values are different. For $R_1 = R_2 = R$ and $C_1 = C_2 = C$, the poles of the interacting two-stage RC integrator are *not* a double pole at $1/R \cdot C$ because the second stage loads the first stage. The transfer function is instead

$$\frac{v_o}{v_i} = \frac{1}{s^2 \cdot (R_1 \cdot C_1 \cdot R_2 \cdot C_2) + s \cdot [R_1 \cdot C_1 + (R_1 + R_2) \cdot C_2] + 1}$$

The quadratic pole factor that is the denominator differs from the isolated circuit by the extra $R_1 \cdot C_2$ term in the coefficient of s. If the first stage is thevenized, then its Thevenin series resistance of R_1 adds to R_2 of the second stage, changing its time constant. The same kind of interactive effect occurs for cascade BJT amplifier stages. What is difficult to determine is the interactive aspect in the transfer functions.

The roots of $D(s)$ can be found by solving for them as expressed in

$$\tau(a, b) = R_1 \cdot C_1$$

$$\tau(a) = R_2 \cdot C_2$$

$$\tau(b) = \tau(a) + R_1 \cdot C_2 = \tau(a) + \tau_M(b)$$

where $\tau(a, b)$ is an OCTC common to both coefficients a and b, $\tau(a)$ appears only as a SCTC in a, and $\tau(b)$ as a time constant only in b. $\tau(a)$ is not a time constant in b but is a *time-constant component* of $\tau(b)$, and thus $\tau(b) \geq \tau(a)$. The other component of $\tau(b)$ is the *interaction component*, $\tau_M(b)$, that later is found to be associated with the Miller effect in transistor stages. If it is zero, then $\tau(a, b)$ and $\tau(a)$ are the circuit poles. As it increases from zero, the poles diverge from the OCTCs.

Pole interaction is made more explicit by expressing ζ in the OCTC *separation factor*, ξ, and *interaction factor*, k;

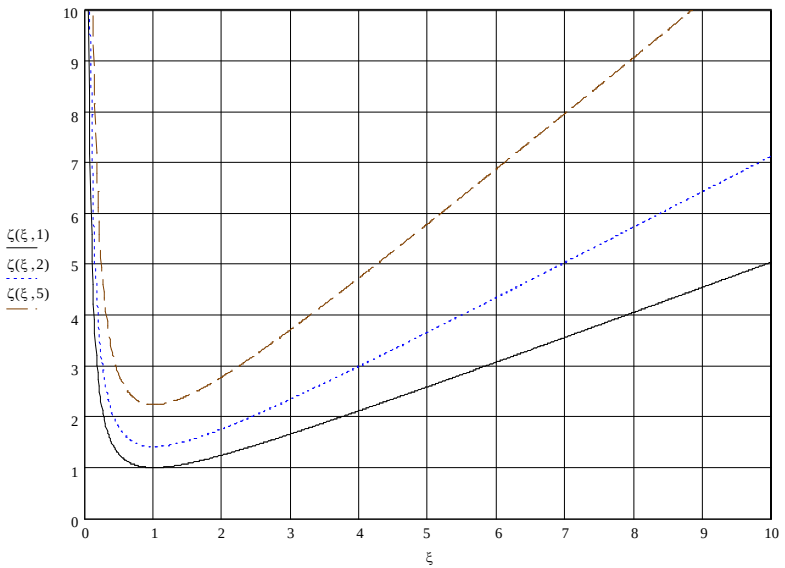
$$\text{separation, } \xi = \sqrt{\frac{\tau(a,b)}{\tau(b)}} \quad ; \quad \text{interaction, } k = \frac{\tau(b)}{\tau(a)} = 1 + \frac{\tau_M(b)}{\tau(a)}$$

In time constants and components, the pole-pair magnitude and damping are

$$\tau_n = \sqrt{a} = \sqrt{\tau(a,b) \cdot \tau(a)} = \sqrt{k} \cdot \xi \cdot \tau(a) = \frac{\xi}{\sqrt{k}} \cdot \tau(b) \quad ;$$

$$\zeta = \frac{b}{2 \cdot \sqrt{a}} = \frac{1}{2} \cdot \frac{\tau(a,b) + \tau(b)}{\tau_n} = \frac{\sqrt{k}}{2} \cdot \left(\xi + \frac{1}{\xi} \right)$$

$\zeta(\xi, k)$ is plotted below for k = 1, 2, and 5.



Given the ability to find OCTCs from circuit structure, the question of interest is where the poles are relative to the OCTCs. The poles are expressed in the interaction and OCTC separation factors by substituting for τ_n and ζ to result in the *pole-pair formula*,

$$p_{1,2} = -\frac{1}{2\cdot\tau(a)}\cdot\left[(1+\xi^{-2}) \pm \sqrt{1+2\cdot\left(1-\frac{2}{k}\right)\cdot\xi^{-2}+\xi^{-4}}\right]$$

In this one equation is captured the effect on pole location of both interaction caused by $\tau_M(b)$ in k and OCTC separation in ξ. The factor in front, $-1/2\cdot\tau(a) = -k/2\cdot\tau(b)$.

For minimum interaction, $k = 1$, and the expression under the radical factors into

$$(\xi^{-2}-1)^2$$

with the resulting poles at

$$p_1 = -\frac{1}{\tau(a)} = -\frac{1}{\tau(b)},\ p_2 = -\frac{1}{\tau(a,b)}$$

When interaction is minimal, *the poles are at the OCTCs*. It does not matter how separated the OCTCs are; the pole polynomial factors into the OCTCs and they are at the poles. $\tau(a)$ is the time-constant component common to a and $\tau(b)$, and is an SCTC. For $k = 1$, its value is that of the corresponding OCTC for a given port. Either shorting or opening the other reactance port does not affect the resistance across the given port and it is isolated from the shorted or opened port. Thus, the two reactances do not affect each other and are *isolated* from each other.

For large interaction, k becomes large as τ_M dominates $\tau(b)$, and the expression under the radical factors into

$$(\xi^{-2}+1)^2$$

resulting in poles at

$$p_1 = -\frac{k}{\tau(a,b)}\ ,\ p_2 \to 0$$

The low pole approaches the origin while the other approaches a frequency k times higher than that of the OCTC $\tau(a, b)$. The lower pole dominates dynamic response and some means of approximating

it (other than as 0 in the limit) is needed. For widely separated real poles resulting from a dominant τ_M, the TC values for the poles can be linearly approximated, and the low-frequency TC $\approx b \approx \tau_M$. Having one pole value and knowing that the quadratic coefficient, a, is the product of the real poles, then the high-frequency TC $\approx a/b \approx \tau_n^2/\tau_M$. Or more precisely, $b \approx \tau(b)$ and

$$\frac{\tau_n^2}{\tau(b)} = \frac{k \cdot \xi^2 \cdot [\tau(a)]^2}{\tau(b)} = \xi^2 \cdot \tau(a) = \frac{\tau(a,b)}{\tau(b)} \cdot \tau(a) = \frac{\tau(a,b)}{k}$$

and corresponds to the higher pole, p_1, above. In general, for TCs τ_1 and τ_2 of real poles,

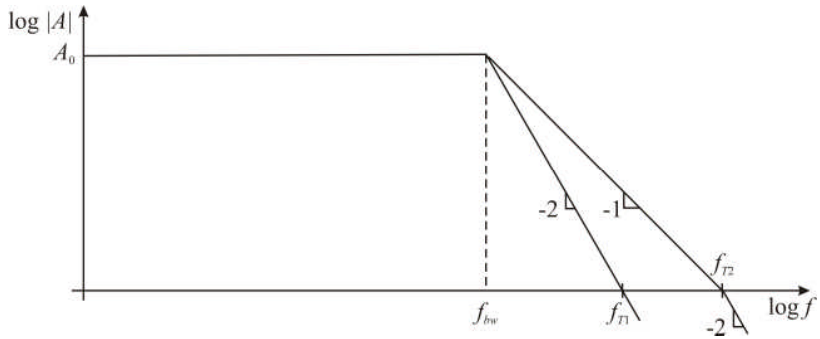
$$\frac{a}{b} = \frac{\tau_1 \cdot \tau_2}{\tau_1 + \tau_2} \approx \frac{\tau_1 \cdot \tau_2}{\tau_1} \approx \tau_2, \; \tau_1 >> \tau_2$$

The second effect is OCTC separation. OCTCs are separated in the pole-pair formula above by letting ξ become large so that $\xi^{-2} \approx 0$. Then in the limiting case,

$$p_1 = -\frac{1}{\tau(a)}, \; p_2 \to 0$$

As ξ increases, the low pole migrates toward the origin while the high pole remains at $-1/\tau(a)$. *As the OCTCs separate, the poles separate.* With widely separated poles, the lower pole can be approximated by $1/\tau(b)$.

Reactance interaction affects pole frequencies, but a lack of OCTC separation ($\xi \approx 1$) causes a different interactive effect. In the cascade RC integrator, stages have real poles that can be widely separated; one pole can be made lower in frequency (the dominant pole) while the other is higher, as shown.

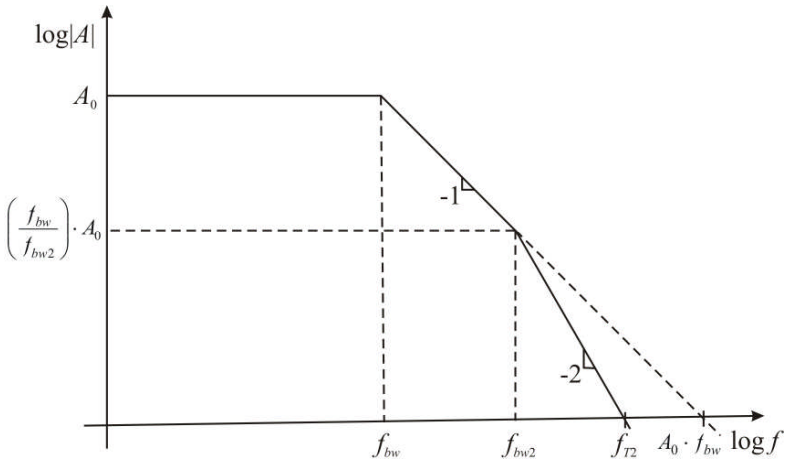


Poles are maximally interactive at the same pole frequency and the output decreases logarithmically (or *rolls off*) at a log-log slope of –2. For lightly-interacting stages, a roll-off beginning at the lower pole with a –1 slope changes to –2 at the higher pole.

The two frequency-response magnitude plots cross a gain of one (at f_T) at different frequencies, as shown. $f_{T2} > f_{T1}$ because

$$f_{T2} = A_0 \cdot f_{bw} > f_{T1} = \sqrt{A_0} \cdot f_{bw} \Rightarrow \frac{f_{T2}}{f_{T1}} = \left(\frac{A_0}{\sqrt{A_0}} \right) > 1$$

Plots are as shown below for two poles, one at f_{bw} and the other at f_{bw2}. Close poles have f_T less than that of separated poles because of the steeper –2-slope roll-off from the higher frequency. The separation in f_T is $f_{T1}/f_{T2} = 1/\sqrt{A_0}$.



From the plots, f_{T2} is at the geometric mean of f_{bw2} and $A_0 \cdot f_{bw}$;

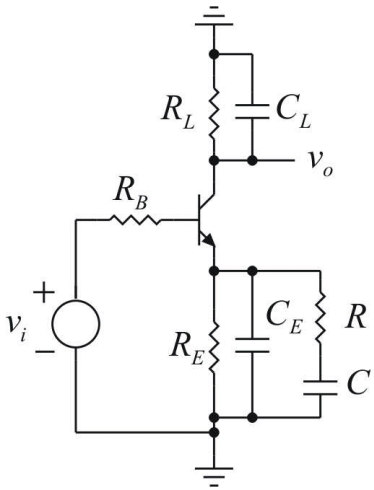
$$f_{T2} = f_{bw2} \cdot \sqrt{\frac{A_0 \cdot f_{bw}}{f_{bw2}}} = \sqrt{(A_0 \cdot f_{bw}) \cdot f_{bw2}}$$

When poles are closer in frequency ratio than the amplifier quasistatic gain, they are not separated enough for f_T to remain unaffected. BJT amplifiers not designed for high bandwidth often have poles that are widely separated so that one is a dominant pole and bandwidth is approximated by the dominant pole.

The general single-stage BJT amplifier and the cascade CE amplifier with fast transistors (to be derived) have real poles. Amplifier stages with all real poles have bandwidths that can be

approximated from OCTCs. These amplifiers do not have bandwidth extension such as inductive peaking or input-impedance control with external emitter capacitances; they have no added reactances. When additional impedance- or bandwidth-modifying reactances are added, poles (and sometimes zeros) are no longer constrained to the real axis and a different category of circuit results with quadratic poles or zeros. For instance, by adding R_f in parallel with C_f in the generalized stage, a model of a *shunt-feedback amplifier stage* results that has complex poles. Those with cubic or higher-degree polynomial(s) are often able to be reduced to quadratics by applying design constraints to them.

Emitter Peaking



The above circuit diagram is typical of the way engineers sometimes simplify drawings of incremental circuit models; biasing is ignored and the incremental equivalent circuit is inferred. The incremental T model or hybrid-π model is now assumed when looking at the BJT symbol as we envision the circuit model in the circuit diagram.

The circuits that can be placed in the emitter branch of the input loop give us at least three possibilities for design. The first is to compensate for the collector time constant with only C_E. For amplifiers with poles and zeros well below f_β, hf considerations can sometimes be ignored. Generally, it is suboptimal to compensate τ_L by

setting it equal to $R_E \cdot C_E$ because emitter peaking is used mostly for frequency compensation of BJT poles in the hf range.

The first design option for poles and zeros in the BJT hf region (as demonstrated in design of circuit 1A) is to use C_E without R and C to compensate for $\alpha(s)$ so that the emitter impedance appears at the base node as capacitance $C_b = \tau_T/R_E$. This is accomplished simply by making $C_E = \tau_T/R_E$ and consequently $C_b = C_E$. While this hf-compensates Z_b(hf) it does nothing for τ_L.

The second option uses R and C without C_E for lf compensation. The lf transfer function of the above CE amplifier with R and C is

$$A_v = -\alpha_0 \cdot \frac{R_L}{r_M} \cdot \frac{s \cdot (R_E + R) \cdot C + 1}{(s \cdot [r_E \parallel R_E + R] \cdot C + 1) \cdot (s \cdot \tau_L + 1)}$$

where $r_E = r_e + R_B/(\beta_0 + 1)$ and the collector time constant, $\tau_L = R_L \cdot C_L$. To compensate by pole-zero cancellation, set

$$(R_E + R) \cdot C = \tau_L$$

To benefit from this compensation, the remaining pole must be much higher in frequency, widely separated from the compensated pole. To achieve this, R must not be made too large yet it must be large enough to dominate $r_E \| R_E$. Hence,

$$r_E \parallel R_E << R << R_E$$

R is maximally separated from both at the geometric mean;

$$R = \sqrt{(r_E \parallel R_E) \cdot R_E}$$

Given the value of R, then C follows immediately;

$$C = \frac{\tau_L}{R_E + R}$$

The remaining pole that determines bandwidth is higher in frequency and has a time constant of

$$\tau_p = (r_E \parallel R_E + R) \cdot C$$

The third design option is the most elegant because it compensates both $\alpha(s)$ and τ_L while allowing the remaining pole-pair damping, ζ, to also be specified within a wide range. The choice of BJT fixes $\alpha_0 \cdot \tau_T$, A_{v0} determines R_E, and the collector circuit determines τ_L; these three parameters are given and hence known. Three free

design parameters result from three free circuit parameters: R, C, and C_E. The emitter-peaking design formulas are first derived for the lf region; then the total-frequency case is considered.

The emitter impedance of the external network is

$$Z_E = R_E \parallel (1/s \cdot C_E) \parallel (R + (1/s \cdot C))$$

Expanded in normalized form,

$$Z_E = R_E \cdot \frac{s \cdot R \cdot C + 1}{s^2 \cdot (R_E \cdot C_E \cdot R \cdot C) + s \cdot [R_E \cdot (C_E + C) + R \cdot C] + 1}$$

Z_E is in series with and thus adds to the emitter-port resistance,

$$r_E = r_e + R_B/(\beta_0 + 1)$$

so that the total lf input-loop transimpedance at the emitter node is

$$\begin{aligned} z_M &= r_E + Z_E \\ &= r_M \cdot \frac{s^2 \cdot (r_E \parallel R_E) \cdot C_E \cdot (R \cdot C) + s \cdot [(r_E \parallel R_E) \cdot (C_E + C) + R \cdot C] + 1}{s^2 \cdot R_E \cdot C_E \cdot (R \cdot C) + s \cdot [R_E \cdot C_E + (R_E + R) \cdot C] + 1} \end{aligned}$$

where input-loop transresistance is

$$r_M = r_E + R_E$$

We can simplify the above z_M for the case where $C_E = 0$ pF to result in

$$z_M \big|_{C_E = 0\ \text{pF}} = r_M \cdot \frac{s \cdot [(r_E \parallel R_E) + R] \cdot C + 1}{s \cdot (R_E + R) \cdot C + 1}$$

Because z_M is in the denominator of $A_v(s)$, the z_M pole time constant cancels the τ_L pole in A_v while the z_M numerator time constant sets the new (higher) pole frequency.

Continuing with C_E, the lf voltage gain is found by substituting z_M into A_v to result in

$$\begin{aligned} A_v = &-\frac{\alpha_0}{s \cdot \alpha_0 \cdot \tau_T + 1} \cdot \frac{R_L}{r_M} \cdot \frac{1}{s \cdot \tau_L + 1} \\ &\cdot \frac{s^2 \cdot R_E \cdot C_E \cdot (R \cdot C) + s \cdot [R_E \cdot C_E + (R_E + R) \cdot C] + 1}{s^2 \cdot (r_E \parallel R_E) \cdot C_E \cdot (R \cdot C) + s \cdot [(r_E \parallel R_E) \cdot C_E + [(r_E \parallel R_E) + R] \cdot C + 1} \end{aligned}$$

To compensate both $\alpha(s)$ and τ_L, the numerator, $N(s)$, is constrained to have time constant factors of $\alpha_0 \cdot \tau_T$ and τ_L:

$$s^2 \cdot a + s \cdot b + 1 = (s \cdot \alpha_0 \cdot \tau_T + 1) \cdot (s \cdot \tau_L + 1)$$
$$= s^2 \cdot \alpha_0 \cdot \tau_T \cdot \tau_L + s \cdot (\alpha_0 \cdot \tau_T + \tau_L) + 1$$

The polynomial coefficients of the two undesirable poles are equated to those of $N(s)$ for compensation:

$$a = \alpha_0 \cdot \tau_T \cdot \tau_L = R_E \cdot C_E \cdot R \cdot C \quad (1)$$

$$b = \alpha_0 \cdot \tau_T + \tau_L = R_E \cdot C_E + R_E \cdot C + R \cdot C \quad (2)$$

With three free parameters (C_E, R, C) and two time constants as constraints (τ_T, τ_L), an additional damping constraint can be applied to the pole-pair;

$$\zeta = \frac{b}{2 \cdot \sqrt{a}} = \frac{(r_E \parallel R_E) \cdot (C_E + C) + R \cdot C}{2 \cdot \sqrt{(r_E \parallel R_E) \cdot C_E \cdot R \cdot C}} \quad (3)$$

Solve eqn (1) for $R \cdot C$;

$$R \cdot C = \frac{\alpha_0 \cdot \tau_T \cdot \tau_L}{R_E \cdot C_E} \Rightarrow C_E = \frac{\alpha_0 \cdot \tau_T \cdot \tau_L}{R_E \cdot (R \cdot C)}$$

Solve (2) for

$$C_E + C = \frac{\alpha_0 \cdot \tau_T + \tau_L - R \cdot C}{R_E}$$

Substitute the expressions for $R \cdot C$ and $(C_E + C)$ into (3);

$$\zeta = \frac{\left(\frac{r_E}{r_E + R_E}\right) \cdot (\alpha_0 \cdot \tau_T + \tau_L - R \cdot C) + R \cdot C}{2 \cdot \sqrt{\left(\frac{r_E}{r_E + R_E}\right) \cdot (\alpha_0 \cdot \tau_T \cdot \tau_L)}}$$

This reduces to

$$\zeta = \frac{\left(\frac{r_E}{r_M}\right) \cdot (\alpha_0 \cdot \tau_T + \tau_L) + \left(\frac{R_E}{r_M}\right) \cdot (R \cdot C)}{2 \cdot \sqrt{\left(\frac{r_E}{r_M}\right) \cdot (\alpha_0 \cdot \tau_T \cdot \tau_L)}}$$

This linear equation is solved for the first design formula and reduces to the preferred form for calculation:

$$\Rightarrow \quad R \cdot C = 2 \cdot \zeta \cdot \sqrt{\left(\frac{r_M \cdot r_E}{R_E^2}\right) \cdot (\alpha_0 \cdot \tau_T \cdot \tau_L) - \left(\frac{r_E}{R_E}\right) \cdot (\alpha_0 \cdot \tau_T + \tau_L)}$$

Then solve for the remaining circuit element values:

$$C_E = \frac{\alpha_0 \cdot \tau_T \cdot \tau_L}{R_E \cdot (R \cdot C)} \Rightarrow C = \frac{\alpha_0 \cdot \tau_T + \tau_L - R \cdot C}{R_E} - C_E \Rightarrow R = \frac{R \cdot C}{C}$$

The time constant, $R{\cdot}C$, lies between $\alpha_0{\cdot}\tau_T$ and τ_L and varies with the damping of the pole-pair. For a physically realizable $R{\cdot}C > 0$ s,

$$\zeta \geq \frac{1}{2} \cdot \sqrt{\frac{r_E}{r_M} \cdot \left(\frac{\alpha_0 \cdot \tau_T}{\tau_L} + \frac{\tau_L}{\alpha_0 \cdot \tau_T} + 2\right)}$$

For $r_E << R_E$, $r_E/r_M \approx 0$; the lower limit of ζ is below that of wideband amplifier damping and is not restrictive. The lower bound on ζ is minimum when $\alpha_0{\cdot}\tau_T = \tau_L$ and is $\zeta \geq \sqrt{r_E / r_M}$. Design-range damping is only limited when r_E/R_E is large; a large R_B produces a large r_E and also a slow amplifier.

Emitter peaking in the lf region is applicable for amplifiers having fast transistors relative to dominant poles in the lf region. More often, parasitic BJT capacitances limit bandwidth, and emitter peaking formulas that include hf effects are applicable.

Total-Frequency Emitter Peaking

The total-frequency derivation of emitter-peaking design formulas follows the lf derivation of the previous section, though with the added complication of hf effects. Derivation could be simplified by limiting it to the hf region only, but total-frequency results are not much harder to derive and result in the most general formulas.

The cascode CE voltage gain for emitter peaking that includes the hf region is

$$A_v = -Z_L \cdot \frac{i_c(s)}{v_i(s)} = -\alpha(s) \cdot Z_L \cdot \frac{i_e(s)}{v_i(s)} = -\alpha(s) \cdot Z_L \cdot \frac{i_e}{v_e} \cdot \frac{v_e}{v_i} = -\alpha(s) \cdot \frac{Z_L}{z_M}$$

where

$$\frac{i_e}{v_e} = \frac{1}{Z_E} \quad ; \quad \frac{v_e}{v_i} = \frac{Z_E}{z_E + Z_E}$$

$$Z_L = R_L \parallel (1/s \cdot C_L) = \frac{R_L}{s \cdot \tau_L + 1}, \tau_L = R_L \cdot C_L$$

To extend the emitter-peaking derivation to include the hf region, the lf

$$z_M = r_E + Z_E$$

is modified by extending r_E to include emitter-port impedance,

$$z_E = \frac{r_e}{s \cdot \alpha_0 \cdot \tau_T + 1} + \frac{Z_B}{\beta(s) + 1}$$

to become

$$z_M = z_E + Z_E$$

For $Z_B = R_B$, z_E becomes

$$z_E = \frac{r_e}{s \cdot \alpha_0 \cdot \tau_T + 1} + \frac{R_B}{\beta_0 + 1} \cdot \frac{s \cdot \tau_\beta + 1}{s \cdot \alpha_0 \cdot \tau_T + 1}, Z_B = R_B$$

and in normalized form is

$$z_E = r_E \cdot \frac{s \cdot \alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_E} + 1}{s \cdot \alpha_0 \cdot \tau_T + 1}, \; r_E = r_e + \frac{R_B}{\beta_0 + 1}$$

Combining with Z_E,

$$z_M = r_E \cdot \frac{s \cdot \alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_E} + 1}{s \cdot \alpha_0 \cdot \tau_T + 1} + Z_E$$

The algebra of z_M is simplified by expressing Z_E in a general form, as time constants of a zero and two poles and quasistatic R_E:

$$Z_E = R_E \cdot \frac{s \cdot \tau_R + 1}{D(s)}$$

where $\tau_R = R \cdot C$ and

$$D(s) = (s \cdot \alpha_0 \cdot \tau_T + 1) \cdot (s \cdot \tau_L + 1) = s^2 \cdot [\alpha_0 \cdot \tau_T \cdot \tau_L] + s \cdot [\alpha_0 \cdot \tau_T + \tau_L] + 1$$

Substituting this into z_M results in

$$z_M = r_E \cdot \frac{s \cdot \alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_E} + 1}{s \cdot \alpha_0 \cdot \tau_T + 1} + R_E \cdot \frac{s \cdot \tau_R + 1}{D(s)} = r_M \cdot \frac{N(s)}{D(s)}$$

where

$$N(s) = s^2 \cdot [\alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_M} \cdot \tau_L] + s \cdot [\alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_M} + \tau_L \cdot \frac{r_E}{r_M} + \tau_R \cdot \frac{R_E}{r_M}] + 1$$

The coefficients of $D(s)$ can be equated to those from the circuit derivation of Z_E. The results are the same design formulas as for the lf emitter-peaked circuit:

$$\alpha_0 \cdot \tau_T \cdot \tau_L = R_E \cdot C_E \cdot \tau_R;\ \alpha_0 \cdot \tau_T + \tau_L = R_E \cdot (C_E + C) + \tau_R$$

Solving, as before,

$$C_E = \frac{\alpha_0 \cdot \tau_T \cdot \tau_L}{R_E \cdot \tau_R} \Rightarrow C = \frac{\alpha_0 \cdot \tau_T + \tau_L - \tau_R}{R_E} - C_E \Rightarrow R = \frac{\tau_R}{C}$$

Substitute z_M into the voltage gain to result in

$$A_v = -\alpha(s) \cdot \frac{Z_L}{z_M} = -\frac{\alpha_0}{s \cdot \alpha_0 \cdot \tau_T + 1} \cdot \frac{R_L}{s \cdot \tau_L + 1} \cdot \frac{(s \cdot \alpha_0 \cdot \tau_T + 1) \cdot (s \cdot \tau_L + 1)}{r_M \cdot N(s)}$$

The poles of $D(s)$ (now in the numerator) become zeros in A_v and cancel those of $\alpha(s)$ and Z_L, reducing the gain expression to

$$A_v = -\alpha_0 \cdot \frac{R_L}{r_M} \cdot \frac{1}{N(s)}$$

$N(s)$ can be expressed in the general form of a pole-pair;

$$N(s) = s^2 \cdot \tau_n^2 + s \cdot (2 \cdot \zeta \cdot \tau_n) + 1$$

With this form of $N(s)$, we can substitute a desired damping value, ζ, as a design parameter and calculate the needed τ_R and also the pole-pair magnitude, τ_n. To derive them, equate coefficients:

$$\alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_M} \cdot \tau_L = \tau_n^2$$

$$\alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_M} + \tau_L \cdot \frac{r_E}{r_M} + \tau_R \cdot \frac{R_E}{r_M} = 2 \cdot \zeta \cdot \tau_n$$

where

$$r_E = r_e + \frac{R_B}{\beta_0 + 1} \; ; \; r_M = r_E + R_E$$

The solutions for τ_n and τ_R follow readily;

$$\tau_R = 2 \cdot \zeta \cdot \sqrt{\alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_M} \cdot \tau_L} - \left(\alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_M} + \tau_L \cdot \frac{r_E}{r_M} \right)$$

$$\tau_n = \sqrt{\alpha_0 \cdot \tau_T \cdot \frac{R_B}{r_M} \cdot \tau_L}$$

Some simplification of these formulas occurs when the hf approximation is applied: $\alpha_0 = 1$, $r_e = 0\ \Omega$. Then

$$\text{hf } \tau_R = 2 \cdot \zeta_{hf} \cdot \sqrt{\tau_T \cdot \frac{R_B}{R_E} \cdot \tau_L} - \tau_T \cdot \frac{R_B}{R_E} \; ; \text{hf } \tau_n = \sqrt{\tau_T \cdot \frac{R_B}{R_E} \cdot \tau_L}$$

The lower limit on damping is

$$\zeta_{hf} \geq \frac{1}{2} \cdot \sqrt{\frac{\tau_T}{\tau_L} \cdot \frac{R_B}{R_E}}$$

Unless τ_L is close to the upper end of the hf range, $\tau_T/\tau_L << 1$, and ζ is too low to limit design. Similarly, for $R_B << R_E$, the lower limit of ζ is lower than what is desired for amplifier design and is not restricted.

As other elements are added to Z_B, $D(s)$ remains unchanged and $N(s)$ becomes a higher-degree polynomial. The problem is not so much design of Z_E but bandwidth estimation. With $Z_B = R_B$, the dominant pole of the compensated amplifier is at $1/\tau_R$.

Another way of handling the complexity of pole-zero polynomials is to abandon algebra in favor of reactance plots. The hf equivalent plots of $(Z_B + Z_\pi)/(\beta(s) + 1)$ combined with those of Z_E can reveal the shape of z_M and the locations of poles and zeros. It is a graphic method of solving for polynomial roots and is demonstrated in the next chapter. (See cascode amplifier 3.)

The following example continues from the previous refinement of the CE design 1A (page 196) using the same circuit but with emitter

bypass compensation. Choose a pole angle of 30° (MFED response) for which $\zeta = \sqrt{3}/2$. Because $r_E << R_E$, the hf design formulas can be applied as closely approximate;

$$R \cdot C = 2 \cdot \zeta \cdot \sqrt{\frac{\tau_T \cdot R_B}{R_E} \cdot \tau_L} - \frac{\tau_T \cdot R_B}{R_E} = \sqrt{\frac{3 \cdot (70.7\,\text{nH})}{1067\,\Omega} \cdot 19.33\,\text{ns}} - \frac{70.7\,\text{nH}}{1067\,\Omega} = 1.89\,\text{ns}$$

$$C_E = \frac{\tau_T}{R_E} \cdot \frac{\tau_L}{R \cdot C} = (0.663\,\text{pF}) \cdot [(19.33\,\text{ns})/1.894\ \text{ns})] = 6.76\,\text{pF}$$

$$C = \frac{\tau_T + \tau_L - R \cdot C - \tau_T \cdot \tau_L / R \cdot C}{R_E} = 10.24\,\text{pF}$$

$$R = R \cdot C / C = 185\,\Omega$$

The remaining two poles of the BJT are from the time constants;

$$\tau_e = \tau_{be} = 0.683\ \text{ns}$$

$$\tau_c = \tau_{cb} + \tau_{cc} = 9.23\ \text{ns}$$

R·C lies between them making τ_c the largest time constant. It determines the bandwidth as $f_{bw} \approx 17$ MHz.

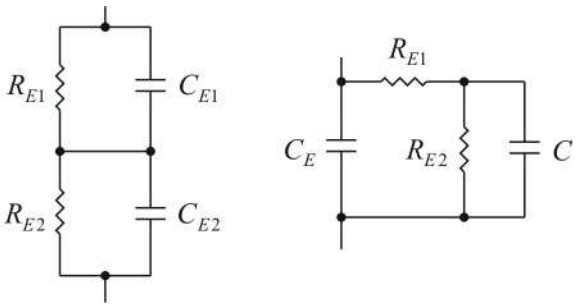
The series RC in the emitter also refers to the base but is included in its effects on Z_b in the above design formulas as α compensation. Without this inclusion, C would cause negative elements to appear at the base. These could otherwise be compensated by the addition of a series RC from base to ground which forms an all-pass network with the referred –R, –C from the emitter. (The details are presented in *Designing Dynamic Circuit Response* in the references.)

The Z_E circuit for which design formulas have been derived is not unique. Equivalent passive circuits can be used instead. Some are shown below, with their equivalent values. The given circuit has the advantage that R_E is a single resistor that sets the quasistatic gain. The other three components can be modified to optimize the dynamic response.

Left circuit:

$$Z_E = R_E \cdot \frac{s \cdot (R_{E1} \parallel R_{E2}) \cdot (C_{E1} + C_{E2}) + 1}{(s \cdot R_{E1} \cdot C_{E1} + 1) \cdot (s \cdot R_{E\,2} \cdot C_{E2} + 1)}$$

$$= R_E \cdot \frac{s \cdot (R_{E1} \parallel R_{E2}) \cdot (C_{E1} + C_{E2}) + 1}{s^2 \cdot [R_{E1} \cdot C_{E1} \cdot R_{E2} \cdot C_{E2}] + s \cdot [R_{E1} \cdot C_{E1} + R_{E2} \cdot C_{E2}] + 1}$$



Set $R_{E1}\cdot C_{E1} = \tau_T$, $R_{E2}\cdot C_{E2} = \tau_L$. Then

$$R_{E1} = R_E \cdot \frac{\tau_T - \tau_R}{\tau_T - \tau_L};\; R_{E2} = R_E \cdot \frac{\tau_R - \tau_L}{\tau_T - \tau_L}$$

$$C_{E1} = \frac{\tau_T}{R_E} \cdot \frac{\tau_T - \tau_L}{\tau_T - \tau_R}\;;\; C_{E2} = \frac{\tau_L}{R_E} \cdot \frac{\tau_T - \tau_L}{\tau_R - \tau_L}$$

$$R_E = R_{E1} + R_{E2}$$

Right circuit:

$$Z_E = \frac{1}{s \cdot C_E} \| (R_{E1} + \frac{R_{E2}}{s \cdot R_{E2} \cdot C + 1}$$

$$\Rightarrow Z_E = R_E \cdot \frac{s \cdot \tau_R + 1}{s^2 \cdot [\tau_T \cdot \tau_L] + s \cdot [\tau_T + \tau_L] + 1}$$

$$\tau_R = (R_{E1} \| R_{E2}) \cdot C\;;\; \tau_T \cdot \tau_L = R_{E1} \cdot R_{E2} \cdot C_E \cdot C$$

$$\tau_T + \tau_L = R_E \cdot C_E + R_{E2} \cdot C\;;\; R_E = R_{E1} + R_{E2}$$

$$R_{E1} = R_E \cdot \frac{\tau_R}{\tau_T + \tau_L - \tau_T \cdot \tau_L / \tau_R}\;;\; R_{E2} = R_E - R_{E1}$$

$$C_E = \frac{\tau_T \cdot \tau_L}{\tau_R \cdot R_E}\;;\; C = \frac{\tau_T + \tau_L - \tau_T \cdot \tau_L / \tau_R}{R_{E2}}$$

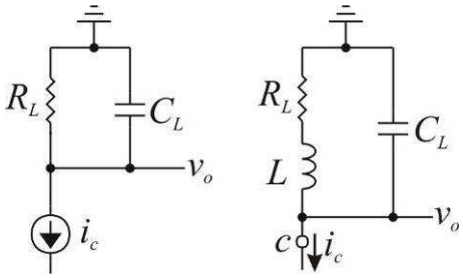
For additional bandwidth extension, shunt and series inductive peaking can be introduced into the collector circuit. By adding an inductor in series with R_L to implement *shunt inductive peaking*, then both a zero and a pole-pair appears, formed by the collector inductance, L, with C_o. The zero is at $-1/(L/R)$, where R is the collector-node resistance, R_L. If L is chosen so that the pole-pair has a

pole angle of 30° (MFED) then the collector pole is extended about 1.6 times to about 32 MHz.

Shunt Inductive Peaking

The collector-node time constant can be effectively reduced by adding an inductance that resonates with the unintended and undesirable *parasitic* capacitance to ground, C_L. This includes the input capacitance of the next stage. The scheme in general is referred to as *inductive peaking*. The inductor is in parallel with C_L for *shunt peaking* and in series for *series peaking*, though for both, the resonance formed is a series resonance.

Shunt peaking compensates for the effect of C_L on bandwidth by giving R_L a reactive component, making it Z_L.



The load impedance of the stage is its transfer function:

$$Z_L = \frac{v_o(s)}{i_c(s)}$$

For the uncompensated circuit on the left,

$$Z_L = R_L \parallel (1/s \cdot C_L) = \frac{R_L}{s \cdot \tau_L + 1} = \frac{R_L}{s/\omega_{bw} + 1}$$

where ω_L is the collector pole, $\omega_{bw} = 1/\tau_{bw} = 1/R_L \cdot C_L$.

Shunt peaking extends the bandwidth by modifying Z_L. With shunt peaking,

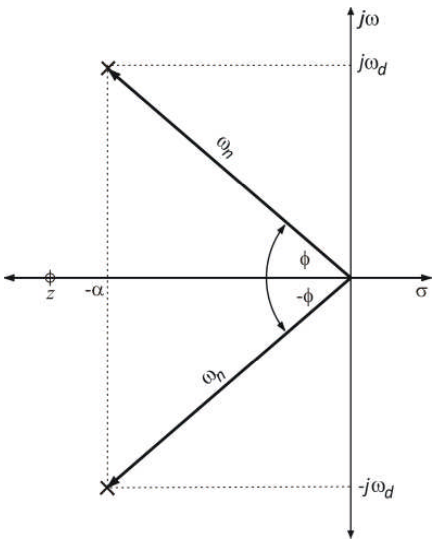
$$Z_L = (R_L + s \cdot L) \parallel (1/s \cdot C_L) = R_L \cdot \frac{s \cdot (L/R_L) + 1}{s^2 \cdot L \cdot C_L + s \cdot R_L \cdot C_L + 1}$$

Two changes to Z_L have occurred. A zero has been added and the pole factor is now quadratic, with possibilities that its roots are complex and that the response can be underdamped.

As we have seen, slightly underdamped resonances have faster risetime and higher bandwidth than equal real poles. The step response can be found after the denominator has been factored and the poles and zeros known. A second-degree or quadratic polynomial will have two first-degree (linear) factors and two roots which are the poles. If the poles are complex, they will be conjugates of each other, as shown on the *s*-domain plot.

The quadratic factor can be set to zero and the roots found using the quadratic formula from algebra. In a normalized transfer function, quadratic factors, whether poles or zeros, appear as

$$a \cdot s^2 + b \cdot s + 1 = \left(\frac{s}{\omega_n}\right)^2 + (2 \cdot \zeta) \cdot \left(\frac{s}{\omega_n}\right) + 1$$



The roots are

$$s = -\alpha \pm j \cdot \omega_d = -\zeta \cdot \omega_n \pm j \cdot \omega_n \cdot \sqrt{1-\zeta^2}$$

where the real and imaginary components are

$$\alpha = \zeta \cdot \omega_n = \omega_n \cdot \cos\phi \ , \ \omega_d = \omega_n \cdot \sqrt{1-\zeta^2} = \omega_n \cdot \sin\phi$$

On the above pole-zero plot, the complex poles show a vector magnitude - the length of the hypotenuse of a right triangle or radius of a circle - which is ω_n, a pole angle of ϕ, a real component of length α (adjacent side of triangle) and an imaginary component of ω_d (opposite side). The $\sin\phi$ and $\cos\phi$ factors of ω_n in the complex components follow from basic trigonometry.

This form of the quadratic factor reveals the effect of its terms. From the coefficient of the quadratic (squared) term, the natural or resonant frequency can be found;

$$\omega_n = \frac{1}{\sqrt{a}}$$

Then from the linear term, the damping can be found;

$$b = \frac{2 \cdot \zeta}{\omega_n} \Rightarrow \zeta = \tfrac{1}{2} \cdot b \cdot \omega_n$$

Substituting for ω_n,

$$\zeta = \frac{b}{2 \cdot \sqrt{a}}$$

With these general results, we can find the locations of s_1, s_2 for quadratic factors in the *s*-plane.

From the transfer function for the circuit, we can derive the values of the poles and zeros:

$$\omega_n = 2 \cdot \pi \cdot f_n = \frac{1}{\sqrt{L \cdot C_L}}$$

$$\zeta = \frac{R_L \cdot C_L}{2 \cdot \sqrt{L \cdot C_L}} = \frac{1}{2} \cdot \frac{R_L}{\sqrt{L / C_L}}$$

The denominator of ζ is in the form of an impedance, and ζ simplifies further to

$$\text{series resonant } \zeta = \frac{1}{2} \cdot \frac{R_L}{Z_n}$$

As R_L increases, damping increases. This is the formula for ζ for series resonance. (For parallel resonance, exchange R_L and Z_n.)

The time constant of the zero is

$$\frac{L}{R_L} = \frac{L \cdot C_L}{R_L \cdot C_L} = \frac{1/\omega_n^2}{2 \cdot \zeta / \omega_n} = \frac{1}{2 \cdot \zeta \cdot \omega_n}$$

The bandwidth of the shunt peaking output node is found by setting

$$\| Z_L(j \cdot \omega_{bw}) \|^2 = \tfrac{1}{2}$$

When this is solved (after some complex-number algebra),

$$\frac{\omega_{bw}}{\omega_n} = \sqrt{\left(1 + \frac{1}{(2 \cdot \zeta)^2} - \frac{(2 \cdot \zeta)^2}{2}\right) + \sqrt{\left(1 + \frac{1}{(2 \cdot \zeta)^2} - \frac{(2 \cdot \zeta)^2}{2}\right)^2 + 1}}$$

Q is the magnitude of a resonant pole-pair at ω_n and is

$$\|Z_L(j \cdot \omega_n)\| = Q \equiv 1/2\zeta$$

Rewriting the above equation in Q instead of ζ,

$$\frac{\omega_{bw}}{\omega_n} = \sqrt{\left(1 + Q^2 - \frac{1}{2 \cdot Q^2}\right) + \sqrt{\left(1 + Q^2 - \frac{1}{2 \cdot Q^2}\right)^2 + 1}}$$

Furthermore, the uncompensated bandwidth is the pole frequency,

$$\omega_p = \frac{1}{R_L \cdot C_L}$$

or from Z_L, $\omega_p = 1/b = 1/(2 \cdot \zeta \omega_n)$. Then

$$\frac{\omega_{bw}}{\omega_p} = (2 \cdot \zeta) \cdot \left(\frac{\omega_{bw}}{\omega_n}\right)$$

With these equations we can select a pole magnitude and damping, then solve for the component values. The table below has some common pole angles.

ϕ, °	ζ	$\frac{f_{bw}}{f_n}$	$\frac{f_{bw}}{f_p}$	M_p, %	Shunt Peaking Response
0	1	0.7071	1.414	0	critical damping
30	0.866	0.9204	1.594	0.620	MFED
45	0.707	1.272	1.799	6.70	MFA
60	0.5	1.817	1.817	29.8	

The ratio of bandwidth to pole magnitude, f_{bw}/f_n, increases with pole angle but so does the undesirable overshoot, M_p, of a step response. The bandwidth improvement factor, f_{bw}/f_p, is the ratio of shunt-peaking bandwidth to the uncompensated, single-pole circuit bandwidth. Shunt peaking can increase bandwidth 60 % to 80 %, a substantial improvement in circuit speed.

By solving the derived equations for component values as a function of performance parameters, we obtain design equations for shunt peaking. The circuit is first designed quasistatically and R_L is determined to give the desired quasistatic gain. C_L is estimated from C_c of the BJT and C_o of whatever is loading the collector node. Then with R_L and C_L given, the dynamic response parameter is chosen. Either M_p or ζ determines the pole angle which is chosen to give the desired step response. Then from the bandwidth extension, f_{bw}/f_n, in the table, f_n can be found with bandwidth as a design parameter. Given both f_n and ζ, first find

$$Z_n = \frac{R_L}{2 \cdot \zeta} = \sqrt{\frac{L}{C_L}}$$

Solving for L,

$$L = Z_n^2 \cdot C_L = \left(\frac{R_L}{2 \cdot \zeta} \right)^2 \cdot C_L$$

With this one equation and the above table, quasistatic amplifier designs can be given shunt peaking.

Which capacitances should be included in C_L? From the general single-stage BJT model, a "dual" of Miller's Theorem falls out from a different factorization of the open-circuit resistance of C_c, where b-c current gain is defined as

$$K_i = \alpha_0 \cdot \frac{R_b}{r_e + R_E} = \frac{R_b}{R_L} \cdot K_v$$

$$R_{bc} = R_L + R_b \cdot (1 + K_v) = R_L + R_b \cdot \left(1 + \alpha_0 \cdot \frac{R_L}{r_e + R_E} \right)$$

$$= R_b + R_L \cdot \left(1 + \alpha_0 \cdot \frac{R_b}{r_e + R_E} \right) = R_b + R_L \cdot (1 + K_i)$$

The factorization, $R_L + R_b \cdot (1 + K_v)$, views R_{bc} from the base as the Miller multiplier applied to the base resistance with the collector resistance, R_L, in series with it. The alternative factorization, $R_b + R_L \cdot (1 + K_i)$, is a collector-side view, where a "dual Miller multiplier", $(1 + K_i)$, is applied to the collector resistance with R_b added to it. K_i is a meaningful gain that is often used in fast amplifier design because stages are usually driven by current sources with a Norton equivalent input resistance of R_b.

Now consider the linear pole coefficient of the general single-stage BJT amplifier and factor it in three different ways by collecting terms according to R_L, R_b, and the capacitances:

$$\text{base node, } R_b\text{: } R_b \cdot [C_e + (1 + K_v) \cdot C_c] + R_L \cdot (C_L + C_c)$$

$$\text{collector node, } R_L\text{: } R_L \cdot [C_L + (1 + K_i) \cdot C_c] + R_b \cdot (C_e + C_c)$$

$$\text{OCTCs, capacitors: } \begin{aligned} &(R_L + R_b \cdot (1 + K_v)) \cdot C_c + R_L \cdot C_L + R_b \cdot C_e \\ &= (R_b + R_L \cdot (1 + K_i)) \cdot C_c + R_L \cdot C_L + R_b \cdot C_e \end{aligned}$$

In the first two expressions, the terms are differentiated by node. The Miller multiplier has K_v for base-referred C_c and K_i for collector-referred C_c. It might seem that $[C_L + (1 + K_i) \cdot C_c]$ is the capacitance in inductive peaking at the output node, but the correct value is found in the OCTC equation associated with R_L of C_L. C_c forms its own pole with R_{bc} that can be expressed equivalently as referred by the Miller multiplier to either base or collector node.

Amplifier bandwidth is increased for a given R_L by keeping R_b small. Then τ_{cb} is small (despite the Miller effect at the base), C_L, C_c are approximately in parallel, and are combined at the collector into a single pole as

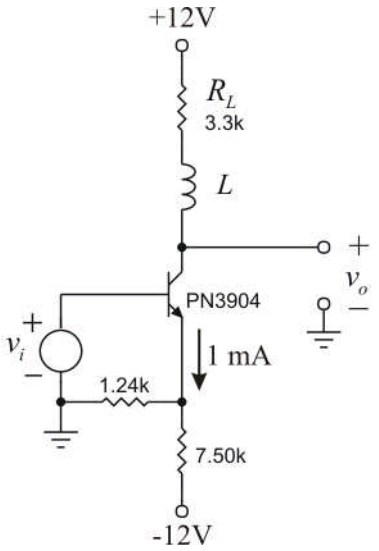
$$R_L \cdot (1 + K_i)) \cdot C_c + R_L \cdot C_L \cong R_L \cdot (C_c + C_L),\ K_i \approx 0$$

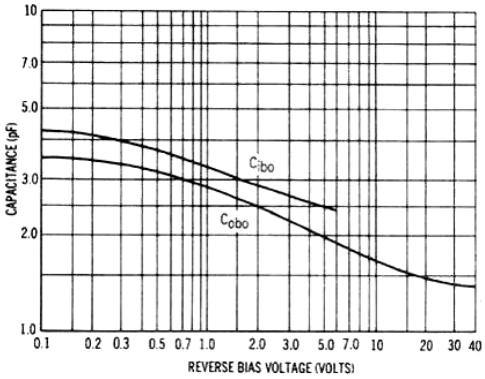
$K_i = 0$ whenever $R_b = 0\ \Omega$. Thus, $C_L \approx C_c + C_o$, $R_b \approx 0\ \Omega$, where C_o is the output capacitance, usually the stage loading the collector node.

As an example, the CE amplifier from the previous design (amplifier 1) is shown with shunt inductive peaking.

The Motorola data for a 2N3904 BJT shows the following plots (*Small-Signal Transistor Data*, DL-126, p. 2-3). The plots are of capacitance as a function of V_{BC}, though the quantities given are not denoted as C_c or C_e. C_{obo} is the collector capacitance to the base with emitter open, or C_c, and C_{ibo} is the reverse-biased (nonconducting)

emitter-to-base capacitance with collector open, or the junction capacitance of C_e. Then $C_c \approx C_{obo}$ and at V_{CB} = 8.8 V, $C_c \approx 1.8$ pF.





The additional capacitive loading on the collector node from a 100 MHz oscilloscope ×10 probe is about 15 pF. Then $C_L \approx 15$ pF when the output is probed for measurement.

The uncompensated single-pole bandwidth has OCTCs:

$$\tau_e = \alpha_0 \cdot \tau_T \cdot \frac{R_B + R_E}{r_M} = (0.993) \cdot (707 \text{ ps}) \cdot \frac{50\,\Omega + 1064\,\Omega}{(1064 + 26 + 50/151)\,\Omega} = 0.718 \text{ ns}$$

$$\Rightarrow 222 \text{ MHz}$$

$$\tau_c = [R_L + R_b \cdot (1 + K_v)] \cdot C_c = [1064\ \Omega + (50\ \Omega) \cdot (4)] \cdot (2 \text{ pF}) = 2.53 \text{ ns}$$

$$\Rightarrow 63 \text{ MHz}$$

The dominant collector pole is at

$$\tau_{L1} = R_L \cdot C_L = (3.3\ \text{k}\Omega)\cdot(15\ \text{pF}) = 49.5\ \text{ns} \ \Rightarrow 3.215\ \text{MHz}$$

The OCTC bandwidth is then

$$\tau_{bw1} \approx \sqrt{\tau_e^2 + \tau_c^2 + \tau_{L1}^2} = 49.6\,\text{ns} \ \Rightarrow f_{bw} = 3.21\ \text{MHz}$$

The response for this design is chosen to be MFA. From the table, $\zeta \approx 0.707$ and the extended pole bandwidth is

$$f_{bw} \approx (1.8)\cdot(3.215\ \text{MHz}) \approx 5.78\ \text{MHz}$$

The series-resonant impedance needs to be

$$Z_n = \frac{R_L}{2\cdot\zeta} = 2.33\,\text{k}\Omega$$

and

$$L = Z_n^2 \cdot C_L = \left(\frac{R_L}{2\cdot\zeta}\right)^2 \cdot C_L = 81.7\ \mu\text{H}$$

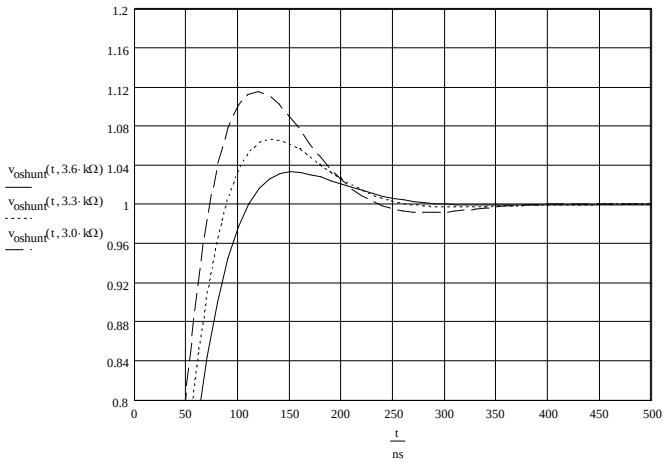
The location of the zero frequency is at $z = -1/(L/R_L) = 6.43$ MHz in shunt peaking and is greater than $f_n = 4.55$ MHz. The extended bandwidth from the table must then be $(1.272)\cdot f_n = 5.78$ MHz.

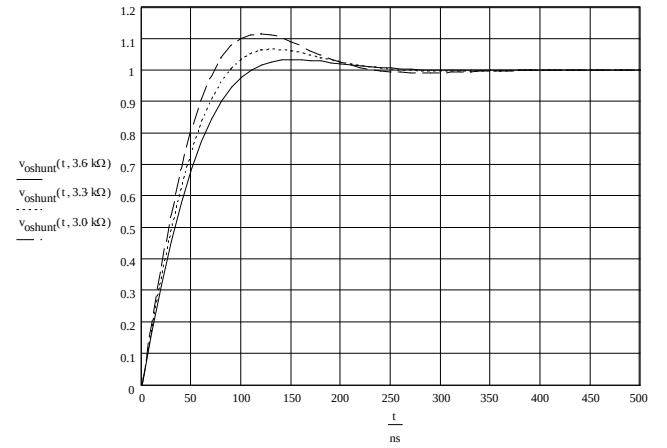
Some step responses for shunt peaking near the top corner of the step are shown below for R_L and values ±5 % around its value. A 5 % variation in R_L can make a big difference in the damping. For $R_L = 3.3$ kΩ, the overshoot is the value in the table: 6.7 % for a peak at 1.067 V with a voltage step of 1 V. Because of parts tolerances, L is usually made a variable (adjustable) component so that the quasistatic gain is unaffected. The complete range of the steps is shown below for risetime calculation.

Amplifier dynamic performance is specified either as risetime or bandwidth. For a single-pole response having $\omega_{bw} = \omega_p$, the relationship between the two from the risetime formula is simply

$$t_r \approx 2.2\cdot\tau = \frac{2.2}{\omega_{bw}} = \frac{2.2}{2\cdot\pi\cdot f_{bw}} \approx \frac{0.35}{f_{bw}}$$

where for a single pole, $p = \omega_{bw} = 1/\tau$. A 10 MHz bandwidth corresponds to a 35 ns risetime for a single-pole amplifier. For complex pole-pairs and a zero, it is more complicated.





From the lower plot, $t_r \approx 69\text{ ns} - 6\text{ ns} = 63\text{ ns}$ for the example circuit. The single-pole bandwidth calculated from single-pole risetime, is

$$0.35/63\text{ ns} = 5.56\text{ MHz}$$

or about −11 % from the theoretical bandwidth. The zero both undamps and extends the bandwidth somewhat. A prototype of amplifier 2 had a measured uncompensated (no L) $f_{bw} = 2.85$ MHz. (Note the disparity in value from that of unbuffered amplifier 1A of the same circuit, though a separate unit.) Adding an 87 μH inductor resulted in $f_{bw} = 5.38$ MHz with a flat response. The improvement in bandwidth is $f_{bw}/f_p = 1.89$. The slightly higher (than 81 μH) inductance of 7.5 % had no observable peaking in the frequency

response but did extend the bandwidth by 5 % more than the 81 μH calculated for MFA response. Even so, the measured bandwidth falls short of the predicted value by about 7 %. Some of this might be explained as parasitic resistance of the inductor, either in the winding as series electrical loss or as referred to the winding by the magnetic field as magnetic power dissipation.

An inductor with a magnetic path of air and a NiZn ferrite "slug" for inductance adjustment is preferred over MnZn ferrites or powdered-iron magnetic materials that dissipate excessive power over 1 MHz. This *core power loss* is an equivalent shunt resistance across the inductance. It increases damping and varies with frequency.

Z_n of the previous circuit is somewhat large (2.33 kΩ) for a design of this kind. What makes it large is that in practice, L is made large and a 100 μH inductor will have multiple turns with parasitic capacitance between the turns. Worse yet, it might have multiple layers that significantly increase skin and proximity effects. These effects crowd conduction toward the surface of the conductor cross-section so that the effective conductor area is reduced and resistance is increased. These effects increase with frequency, causing increased series resistance. Multiple layers also increase parasitic capacitance. The shunt parasitic capacitance across the inductor as a component can be comparable to C_L or even exceed it.

These practical considerations limit the value of R_L, and for the same gain, limit R_E. If each is reduced so that R_L = 1.0 kΩ, R_E = 430 Ω, and L = 15 μH, then measured f_{bw} = 16.0 MHz, measured M_m = 1.03 and f_m = 5.1 MHz. The step response has no observable overshoot, the generator $t_r \approx 6$ ns and the measured t_r = 24 ns. Then the corrected t_r = 23.2 ns and 0.35/23.2 ns ≈ 15.1 MHz.

Frequency Response Characterization

The design *criteria* (the parameters on which the design is based) for shunt-peaking dynamic response emphasized the time-domain: step risetime and overshoot. In some applications, a "flat" frequency response is more important. A *flat response* is one in which the frequency-response magnitude remains constant out to or near the bandwidth. Then sine-waves of differing frequencies below bandwidth with the same input amplitude will have the same output amplitude. In other words, sine-waves will have the same gain up to the bandwidth of the amplifier. This is important in sine-wave instrumentation. The amplifier for the vertical axis of an oscilloscope

needs to have constant gain over its bandwidth so that waveforms have the same scale factor in volts per division (V/div) on the display. Audio and video amplifiers also generally require that all frequencies within the amplified frequency band be amplified equally.

Single-pole amplifiers have a flat response to bandwidth and, as for resonance, are not underdamped. For circuits with a complex pole-pair in the transfer function, as the resonant frequency, ω_n (or f_n) is approached, the gain increases. The resonant impedance plots of the "Resonance" section show this peaking for a parallel resonant circuit. Underdamped circuits such as the shunt peaking circuit have a peak magnitude at

$$f_m = f_n \cdot \sqrt{1-2\cdot\zeta^2}\ ,\ \zeta < \frac{\sqrt{2}}{2} \approx 0.707$$

For $\zeta \geq \sqrt{2}/2 \approx 0.707$, the frequency response is flat and there is no peak frequency. The normalized peak magnitude at ω_m is

$$M_m = \frac{1}{2\cdot\zeta\cdot\sqrt{1-\zeta^2}}\ ,\ \zeta < \frac{\sqrt{2}}{2} \approx 0.707$$

For highly underdamped (or nearly undamped) circuits,

$$M_m \approx 1/\zeta,\ \zeta << 1$$

At the resonant frequency (or pole vector magnitude) ω_n,

$$M_n(\omega_n) = \frac{1}{2\cdot\zeta} \equiv Q$$

For critical damping, $Q = 0.5$. The quantity Q is widely used in frequency-domain or *narrow-band* electronics such as radio or wireless communications as an alternative to damping, ζ. In these applications, bandwidth is relatively narrow and is between the lower and upper frequencies at the edges of the band. The emphasis in this book is on *wideband amplifiers*, those having a frequency band from 0 Hz to f_{bw} and beyond.

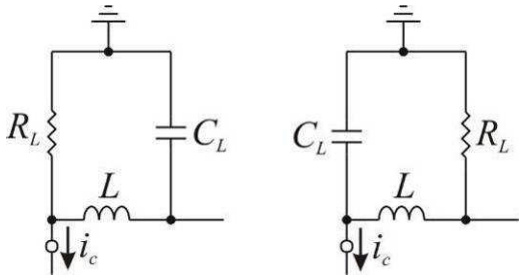
For a pole-pair (without zero) from a measured M_m, damping is

$$\zeta = \sqrt{\frac{1}{2}\cdot\left(1-\sqrt{1-\frac{1}{M_m^2}}\right)}$$

This useful formula enables calculation of pole-pair damping from a measured value of M_m.

Series Inductive Peaking

A variation in the previous inductive peaking circuit results in the *series peaking* circuit. Two circuits are shown that have the same series-peaking transfer function.



The transfer function for series peaking is found by noting first that for the circuit on the left, i_c develops v_c across

$$Z_C = R_L \parallel (s \cdot L + 1/s \cdot C_L) = \frac{R_L \cdot (s \cdot L + 1/s \cdot C_L)}{R_L + (s \cdot L + 1/s \cdot C_L)}$$

Then this voltage is divided, resulting in the transfer function,

$$Z_L = \frac{v_o}{i_c} = \frac{v_c}{i_c} \cdot \frac{v_o}{v_c} = Z_C \cdot \frac{1/s \cdot C_L}{s \cdot L + 1/s \cdot C_L} = \frac{R_L \cdot (1/s \cdot C_L)}{R_L + (s \cdot L + 1/s \cdot C_L)}$$

Put into normalized form,

$$Z_L = \frac{v_o}{i_c} = R_L \cdot \frac{1}{s^2 \cdot L \cdot C_L + s \cdot R_L \cdot C_L + 1}$$

The quadratic pole is the same as for shunt peaking. What is different is that the zero is missing.

For the above circuit on the right,

$$Z_L = [(1/s \cdot C_L) \parallel (s \cdot L + R_L)] \cdot \left(\frac{R_L}{s \cdot L + R_L} \right)$$

$$= \frac{s \cdot L + R_L}{s^2 \cdot L \cdot C_L + s \cdot R_L \cdot C_L + 1} \cdot \frac{R_L}{s \cdot L + R_L}$$

This simplifies to the same transfer function as the circuit on the left:

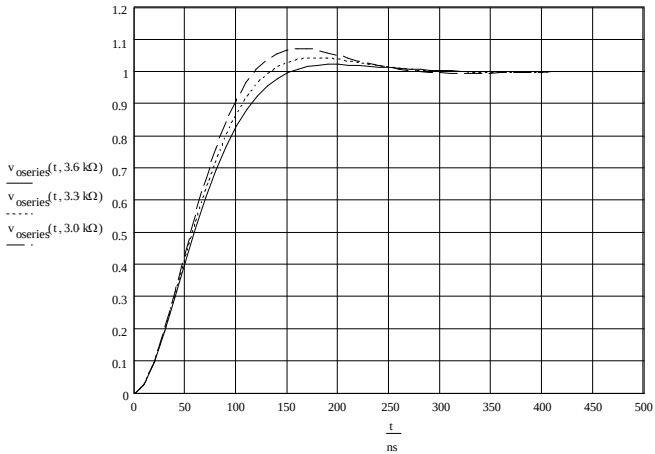
$$Z_L = R_L \cdot \frac{1}{s^2 \cdot L \cdot C_L + s \cdot R_L \cdot C_L + 1}$$

The design formula for L is the same as for shunt peaking. Both have a series resonance. The table given below is for series peaking.

ϕ, °	ζ	$\frac{\omega_{bw}}{\omega_n}$	$\frac{\omega_{bw}}{p}$	M_p, %	Series Peaking Response
0	1	0.644	1.288	0	critical damping
30	0.866	0.786	1.361	0.433	MFED
45	0.707	1.000	1.414	4.32	MFA
60	0.5	1.272	1.272	16.3	M_m = 1.155

For *precision amplification* - that which attempts to only scale up the input waveform without changing it in any other way - a pole angle of 30° (MFED response) is generally considered optimal. It has 0.43 % overshoot and the phase is linear.

The series-peaked circuit does not have as much overshoot for the same parameters as shunt peaking though it also has lower bandwidth and slower risetime for the same pole angle. Whereas shunt peaking increases bandwidth by 60 % to 80 %, series peaking increases it by only 35 % to 40 %.



Series peaking of the previous uncompensated circuit that was used for shunt peaking results in the plotted step responses, with R_L as

parameter. From the plot, the risetime for R_L = 3.3 kΩ is about 85 ns, about 35 % slower than shunt peaking.

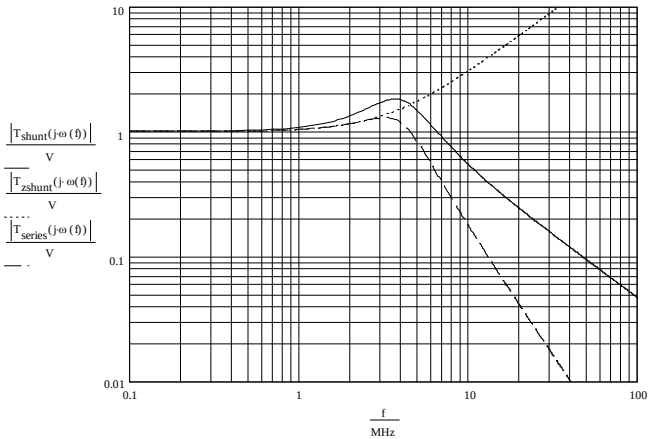
Series peaking has the same damping and resonant frequency as the shunt-peaked circuit. The general formulas apply for calculating f_m and M_m. To illustrate series peaking with a substantial M_m, reduce R_L to 2.0 kΩ. Then $\zeta = 0.428$, and the pole angle is increased to $\phi = 64.6°$.

$$f_m = f_n \cdot \sqrt{1 - 2 \cdot \zeta^2} = (4.55\,\text{MHz}) \cdot (0.796) = 3.62\ \text{MHz}$$

The peaking in the frequency response is

$$M_m = \frac{1}{2 \cdot \zeta \cdot \sqrt{1 - \zeta^2}} = \frac{1}{0.774} = 1.29$$

These values agree with the T_{series} plot, shown along with the shunt-peaked transfer function and its zero. The numerator and denominator of the transfer function, denoted on the graph as $T_{shunt}(j \cdot \omega)$, are separated so that the effect of each on $|T_{shunt}(j \cdot \omega)|$ can be seen. The zero is T_{zshunt} and the pole-pair, T_{series}.



Cochrun-Grabel Method

An improvement in accuracy over the OCTC method was introduced by Cochrun and Grabel in 1973 ("A Method for the Determination of the Transfer Function of Electronic Circuits", Basil L. Cochrun, Arvin Grabel, *IEEE Transactions on Circuit Theory*, Vol. CT-20, No. 1, JAN73). It proceduralizes determination of the polynomial coefficients of $D(s)$ for circuits with n capacitors but does

not include zeros. It is based on an expansion of $D(s)$ as in the OCTC method, where the linear term is the sum of the OCTCs. However, the method does not stop there and includes the coefficients of all the powers of s. In normalized form,

$$a_0 = 1$$

Designate the OCTCs as τ_i, $i = 1, \ldots, n$. Then the coefficient of the linear term is

$$a_1 = \tau_1 + \tau_2 + \cdots + \tau_n$$

The coefficient of s^2 is

$$a_2 = \tau_1 \cdot (\tau_{2;1} + \tau_{3;1} + \cdots + \tau_{n;1}) + \tau_2 \cdot (\tau_{3;2} + \tau_{4;2} + \cdots + \tau_{n;2}) + \cdots + \tau_{n-1} \cdot \tau_{n;n-1}$$

where $\tau_{m;k}$ is τ of the C_m port with the port of C_k shorted: $\tau_{m;\text{shorted}}$ with the remaining ports open. For the τ_1 factor, find the time constants of the other ports with port 1 shorted. Then short port 2 and do the rest of the ports. For each successive τ_i, the number of product terms decreases by one. For $n = 2$, it is the product of two time constants, an OCTC and the other with the OCTC port shorted.

$$a_2 = \tau_1 \cdot \tau_{2;1}$$

For $n = 3$,

$$a_2 = \tau_1 \cdot (\tau_{2;1} + \tau_{3;1}) + \tau_2 \cdot \tau_{3;2}$$

By the third degree, the nth-degree polynomial coefficient is quite cumbersome and nonintuitive:

$$\begin{aligned} a_3 = {} & [\tau_1 \cdot \tau_{2;1} \cdot (\tau_{3;1,2} + \tau_{4;1,2} + \cdots + \tau_{n;1,2}) + \\ & \tau_1 \cdot \tau_{3;1} \cdot (\tau_{4;1,3} + \cdots + \tau_{n;1,3}) + \cdots + \\ & \tau_1 \cdot \tau_{n-1;1} \cdot \tau_{n;1,n-1}] + \\ & [\tau_2 \cdot \tau_{3;2} \cdot (\tau_{4;2,3} + \tau_{5;2,3} + \cdots + \tau_{n;2,3}) + \\ & \tau_2 \cdot \tau_{3;2} \cdot (\tau_{5;2,4} + \cdots + \tau_{n;2,4}) + \cdots + \\ & \tau_2 \cdot \tau_{n-1;2} \cdot \tau_{n;2,n-1}] + \cdots \\ & \tau_{n-2} \cdot \tau_{n-1;n-2} \cdot \tau_{n;n-1,n-2} \end{aligned}$$

For $n = 3$, it is not nearly so daunting;

$$a_3 = \tau_1 \cdot \tau_{2;1} \cdot \tau_{3;1,2}$$

For a large (> 3) value of n, keeping track of the combinations of terms in each coefficient can be arduous, and Sol Rosenstark streamlined the method by tabulating it in his book, *Feedback Amplifier Principles* (Macmillan, 1986). (I also describe it in *Designing High-Performance Amplifiers*.) The use of *Rosenstark tables* is central to a refined Cochrun-Grabel procedure. The tables are triangular in form; the general second- and third-degree tables are shown below.

a_1	a_2
τ_1	$\tau_{2;1}$
τ_2	

a_1	a_2	a_3
τ_1	$\tau_{2;1}$	$\tau_{3;1,2}$
	$\tau_{3;1}$	
τ_2	$\tau_{3;2}$	
τ_3		

The columns are added to give the a_m after multiplying each column entry by the time constants in the columns to its left. The first (leftmost) column is a_1, the sum of the OCTCs. The second collects the combinations of a_2. Each column entry is then multiplied by the time constants in each of the columns to its left before adding it to the other (similarly-multiplied) column entries. The sums for columns a_2 and a_3 for $n = 3$ are

$$a_2 \text{ column: } \tau_{2;1}\cdot\tau_1 + \tau_{3;1}\cdot\tau_1 + \tau_{3;2}\cdot\tau_2$$
$$a_3 \text{ column: } \tau_{3;1,2}\cdot\tau_{2;1}\cdot\tau_1$$

Rosenstark truncates the tables at a_2 or a_3 as approximations to $D(s)$.

Textbook CE Stage Poles

While the Cochrun-Grabel method is not difficult to learn using Rosenstark tables, it is easy to retain intuitively in the case of quadratic or cubic polynomials. Most of the work is in finding the OCTCs. Because the order of choosing capacitors is arbitrary there are two possible ways of computing the quadratic coefficient and they are equivalent. It is demonstrated here for the textbook CE stage. First, find the OCTCs for both C_e and C_c:

$$R_{be} = R_B \| r_\pi \; ; R_{bc} = R_L + R_{be}\cdot(1 + R_L/r_m)$$

Now short C_c and find

$$R_{ex} = R_{be} \| (R_L \| r_m)$$

A shorted C_c places the dependent current source across the voltage upon which it is dependent. By the substitution theorem, if a current source of v_{be}/r_m is placed across v_{be}, then it is equivalent to a resistance of $v_{be}/(v_{be}/r_m) = r_m$.

Next, short C_e instead and find the resistance across the *b*-*c* port with *b*-*e* port shorted:

$$R_{c;e} = R_L$$

We only need to short one of the two capacitances for the procedure, but because the order of the capacitors does not matter, there are two ways of working the method and both are given here.

The linear coefficient, *b*, is the sum of the OCTCs:

$$b = R_{be} \cdot C_e + R_{bc} \cdot C_c$$

The quadratic coefficient can be found in two ways:

$$a = (R_{be} \cdot C_e) \cdot (R_{c;e} \cdot C_c) = (R_{be} \cdot C_e) \cdot (R_L \cdot C_c)$$

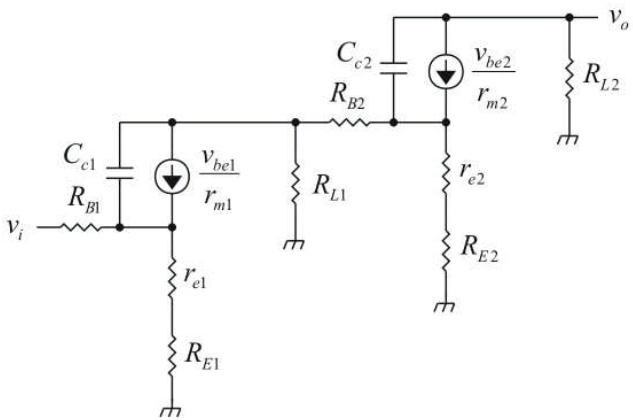
or

$$a = (R_{e;c} \cdot C_e) \cdot (R_{bc} \cdot C_c) = ([R_{be} \parallel (R_L \parallel r_m)] \cdot C_e) \cdot ([R_L + R_{be} \cdot (1 + R_L / r_m)] \cdot C_c)$$

After much algebra, the two expressions for *a* are found to be equal. In both cases, the OCTC of the first C is multiplied by the SCTC of the second capacitor with the first C shorted. For either alternative, the resulting pole polynomial is

$$s^2 \cdot (\tau_e \cdot \tau_{c;e}) + s \cdot (\tau_e + t_c) + 1 = s^2 \cdot a + s \cdot b + 1$$

Cascade CE Stage Interaction



Continuing the topic of interacting cascade stages, a simplified two-stage cascade CE amplifier incremental model is shown. The BJTs are ideally fast (τ_T = 0 s) with C_e = 0 pF. In simplifying the circuit model with fast transistors, we can gain some insight into stage interaction caused by C_c and how pole locations are affected by them - and also illustrate use of the Cochrun-Grabel method.

By defining

$$R_{L1}' = R_{L1} \| [R_{B2} + (\beta_0 + 1) \cdot r_{M2}] \ , \ r_{M2} = r_{e2} + R_{E2}$$

then the OCTC resistances of the two C_c are

$$R_{c1} = R_{L1}' + R_{b1} \cdot \left(1 + \alpha_{01} \cdot \frac{R_{L1}'}{r_{M1}} \right) \ , \ r_{M1} = r_{e1} + R_{E1} ,$$

$$R_{b1} = R_{B1} \| (\beta_{01} + 1) \cdot r_{M1}$$

$$R_{c2} = R_{L2} + R_{b2} \cdot \left(1 + \alpha_{02} \cdot \frac{R_{L2}}{r_{M2}} \right) \ , \ r_{M2} = r_{e2} + R_{E2} ,$$

$$R_{b2} = (R_{L1} + R_{B2}) \| (\beta_{02} + 1) \cdot r_{M2}$$

The OCTCs can now be used in the OCTC bandwidth formula, though we will venture somewhat farther to demonstrate for quadratic factors how to find the quadratic pole coefficient from circuit port resistances. With it we can also determine the damping.

The capacitor port resistances with the other capacitor shorted are, first for the C_{c1} port,

$$R_{c1sc} = R_{L1;2} + R_{b1} \cdot \left(1 + \alpha_{01} \cdot \frac{R_{L1;2}}{r_{M1}} \right) \ , \ C_{c2} \to \infty \text{ (shorted)}$$

where

$$R_{L1;2} = R_{L1} \| (R_{B2} + R_{L2} \| r_{M2}) \ , \ C_{c2} \to \infty \text{ (shorted)}$$

For the C_{c2} port,

$$R_{c2;1} = R_{L2} + R_{b2;1} \cdot \left(1 + \alpha_{02} \cdot \frac{R_{L2}}{r_{M2}} \right) \ , \ C_{c1} \to \infty \text{ (shorted)}$$

with

$$R_{b2;1} = (\beta_{02} + 1) \cdot r_{M2} \| [R_{B2} + (R_{L1} \| r_{M1} \| R_{B1})] \ , \ C_{c1} \to \infty \text{ (shorted)}$$

where subscripts *m*;*k* apply to the *m*th port with the *k* port(s) shorted and all the rest open.

As we already know, the coefficient of *s* in *D*(*s*) is the sum of the OCTCs;

$$b = R_{c1} \cdot C_{c1} + R_{c2} \cdot C_{c2}$$

The s^2 coefficient *a*, can be derived from either of the following expressions:

$$a = (R_{c1} \cdot C_{c1}) \cdot (R_{c2;1} \cdot C_{c2}) = (R_{c2} \cdot C_{c2}) \cdot (R_{c1;2} \cdot C_{c1})$$

In both expressions, the OCTC of one capacitor is multiplied by the *short-circuit time constant* (SCTC) of the other with the OCTC capacitor shorted.

To simplify the circuit somewhat more, let $\beta_{01}, \beta_{02} \to \infty$. Then

$$R_{L1}' = R_{L1}$$

$$R_{L1;2} \approx R_{L1} \parallel (R_{B2} + R_{L2})$$

$$R_{b1} \approx R_{B1} \; ; \; R_{b2} \approx R_{L1} + R_{B2}$$

$$R_{b2}' \approx R_{B2} + (R_{L1} \parallel r_{E1} \parallel R_{B1})$$

The previous resistances simplify to

$$R_{c1} \approx R_{L1} + R_{B1} \cdot \left(1 + \frac{R_{L1}}{r_{M1}}\right) \; ; \; R_{c2} \approx R_{L2} + (R_{L1} + R_{B2}) \cdot \left(1 + \frac{R_{L2}}{r_{M2}}\right)$$

and for the SCTCs, the resistances are

$$R_{c2;1} \approx R_{L2} + [R_{B2} + (R_{L1} \parallel r_{M1} \parallel R_{B1})] \cdot \left(1 + \frac{R_{L2}}{r_{M2}}\right) \; ;$$

$$R_{c1;2} = R_{L1sc} + R_{B1} \cdot \left(1 + \frac{R_{L1;2}}{r_{M1}}\right)$$

Then the coefficient of s^2 for poles is

$$a \approx \left[R_{L1} + R_{B1} \cdot \left(1 + \frac{R_{L1}}{r_{M1}}\right)\right] \cdot C_{c1} \cdot \left[R_{L2} + [R_{B2} + (R_{L1} \parallel r_{M1} \parallel R_{B1})] \cdot \left(1 + \frac{R_{L2}}{r_{M2}}\right)\right] \cdot C_{c2}$$

or

$$a \approx \left[R_{L1} \parallel (R_{B2} + R_{L2}) + R_{B1} \cdot \left(1 + \frac{R_{L1} \parallel (R_{B2} + R_{L2})}{r_{M1}}\right)\right] \cdot C_{c1} \cdot \left[R_{L2} + (R_{L1} + R_{B2}) \cdot \left(1 + \frac{R_{L2}}{r_{M2}}\right)\right] \cdot C_{c2}$$

Much algebra is needed to show that the two expressions for a are equal, and the choice of which to derive can also simplify or complicate physical interpretation of the result.

Stage interaction is evident in the difference between the short- and open-circuit resistances. If they were equal, the stage poles would be equal (and isolated), but instead $R_{sc} < R_{oc}$. The pole polynomial has the general form:

$$s^2 \cdot (\tau_{c1} \cdot \tau_{c2;1}) + s \cdot (\tau_{c1} + \tau_{c2}) + 1 \ , \ \tau_{c2} > \tau_{c2;1}$$

By comparing R_{c2} with $R_{c2;1}$, the difference is that in $R_{c2;1}$, R_L is shunted by $r_{M1} \| R_{B1}$. This causes a to be less than is required to form a perfect square that can be factored into two equal poles. Instead, the two poles are real but separated on the real axis by the value of equal poles, $-b/2 \cdot a$. (See "Pole Separation".) From a and b, ζ and ω_n can be calculated and the pole locations determined. Pole separation causes the dominant pole to decrease the amplifier gain at a –1 slope while an approximation that results in equal poles will roll it off at a –2 slope and have a lower value of f_T. These general equations can be used to determine a and b so that the two poles can be found using the quadratic formula.

The damping is

$$\zeta(\xi,k) = \frac{b}{2 \cdot \sqrt{a}} = \frac{1}{2} \cdot \left(\frac{\tau_{c1} + \tau_{c2}}{\sqrt{\tau_{c1} \cdot \tau_{c2;1}}} \right) = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_{c1}}{\tau_{c2;1}}} + k \cdot \sqrt{\frac{\tau_{c2;1}}{\tau_{c1}}} \right) = \frac{\sqrt{k}}{2} \cdot \left(\xi + \frac{1}{\xi} \right)$$

The results and plot of ζ are the same as for the interacting cascade RC integrators in "Cascade Stage Interaction" on page 215.

Minimum ζ occurs at $\sqrt{k} \geq 1$ at $\xi = 1$. For the cascade CE stages,

$$k = \frac{\tau_{c2}}{\tau_{c2;1}} = \frac{R_{c2}}{R_{c2;1}} = \frac{R_{L2} + (R_{L1} + R_{B2}) \cdot (1 + R_{L2} / r_{M2})}{R_{L2} + [R_{L1} \| R_{B1} \| r_{M1} + R_{B2}] \cdot (1 + R_{L2} / r_{M2})}$$

$$= \frac{1 + (R_{L1} + R_{B2}) / (R_{L2} \| r_{M2})}{1 + [R_{L1} \| R_{B1} \| r_{M1} + R_{B2}] / (R_{L2} \| r_{M2})}$$

The distinction between numerator and denominator of k is that in the denominator, R_{L1} is shunted by $R_{B1} \| r_{M1}$. For R_{B1}, $r_{M1} \to \infty$, then $k = 1$ and $\zeta = 1$. For $R_{B1} \| r_{M1} = 0\ \Omega$, then

$$k = 1 + \frac{R_{L1}}{R_{L2} \| r_{M2} + R_{B2}} = 1 + \frac{R_{L1}}{R_{i2;2}}$$

where $R_{i2;2}$ is the input resistance of the second stage with C_{c2} shorted. For $R_{L1} = R_{i2;2}$, then $k = 2$ and min $\zeta = \sqrt{2}$.

What the above analysis shows is that cascade CE stages with fast transistors and no other capacitances than C_c always have real poles. The general BJT stage also has only real poles. For amplifiers with only real poles, the OCTC bandwidth method is applicable.

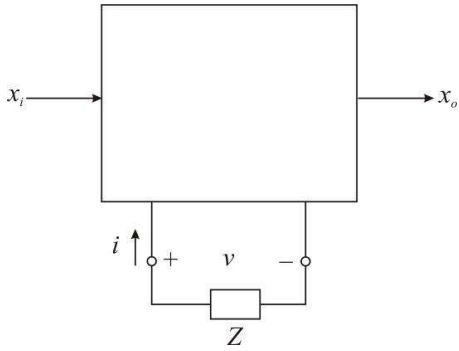
Amplifiers with additional stage reactances that can resonate are often approximated by a pole-pair and quadratic response. When there is pole interaction causing a complex pole-pair, bandwidth is more accurately calculated using the quadratic bandwidth formula. Three independent, interacting reactances produce a cubic pole polynomial, and with design constraints applied, can often be reduced to a quadratic polynomial and design formulas.

Both the OCTC and Cochrun-Grabel methods provide only poles and no transfer function zeros which also influence dynamic response. The Cochrun-Grabel method results in accurate poles, though a cubic or higher-degree polynomial must still be factored. A more refined method developed by R. David (or Robert D.) Middlebrook combines and extends linear-system port techniques by allowing for incremental expansion of transfer functions, one reactance at a time.

Extra Element Theorem (EET)

The *extra element theorem* (EET) was developed by R. D. Middlebrook as a refinement of a long history of related methods. The EET combines ideas that are found in the work of Gray and Searle at MIT (*Electronic Principles: Physics, Models, and Circuits*, Wiley, 1969) and in Blackman's Impedance Theorem. More can be determined from port analysis by subjecting the ports to different conditions than is at first apparent, and this is in part a consequence of the properties of linear systems. (The Cochrun-Grabel method and EET are based on linearized circuit variables that vary incrementally around a static operating-point.)

The EET is based on the diagram of a circuit (block) with input and output ports, x_i and x_o and an additional port somewhere in the circuit with external impedance Z across it having port voltage v and current i. The Z port is that of a circuit element - the "extra" element.



With Z attached to the circuit, the port *v*-*i* relationship is

$$v = -Z \cdot i$$

The negative sign signifies that Z is external to the port. By port convention, the port *driving-point* impedance, v/i, is that looking into the circuit port. The current direction is reversed (its polarity changed) to refer to Z. The port equations of the EET are

$$x_o = A_{oc} \cdot x_i + T_i \cdot i$$
$$v = T_v \cdot x_i + Z_D \cdot i$$

The four port parameters can be found:

$$A_{oc} = \frac{x_o}{x_i}\ ,\ i = 0 \text{ (port open)}\ ;\ T_i = \frac{x_o}{i},\ x_i = 0\ ;$$

$$T_v = \frac{v}{x_i}\ ,\ i = 0 \text{ (port open)}\ ;\ Z_D = \frac{v}{i},\ x_i = 0$$

A_{oc} is the gain from x_i to x_o with the Z port open-circuited. Z_D is the Z-port driving-point impedance, the impedance of the circuit from the port without the external Z and with the condition on Z_D that the input be set to zero.

From the port equations, substituting for v and solving for i,

$$-Z \cdot i = T_v \cdot x_i + Z_D \cdot i \ \Rightarrow\ i = -\frac{T_v}{Z + Z_D} \cdot x_i$$

Substituting for i in the port equation for x_o,

$$x_o = A_{oc} \cdot x_i - \frac{T_i \cdot T_v}{Z + Z_D} \cdot x_i \Rightarrow \frac{x_o}{x_i} = A_{oc} - \frac{T_i \cdot T_v}{Z + Z_D}$$

While this gain expression is in port parameters, it can be made more useful by eliminating T_i and T_v. To do this, another condition is imposed on the amplifier: x_o is *nulled*, or made to be zero. *Nulling* does not mean forcing x_o to be zero by shorting its node or opening its loop. A dependent variable such as x_o can be nulled only by adjusting circuit conditions so that $x_o = 0$. To do this, *assume* $x_o = 0$ and solve for the impedance at the *Z*-port that makes it so. This might sound difficult but it is usually easier than finding Z_D. Often the assumption of zero output ripples backwards through the circuit, reducing analysis significantly. The first port equation becomes

$$x_o = 0 = A_{oc} x_i + T_i \cdot i$$

Solving for *i* with the output nulled,

$$i = -\frac{A_{oc}}{T_i} \cdot \left(\frac{v - Z_D \cdot i}{T_v} \right) = \frac{-\frac{A_{oc}}{T_i \cdot T_v} \cdot v}{1 + \frac{A_{oc}}{T_i \cdot T_v} \cdot Z_D}$$

The *Z*-port impedance is then

$$Z_N = \frac{v}{i} = Z_D - \frac{T_i \cdot T_v}{A_{oc}}, \; x_o = 0$$

where Z_N is the *Z*-port output-nulled impedance. We can find both Z_D and Z_N by imposing conditions on the circuit. Solving the Z_N equation for the superfluous port parameters,

$$T_i \cdot T_v = A_{oc} \cdot (Z_D - Z_N)$$

They are eliminated by substituting them into

$$\frac{x_o}{x_i} = A_{oc} - \frac{T_i \cdot T_v}{Z + Z_D} = A_{oc} - \frac{A_{oc} \cdot (Z_D - Z_N)}{Z + Z_D} = A_{oc} \cdot \frac{Z + Z_N}{Z + Z_D}$$

Finally, the EET formula is

$$\frac{x_o}{x_i} = A_{oc} \cdot \left(\frac{1 + \dfrac{Z_N}{Z}}{1 + \dfrac{Z_D}{Z}} \right)$$

The factor in parentheses is called the *correction factor* because it modifies the otherwise unmodified gain A_{oc} by the effect of Z on the circuit. It adds poles or zeros to the existing transfer function, A_{oc}, to account for the added impedance - the "extra element".

What the EET enables us to do is to find the gain of the circuit when it is affected by the addition of Z. With Z removed (port open), the gain is A_{oc}. When Z is included, the modified gain can be found by finding the correction factor by finding Z_D and Z_N.

EET Procedure

1. Find A_{oc} with Z removed.
2. Find Z_D by setting the input to zero and deriving the port impedance at Z.
3. Find Z_N by nulling the output (with x_i applied) and find the port impedance.
4. Substitute Z_D and Z_N into the correction factor and solve for the modified gain.

The EET has a dual theorem that is expressed by exchanging v and i. It is derived from the dual of the above port equations and applies when an internal port is normally shorted (as by a circuit-board trace) and is opened to insert Z. For it,

$$A_{sc} = \frac{x_o}{x_i},\ v = 0 \text{ (port shorted)}$$

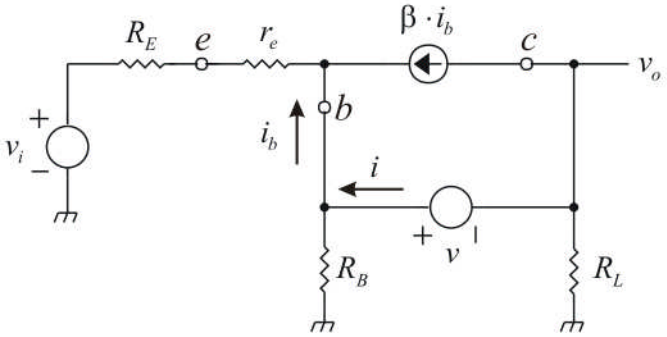
The short-circuit port EET formula is

$$\frac{x_o}{x_i} = A_{sc} \cdot \left(\frac{1 + \dfrac{Z}{Z_N}}{1 + \dfrac{Z}{Z_D}} \right)$$

The open- and short-circuit EET formulas differ only by the condition on A and by the Z ratios in the correction factor. To remember whether Z is in the numerators or denominators of the Z ratios, when

Z is shorted, the correction factor reduces to 1, leaving A_{sc}, the short-circuit gain. Thus Z for the short-circuit formula must be in the numerators. The open-circuit formula has Z in the denominators, and it must be infinite (open port) to cause the correction factor to be 1. Additionally, although the Z subscripts D and N stand for "driving-point" and "nulled", they can equally stand for "denominator" and "numerator", where they are found in both (dual) formulas.

Single CB Stage C_c



An example of the use of the EET is to find the gain of the CB stage with and without C_c. Let $C_e = 0$ pF, the ideal fast transistor without hf $\beta(s)$ effects. Z becomes $1/s{\cdot}C_c$ and the b-c port is the port across which C_c is placed to modify the circuit, shown with an external voltage source, v, applied to the b-c Z-port.

The quasistatic voltage gain is found from inspection with the C_c port open using the transresistance gain method;

$$A_{v0oc} = \alpha_0 \cdot \frac{R_L}{r_e + R_E + R_B/(\beta_0 + 1)}$$

Next, $Z_D = R_D$ is found by opening the b-c port and finding the open-circuit resistance. We have already done this for the general single-stage circuit; $R_D = R_{bc}$. To find R_N, assume $v_o = 0$ V. Then the current in R_L must be zero and by KCL at the collector node,

$$i = -\beta_0 \cdot i_b$$

The base voltage is

$$v_b = -(i_b - i) \cdot R_B = -(\beta_0 + 1) \cdot i_b \cdot R_B$$

Then

$$R_N = \frac{v}{i} = \frac{v_b}{-\beta_0 \cdot i_b} = \frac{-(\beta_0 + 1) \cdot i_b \cdot R_B}{-\beta_0 \cdot i_b} = \frac{R_B}{\alpha_0}$$

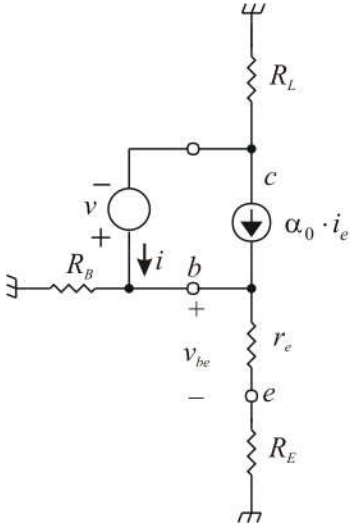
Combining these resistances along with the extra element, $Z = 1/s \cdot C_c$, into the open-circuit formula,

$$A_v = A_{v0_{oc}} \cdot \frac{1 + \frac{Z_N}{Z}}{1 + \frac{Z_D}{Z}} = A_{v0_{oc}} \cdot \frac{s \cdot (R_B / \alpha_0) \cdot C_c + 1}{s \cdot R_{bc} \cdot C_c + 1}$$

The EET method provides a complete transfer function by including the zero at $z = -1/(R_B/\alpha_0) \cdot C_c$ in addition to the open-circuit pole that the OCTC method produces. Thus the EET is a more powerful and complete circuit theorem that in this case required only a little more effort than the previous pole-only method, to find Z_N for the zero.

Single CE Stage C_c

The EET can also be used to find the RHP zero of the single CE stage.



Given the quasistatic $A_v = A_{voc}$ and the extra element, $Z = 1/s \cdot C_c$, then $Z_D = R_{bc}$. The nulled-output impedance, Z_N, is derived as follows.

$$i = -\alpha_0 \cdot i_e$$

$$v = v_b = r_M \cdot i_e = -i \cdot \frac{r_M}{\alpha_0}\ ,\ r_M = r_e + R_E$$

Then

$$Z_N = \frac{v}{i} = -\frac{r_M}{\alpha_0}$$

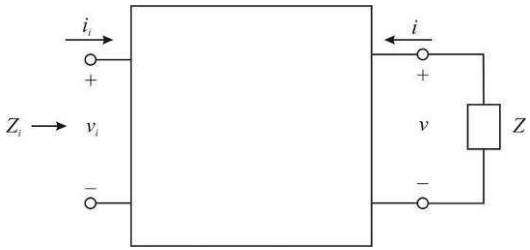
The gain with C_c is thus

$$A_v = A_{voc} \cdot \frac{1 + \dfrac{Z_N}{Z}}{1 + \dfrac{Z_D}{Z}} = A_{voc} \cdot \frac{1 + \dfrac{-r_M / \alpha_0}{1/s \cdot C_c}}{1 + \dfrac{R_{bc}}{1/s \cdot C_c}} = A_{voc} \cdot \frac{-s \cdot \left(\dfrac{r_M}{\alpha_0}\right) \cdot C_c + 1}{s \cdot R_{bc} \cdot C_c + 1}$$

The RHP zero time-constant resistance is the resistance of the emitter circuit whereas for the CB, it was of the base circuit.

Impedance Extra Element Theorem

A form of the EET for finding port impedances is the impedance EET (or ZEET) and is derived from the following diagram showing two ports.



The goal is to determine how a port of a circuit (on the right) affects the input impedance, Z_i, of the left-side input port. By opening and shorting the input port, the effect of Z can be determined on Z_i. The circuit port, for our interest, is chosen to be across a circuit capacitance such as C_c.

The first of three port equations is

$$v = -i \cdot Z$$

The negative sign signifies that Z is external to the port. By port convention, the port impedance looking into the box (that is the

circuit) is v/i. The current direction must be reversed (its polarity changed) to refer to Z. The other two equations are port parameter equations:

$$v_i = Z_{ioc} \cdot i_i + Z_c \cdot i \;;\; v = Z_v \cdot i_i + Z_D \cdot i$$

These equations are port functions of the form $v_i(i_i, i)$ and $v(i_i, i)$. The circuit is assumed linear and by superposition the effects of sources at the ports add. The coefficients must be impedances and are so designated, though they have yet to take on meaning. If the circuit (right) port is opened, $i = 0$ A and

$$Z_{ioc} = \frac{v_i}{i_i} \;,\; i = 0 \text{ A} \Rightarrow \text{circuit port opened}$$

It is the open-circuit Z_i. For the other parameters,

$$Z_c = \frac{v_i}{i};\; Z_v = \frac{v}{i_i},\; Z_D = \frac{v}{i}, i_i = 0 \text{ A} \Rightarrow \text{port opened}$$

For Z_v, the circuit port is opened. Z_D is the circuit-port impedance with the input port open. When it is shorted, the circuit-port impedance can be found by setting $v_i = 0$ V in the port equations:

$$0 \text{ V} = Z_{ioc} \cdot i_i + Z_c \cdot i \;;\; v = Z_v \cdot i_i + Z_D \cdot i$$

Solving for i_i in the first equation and substituting for i_i in the second,

$$v = Z_v \cdot \left(-\frac{Z_c}{Z_{ioc}} \right) \cdot i + Z_D \cdot i$$

Then the right-port Z with the input port shorted is

$$Z_N = \frac{v}{i} = Z_D - \frac{Z_v \cdot Z_c}{Z_{ioc}}$$

Now solve the two port equations without constraints applied to the ports for input impedance, v_i/i_i. Substitute for v from the first equation into the third,

$$-i \cdot Z = Z_v \cdot i_i + Z_D \cdot i \Rightarrow i = -\frac{Z_v}{Z + Z_D} \cdot i_i$$

then solve for i and substitute it into the second equation:

$$v_i = Z_{ioc} \cdot i_i + Z_c \cdot \left(- \frac{Z_v}{Z + Z_D} \right) \cdot i_i$$

The input impedance is

$$Z_i = \frac{v_i}{i_i} = Z_{ioc} - \frac{Z_c \cdot Z_v}{Z + Z_D}$$

This result can be put into a better form for use as a general method by writing it as

$$Z_i = Z_{ioc} \cdot \left(1 - \frac{Z_c \cdot Z_v}{Z_{ioc}} \cdot \frac{1}{Z + Z_D} \right) = Z_{ioc} \cdot \left(1 + \frac{Z_N - Z_D}{Z + Z_D} \right)$$

This simplifies to its working form as

$$Z_i = Z_{ioc} \cdot \frac{Z + Z_N}{Z + Z_D} = Z_{ioc} \cdot \frac{1 + \dfrac{Z_N}{Z}}{1 + \dfrac{Z_D}{Z}} \quad \text{(ZEET, open-circuit port)}$$

If the circuit port is opened, then $Z \to \infty$ and $Z_i = Z_{ioc}$, the original open-circuit Z_i. The effect of Z on Z_i is found by finding the three parameters. Z_N is the circuit-port Z when the input port is shorted, and Z_D is its impedance when the input-port is opened. It might be easier to remember the theorem in the following form:

$$Z_i = Z_{ioc} \cdot \frac{Z + Z_N}{Z + Z_D} = Z_{ioc} \cdot \frac{1 + \dfrac{Z_{sc}}{Z}}{1 + \dfrac{Z_{oc}}{Z}} \quad \text{(ZEET, open-circuit port)}$$

The rational factor after Z_{ioc} is called the *correction factor* because it changes the original circuit port impedance, Z_{ioc}, to account for Z.

The dual of the open-circuit ZEET has a circuit port that is normally a connection or short that is opened. The formula is

$$Z_i = Z_{isc} \cdot \frac{1 + \dfrac{Z}{Z_{sc}}}{1 + \dfrac{Z}{Z_{oc}}} \quad \text{(ZEET, short-circuit port)}$$

The two are easy to remember in that for the open-circuit port, as $Z \to \infty$, $Z_i \to Z_{ioc}$ and for the short-circuit port, setting $Z = 0\ \Omega$ results in the original circuit condition, that $Z_i = Z_{isc}$. In both cases, the correction factor becomes 1.

Stability of Feedback Circuits

The major impediment of feedback for fast circuits is the loop delay. As phase delay decreases to –180°, an inverted waveform is returned as x_B relative to x_i and it reinforces x_i by adding to it with the same polarity. This is *positive feedback*. It results in runaway behavior whereby x_o increases in amplitude until linear range limits are encountered.

The condition for stability of feedback amplifiers is that

$$||G(s)\cdot H(s)|| < 1 \text{ whenever } \angle\, G(s)\cdot H(s) = -180° = -\pi/2$$

On a Bode plot, the frequency at which the gain of $||G\cdot H|| = 1$ is the *unity-gain frequency*, f_T. The phase difference at f_T on the corresponding phase plot from –180° is the *phase margin*, a measure of the stability of the feedback loop. For a feedback loop with a complex pole-pair such as series peaking, the closed-loop damping can be approximated as

$$\zeta_{cl} \approx \frac{PM}{100}\ ,\ PM \text{ in degrees},\ \zeta_c < 0.7,\ PM < 64°$$

where PM is the phase margin:

$$PM = 180° - |\angle\, G(s)\cdot H(s)|$$

Single-pole loops are unconditionally stable because their phase decreases to no more than –90°. Two-pole loops asymptotically approach –180° and are theoretically unconditionally stable. However, real circuits always have higher-frequency poles that might cause the phase to continue to decrease somewhat and cross –180°. If this occurs for loop gain greater than one - in other words, at a frequency less than f_T - the loop will oscillate at the frequency at which the phase crosses –180°.

Control theory is a topic in itself and is a sub-discipline within electrical engineering. Techniques for stabilizing feedback circuits are many and varied. Below are a few of the most common methods.

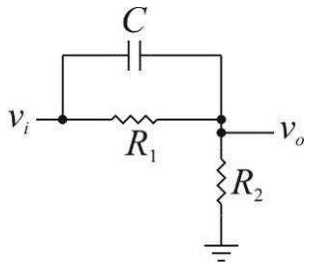
Dominant-pole compensation: Because a one-pole loop is stable, make that pole occur at a frequency far below the next higher-frequency pole. This lets the gain (magnitude) decrease to one before the higher poles contribute much phase delay.

Separate poles: More than two poles of about the same frequency cause excessive phase delay at that frequency. Spread the poles out, if possible. The Miller effect does this by causing two amplifier poles to move apart or "split", one toward 0 Hz and the other toward ∞.

Reduce quasistatic loop gain: If the loop has less gain at low frequencies, it will decrease to one at a lower frequency before the phase has decreased excessively.

Add zeros: Zeros in the loop transfer function will add phase lead (positive phase) and pull the phase angle upward, away from –180°. Circuits that add zeros are *phase-lead compensators.* Zeros cannot be added without poles. The RC differentiator is a kind of phase lead circuit. The classic phase-lead circuit is shown below. The resistors must have very different values to separate the zero from the pole so that the pole is at a much higher frequency, preferably greater than the feedback-loop (loop gain) f_T. The transfer function is a voltage divider in the *s*-domain:

$$\frac{v_o}{v_i} = \frac{R_2}{R_1 \parallel (1/s \cdot C) + R_2} = \frac{R_2}{R_1 + R_2} \cdot \frac{s \cdot R_1 \cdot C + 1}{s[(R_1 \parallel R_2) \cdot C] + 1}$$



The zero and pole, in frequencies of s^{-1} (or radian frequencies), are at

$$z = -\frac{1}{R_1 \cdot C}\ ,\ p = -\frac{1}{(R_1 \parallel R_2) \cdot C}$$

This circuit also reduces loop gain by the static attenuation of the resistive divider - the factor in front of the rational function in *s*. If R_2 is not much less than R_1, then $z \approx p$ and little phase lead will be added to the loop. A decade separation will cause the quasistatic gain to be reduced by about a decade, and at 45°/dec of phase lead, the loop gains 45°. Both effects contribute stability, though the benefits of feedback require sufficiently high loop gain.

Feedback Increases Bandwidth

The advantages of feedback extend to circuit dynamics. Consider a single-pole forward path,

$$G = \frac{G_0}{\frac{s}{\omega_{bw}} + 1}$$

where ω_{bw} is the pole frequency of *G*, the open-loop bandwidth of the amplifier. The closed-loop gain, from the feedback formula, is

$$A = \frac{G}{1 + G \cdot H}$$

Let *H* be (as it often is) a fixed resistive divider and frequency-independent; $H = H_0$. (It has no poles or zeros and is quasistatic.) Then substituting for *G* and *H* into the closed-loop formula,

$$A = \frac{\frac{G_0}{\frac{s}{\omega_{bw}} + 1}}{1 + \left(\frac{G_0}{\frac{s}{\omega_{bw}} + 1} \right) \cdot H_0} = \frac{G_0}{\frac{s}{\omega_{bw}} + (1 + G_0 \cdot H_0)}$$

This reduces to a closed-loop gain of

$$A = \frac{G_0}{1 + G_0 \cdot H_0} \cdot \frac{1}{\frac{s}{\omega_{bw} \cdot (1 + G_0 \cdot H_0)} + 1} = A_0 \cdot \frac{1}{\frac{s}{\omega_{bw}(\text{cl})} + 1}$$

The first factor is the closed-loop quasistatic gain, A_0. The second has a single closed-loop pole magnitude of

$$\omega_{bw}(\text{cl}) = (1 + G_0 \cdot H_0) \cdot \omega_{bw}$$

The bandwidth of the closed-loop amplifier has been extended from the open-loop ω_{bw} of G by the quasistatic feedback factor.

For single-pole rolloff, $G = 1$ at ω_T;

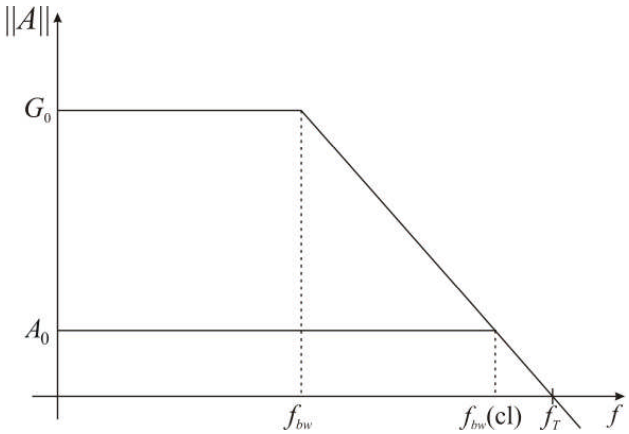
$$\omega_T = G_0 \cdot \omega_{bw}$$

From the quasistatic value of G_0 at ω_{bw}, $G(j \cdot \omega)$ rolls off until it reaches 1 (0 on a log plot) at $G_0 \cdot \omega_{bw}$. Then the closed-loop pole magnitude (which is the closed-loop bandwidth of A) can be expressed as

$$\omega_{bw}(\text{cl}) = (1 + G_0 \cdot H_0) \cdot \omega_{bw} = (1 + G_0 \cdot H_0) \cdot \left(\frac{\omega_T}{G_0} \right) = \left(\frac{1 + G_0 \cdot H_0}{G_0} \right) \cdot \omega_T = \frac{\omega_T}{A_0}$$

The closed-loop bandwidth is reduced from ω_T of G by the quasistatic closed-loop gain. A single-pole open-loop amplifier, G, with $f_T = 5$ MHz will have a closed-loop bandwidth, $f_{bw}(\text{cl}) = 1$ MHz for a closed-loop gain, $A_0 = 5$.

The following magnitude plot shows the open- and closed-loop gains and bandwidths of G. As the gain of G decreases, the log-difference between the open- and closed-loop gains also decreases. As frequency increases, the benefit of feedback diminishes. High precision can be achieved at low frequencies with high loop gain, but with increasing frequency, loop gain and precision decrease.



For single-pole rolloff, the closed-loop bandwidth, $f_{bw}(\text{cl}) > f_{bw}$ of the open loop by the ratio, or log difference, of their loop gains, $G_0/A(\text{cl})$;

$$f_{bw}(\text{cl}) = \frac{G_0}{A_0} \cdot f_{bw}$$

The closed-loop bandwidth can be extended to the open-loop f_T but only by decreasing closed-loop gain, for the closed-loop gain cannot exceed the open-loop gain and the closed-loop magnitude plot must remain within (below) the open-loop gain plot.

The single-pole constraint on G can be removed for the general case. On a log-log plot of $G{\cdot}H$,

$$\log\|G \cdot H\| = \log\|G\| + \log\|H\| = \log\|G\| - \log\left\|\frac{1}{H}\right\| \approx \log\|G\| - \log\|A\|$$

where $A \approx 1/H$ whenever $G{\cdot}H >> 1$. At $\omega_{bw}(\text{cl})$, $\|GH\| \approx 1$, $\log(1) = 0$, and

$$\log\|G\| \approx \log\|A\|$$

The magnitude plots of G and A cross at $\omega_{bw}(\text{cl})$.

Most integrated-circuit op-amps have a high open-loop gain of over 20k and low bandwidth, often less than 100 Hz. As open-loop amplifiers, they are essentially integrators and can be modeled with the hf approximation, by removing their f_{bw} breakpoint and allowing the gain below bandwidth to increase upward to infinity, resulting in a pole at the origin. Then the open-loop gain is

$$G(s) \approx \frac{1}{\left(\frac{s}{\omega_T}\right)}$$

The bandwidth is

$$\text{hf } A_v = \frac{G}{1 + G \cdot H} \approx \frac{\omega_T / s}{1 + (\omega_T / s)/A_{v0}} = A_{v0} \cdot \frac{1}{\frac{s}{\omega_T / A_{v0}} + 1}$$

For op-amps with one pole at the origin, $f_{bw}(\text{cl}) = f_T/A_{v0}$.

For n repeated poles at the origin in G,

$$\text{hf } A_v = \frac{G}{1 + G \cdot H} \approx \frac{(\omega_T / s)^n}{1 + (\omega_T / s)^n / A_{v0}} = A_{v0} \cdot \frac{1}{\left(\frac{s}{\omega_T / \sqrt[n]{A_{v0}}}\right)^n + 1}$$

Applying the definition of bandwidth to $A_v(s)$,

$$\left(\frac{A_v}{A_{v0}}\right)^2 = \frac{1}{2} \Rightarrow \left(\frac{A_{v0}}{A_v}\right)^2 = 2 = \left(\frac{\omega_{bw}(\text{cl})}{\omega_T / \sqrt[n]{A_{v0}}}\right)^n + 1$$

or

$$f_{bw}(\text{cl}) = \frac{f_T}{\sqrt[n]{A_{v0}}}$$

For feedback amplifiers with a two-pole roll-off and a –2 slope, substitute $\sqrt{G_0/A_0}$ for G_0/A_0, and in general, for a slope of $-n$,

$$\left(\frac{f_{bw}(\text{cl})}{f_{bw}}\right)^n = \frac{G_0}{A_0} \Rightarrow \frac{f_{bw}(\text{cl})}{f_{bw}} = \left(\frac{G_0}{A_0}\right)^{1/n}$$

Quadratic Feedback-Loop Bandwidth

We can take the effect of loop gain on pole location another step by considering a loop with a quadratic pole-pair. For two real poles with OCTCs τ_1 and τ_2,

$$G(s) = \frac{G_0}{(s \cdot \tau_1 + 1) \cdot (s \cdot \tau_2 + 1)} = \frac{G_0}{s^2 \cdot (\tau_1 \cdot \tau_2) + s \cdot (\tau_1 + \tau_2) + 1}$$

and

$$\tau_n = \sqrt{a} = \sqrt{\tau_1 \cdot \tau_2}$$

$$\zeta = \frac{b}{2 \cdot \sqrt{a}} = \frac{1}{2} \cdot \frac{\tau_1 + \tau_2}{\tau_n} = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_1}{\tau_2}} + \sqrt{\frac{\tau_2}{\tau_1}}\right)$$

ζ is minimum for $\tau_1 = \tau_2$, and $\zeta = 1$.

$G(s)$ is expressed in the quadratic form of a pole-pair as

$$G(s) = \frac{G_0}{s^2 \cdot \tau_n^2 + 2 \cdot \zeta \cdot \tau_n + 1}; \; H = H_0$$

The closed-loop gain is

$$A = \frac{G}{1 + G \cdot H} = \frac{G_0}{(s^2 \cdot \tau_n^2 + 2 \cdot \zeta \cdot \tau_n + 1) + G_0 \cdot H_0}$$

This reduces to a normalized form as

$$\frac{G_0}{1 + G_0 \cdot H_0} \cdot \frac{1}{s^2 \cdot \left(\frac{\tau_n}{\sqrt{1 + G_0 \cdot H_0}} \right)^2 + s \cdot \left(\frac{\zeta}{\sqrt{1 + G_0 \cdot H_0}} \right) \cdot \left(\frac{\tau_n}{\sqrt{1 + G_0 \cdot H_0}} \right) + 1}$$

In this closed-loop response, the factor on the left is the quasistatic A_0. The closed-loop pole-pair has the parameters

$$\tau_{ncl} = \frac{\tau_n}{\sqrt{1 + G_0 \cdot H_0}}; \; \zeta_{cl} = \frac{\zeta}{\sqrt{1 + G_0 \cdot H_0}}$$

The pole-pair is not in the same location as in G. Its magnitude has increased in frequency by a factor of $\sqrt{1 + G_0 \cdot H_0}$ and ζ has decreased by the same factor, causing the pole-pair to be less damped and making the pole angle for an underdamped pole-pair greater. For real poles, as loop gain increases, the poles move toward each other on the LHP real axis until they meet and split off the axis to become complex. A pole-pair is complex whenever

$$\zeta = \frac{b}{2 \cdot \sqrt{a}} < 1 \Rightarrow \left(\frac{b}{2} \right)^2 < a$$

Feedback dynamics is an extensive subject, that of *control theory*. One of the topics in continuous (analog) control theory traces the movement in the *s*-plane of the poles of a feedback circuit as the quasistatic loop gain is increased. In a simple way, we have seen that a single-pole amplifier has a different (faster) closed-loop pole, and that as the feedback factor or loop gain is increased, the open-loop pole "migrates" to the left on the real axis of the *s*-plane, and its magnitude (which is the bandwidth) increases. The plot of migrating poles in the *s*-plane is referred to as the *root locus* of the feedback circuit. While the above closed-loop expression accounts for only two poles, root-locus analysis is not limited by the number of poles, though for over about 5 poles, tracing the loci graphically can become unwieldy, just as it is on frequency-response plots.

The reader is referred for control theory to the references. Closed-loop root-locus plots of $G{\cdot}H(s)$ are covered in more detail in

Designing Dynamic Circuit Response. Control theory methods are both powerful and essential for a mastery of transistor amplifier design. Unlike many circuits textbooks, the feedback notation used in this book follows that of control textbooks. In circuits books, G is typically replaced by A and H by β, a symbol already in frequent use as a BJT parameter. By using control notation, the reader can more easily access the control literature and its more complete and rigorous coverage of control as it applies to circuits.

Feedback Amplifier Output Impedance

High-frequency modeling can be extended to feedback amplifiers. Consider the simple but common case of an amplifier G path having a single-pole response with an open-loop voltage gain,

$$G = \frac{G_0}{s \cdot \tau_{bw} + 1}$$

where $1/\tau_{bw} = \omega_{bw}$ is the incremental open-loop bandwidth. If you prefer working in f for frequency and not ω, then in transfer-function normalized form, all occurrences of s are in frequency ratios such as s/ω_{bw}. These ratios can be replaced by s/f_{bw} where s is now in Hz.

In the lf region, the open-loop output resistance, r_o, is reduced by feedback by the feedback factor, $1 + G{\cdot}H$. The resulting closed-loop output impedance for a voltage-output amplifier is

$$Z_o(\text{cl}) = \frac{r_o}{1 + G \cdot H} = \frac{r_o}{1 + G_0 \cdot H_0} \cdot \frac{s \cdot \tau_{bw} + 1}{s \cdot (\tau_{bw} / (1 + G_0 \cdot H_0)) + 1}$$

where $H = H_0$ and $G_0{\cdot}H_0$ is the quasistatic loop gain and is constant. Then it is put in continued-fraction form to find the equivalent circuit;

$$Z_o(\text{cl}) = \frac{r_o \cdot (s \cdot \tau_{bw} + 1)}{s \cdot \tau_{bw} + (1 + G_0 \cdot H_0)} = \frac{r_o \cdot (s \cdot \tau_{bw} + 1)}{(s \cdot \tau_{bw} + 1) + G_0 \cdot H_0}$$

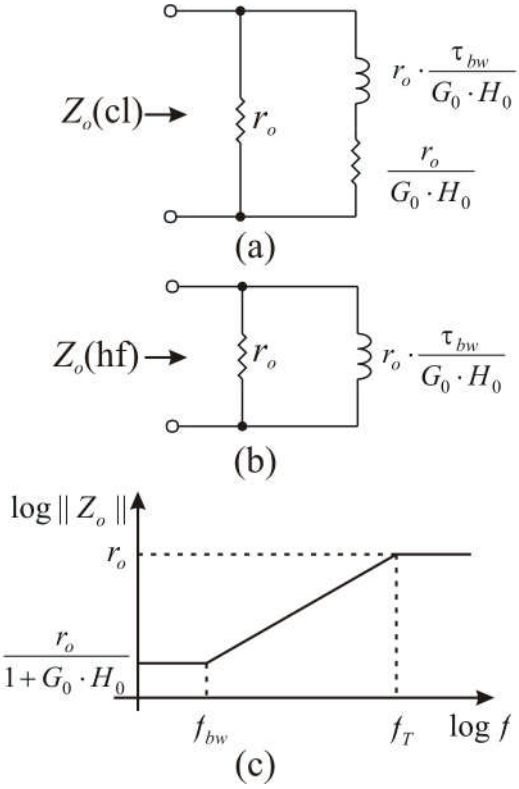
Dividing numerator and denominator by the numerator,

$$Z_o(\text{cl}) = \frac{1}{\dfrac{1}{r_o} + \dfrac{1}{s \cdot \tau_{bw} \cdot (r_o / G_0 \cdot H_0) + (r_o / G_0 \cdot H_0)}}$$

In continued-fraction form, the corresponding circuit topology of Z_o(cl) is made explicit. Rewritten more simply,

$$Z_o(\text{cl}) = r_o \,\|\, [s \cdot \tau_{bw} \cdot (r_o / G_0 \cdot H_0) + (r_o / G_0 \cdot H_0)]$$

The total-frequency model is shown below in (a), the hf model in (b), and the impedance magnitude versus frequency plot in (c).



The lf closed-loop resistance,

$$r_o/(1 + G_0 \cdot H_0)$$

gyrates +90° at the open-loop bandwidth to appear inductive out to the frequency,

$$f_{bw} \cdot (1 + G_0 \cdot H_0) = \left(\frac{f_T}{G_0 \cdot H_0} \right) \cdot (1 + G_0 \cdot H_0) = f_T / \left(\frac{G_0 \cdot H_0}{1 + G_0 \cdot H_0} \right)$$

where $G_0 \cdot H_0/(1 + G_0 \cdot H_0)$ corresponds to α_0 in f_T/α_0 of BJT $\beta(s)$. For $G_0 \cdot H_0 >> 1$, above the transverse or *unity-gain* frequency,

$$f_T = (G_0 \cdot H_0) \cdot f_{bw}$$

Z_o(cl) reverts to r_o. By analogy, f_{bw} corresponds to the BJT f_β and $G_0 \cdot H_0$ to β_o. The simplified hf equivalent output impedance is derived by letting $G_0 \cdot H_0 \to \infty$ as $f \to 0$ Hz, with resulting output impedance corresponding to $Z_e(R_B)$. When r_o is generalized to Z_o, the corresponding BJT models readily apply. Z_o(hf) is the hf equivalent circuit of Z_o(cl), and like the BJT model is only valid above f_{bw}.

The general principles of behavior of transistors and amplifiers (and any single-pole active devices) in their hf regions apply to other devices including JFETs and MOSFETs. The reactance values of gyrating $s \cdot \tau_T$ and $1/s \cdot \tau_T$ remain as hidden effects not found on the circuit diagram. Be cognizant of the hf range, and especially the low end, f_β. With some practice, you can know when to be alert to hf effects. For instance, if R_E is large, capacitance τ_T/R_E will be miniscule and can be ignored. As R_E is reduced (as happens when amplifiers are made faster), it begins to become significant. Similarly, a small R_B viewed from the emitter makes for a miniscule inductance, $\tau_T \cdot R_B$, that resonates with parasitic capacitance at a frequency well above f_T. When R_B becomes larger, it becomes significant when its resonances are within the hf region.

For feedback amplifiers with a single-pole loop gain, as open-loop output resistance increases, so does the gyrated inductance. This is another reason for why it is good to design the open-loop feedback amplifier as close to the ideal (low r_o) as is feasible before closing the loop. The hf inductance of a feedback amplifier will resonate with load capacitance and can cause oscillation. It is for the same basic reason that a capacitance-loaded emitter will cause a BJT to oscillate.

Both C_E and L_B can produce hf resonances. C_E gyrates when referred to the base as a series $-\tau_T/C_E$, $-\alpha_0 \cdot C_E$ that can be compensated with a series $+\tau_T/C_E$, $+\alpha_0 \cdot C_E$ at the base node. The dual for L_B is a shunt RL, $-L_B/\tau_T$, $-\alpha_0 \cdot L_B$, which can be compensated in the emitter by a shunt RL: $+L_B/\tau_T \,||\, +\alpha_0 \cdot L_B$. The same applies to voltage-output feedback amplifiers.

To avoid encountering a negative-element gyration in hf analysis, emitter (or output load) capacitance can often be analyzed more easily from the emitter if there is no base (or open-loop Z_o) inductance by referring Z_B to the emitter. Base resistance refers as gyrated inductance from the base and combines with emitter capacitance, both having positive values. Similarly, L_B is often more easily analyzed at the base by referring emitter resistance as capacitance to the base. The hf analyses at either base or emitter produce equivalent results though the avoidance of negative elements simplifies hf algebra.

Feedback Amplifier Port Impedances

Feedback amplifier output impedance is a special case of a more general principle that can be applied to both feedback amplifier input and output ports. In general, a feedback amplifier with voltage output has an output port impedance of

$$Z_o(\text{cl}) = \frac{Z_o}{1+G\cdot H}\text{, } v_o$$

A feedback amplifier with current input has an input port impedance that is of the same form as a feedback amplifier with voltage output;

$$Z_i(\text{cl}) = \frac{Z_i}{1+G\cdot H}\text{, } i_i$$

The other two cases, of input voltage and output current, have the same form:

$$Z(\text{cl}) = (1+G_0\cdot H_0)\cdot Z\text{, } v_i, i_o$$

Consequently, generalized equivalent port impedances can be derived. For feedback amplifiers with single-pole loop gains, the feedback factor is

$$1+G\cdot H = 1+\frac{G_0\cdot H_0}{s\cdot\tau_{bw}+1} = (1+G_0\cdot H_0)\cdot\frac{s\cdot\left(\dfrac{\tau_{bw}}{1+G_0\cdot H_0}\right)+1}{s\cdot\tau_{bw}+1}$$

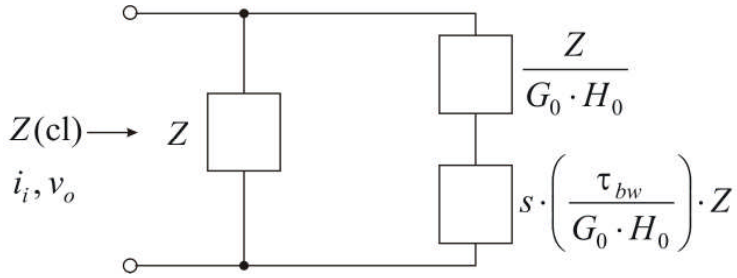
Then the general feedback port impedance for i_i and v_o is

$$Z(\text{cl}) = \frac{Z}{1+G\cdot H} = \frac{Z}{1+G_0\cdot H_0}\cdot\frac{s\cdot\tau_{bw}+1}{s\cdot\left(\dfrac{\tau_{bw}}{1+G_0\cdot H_0}\right)+1}$$

This can be rewritten by distributing the feedback factor in the denominator as

$$Z(\text{cl}) = \frac{Z\cdot(s\cdot\tau_{bw}+1)}{s\cdot\tau_{bw}+1+G_0\cdot H_0} = \frac{1}{\dfrac{1}{Z}+\dfrac{1}{s\cdot\left(\dfrac{\tau_{bw}}{G_0\cdot H_0}\right)\cdot Z+\dfrac{Z}{G_0\cdot H_0}}}$$

The equivalent circuit is easily drawn from the continued-fraction form of Z(cl), as shown.



A transimpedance amplifier ($Z_m = v_o/i_i$) has input and output impedance of the same form. For $Z = R$, it is R shunting a series RL. The combined quasistatic resistance is

$$R \| \frac{R}{G_0 \cdot H_0} = \frac{R}{1 + G_0 \cdot H_0}$$

The series inductance forms a time constant in the series RL branch with $R/(G_0 \cdot H_0)$ of τ_{bw}.

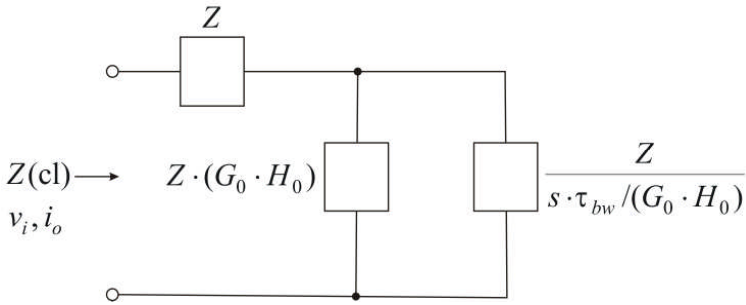
The dual case of v_i and i_o has a closed-loop port impedance in the general form of

$$Z(\text{cl}) = Z \cdot (1 + G \cdot H) = Z \cdot (1 + G_0 \cdot H_0) \cdot \frac{s \cdot \left(\frac{\tau_{bw}}{1 + G_0 \cdot H_0} \right) + 1}{s \cdot \tau_{bw} + 1}$$

Some "dual algebra" results in a topologically-explicit form;

$$Z(\text{cl}) = Z + \frac{Z \cdot G_0 \cdot H_0}{s \cdot \tau_{bw} + 1} = Z + \frac{1}{\frac{1}{\left(\frac{Z}{s \cdot \tau_{bw} / (G_0 \cdot H_0)} \right)} + \frac{1}{Z \cdot (G_0 \cdot H_0)}}$$

The equivalent circuit is shown below and is analogous to the impedance of a BJT base port with emitter impedance Z. For $Z = R$, the quasistatic port impedance has two resistances in series having a combined resistance of $R \cdot (1 + G_0 \cdot H_0)$. The gyrated R is a capacitive reactance with capacitance $\tau_{bw}/(R \cdot G_0 \cdot H_0)$. It forms a parallel RC with $R \cdot G_0 \cdot H_0$ having a time constant of τ_{bw}.



The analogy with BJT hf behavior allows us to refer for the three elemental cases of Z to the previous BJT derivations in chapter 5 for total-variable and hf equivalent circuits. By knowing the BJT hf impedance gyrations, we know the feedback port impedance gyrations above loop bandwidth.

The close analogy between single-pole feedback loops and BJTs leads to the realization that BJT stages can be interpreted as feedback amplifiers with the following analogs:

$$G_0 \cdot H_0 \Leftrightarrow \beta_0 \; ; \; A_0 \cdot H_0 = \frac{G_0}{1 + G_0 \cdot H_0} \cdot H_0 \Leftrightarrow \alpha_0$$

$$\tau_{bw} \Leftrightarrow \tau_\beta = \beta_0 \cdot \tau_T$$

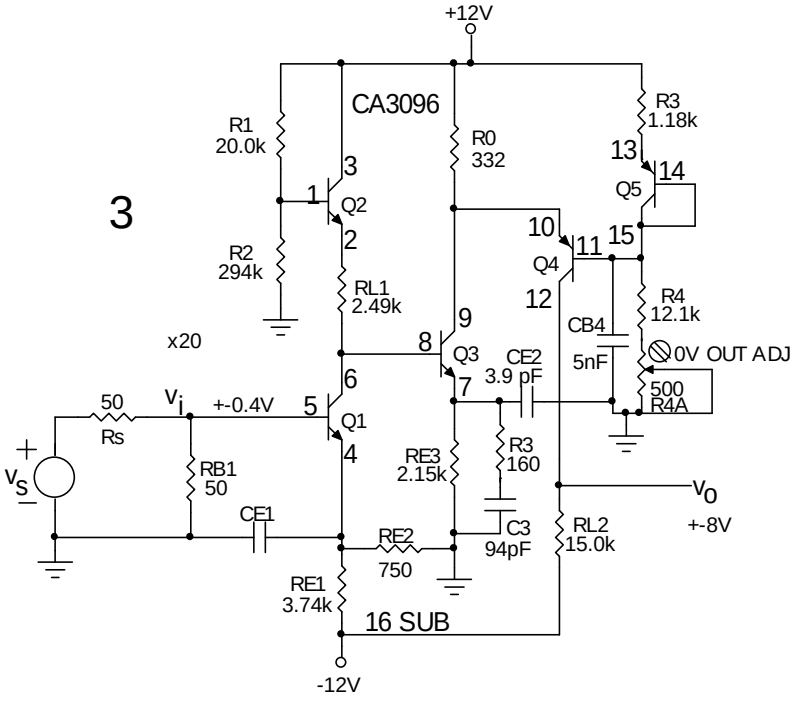
$$\frac{\tau_{bw}}{G_0 \cdot H_0} \Leftrightarrow \tau_T \; ; \; \tau_{bw}(\text{cl}) = \frac{\tau_{bw}}{1 + G_0 \cdot H_0} \Leftrightarrow \alpha_0 \cdot \tau_T$$

The series R_E of the general BJT stage can be viewed as providing voltage or "series" (in a loop) feedback to the input. Then the quasistatic base port r_B is $1 + G_0 \cdot H_0$ or $1 + \beta_0$ times larger than R_E and has a capacitive component corresponding to the rightmost element in the above general equivalent circuit, or a capacitance of $(\tau_{bw}/\beta_0)/R_E = \tau_T/R_E$. This capacitance is in series with $Z = R_E$, as we found for the BJT in the hf region.

Cascode Amplifiers

Cascode Amplifier

This detailed design exercise of amplifier 3, based on a CA3096 BJT array, applies shunt inductive peaking to the input stage followed by a complementary cascode stage with emitter peaking. The input-stage inductor is synthesized as a BJT hf-gyrated base resistance. The component values are given in the circuit diagram as they are derived in the design exercise.



The Q1 CE stage is shunt-peaked by Q2 in the collector circuit. The following CE stage, Q3, has emitter peaking to compensate for the output pole of Q4 and $\alpha_3(s)$. The *complementary cascode* second stage BJTs are NPN CE (Q3) and PNP CB (Q4). Q5 matches Q4 for stable biasing of the Q4 emitter current by maintaining constant voltage across R_0.

Amplifier Design Strategy

The set of defining statements that is the goal for a design is its *specification.* The amplifier 3 design is specified to have 0 V output for 0 V input, a voltage gain of 20 with 8-bit linearity (< 0.4 %) and an output range of ±8 V at a bandwidth of at least 3.5 MHz (sufficient for NTSC monochrome video) when loaded by an oscilloscope ×10 probe (15 pF, 10 MΩ), and with a flat (constant magnitude) frequency response.

We are forced at the outset of the design to consider dynamics in choosing quasistatic stage gains. To achieve the given gain while maximizing bandwidth, apportion the gains of the two stages so that they are nearly equal, with somewhat more for the cascode stage. The first stage has a chosen gain of –4 and the cascode has –5. The maximum input range for the given output range is thus ±8 V/20 or ±0.4 V. The first-stage output range is ±8 V/5 = ±1.6 V.

To maintain bandwidth while minimizing power, the NPN BJTs are given a collector current of 2 mA and the PNPs 0.8 mA. The CA3096 NPNs are biased somewhat less than but quite near the broad peak of the maximum $f_T(I_C)$ curve at 4 mA so that they have close to the maximum f_T of 335 MHz (τ_T = 475 ps). The CA3096 PNPs are not like discrete PNP BJTs in that their β_0 value of 47 at 1 mA decreases quickly above it to a low value of 15 at 4 mA. The PNP $f_T(I_C)$ function also does not increase to a maximum at some current, then decrease, but has a maximum near 0 mA and decreases as current increases. The CA3096 dynamic data needed for this design is

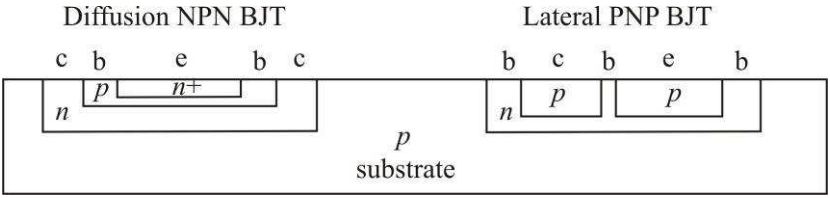
NPN f_T = 335 MHz (τ_T = 475 ps), 2 to 6 mA, 5 V; typical β_0 = 390
280 MHz (568 ps), 1 mA, 5 V
PNP f_T = 6.8 MHz (τ_T = 23.4 ns), 0.1 mA, 5 V;
6.5 MHz (24.5 ns), 0.8 mA, 5 V; typical β_0 = 50
6.5 MHz (24.5 ns), 1 mA, 5 V; typical β_0 = 47

NPN C_c = 0.5 pF, 3 V to 10 V; C_{cs} = 2 pF, 5 V
PNP C_c = 2 pF, 5 V; C_{bs} = 2.5 pF, 5 V

The currents for f_T are $|I_C|$ and voltages for C_c are $|V_{CB}|$. The diced-wafer substrate that is the foundation of integrated circuits introduces additional parasitic capacitances. C_{cs} is NPN collector-to-substrate capacitance; C_{bs} is PNP base-to-substrate capacitance.

Monolithic BJT Structure

Why *c-s* capacitance is given for NPN and *b-s* for PNP is evident from a vertical cross-section of a silicon wafer, as shown.



NPN BJTs are made by thermally diffusing or ion-implanting n and p material in stacked layers into the p-type substrate. The NPN base width - the vertical distance between collector and emitter - is thin and the emitter n concentration is higher (n+) than the collector. This results in more emitter electrons being injected into the base than base holes being recombined in the emitter for higher *emitter injection efficiency*. With fewer base holes lost, base current is less and β is increased. The consequence of high n concentration in the emitter is that the *b-e* diode reverse breakdown voltage is low, typically 5 to 7 volts. The higher concentration also reduces V_{BE} somewhat for the same current.

The PNP is made as a *lateral* (sideways) BJT, with an *n well* in the substrate as base. Two p regions that are close together are diffused into the base well to form the active base region between them. The lateral (horizontal) distance between the emitter and collector sidewalls is the base width. NPN diffusion or implantation depth is more accurate to control than the placement of PNP sidewalls. Consequently, PNP β is lower than NPN β because it is more difficult to produce a thin base in a lateral BJT. The base is thus made wider, and it takes more time for minority carriers from the emitter to cross to the collector, causing f_T to be lower. More time for minority-carrier recombination with base majority carriers results in more base current and lowers β. Adjacent to the substrate in a NPN is the collector and the base in a PNP. Each forms a junction with the substrate which is kept reverse-biased in operation by connecting it to the negative supply. Each has parasitic capacitance.

Unlike discrete BJTs, the lateral PNP f_T decreases with increasing current. Ordinarily, f_T increases with I_C. This is evident from the BJT model in that τ_β decreases as r_e (and hence r_π) decreases with I_C. C_e

also increases, though not as much, because of greater diffusion charge on each side of the *b-e* junction at higher current.

A forward-biased p-n junction has two components of its capacitance: diffusion and junction. *Diffusion capacitance* results from the excess charge at both junction edges caused by conduction. It increases with increasing current. *Junction capacitance* results from the p-n junction as a capacitor dielectric. It varies in width with reverse junction voltage and remains nearly constant in forward bias. Thus junction capacitance increases with decreasing junction voltage. For the forward-biased *b-e* junction, f_T increases until the current is high enough to cause significant voltage drop in the series collector resistance and this causes the effective base width to widen into the collector. A wider base takes charge carriers longer to transit from emitter to collector, and f_T decreases.

Source Characteristics

The input circuit of the amplifier shows a 50 Ω base resistance in series with the voltage input source. This represents the unterminated output of the source generator of v_i. To terminate the input from a coax cable to avoid time-domain reflections or frequency-domain reactance variation with frequency, the 50 Ω (51 Ω, 5 %) base resistor terminates the cable of Z_n = 50 Ω. Then the equivalent source R_B from the Q1 base is the 51 Ω terminating R_{B1} in parallel with the 50 Ω cable, resulting in 25 Ω at the base port of Q1. The Thevenin equivalent input circuit has $v_i = v_g/2$ "behind" - that is, in series with - 25 Ω, where v_g is the generator open-circuit output voltage.

Function, pulse, and sine generator instruments are waveform sources, usually designed to have an output circuit that is a voltage source in series with 50 Ω. (Audio and video equipment are exceptions, using 600 Ω for audio and 75 Ω for video.) At the generator output connector, when open-circuited (no load), the voltage is twice what it is when the output is terminated in the cable impedance. This kind of source output is convenient for sending high-speed waveforms to a widely separated load via a terminated cable without appreciable degradation in speed or waveshape. For low-frequency outputs (cable length $<< \lambda$), cable propagation time is negligible and the cable can be left unterminated and used as a shielded-wire pair. This has the advantage of twice the output voltage from the source.

Subsystem interconnections sometimes use conductors in a planar or "ribbon" cable with alternating grounded conductors. $Z_n \approx 75\ \Omega$ to $100\ \Omega$ and can be used for short (< 50 cm) connections in circuits with bandwidths not exceeding 100 MHz. Test and measurement (T&M) instruments usually use coaxial cables for interconnection.

First-Stage Static & Quasistatic Design

We now attend to the static and quasistatic design by finding some required design values. For CA3096 NPN BJTs at 0.2 mA, V_{BE} = 0.65 V and for PNP BJTs V_{BE} = 0.60 V at 0.1 mA. For NPNs operating at a collector current of 2 mA,

NPN V_{BE} = 0.65 V + (60 mV/dec)·log(2 mA/0.2 mA) = 0.710 V

We will also need α_{NPN} = 0.9974 and α_{PNP} = 0.9804 corresponding to β_{NPN} = 390 and β_{PNP} = 50. The NPN BJT r_e = 26 mV/2 mA = 13 Ω and PNP r_e = 32.5 Ω. The specified manufacturer tolerances on the above BJT parameters are wider than the number of significant figures. They are retained to minimize arithmetic error and to maintain consistency in calculations for the fixed nominal values.

Some CA3096 parts were measured for β_0. The part data for β_{PNP} agreed well but β_{NPN} was usually much lower than the typical specification - from 175 to 250 - though some BJTs measured close to 390. Few exceeded it.

Input Linearity

The base voltage of Q2 is set by voltage divider R_1, R_2 which indirectly affects the stage linearity and thermal distortion. To minimize thermal distortion, the Q1 operating-point is set at maximum Q1 power. This occurs when

$$I_{C1} = \frac{V_{C1}}{R_{L1}} = \frac{V_{CC1}/2}{R_{L1}} = \alpha_0 \cdot \frac{V_{RE}}{R_E}$$

where $V_{CC1} = V_{E2}$ is effectively the supply voltage for Q1, R_E is the Thevenin resistance of $R_{E1}||R_{E2}$, and V_{RE} is the voltage across R_E;

$$V_{RE} = I_{E1} \cdot R_E \ ; \ R_E = R_{E1}||R_{E2}$$

The thevenized emitter supply voltage is

$$-V_{EE1} = \frac{R_E}{R_{E1}} \cdot (-12\ \text{V})$$

The voltage gain is approximately

$$-A_{v1} \approx \alpha_0 \cdot \frac{R_{L1}}{R_E} = \frac{V_{CC1}/2}{V_{RE}}$$

The rightmost expression results from substitution from the I_{C1} equation above. Thus we see that quasistatic gain and minimization of thermals are related through this equation and that consequently

$$R_{L1} = \frac{V_{CC1}/2}{I_{C1}}$$

With a given A_v and with R_{L1} determined, R_E is largely determined. However, we have not yet chosen a value of V_{CC1}. It will be determined instead by R_{L1} which will be determined in turn by A_v, and it in turn by R_E which is determined by V_{RE} which, as we will see, affects linearity. The causal design flow is

$$\text{linearity} \rightarrow V_{RE} \rightarrow R_E \rightarrow R_{L1} \rightarrow V_{CC1}$$

Linearity is specified, given as 8 bits or $1/256 \approx 0.4\,\%$. In the chapter "Matched-Transistor Circuits", the section "BJT Input-Loop Linearity" derived the basic design formulas we need here. The same direction will be taken in applying those concepts to this circuit.

The input loop of the BJT stage consists of the input voltage source in series with the *b-e* junction, R_E and $-V_{EE1}$, the Thevenin emitter supply voltage. Let the static, no-input, or *quiescent* ($v_i = 0$ V) $i_E = I_0$. A change in input voltage, Δv_I, is the sum of the voltage changes across the other series elements in the loop;

$$\Delta v_I = R_E \cdot \Delta i_E + \Delta v_{BE} = \Delta v_{RE} + \Delta v_{BE}$$

$\Delta v_I \neq v_i$ because it is large enough to include the nonlinear effects and is not incremental. From the diode *v-i* equation and $i_E \approx i_C$,

$$\Delta v_{BE} = V_T \cdot \ln\left(\frac{i_E}{I_S}\right) - V_T \cdot \ln\left(\frac{I_0}{I_S}\right) = V_T \cdot \ln\left(\frac{i_E}{I_0}\right) = V_T \cdot \ln\left(1 + \frac{\Delta i_E}{I_0}\right)$$

The linear voltage change with Δi_E occurs across R_E and the fixed operating-point $r_e(I_0) = V_T/I_0$. The transresistance when $\Delta i_E = 0$ mA is

$$r_{M0} = R_E + r_{e0} = R_E + \frac{V_T}{I_0}$$

The total linear voltage change from Δv_I is $r_{M0} \cdot \Delta i_E$. The fraction of nonlinear voltage change around the input loop to the linear change is

$$\frac{\Delta v_I - r_{M0} \cdot \Delta i_E}{r_{M0} \cdot \Delta i_E} = \frac{R_E \cdot \Delta i_E + V_T \cdot \ln\left(1 + \frac{\Delta i_E}{I_0}\right)}{\left(R_E + \frac{V_T}{I_0}\right) \cdot \Delta i_E} - 1 = \left(\frac{V_T}{R_E \cdot I_0 + V_T}\right) \cdot \left(\frac{\ln\left(1 + \frac{\Delta i_E}{I_0}\right)}{\left(\frac{\Delta i_E}{I_0}\right)} - 1\right)$$

This is the fractional voltage error caused by *b-e* nonlinearity.

The dynamic range has a significant effect on linearity for a given gain. In this design, the full-scale input dynamic range is

$$\pm\Delta v_I = \pm\Delta v_O / A_v = \pm 8 \text{ V} / 20 = \pm 0.4 \text{ V}$$

Linearity requires a certain minimum V_{RE}. With I_E chosen, R_E has a minimum. V_{RE} is limited by the static condition that it must exceed the maximum voltage change across it - the input voltage range of ±0.4 V - to avoid *clipping*, nonlinearity with waveform slope discontinuity. V_{RE} is limited at the static operating-point to

$$V_{RE} = V_{EE1} - V_{BE1} = V_{EE1} - 0.710 \text{ V}$$

To maximize linearity, make V_{CC1} large but without saturating Q2. If we choose $V_{CC1} \approx 10.4$ V, C_{c2} is relatively constant with v_{CB}. Apply thermal-distortion minimization:

$$V_{RE} \approx \frac{V_{CC1} / 2}{|A_{v1}|} = \frac{10.4 \text{ V} / 2}{4} = 1.3 \text{ V}$$

Choose $V_{EE1} = 2.0$ V so that $V_{RE} = 2.0$ V – 0.71 V = 1.29 V, and solve the circuit, including for linearity. Given $\Delta v_{RE} \approx \Delta v_I$, then

$$\Delta i_E / I_0 = \Delta v_{RE} / V_{RE} \approx \Delta v_I / V_{RE} = 0.4 \text{ V} / 1.29 \text{ V} = 0.310$$

Applying the error formula given above,

$$|\text{stage-1 input-loop voltage error}| = \frac{1}{7.754} \cdot \frac{1}{50.62} = \frac{1}{392.5} < \frac{1}{256}$$

The choice of $V_{EE1} = 2.0$ V yields $\log_2(392.5) = 8.6$ bits of linearity.

Using the $|A_v|$ thermal formula,

$$V_{CC1} = 2 \cdot |-4| \cdot (1.29 \text{ V}) = 10.32 \text{ V}$$

$$R_{L1} = \frac{(10.32\ \text{V})/2}{2.07\ \text{mA}} = 2490\ \Omega \rightarrow 2490\,\Omega, 1\,\%$$

The total collector resistance depends on R_{B2} which depends on the shunt-peaking design. The static design cannot be completed without the dynamic design. Analog design is not a linear list of steps but can have iterative loops, like an artist refining a rough sketch. Looking ahead in the dynamic design section, we find that the required value of $R_{B2} = 18.73\ \text{k}\Omega$. The collector resistance across which i_{c1} develops v_{c1} is thus

$$R_{L1}' = R_{L1} + r_{e2} + \frac{R_{B2} + r_b'}{\beta_{NPN} + 1} = 2490\ \Omega + 13\ \Omega + \frac{18.73\ \text{k}\Omega + 50\ \Omega}{391} = 2551\Omega$$

The third term is the Q2 base resistance referred to the emitter, which has a value of 48 Ω. The Q1-stage transresistance is

$$r_{M1} = r_{e1} + \frac{R_{B1}'}{\beta_{NPN} + 1} + R_E = 13\ \Omega + \frac{75\ \Omega}{391} + R_E = 13.19\Omega + R_E$$

Solving for R_E,

$$R_E = \alpha_{NPN} \cdot \frac{R_{L1}'}{|A_{v1}|} - \left(r_{e1} + \frac{R_{B1}'}{\beta_{NPN} + 1} \right) = (0.9974) \cdot \frac{(2551\Omega)}{4} - 13.19\ \Omega = 623\,\Omega$$

Solving for R_{E1}, R_{E2} given Thevenin voltage V_{EE1} and resistance R_E,

$$\frac{R_E}{R_{E1}} = \frac{V_{EE1}}{V_{EE}} = \frac{2\ \text{V}}{12\ \text{V}} = \frac{1}{6} \Rightarrow R_{E1} = 6 \cdot R_E = 3736\ \Omega \rightarrow 3.74\ \text{k}\Omega,\ 1\ \%$$

$$R_{E2} = 747\ \Omega \rightarrow 750\ \Omega,\ 1\ \%$$

Now that the actual values are determined, the more accurate values of the circuit parameters are found from them;

$$R_E = 624.7\ \Omega\ ;\ V_{EE1} = 2.004\ \text{V}\ ;\ I_{E1} = \frac{2.004\ \text{V} - 0.710\ \text{V}}{624.7\ \Omega} = 2.071\ \text{mA}$$

The static voltage across R_{L1} is

$$V_{RL} = (I_{C1} + I_{B3}) \cdot R_{L1} \approx I_{E1} \cdot R_{L1} = (2.071\ \text{mA}) \cdot (2.49\ \text{k}\Omega) = 5.157\ \text{V}$$

The Q1 static collector voltage is thus

$$V_{C1} = V_{B3} = V_{CC1}/2 = 5.16\ \text{V}$$

$$V_{CC1} = V_{E2} = V_{C1} + V_{RL} = 5.16\ \text{V} + 5.157\ \text{V} = 10.32\ \text{V}$$

$$V_{B2} = V_{CC1} + 0.710\ \text{V} = 11.14\ \text{V}$$

The Q2 base divider resistance calculations include the loading effect of I_{B2} on V_{B2} which is reduced by

$$[I_{E2}/(\beta_0 + 1)]\cdot R_{B2}' = (5.12\ \mu\text{A})\cdot(18.78\ \text{k}\Omega) = 96\ \text{mV}$$

from the open-circuit V_{B2};

$$V_{B2}(\text{oc}) = 11.14\ \text{V} + 96\ \text{mV} = 11.24\ \text{V}$$

The Q2 divider resistors can now be calculated after subtracting r_b':

$$\frac{R_2}{R_1 + R_2} = \frac{11.24\ \text{V}}{12\ \text{V}} = 0.9363$$

$$R_1 = \frac{R_{B2}}{0.9363} = 19.96\ \text{k}\Omega \ \Rightarrow 20.0\ \text{k}\Omega,\ 1\ \%$$

$$R_2 = \left(\frac{0.9363}{1 - 0.9363}\right)\cdot R_1 = 14.70\cdot R_1 = 293.4\ \text{k}\Omega \Rightarrow 294\ \text{k}\Omega,\ 1\ \%$$

The recalculated $R_{B2} = 18.73\ \text{k}\Omega$ and with $r_b' = 50\ \Omega$ included, the total $R_{B2}' = 18.78\ \text{k}\Omega$. The open-circuit $V_{B2} = 11.24$ V. Then

$$V_{CC1} = V_{E2} = V_{B2} - 96\ \text{mV} - 0.710\ \text{V} = 10.43\ \text{V}$$

The transresistance is

$$r_{M1} = 13.19\ \Omega + 623\ \Omega = 636.2\ \Omega$$

The voltage gain of the first stage (disconnected from Q2) is

$$A_{v1} = -(0.997)\cdot\frac{2551\,\Omega}{636.2\,\Omega} = -4.00$$

The value is that specified (–4) and the stage-1 quasistatic design is verified. However, there is a quasistatic BJT parameter that spoils it somewhat: r_o.

In applying the formula for r_o,

$$r_o \equiv \frac{|V_{CE}| + V_A}{|I_C|}$$

the CA3096 NPN $V_A \approx 75$ V and PNP $V_A \approx 65$ V for collector currents between 20 μA and 10 mA. At 2 mA for Q1, and $V_{CE1} \approx 5$ V,

$$r_{o1} = \frac{75\text{ V} + 5\text{ V}}{2\text{ mA}} = 40\text{ k}\Omega$$

From "BJT Model Parameter r_o" (page 80), for large r_o, CE voltage gain,

$$\text{CE } A_v \approx A_v\big|_{r_o \to \infty} + F \;\;,\; F = \frac{R_L}{R_L + r_o}$$

where the first term, which is negative for the CE stage, is the voltage gain without r_o. F reduces gain magnitude.

When the approximate circuit gain with r_o is evaluated, the first-stage gain is –3.99. The more approximated A_v is –3.94. The effect of r_o is significant when the collector node resistance is high.

Second-Stage Static & Quasistatic Design

Progressing to the second stage and Q3, the static emitter voltage calculates to be

$$V_{E3} = V_{C1} - V_{BE3} = 5.16\text{ V} - 0.710\text{ V} = 4.45\text{ V}$$

For I_{E3} = 2.07 mA (for which I_{C3} = 2.06 mA), then

$$R_{E3} = \frac{4.45\text{ V}}{2.07\text{ mA}} = 2225\,\Omega \to 2.15\text{ k}\Omega, 1\%$$

The second-stage transresistance (of Q3) is

$$r_{M3} = 13\,\Omega + 2.15\text{ k}\Omega + \frac{R_{L1}'}{\beta_{NPN} + 1} = 2163\,\Omega + \frac{2551\,\Omega}{391} = 2170\,\Omega$$

The cascode stage can be decomposed into a two-stage amplifier (stages 2a and 2b) where stage 2a is the CE with the Q4 emitter disconnected. Then the CE load resistor is R_0 and ignoring r_o for now, the quasistatic voltage gains are

$$A_{v2a} = -\alpha_{NPN} \cdot \frac{R_0}{r_{M3}} \;;\; A_{v2b} = +\alpha_{PNP} \cdot \frac{R_{L2}}{r_{M4}}$$

where

$$r_{M4} = R_0 + r_{e4} + R_{B4}'/(\beta_{PNP} + 1) \;= 332\,\Omega + 54\,\Omega = 386\,\Omega$$

The interstage loading effect appears in A_{v2b}. The Q3 collector current source and R_0 form a Norton circuit which is thevenized. Then the Q4 emitter is driven by this Thevenin source of R_0 in series with the open-circuit (unloaded) voltage of $A_{v2a} \cdot v_{C1}$(oc). When the gains are combined, the loading effect becomes a current divider;

$$A_{v2} = A_{v2a} \cdot A_{v2b} = -(\alpha_{NPN} \cdot \alpha_{PNP}) \cdot \left(\frac{R_0}{r_{M4}} \right) \cdot \frac{R_{L2}}{r_{M3}}$$

The gain is expressed as the α loss factor followed by an interstage current-divider attenuation factor,

$$T_0 = \frac{R_0}{r_{M4}} = \frac{R_0}{R_0 + r_{e4} + R_{B4}'/(\beta_{PNP} + 1)}$$

followed by the transresistance gain of an ideal cascode amplifier. T_0 is the fraction of i_{c3} that is output by the divider as i_{e4} emitter current. The Q4 emitter resistance,

$$r_{E4} = r_{e4} + \frac{R_{B4}'}{\beta_{PNP} + 1}$$

is in parallel with R_0 and some current is lost to R_0. The output fraction is T_0.

The static output voltage is specified to be 0 V. This constrains R_{L2} to be V_{EE}/I_{C4} unless an additional resistor to ground is added and the supply thevenized. From the $\beta(I_C)$ and $f_T(I_C)$ data for CA3096 PNP BJTs, the operating-point is chosen to be around

PNP I_C = 0.8 mA $\Rightarrow$ V_{BE} = 0.654 V, β_{PNP} = 50, r_e = 32 Ω, r_o = 94 kΩ

with r_b' = 50 Ω. The value of I_{E4} is not constrained but 1 % resistor values are. Thus, choose R_{L2} = 15.0 kΩ, 1 %, and

$$I_{C4} = 12 \text{ V}/15 \text{ k}\Omega = 0.8 \text{ mA} \Rightarrow I_{E4} = 0.816 \text{ mA}$$

At the PNP collector current of 0.8 mA for which the typical β_0 = 50 and emitter current is 0.816 mA,

PNP V_{BE} = 0.6 V + (60 mV/dec)·log(0.8 mA/0.1 mA) = 0.654 V

From I_E the parameters β_0, r_e, and r_o are all found. Define the current through R_0 as

$$I_0 = I_{C3} + I_{E4} = 2.06 \text{ mA} + 0.816 \text{ mA} = 2.876 \text{ mA}$$

In the Q4 base bias divider, Q5 compensates statically for V_{BE4} and requires for V_{BE} matching that $I_{E5} = I_{E4} = 0.816$ mA.

We can reduce R_{E3} to adjust A_{v2} to –5 because $T_0 < 1$, but that will disturb the completed static design of the Q3 input loop. One other place in the cascode where gain can be adjusted is in the interstage divider, T_0, itself. R_0 can be adjusted in value by changing the voltage $V_0 = I_0 \cdot R_0$ across it by the Q4 base divider. We can find R_0 by first solving for

$$T_0 = \frac{|A_{v2}| \cdot \left(\dfrac{r_{M3}}{\alpha_{NPN} \cdot \alpha_{PNP}} \right)}{R_{L2}} = \frac{(5) \cdot (2219\Omega)}{15.0\ \text{k}\Omega} = 0.7397$$

From the current-divider formula,

$$T_0 = \frac{R_0}{R_0 + r_{E4}} \Rightarrow R_0 = \frac{T_0}{1 - T_0} \cdot r_{E4} = \frac{T_0}{1 - T_0} \cdot \left(r_{e4} + \frac{R_{B4}'}{\beta_{PNP} + 1} \right)$$

r_{E4} depends on Q4 base resistance R_{B4}' which depends on R_0. The solution for R_0 is iterated by starting with $r_{E4} = 54.66\ \Omega$. Then it converges to $r_{E4} = 53.95\ \Omega$; $R_0 = 332\ \Omega$, 1 %, $V_0 = 0.955$ V, $R_3 = V_0/I_{E4} = 1170\ \Omega \rightarrow 1.18$ kΩ, 1 %; then

$$r_{E4} = 54.05\ \Omega \rightarrow 54\ \ \Omega$$

The Q4 Thevenin equivalent base resistance, including r_b' = 50 Ω, is

$$R_{B4}' = 1128\ \Omega$$

$$V_{B4} = 12\ \text{V} - 0.955\ \text{V} - 0.654\ \text{V} = 10.39\ \text{V}$$

The Q4 base bias resistor to ground has the value,

$$R_4 = \frac{V_{B4}}{I_{B4} + I_{E5}} = \frac{V_{B4}}{(1 + 1/\beta_{PNP}) \cdot I_{E4}} = 12.48\ \text{k}\Omega$$

This bias voltage affects I_{C4} but the gain very little. It can be used to null the output offset voltage. A 500 Ω trim-pot in series with R_4 can be used for this adjustment. The design center for potentiometers is their center value, in this case 250 Ω. Thus the R_4 value is about 250 Ω less or $R_4 = 12.1$ kΩ, 1 %. Then the total is $R_4' = 12.48$ kΩ.

Checking the second-stage gain (with $r_o \rightarrow \infty$),

$$A_{v2} = -(\alpha_{NPN} \cdot \alpha_{PNP}) \cdot \left(\frac{R_0}{R_0 + r_{e4} + R_{B4}/(\beta_{PNP}+1)} \right) \cdot \frac{R_{L2}}{r_{M3}}$$

$$= -(0.9974) \cdot (0.9804) \cdot (0.7397) \cdot \frac{15.0\,\text{k}\Omega}{2170\,\Omega} = -5.000$$

Thermal Design

A consideration that is usually made a part of static design (because it happens so slowly) is the *thermal drift.* As temperature varies, so does v_{BE} (which at the time-scale of seconds is approximately V_{BE}). This is not the same as thermal distortion, which varies with the amplified waveform, but can be thought of as the static component of thermal effects.

The silicon (Si) p-n junctions of CA3096 BJTs are specified as having a temperature coefficient of –1.8 mV/K (= –1.8 mV/°C). This is within the typical Si p-n junction range of –1.8 to –2.2 mV/K. For a given ΔT, the resulting ΔV_{BE} can be regarded as an input waveform added to the static V_{BE} with its polarity inverted. A positive ΔT causes v_{BE} to be reduced, causing more emitter current and a lower collector voltage. The effect is as though +1.8 mV/K were amplified. For Q1, that would be (–4)·(1.8 mV/K) = –7.2 mV/K. The Q2 v_{BE} is in series with the collector and raises the collector voltage, causing the net thermal effect of Q1 and Q2 to be (–3)·(1.8 mV/K) = –5.4 mV/K. This is in series with the Q3 v_{BE} which reduces the combined effect by another *b*-*e* junction drift to –3.6 mV/K. Then this is amplified by –5 for an output voltage drift of +18 mV/K.

The actual voltage change at the output depends, of course, on ΔT. The thermal Ohm's Law analogy (where $T \Leftrightarrow v$ and $p_D \Leftrightarrow i$) is

$$\Delta T = R_{JA} \cdot \Delta p_D$$

where p_D is the BJT power dissipation and R_{JA} is the thermal resistance from the BJT junction to the ambient (surrounding air) heat reservoir. The CA3096 is packaged in a 16-pin dual-in-line plastic (DIP) or DIP16 package having $R_{JA} \approx 150$ K/W. Then

$$\frac{dv_{BE}}{dp_D} = \frac{dv_{BE}}{dT} \cdot \frac{dT}{dp_D} = (-1.8\,\text{mV/K}) \cdot R_{JA} = -0.27\,\frac{\text{mV}}{\text{mW}}$$

Now we need to know Δp_D and we can find Δv_{BE}. Power analysis of each of the transistors is performed at the operating point and at each extreme from it. The Δi_C and Δv_{CE} of each transistor is calculated and then added and subtracted from the operating-point value. Because v_{BE} varies little, $\Delta v_{CE} \approx \Delta v_{CB}$ and v_{CB} is used instead. The *c-b* junction power dissipation is most of p_D and almost all of Δp_D. For Q1,

$$\Delta i_{E1} = \frac{v_i}{r_{M1}} = \frac{0.4\text{ V}}{630\,\Omega} = 0.635\text{ mA}$$

Corresponding to the input voltage range,

$$v_I \in [-0.4\text{ V}, 0\text{ V}, +0.4\text{ V}]$$

is $i_{E1} \in$ [1.436 mA, 2.071 mA, 2.706 mA].

Δv_{C1} is easy to find from the range and gain values. The maximum input range of ±0.4 V times the gain magnitude of 4 results in $\Delta v_{C1} = \pm 1.6$ V. Then $v_{C1} \in$ [6.76 V, 5.16 V, 3.56 V] and

$$v_{CB1} \in [7.16\text{ V}, 5.16\text{ V}, 3.16\text{ V}]$$

Finally, the p_{D1} range is

$$p_{D1} \in [10.3\text{ mW}, 10.7\text{ mW}, 8.55\text{ mW}]$$

From the three power points, it is evident that Δp_D each side of the center operating-point is not equal and minimum thermal distortion is not achieved. However, the difference between $+\Delta p_D$ and $-\Delta p_D$ are not huge and both are less than the op-pt value of 10.7 mW, indicating that the maximum p_D value is within this range.

A similar exercise carried out for Q2 results in

$$p_{D2} \in [1.24\text{ mW}, 1.78\text{ mW}, 2.33\text{ mW}]$$

The values are monotonic and maximum power does not occur within the interval because v_{CB2} does not decrease with increasing current. The p_{D2} values are small relative to the Q1 values and do not track to cancel Q1 values.

The temperature drift of V_{BE5} will cause a voltage change at the Q4 base that subtracts from the same change that occurs in V_{BE4} and their V_{BE} values track. This makes V_0 independent of V_{BE} and ΔT and no static thermal effect is added to the output by Q4 and Q5. Given V_0, for Q3, $V_{C3} = 12\text{ V} - V_0 = 11.05$ V. The Q3, Q4 voltage, current, and power ranges are calculated as for the first stage:

$$i_{E3} \in [1.313\ \text{mA}, 2.051\ \text{mA}, 2.788\ \text{mA}]\ ;\ \Delta i_{E3} = \pm 0.738\ \text{mA}$$

$$v_{CB3} \in [4.285\ \text{V}, 5.885\ \text{V}, 7.485\ \text{V}]\ ;\ \Delta v_{CB3} = \pm 1.6\ \text{V}$$

$$p_{D3} \in [5.63\ \text{mW}, 12.1\ \text{mW}, 20.9\ \text{mW}]$$

Thermal distortion for Q3 is not minimized because its collector voltage remains nearly constant with Δi_{E3}. In a more refined design, a shunt RC is placed in series with the Q3 collector, with R chosen for thermal balance and C for dynamic bypassing of R. For Q4,

$$i_{E4} \in [0.1812\ \text{mA}, 0.816\ \text{mA}, 1.451\ \text{mA}]\ ;\ \Delta i_{E4} = T_0 \cdot i_{E3} = 0.6348\ \text{mA}$$

$$v_{CB4} \in [18.39\ \text{V}, 10.39\ \text{V}, 2.39\ \text{V}]\ ;\ \Delta v_{CB4} = \pm 8\ \text{V}$$

$$p_{D4} \in [3.33\ \text{mW}, 8.48\ \text{mW}, 3.47\ \text{mW}]$$

Q4 is thermally balanced because at midrange (0 V input), its output of 0 V is midway between the supplies.

The thermal resistance of silicon is low enough and the IC area small enough that monolithic BJTs can be assumed to be at the same static (thermal steady-state) temperature and that they share a common ΔT. They each contribute to it so that for the total effect the p_D values of each of the BJTs is added.

Now that all the BJT power dissipations are known, for the amplifier, they are added to result in

$$p_D \in [20.50\ \text{mW}, 33.06\ \text{mW}, 35.25\ \text{mW}]$$

The temperature coefficient (TC) of V_O is

$$\text{TC}(V_O) = \text{TC(amplifier)} \cdot R_{JA} = (18\ \text{mV/K}) \cdot (150\ \text{K/W}) = 2.7\ \text{V/W}$$

At power-on, the thermal heating from zero power dissipation to the operating-point power causes a thermal transient that can be monitored with a voltmeter. For amplifier 3,

$$\text{thermal}\ \Delta V_O = (2.7\ \text{V/W}) \cdot (33.06\ \text{mW}) = 89.3\ \text{mV}$$

The amplifier 3 prototype was adjusted for zero output voltage when in thermal equilibrium, then allowed to cool and measured from power-on. After 5 minutes, $\Delta V_O \approx 95$ mV.

Thermal $v_O(t)$ is not exponential but is the complementary error function (cerf(t)). Because of spatial distribution, thermal systems are like transmission lines and instead of exponential response have the cerf response. Multiple series RC circuits are commonly used in amplifiers to approximate compensation for thermal drift. At power-

on, drift is greatest because of the Δp_D thermal step input applied to the BJTs, and the value of Δv_O depends on how close to zero time (the instant of power-on) the initial reading is taken. For T to reach the thermal operating-point can take a few minutes.

Thermal Feedback

The total change in power dissipation of the BJTs of the CA3096 array affects the temperature of all of them. The effect is *thermal feedback* because a change from the operating-point, where total power dissipation is p_{D0}, to a different p_D is then converted to a ΔT by R_{JA} and then to Δv_O by TC(v_O). ΔT also causes a change in V_{BE} of the BJTs (–1.8 mV/K) that is amplified by a conversion factor of $\Delta p_D/\Delta V_{BE}$. The p_D combined range for the power-changing BJTs was

$$p_D \in [20.50\text{ mW}, 33.06\text{ mW}, 35.25\text{ mW}]$$

Taking the extreme positive and negative changes from the center (op-pt) value and dividing by the corresponding input voltage changes of ±0.4 V, then the change in Δp_D caused by a change in voltage in the Q1 input loop is

$$\frac{\Delta p_D}{\Delta v_I} = \begin{cases} -5.475\text{ mW/V} \\ +31.4\text{ mW/V} \end{cases}$$

Whether the change is applied as an input or as $-\Delta v_{BE1}$ has the same effect. The other BJT v_{BE} values are similarly affected by Δp_D though with some cancellation of opposing p_D effects. The above values thus constitute a worst-case estimate of maximum thermal effect.

The Δp_{DI} caused by Δv_I adds to the op-pt input power (from static electrical and ambient temperature) of a thermal feedback loop. The resulting Δp_D, as the thermal loop error quantity, develops ΔT across R_{JA} that produces Δv_{BE1} from $TC(V_{BE}) = -1.8$ mV/K. If for positive feedback, loop-gain magnitude is not less than one, then an unstable thermal loop exists. Worst-case loop gain magnitude is

$$G \cdot H = R_{JA} \cdot TC(v_{BE}) \cdot \frac{\Delta p_D}{\Delta v_{BE}} = (150\text{ K/W}) \cdot (-1.8\text{ mV/K}) \cdot \begin{cases} -5.475\text{ mW/V} \\ +31.4\text{ mW/V} \end{cases}$$

This reduces for amplifier 3 to

$$G \cdot H = \begin{cases} 1.478 \cdot 10^{-3} \\ -8.48 \cdot 10^{-3} \end{cases}$$

The positive extremum has a positive $G{\cdot}H$ so that with the negative sign in the conventional feedback block diagram, this is negative feedback, is much less than one, and is unconditionally stable. The case of the negative extremum is that of positive feedback, and again the magnitude of the loop gain is far less than one. We can conclude that the thermal feedback on the monolithic IC is weak, insufficient to cause a "thermal runaway" instability (which is otherwise more commonly found in power transistors).

Nevertheless, a shift in the power equilibrium of the circuit requires that the thermal loop settle before the electrical variables will be settled. This is evident when adjusting for zero output voltage. The adjustment seems to have a hysteretic effect that is caused by the thermal response. The final value of Δp_D when the thermal loop settles is that of Δp_{DI} (caused by v_I or an ambient temperature change) times $1/(1 + G{\cdot}H)$.

To illustrate this with numbers, a reduction of circuit power by 20 mW (approximately that caused by a –0.4 V Δv_I) reduces T by 3 K and the output voltage decreases by 54 mV. The change in v_{BE} of the BJTs is 5.4 mV. Multiplying this number by $\Delta p_D/\Delta v_{BE}$ = 31.4 mW/V, the additional reduction in power dissipation is only 0.17 mW. Even so, when millivolts of output offset matter, this open-loop amplifier can drift by tens of mV.

Static precision circuits are often shielded from airflow that can cause ambient temperature changes in the circuit by placing them in an enclosure. Adjustments are made through holes in the enclosure for screwdrivers (and spudgers or other adjustment tools) so that the circuit can remain enclosed while adjusted. To demonstrate the effects of ambient temperature change, blow on the powered CA3096 of amplifier 3 and watch the output voltage decrease from the cooling effect of air expanded through a nozzle (pursed lips).

β Sensitivity

In the category of quasistatic analysis, another consideration is the effect of BJT β variation on amplifier performance. Measurement of β of some BJTs in different CA3096 units showed a wide variation of β_{NPN}, from around 200 to 350, but none as high (in the sample

population) as the specified typical value of 390. The β_{PNP} values were within a smaller range, varied from about 45 to 55, and are well-characterized with a typical value of 50 at 1 mA.

The first-stage gain is affected by α_{NPN} which varies minutely with β_{NPN} because α is insensitive to β for large β. The Q1 load resistance has a β-dependent term;

$$R_{L1}' = R_{L1} + r_{e2} + \frac{R_{B2}}{\beta_{NPN} + 1}$$

The stable R_{L1} must be much greater than $r_{e2} + R_{B2}/(\beta_{NPN} + 1)$ for a constant R_{L1}'.

The second stage α_3 and α_4 are also stable with large β. β_{PNP} (50) is not nearly so large as β_{NPN} and the change in α with β is

$$\frac{d\alpha}{d\beta} = \frac{d}{d\beta}\left(\frac{\beta}{\beta+1}\right) = \frac{1}{\beta+1} - \frac{\beta}{(\beta+1)^2} = \frac{1}{(\beta+1)^2}$$

For β = 50, it is about 0.0385 % or 385 ppm, not very significant in many applications. There is another β-sensitive gain in this amplifier:

$$T_0 = \frac{R_0}{R_0 + r_{E4}},\ r_{E4} = r_{e4} + \frac{R_{B4}}{\beta_{PNP} + 1}$$

As with R_{L1}', if $R_0 >> r_{E4}$, then the effect of β is negligible. Overall, this amplifier has no significant gain sensitivity to β.

Power Supply Rejection

Yet another amplifier performance consideration is its insensitivity to power supply voltage variations, or *power supply rejection* (PSR). Supply voltages can vary because of supply tolerances, temperature, component aging, or loading. The ability of an amplifier to reject power supply changes is characterized as a performance parameter, the *power supply rejection ratio* (*PSRR*):

$$PSRR = 20 \cdot \log(v_{ps} / v_o)$$

The *PSRR* is simply the dB-scaled ratio of change in output voltage, v_o, with a corresponding change in power-supply voltage, v_{ps}. The logarithmic scale is used because rejection is usually effective and the numbers are much less than one. A number like 1/10,000 = 0.0001 is

sometimes more easily expressed as –80 dB, where dB indicates the scaling as logarithmic and is not a unit. *PSRR* is unitless.

PSR is analyzed by considering each supply as an input port and finding the gain from it to the output. No new concepts are required to perform this kind of quasistatic analysis. Measurements of the prototype of amplifier 3 resulted in the following PSR data:

$$\frac{v_o}{v_+} = \frac{-0.93\ \text{V}}{0.5\ \text{V}} = -1.86\,;\ \frac{v_o}{v_-} = \frac{1.065\ \text{V}}{-0.5\ \text{V}} = -2.13$$

The amplifier amplifies variations in the +12 V supply by a factor of –1.86. It is even more sensitive to –12 V. Variations in either are amplified instead of rejected, and this is a weakness of the circuit.

By inspecting the circuit structure, V_- has a direct (×1) effect through R_{L2} on v_O. For Q1, as V_- decreases slightly, it is attenuated by the R_E divider and then amplified. The attenuation is

$$\frac{750\,\Omega}{750\,\Omega + 3740\,\Omega} \approx \frac{1}{6}$$

The gain magnitude of Q1 as a CB stage is the same as the CE including the above attenuation. Thus the overall gain of V_- through Q1 to v_O is about –20/6 = –3.33. The sign opposes the positive change through R_{L2} and the combined gain is thus –3.33 + 1 = –2.33, not too different from the measured value.

A similar exercise can be performed on V_+. The Q2 base divider has an open-circuit voltage attenuation of 0.936. This voltage is "behind" (in series with) R_{L1}', which is taken into account in the gain of the second stage, in r_{M3}. The resulting $(0.936)\cdot v_+$ is amplified by the stage-2 gain of –5 to result in a gain of –4.68.

The gain v_o/v_+ for the Q4, Q5 branches of +12 V is found from the difference voltage applied to the input loop of Q4. The Q4 emitter branch returns directly to v_+ and the Q5 divider path to the base attenuates v_+ by 0.9115. The difference, $0.0885\cdot v_+ = v_+/11.30$, across r_{M4} is amplified by Q4 for a quasistatic gain of

$$\frac{v_o}{v_+} = +\alpha_{PNP}\cdot\frac{R_{L2}}{r_{M\,4}}\cdot\frac{1}{11.30} \approx (0.98)\cdot\frac{15\,\text{k}\Omega}{386\,\Omega}\cdot\frac{1}{11.3} = \frac{38.1}{11.3} = 3.37$$

The opposing contributions add algebraically, and $v_o/v_+ = -1.31$, within range of the measured value. PSR is poor in this amplifier because it lacks stages with high differential rejection of the supply voltages. BJT or FET diff-amp stages are superior for reduced PSR.

Static Voltage and Gain Measurements

The static voltages of amplifier 3 were measured on the prototype unit with the following results. The supplies were first adjusted for +12.00 V and –12.00 V with a 3.5 digit Beckman DM27XL DMM and the output adjusted for 0 V. TDS360 oscilloscope measurements have a specified maximum inaccuracy (otherwise referred to as "accuracy" in much of the literature) of 3 %, a typical value for most analog and digital 'scopes.

CA3906 Pin Number	Voltage	Measured Voltage	Design Voltage
1	V_{B2}	11.10 V	11.14 V
2	V_{E2}	10.42 V	10.43 V
6	V_{C1}, V_{B3}	5.14 V	5.16 V
4	V_{E1}	–0.708 V	–0.710 V
7	V_{E3}	4.44 V	4.45 V
9	V_{C3}, V_{E4}	11.02 V	11.05 V
11	V_{B4}	10.35 V	10.39 V
12	V_{C4}, V_O	≈ 10 mV	0 V
$V(R_0)$	12 V – V_{C3}	0.960 V	0.955 V

Gains are ratios and this reduces error because relative measurement accuracy (*precision*) is better than (absolute) accuracy. The voltage gains for a 1 kHz sine-wave input were measured on a Tektronix TDS360 DSO using the peak-to-peak (pk-pk) volts measurement function. Both channels were used for comparison.

Voltage, V, pk-pk	ch 1	ch 2	ch 1 $\lvert A_v \rvert$	ch 2 $\lvert A_v \rvert$
v_i	0.400	0.394	1	1
v_{c1}	1.58	1.58	3.95	4.01
v_o	7.88	7.84	4.99	4.96

The difference between the static and quasistatic measurements given above and design values are well within the accuracy of the test equipment. In the case of the gains, they are slightly on the low side and though this can be attributed to component tolerances and measurement error, it suggests that r_o has had its effect in reducing the gain slightly.

Dynamic Design: Second Stage CB

It is easier to start at the output and work backwards for the dynamic design; the second stage is considered first. The output node, where a load of $C_{L2} = C_o = 15$ pF is assumed (×10 'scope probe), has $R_{I2} = 15.0$ kΩ, a relatively large value. The quasistatic effect of Q4 r_o on voltage gain is negligible because of emitter bootstrapping (positive feedback), though its value of

$$r_{o4} = \frac{65\text{ V} + 10\text{ V}}{0.8\text{ mA}} = 93.8\text{ k}\Omega$$

is only 6.25 times greater than $R_{L2} = 15.0$ kΩ. The parameter, an incremental resistance, has the same effect for dynamic analysis as for quasistatic and is considered insignificant. Thus, using R_{L2},

$$\tau_{L2} = R_{L2} \cdot C_{L2} = (15.0\,\text{k}\Omega) \cdot (15\text{ pF}) = 225\text{ ns} \Rightarrow 707.4\text{ kHz}$$

Q3 emitter peaking will be used to compensate τ_{L2}.

The calculated large-signal second-stage bandwidth (below which slewing does not occur) is

$$f_{BW2} = \frac{I_{C4}/C_o}{2 \cdot \pi \cdot V_o} = \frac{(0.8\text{ mA})/(15\text{ pF})}{2 \cdot \pi \cdot (8\text{ V})} = 1.06\text{ MHz}$$

Slewing occurs for a full-scale output at $f_{BW} < f_{bw}$ and the full-scale capability will not apply above f_{BW2}. To extend f_{BW}, either I_{C4} can be increased, thereby reducing f_T, or C_o reduced. A CC stage following the cascode as a buffer would reduce C_L and increase f_{BW}.

The C_{c4} OCTC is

$$\tau_{c4} = [R_{B4}' \cdot (1 + K_{v4}) + R_{L2}] \cdot C_{c4}$$

$$= [(1128\ \Omega) \cdot (41.35) + 15.0\text{ k}\Omega] \cdot (2\text{ pF}) = 123.3\text{ ns} \Rightarrow 1.29\text{ MHz}$$

where Q4 base-to-collector Miller-effect voltage gain is

$$K_{v4} = \alpha_{PNP} \cdot \frac{R_{L2}}{r_{e4} + R_0} = (0.979) \cdot \frac{15.0\text{ k}\Omega}{364\ \Omega} = 40.35$$

C_{B4} is intended to reduce base impedance but it does not affect the open-circuit R_{bc4}. However, the C_{c4} OCTC is not isolated from Z_B and the result is not a pole at $1/\tau_{c4}$. We will return to the solution of the

output circuit poles and zeros after the CB gain is found and add C_{c4} using the EET.

Now consider Q4 input-loop dynamics. The PNP f_T at 0.8 mA is 6.5 MHz and $\tau_{TPNP} = r_{m4} \cdot C_{e4} = 24.5$ ns. The OCTC of C_e is

$$\tau_{e4} = \alpha_0 \cdot \tau_T \cdot \frac{R_{B4}' + R_0}{r_{M4}} = (0.980) \cdot (24.5\,\text{ns}) \cdot \frac{1128\,\Omega + 332\,\Omega}{386\,\Omega} = 90.85\,\text{ns}$$

$$\Rightarrow 1.75\ \text{MHz}$$

The Q4 input circuit impedance, z_{E4}, into the emitter port of Q4 is $Z_{\pi 4}$ referred to the emitter in series with the emitter-referred external base impedance. The Q4 base parallel RC impedance is

$$Z_{b4} = R_{B4}' \| (1/s \cdot C_{B4}) = \frac{R_{B4}'}{s \cdot R_{B4}' \cdot C_{B4} + 1} = \frac{R_{B4}'}{s \cdot \tau_{B4} + 1}$$

On the Q3 collector side of the Q4 emitter port, input current i_{c3} develops v_{e4} across the shunt input impedance

$$Z_{i4}(s) = R_0 \| \frac{1}{s \cdot C_{cs3}} = \frac{R_0}{s \cdot R_0 \cdot C_{cs3} + 1} = \frac{R_0}{s \cdot \tau_{E4} + 1}$$

where R_0 is the load resistor of Q3 and also the emitter resistor of Q4.

The dynamic form of A_{v2b} for the input loop is found in much the same way as the quasistatic formula, though with considerably more algebraic effort;

$$A_{v2bi}(s) = \frac{v_o}{i_{c3} \cdot R_0} = \frac{i_{c4} \cdot R_{L2}}{i_{c3} \cdot R_0} = \alpha(s) \cdot \frac{i_{e4} \cdot R_{L2}}{i_{c3} \cdot R_0} = \alpha(s) \cdot T_0(s) \cdot \frac{R_{L2}}{R_0}$$

The work is in finding the current divider;

$$T_0(s) = \frac{i_{e4}}{i_{c3}} = \frac{Z_{i4}(s)}{Z_{i4}(s) + z_{E4}(s)} = \frac{R_0}{R_0 + (s \cdot \tau_{E4} + 1) \cdot z_{E4}}$$

From BJT hf dynamics,

$$\beta(s) + 1 = (\beta_0 + 1) \cdot \frac{s \cdot \alpha_0 \cdot \tau_T + 1}{s \cdot \tau_\beta + 1}$$

Applying the $\beta(s)$ transform to z_{E4},

$$z_{E4} = \frac{Z_{\pi 4} + Z_{b4}}{\beta(s) + 1} = \frac{r_{e4}}{s \cdot \alpha_0 \cdot \tau_T + 1} + \frac{R_{B4}'/(\beta_0 + 1)}{s \cdot \alpha_0 \cdot \tau_T + 1} \cdot \frac{s \cdot \tau_\beta + 1}{s \cdot \tau_{B4} + 1}$$

This simplifies for $\tau_{B4} = \tau_\beta$ to

$$z_{E4}\Big|_{\tau_{B4}=\tau_\beta} = \frac{r_{e4} + R_{B4}'/(\beta_0 + 1)}{s \cdot \alpha_0 \cdot \tau_T + 1} = \frac{r_{E4}}{s \cdot \alpha_0 \cdot \tau_T + 1}$$

When $T_0(s)$ is derived for the general case of z_{E4} and substituted into the gain expression,

$$A_{v2bi} = \alpha_0 \cdot \frac{R_{L2}}{r_{M4}} \cdot \frac{s \cdot \tau_{B4} + 1}{s^2 \cdot a + s \cdot b + 1}$$

where

$$a = \alpha_0 \cdot \tau_T \cdot R_{B4}' \cdot (C_{B4} + C_{cs3}) \cdot \left(\frac{R_0}{r_{M4}} \right) + \tau_{E4} \cdot \tau_{B4} \cdot \left(\frac{r_{E4}}{r_{M4}} \right)$$

$$b = \tau_{e4} + \tau_{B4} \cdot \left(\frac{r_{e4} + R_0}{r_{M4}} \right) + \tau_{E4} \cdot \left(\frac{r_{E4}}{r_{M4}} \right)$$

$$= \tau_{e4} + \left[(R_{B4}'/(\beta_0 + 1)) \,\|\, (r_{e4} + R_0) \right] \cdot [(\beta_0 + 1) \cdot C_{B4}] + (r_{E4} \,\|\, R_0) \cdot C_{cs3}$$

and where $r_{E4} = r_{e4} + R_{B4}'/(\beta_0 + 1) = 54.0\ \Omega$; $r_{M4} = r_{E4} + R_0 = 386\ \Omega$;

$$\tau_{e4} = \alpha_{PNP} \cdot \tau_{T4} \cdot \frac{R_{B4}' + R_0}{r_{M4}} = 90.85 \text{ ns} \text{ ; } \tau_{E4} = R_0 \cdot C_{cs3} = 0.66 \text{ ns} \text{ ;}$$

$$\tau_{B4} = R_{B4}' \cdot C_{B4} = 5.64\ \mu\text{s} \Rightarrow 28.2 \text{ kHz}$$

$$a = (341.42 \text{ ns})^2 + (17.57 \text{ ns})^2 = (341.87 \text{ ns})^2$$

$$b = 90.85 \text{ ns} + 5.319\ \mu\text{s} + 93 \text{ ps} = 5.41\ \mu\text{s}$$

The effects of C_{e4}, C_{cs3}, and C_{B4} each contribute a term to both the a and b coefficients of the pole-pair. The τ_{e4} term dominates a and the τ_{B4} term dominates b. With three independent input-loop capacitances, the pole polynomial would be a cubic, but the capacitors form a loop and are not independent - hence the quadratic polynomial. The resulting pole-pair parameters are

$$f_n = 1/2 \cdot \pi \cdot \sqrt{a} = 465.5 \text{ kHz} \text{ ; } \zeta = b/2 \cdot \sqrt{a} = 7.912$$

The poles are real ($\zeta \geq 1$) and separated by a factor of 248.4, at

$$p_{1,2} = 29.54 \text{ kHz}, 7.34 \text{ MHz}$$

The lower-frequency pole nearly cancels the zero at 28.2 kHz and, because it does not exactly match the zero frequency, forms a

dipole in the response - a slight positive (for $f_z < f_p$) "bump" in both step and frequency responses. The zero causes the gain magnitude to rise at a +1 slope for log(29.54/28.2) ≈ 0.02 decade, and the bump in gain will be +2 % in gain. If that is too much, increase the size of C_{B4} (and τ_{B4}) to increase both a and b and cause the lower pole to become closer to $1/\tau_{B4}$. The dipole was not seen in the prototype amplifier.

The lower pole and zero cancel exactly for the simplifying case of $\tau_{B4} = \tau_{\beta 4}$. The effect is to make $z_{E4} = r_{E4}$ for $f \leq f_T/\alpha_0$. The gain is then

$$A_{v2bi}\Big|_{\tau_{B4}=\tau_{\beta 4}} = \alpha_0 \cdot \frac{R_{L2}}{r_{M4}} \cdot \frac{1}{s \cdot \left[\alpha_0 \cdot \tau_T \cdot \left(\frac{R_0}{r_{M4}} \right) + (r_{E4} \parallel R_0) \cdot C_{cs3} \right] + 1}$$

with a quasistatic gain of $A_{v2bi0} = 38.1$ and

$$\tau_{2b} = 20.66 \text{ ns} + 93 \text{ ps} \approx 20.75 \text{ ns} \Rightarrow 7.67 \text{ MHz}$$

The circuit interpretation of the two terms of τ_{2b} is that the first, from C_e, dominates over the second, from the RC integrator that is Z_i. To simplify the analysis, we could ignore Z_i, but for a general set of template formulas, it is left in. It becomes significant for large r_{E4} and R_0, such as when R_0 is replaced by a current source.

A resistive z_E reduces the gain to a single-pole response but does not compensate for Z_i. The additional constraint that $\tau_{E4} = \alpha_0 \cdot \tau_T$ makes Z_π and Z_i a compensated divider and results in a dynamically ideal amplifier having infinite bandwidth (for the input loop), with no dependence on s:

$$A_{v2bi}\Big|_{\tau_{B4}=\tau_{\beta 4};\tau_{E4}=\alpha_0 \cdot \tau_T} = \alpha_0 \cdot \frac{R_{L2}}{r_{M4}}$$

For this condition to be satisfied, $R_{E4} \cdot C_{E4} = \alpha_0 \cdot \tau_T$, where $R_{E4} = R_0$ and $C_{E4} = C_{cs3}$. Then to meet this condition,

$$C_{E4} = \frac{\alpha_0 \cdot \tau_T}{R_0} = \frac{24.02 \text{ ns}}{332 \, \Omega} = 72.35 \text{ pF}$$

An additional 70.35 pF would need to be added to ground from the Q4 emitter. This value is large because f_{TPNP} is small.

In the design as given, the Q4 base is bypassed with a capacitor that places the frequency of τ_{B4} in the lf region (below f_β). In the limit, a large C_{B4} effectively bypasses the base and makes the stage

dynamically closer to a true CB configuration. In the case of infinite C_{B4}, the zero and lower-frequency pole cancel, leaving

$$A_{v2bi}\Big|_{C_{B4}\to\infty} = \alpha_0 \cdot \frac{R_{L2}}{r_{e4} + R_0} \cdot \frac{1}{s \cdot [(r_{e4} \parallel R_0) \cdot (C_e + C_{cs3})] + 1}$$

This can be rewritten as

$$A_{v2bi}\Big|_{C_{B4}\to\infty} = \alpha_0 \cdot \frac{R_{L2}}{r_{e4} + R_0} \cdot \frac{1}{s \cdot \left[\alpha_0 \cdot \tau_{T4} \cdot \left(\frac{R_0}{r_{e4} + R_0} \right) + \tau_{E4} \cdot \left(\frac{r_{e4}}{r_{e4} + R_0} \right) \right] + 1}$$

When numbers are substituted for parameters,

$$A_{v2bi0}\Big|_{C_{B4}\to\infty} = 40.40$$

$$\tau_{2b} = 21.91 \text{ ns} + 58 \text{ ps} = 21.97 \text{ ns} \Rightarrow 7.24 \text{ MHz}$$

As C_{B4} increases, the zero and lower-frequency pole migrate together toward the origin and disappear from the transfer function. The response is single-pole.

Whenever the time constant of $C_B \cdot R_B' > \tau_\beta$, it is in the lf region where the quasistatic BJT model of the ideally fast BJT ($\tau_T = 0$ s) can be used for analysis. The simplified CB lf gain for $R_B' \cdot C_B > \tau_\beta$ is

$$A_v = \alpha_0 \cdot \frac{R_L}{r_M} \cdot \frac{s \cdot R_B' \cdot C_B + 1}{s \cdot \left[(\beta_0 + 1) \cdot (r_e + R_E) \parallel R_B' \right] \cdot C_B + 1}$$

C_B causes a zero in the gain but if the resistance looking into the base is large, the pole is at about the same frequency as the zero and they cancel. The time constant of the pole is referred to the base, but can be referred to the emitter instead by regrouping:

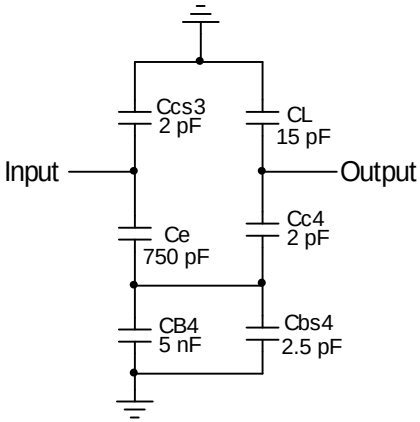
$$\left[(\beta_0 + 1) \cdot (r_e + R_E) \parallel R_B' \right] \cdot C_B = \left[(r_e + R_E) \parallel \frac{R_B'}{\beta_0 + 1} \right] \cdot [(\beta_0 + 1) \cdot C_B]$$

In the lf region, C_B appears $\beta_0 + 1$ times larger from the emitter - a C multiplier. For the complementary cascode, $R_{E4} = R_0$, so that whenever R_0 is much greater than $R_B/(\beta_0 + 1)$, then the pole and zero cancel. For amplifier 3,

$$R_0 = 332\ \Omega >> R_B'/(\beta_0 + 1) = 22.6\ \Omega$$

We now resume at the collector side of Q4 with C_{c4}. The general voltage gain does not include it, having been derived with the input-

loop capacitances in mind. A reactance model of the amplifier has effectively five capacitances, as shown below (where $C_{B4} + C_{bs4}$ is a single capacitance). The input-loop gain formulas apply to the left branch of three. The output pole of C_L, τ_{L2}, has also been calculated.



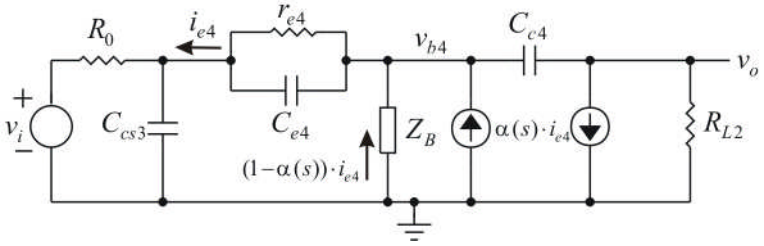
We left the disposition of τ_{c4} until now because we did not yet have the gain formula. We can include the effect of C_{c4} in it by applying the extra-element theorem. This is the kind of problem for which the EET is intended, to incrementally extend the gain expression for an additional reactance - in this case, C_{c4}.

The EET gain formula assumes that $A_{v2bi}(s)$ (without C_{c4} and hence with the EET Z port open-circuited) has already been derived;

$$A_{v2b}(s) = A_{v2bi}(s) \cdot \frac{1 + \frac{Z_N}{Z}}{1 + \frac{Z_D}{Z}}, Z = 1/s \cdot C_{c4}$$

A circuit diagram of the CB stage is shown below. The source-shifting theorem was used on the collector current source, by disconnecting it from the base and connecting it to ground. Then to maintain circuit equivalence, another equal source is connected from ground to the base. The input is in a Thevenin-equivalent form.

We seek to find Z_D and Z_N, in that order. Z_D is the open-circuit impedance into the C_{c4} (Z) port. We have R_{bc} as a template formula but Z_D is Z_{bc}. The difference is in the Miller multiplier, $K_{v4}(s)$, which is the base-collector voltage gain. It is valid as the quasistatic value out to the bandwidth of $K_{v4}(s)$.



Proceeding (with v_i = 0 V),

$$Z_D = R_{L2} + (1 + K_{v4})\cdot Z_{B4} = R_{L2} + (1+ K_{v4})\cdot\left(\frac{R_{B4}'}{s\cdot\tau_B + 1}\right)$$

The denominator of the EET correction factor is

$$1+\frac{Z_D}{Z} = (s\cdot C_{c4}\cdot R_{L2} + 1) + \frac{s\cdot(1+K_{v4})\cdot R_{B4}'\cdot C_{c4}}{s\cdot\tau_B + 1}$$

$$= \frac{s^2\cdot[\tau_B\cdot R_{L2}\cdot C_{c4}] + s\cdot[\tau_B + (R_{L2} + R_{B4}'\cdot(1+K_{v4}))\cdot C_{c4}] + 1}{s\cdot\tau_B + 1}$$

To find Z_N, null v_o to 0 V by letting v_i be whatever value will achieve the null. For the output to be at 0 V, the current in R_L must be zero and this is achieved by forcing the collector current to flow through C_{c4}. Then

$$v_{b4} = -i_{b4}\cdot Z_{B4} = -[(1-\alpha(s))\cdot i_{e4}]\cdot Z_{B4}$$

$\alpha(s)\cdot i_{e4}$ is flowing from left to right through C_{c4} with a voltage from left (+) to right (–) terminal of v_{b4}. Then

$$Z_N = \frac{v_{b4}}{\alpha(s)\cdot i_{e4}} = \frac{-[(1-\alpha(s))\cdot i_{e4}]\cdot Z_{B4}}{\alpha(s)\cdot i_{e4}} = -\frac{Z_{B4}}{\beta(s)} = -\frac{s\cdot\tau_\beta + 1}{\beta_0}\cdot Z_{B4}$$

Combining with Z,

$$1+\frac{Z_N}{Z} = 1-\frac{s\cdot\tau_\beta + 1}{\beta_0}\cdot\frac{R_{B4}'}{s\cdot\tau_B + 1}\cdot(s\cdot C_{c4})$$

$$= \frac{-s^2\cdot[\tau_\beta\cdot(R_{B4}'/\beta_0)\cdot C_{c4}] + s\cdot[\tau_B - (R_{B4}'/\beta_0)\cdot C_{c4}] + 1}{s\cdot\tau_B + 1}$$

The extended voltage gain with the effect of C_{c4} is therefore

$$A_{v2b}(s) = A_{v2bi}(s) \cdot \frac{-s^2 \cdot [\tau_\beta \cdot (R_{B4}'/\beta_0) \cdot C_{c4}] + s \cdot [\tau_B - (R_{B4}'/\beta_0) \cdot C_{c4}] + 1}{s^2 \cdot [\tau_B \cdot R_{L2} \cdot C_{c4}] + s \cdot [\tau_B + (R_{L2} + R_{B4}' \cdot (1 + K_{v4})) \cdot C_{c4}] + 1}$$

We can expect a RHP zero from the negative s^2 coefficient. For $\tau_B = \tau_\beta$, the numerator factors into two zeros with time constants of $\tau_\beta \Rightarrow f_\beta = 130$ kHz and $-(R_{B4}'/\beta_0) \cdot C_{c4} = -45.1$ ps $\Rightarrow 3.5$ GHz. The RHP zero is far outside the design frequency range and f_β is in the lf range. These real zeros are separated so far that even if $\tau_B \neq \tau_\beta$, the $(R_{B4}'/\beta_0) \cdot C_{c4}$ time constant is so small that the quadratic coefficient can be discarded and a zero at τ_B retained as a close approximation.

In the pole polynomial,

$$D(s) = s^2 \cdot [\tau_B \cdot \tau_{cc4}] + s \cdot [\tau_B + \tau_{cc4} + \tau_{cb4}] + 1$$

the time constants of interest are

$$\tau_B \approx R_{B4}' \cdot C_{B4} = (1128\ \Omega) \cdot (5\text{ nF}) = 5.64\ \mu\text{s} \Rightarrow 28.2\text{ kHz}$$

Expanding the previously-calculated τ_{c4} into components,

$$\tau_{cb4} = (1 + K_{v4}) \cdot [R_{B4}' \cdot C_{c4}] = (1 + 40.35) \cdot [(1128\ \Omega) \cdot (2\text{ pF})] = 93.3\text{ ns} \Rightarrow 1.71\text{ MHz}$$

$$\tau_{cc4} = R_{L2} \cdot C_{c4} = 30.0\text{ ns} \Rightarrow 5.31\text{ MHz}$$

$$\tau_{c4} = \tau_{cb4} + \tau_{cc4} = 123.3\text{ ns} \Rightarrow 1.29\text{ MHz}$$

Both τ_B and τ_{cc4} are in a but τ_{cb4} is in b only. τ_B dominates both coefficients, which results in an approximate pole at f_B. The quadratic coefficients are

$$a = \tau_B \cdot \tau_{cc4} = (411.3\text{ ns})^2\ ;\ b = \tau_B + \tau_{c4} = 5.76\ \mu\text{s}$$

for which $f_n = 387.0$ kHz and $\zeta = 7.006$, with values of

$$p_{1,2} = -27.76\text{ kHz}, -5.39\text{ MHz}$$

The lower pole is nearly at f_B and cancels the zero near that frequency. The higher pole is of interest as the contribution of C_{c4} for bandwidth calculation. We are left with two real poles in A_{v2b}, at 7.34 MHz and 5.39 MHz.

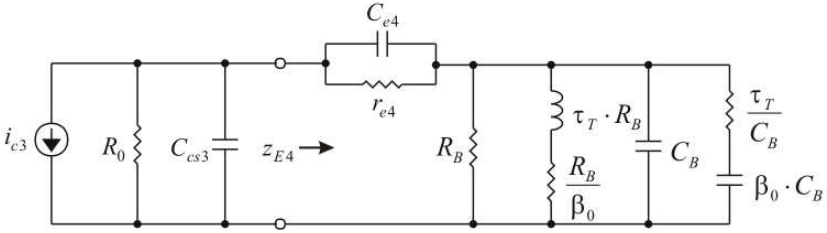
The straightforward mathematical approach to the design of the Q4 stage requires a large amount of algebra to derive the general design formulas. Afterward they can be placed in an engineering notebook or computer file folder and used as *templates* for future designs. This kind of analysis requires skill in interpreting and

reshaping algebraic expressions so that they can be related to circuit structure - so that the circuits (or simplified equivalents of them) can be seen in the math. Algebraic formulas, such as voltage gain, can be built up incrementally using the EET from a simpler model having fewer interacting reactances.

CB Circuit Graphic Design

For those who prefer to avoid so much algebra, there is another way to design - a more graphic method. The difference between the algebraic and graphic methods is mainly in how the $\beta(s)$ transform is applied. In the $A_{v2b}(s)$ derivation, it was applied to the whole input loop; the Z_{b4} block in the circuit diagram was immediately converted to a large algebraic expression. Then the algebra was worked into a form that can be interpreted as an equivalent circuit.

Graphically-oriented design replaces the base circuit elements incrementally with their hf or total-frequency equivalents and the resulting circuit impedances are viewed on a reactance plot. Substitute a total-frequency circuit template for each element of Z_{b4} (referred to the emitter as z_{E4}) and an equivalent circuit model (as was derived from the algebraic expressions) results, shown below.



The $\beta(s)$ transform is applied to each circuit element in the base, one element at a time, using the hf-gyrated circuit templates already derived. Both R_{B4} and C_{B4} are referred from the base to the emitter as shown above in the total-frequency model. The resulting circuit can be compared to the algebraic gain expression to verify consistency.

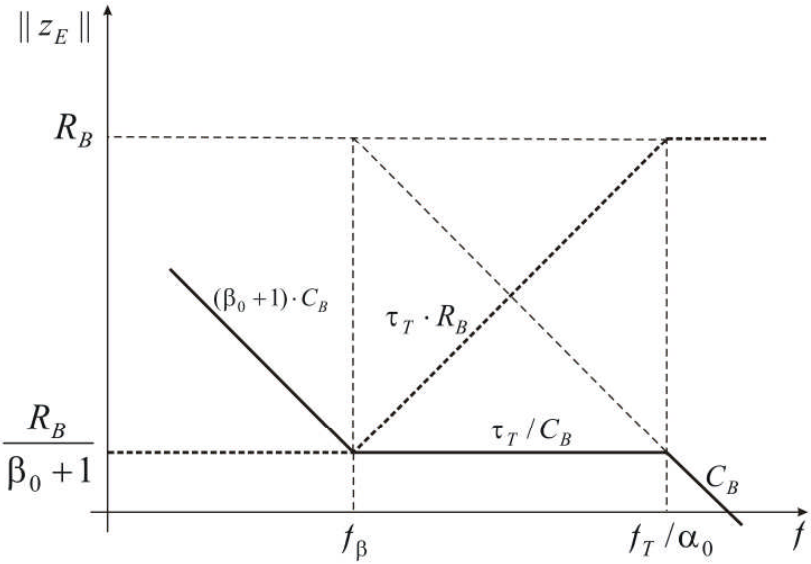
From the emitter, the hf R_B gyrates to form a shunt RL. With L and C in the same circuit, resonance can form. This corresponds to the pole-pair in $A_{v2bi}(s)$ as having an underdamped (resonant, $\zeta < 1$) response.

The Q4 base resistance, referred to the emitter, is $R_{B4}'/(\beta_0 + 1)$. In the hf region, R_B gyrates +90° to become an inductance that from the

template value, $\tau_T \cdot R_B = \tau_{TPNP} \cdot R_{B4}' = 27.64$ μH. Above f_T/α_0, gyration stops and it becomes R_B. On the reactance plot of z_{E4}, the inductive reactance at f_β,

$$X_L = \omega_\beta \cdot (\tau_T \cdot R_B) = (\tau_T \cdot R_B)/\tau_\beta = R_B/\beta_0$$

is less than R_B/β_0 below f_β and series resistance dominates. Combined in parallel with R_B, it is $R_B/(\beta_0 + 1)$ below f_β. The emitter impedance contributed by R_B is shown in the plot as a heavy dashed line.



Similarly, $-X_C = 1/\omega \cdot (\beta_0 \cdot C_B)$ decreases with frequency but in the hf region is dominated by gyrated resistance τ_T/C_B. In the lf region, the β transform refers C_B to the emitter as $(\beta_0 + 1) \cdot C_B$. Above f_β, it gyrates +90° to become a resistance of τ_T/C_B, and is again a capacitance of C_B above the hf region where the BJT lacks the β to gyrate. The impedance plot of C_B is shown as a heavy solid line.

The series RL and series RC branches of z_{E4} combine as a constant impedance of R_B/β_0 throughout the lf and hf regions whenever $\tau_{B4} = \tau_\beta$, or for $C_B = \tau_\beta/R_B$. The two impedances are in parallel; whichever is lower on the impedance plot dominates.

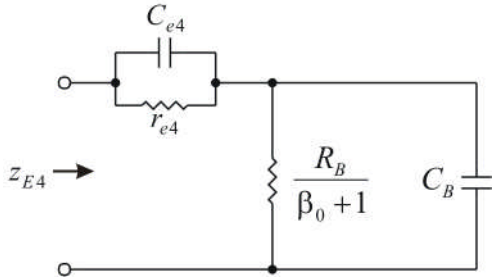
This constant-Z or all-pass condition (which reduced $A_{v2b}(s)$ to a single pole) sets reactances of the branches of z_{E4} equal the resistances at the resonant frequency and the two resistances are set equal to Z_n of the reactances;

$$Z_n = \sqrt{\frac{L}{C}} = \sqrt{\frac{\tau_T \cdot R_B}{\beta_0 \cdot C_B}} \Rightarrow Z_n^2 = \frac{\tau_T \cdot R_B}{\beta_0 \cdot C_B} = \left(\frac{R_B}{\beta_0}\right)^2 = \left(\frac{\tau_T}{C_B}\right)^2$$

or

$$R_B \cdot C_B = \tau_\beta \Rightarrow C_B = \frac{\tau_\beta}{R_B}$$

With the $\tau_B = \tau_\beta$ constraint applied, the equivalent emitter circuit is reduced to that shown below.



The combined gyrated branches of z_{E4} are resistive and when combined further with R_B equal $R_B/(\beta_0 + 1)$ from 0 Hz to the RC time constant of the resulting circuit, or $[R_B/(\beta_0 + 1)]\cdot C_B = \alpha_0 \cdot \tau_T$, which is at a frequency of $f_T/\alpha_0 = (\beta_0 + 1)\cdot f_\beta$. The emitter-referred base impedance is constant throughout both the lf and hf regions whenever

$$\frac{R_B}{\beta_0} \left\| R_B = \frac{R_B}{\beta_0 + 1} = \frac{\alpha_0 \cdot \tau_T}{C_B} \Rightarrow R_B \cdot C_B = \tau_\beta \right.$$

Applied to Q4, for which $R_{B4}' = 1128\ \Omega$,

$$C_{B4} = [(50)\cdot(24.5\text{ ns})]/R_{B4}' = 1.225\ \mu\text{s}/1128\ \Omega = 1.09\text{ nF}$$

For the choice of $C_{B4} = 5$ nF,

$$R_{B4} \cdot C_{B4} = (1128\ \Omega)\cdot(5\text{ nF}) = 5.64\ \mu\text{s} > \tau_\beta = (50)\cdot(24.5\text{ ns}) = 1.225\ \mu\text{s}$$

$\tau_{B4} = R_{B4} \cdot C_{B4}$ exceeds τ_β so that in the hf region, τ_T/C_{B4} appears gyrated at the emitter as a resistance of $4.9\ \Omega$ and is less than $R_{B4}'/(\beta_{PNP} + 1) = 22.1\ \Omega$ by about a factor of 1.09 nF/5 nF. $(\beta_0 + 1)\cdot C_{B4}$ intersects $R_{B4}'/(\beta_0 + 1)$ below f_β, forms a lf pole there, and continues to decrease to f_β, then flattens as it becomes resistive. In the hf region, $\tau_T/C_{B4} < R_{B4}'/(\beta_0 + 1)$ and dominates Z_e.

There is a slight complication caused by r_b', the base spreading resistance, connected between the base terminal, *b*, and the internal base node, *b*'. The Q4 external base impedance is R_{B4} in parallel with C_{B4}. C_{B4} is not connected to the internal base node and does not bypass r_{b4}' to ground. Instead, C_{b4} shunts *b*' and consists of the Miller-multiplied C_{c4} and substrate capacitance, $C_{bs4} \approx 1.75$ pF:

$$C_{b4} = (1 + K_{v4})\cdot C_{c4} + C_{bs4} = (41.4)\cdot(1.8\text{ pF}) + 1.75\text{ pF} = 76.3\text{ pF}$$

where

$$K_{v4} = \alpha_{PNP} \cdot \frac{R_{L2}}{r_{e4} + R_0} = 40.4$$

C_{b4} and r_b' form a time constant of 3.81 ns $\Rightarrow$ 41.7 MHz, about a decade above the expected frequency range of f_{bw}.

The inductance gyrated by r_b' is $\tau_T r_b$', or for Q4 is 1.225 μH, and is dynamically grounded on the external base side through C_B. In the hf region, C_{b4} is in parallel at the Q4 emitter with the Q3 collector load capacitance, $C_{cs3} = 2$ pF, for a resonant capacitance of

$$C_{b4} + C_{cs3} \approx 79\text{ pF}$$

The parallel resonance is at $f_n = 16.18$ MHz and $Z_n = 124.5\ \Omega$. The damping resistance in parallel with LC is $r_b'||(\tau_{TPNP}/C_{b4})||R_0 = 38.2\ \Omega$, and is overdamped ($\zeta = 1.63$) with a pole separation ratio of 8.5 and poles at 5.55 MHz and 47.18 MHz. Without C_{B4}, C_{b4} still sufficiently damps the gyrated r_b'. To have decomposed the base circuit to include *b*' in the algebraic derivation of $A_{v2b}(s)$ would have been a mess, but with the gyrated-element model, the circuit decomposes into localized fragments that can be analyzed for specific behaviors.

Let r_{b4}' be included in R_{B4}' but without C_{B4}. The resonant frequency of $C_{b4} + C_{cs3}$ with gyrated R_{B4}' inductance at the emitter,

$$\tau_{TPNP}\cdot R_{B4}' = (24.5\text{ ns})\cdot(1128\ \Omega) = 27.64\ \mu\text{H}$$

is $f_n = 3.41$ MHz and $Z_n = 592\ \Omega$. Resistance at the Q3 collector is

$$R_{B4}'||(\tau_{TPNP}/C_{b4})||R_0 = (1128\ \Omega)||310\ \Omega||332\ \Omega = 140.4\ \Omega$$

and overdamps the parallel resonance with $\zeta = 2.107$. The poles are at 860 kHz and 13.5 MHz. When C_{B4} is left out, we have a low-frequency pole and need to put it back in.

The total-frequency circuit model can be related to the gain equation. The linear coefficient of the pole polynomial, $D(s)$, is

$$b = \tau_{e4} + \tau_{B4} \cdot \left(\frac{r_{e4} + R_0}{r_{M4}} \right) + \tau_{E4} \cdot \left(\frac{r_{E4}}{r_{M4}} \right)$$
$$= \tau_{e4} + \left[(R_{B4}'/(\beta_0 + 1)) \,\|\, (r_{e4} + R_0) \right] \cdot [(\beta_0 + 1) \cdot C_{B4}] + (r_{E4} \,\|\, R_0) \cdot C_{cs3}$$

and has three OCTCs corresponding to each of the capacitances. The OCTC of C_{e4} is τ_{e4} and was derived (from R_{be}) as a template formula,

$$\tau_e = \alpha_{PNP} \cdot \tau_{TPNP} \cdot \frac{R_{B4}' + R_0}{r_{M4}}$$

The OCTC of C_{B4} is the middle term and is written in the lower expression in a circuit-interpretable form: C_{B4} and $\beta_0 \cdot C_{B4}$ are not independent capacitors and combine with a value of $(\beta_0 + 1) \cdot C_B$. The open-circuit resistance across the C_B terminals is $R_{B4}' \| R_{B4}'/\beta_0 = R_{B4}'/(\beta_0 + 1)$ and is in parallel with $(r_{e4} + R_0)$.

The OCTC of C_{cs3} is found similarly; the open-circuit resistance across it is R_0 in parallel with $r_{E4} = r_{e4} + R_{B4}'/(\beta_0 + 1)$. Consequently, the OCTCs in *b*, put in circuit-recognizable form, correspond to the equivalent circuit, and either can be used to derive design formulas.

In the gain formula, set $\tau_B = \tau_\beta$. Then the first two OCTCs in *b* combine. Distributing within τ_e and recombining the R_B term with the C_B term, we have

$$\alpha_{04} \cdot \tau_{T4} \cdot \frac{R_0}{r_{M4}} + \left(\alpha_{04} \cdot \tau_{T4} \cdot \frac{R_{B4}'}{r_{M4}} + \tau_{\beta 4} \cdot \frac{r_{e4} + R_0}{r_{M4}} \right)$$

Noting that

$$\alpha_0 \cdot \tau_T = \left(\frac{\beta_0}{\beta_0 + 1} \right) \cdot \tau_T = \frac{\tau_\beta}{\beta_0 + 1}$$

the term in parentheses becomes

$$\alpha_{04} \cdot \tau_{T4} \cdot \frac{R_{B4}'}{r_{M4}} + \tau_{\beta 4} \cdot \frac{r_{e4} + R_0}{r_{M4}} = \tau_{\beta 4} \cdot \left(\frac{R_{B4}'/(\beta_0 + 1)}{r_{M4}} + \frac{r_{e4} + R_0}{r_{M4}} \right) = \tau_{\beta 4}$$

Then for $\tau_{B4} = \tau_{\beta 4}$,

$$b = \alpha_{04} \cdot \tau_{T4} \cdot \frac{R_0}{r_{M4}} + \tau_{\beta 4} + \tau_E \cdot \frac{r_{E4}}{r_{M4}}$$

The quadratic coefficient of $D(s)$

$$a = \alpha_{04} \cdot \tau_{T4} \cdot R_{B4}' \cdot (C_{B4} + C_{cs3}) \cdot \left(\frac{R_0}{r_{M4}} \right) + \tau_{E4} \cdot \tau_{B4} \cdot \left(\frac{r_{e4}}{r_{M4}} \right)$$

consists of three terms after C_{B4} and C_{cs3} are distributed. For $\tau_{B4} = \tau_{\beta 4}$ it becomes

$$a = \alpha_{04} \cdot \tau_{T4} \cdot \tau_{\beta 4} \cdot \left(\frac{R_0}{r_{M4}} \right) + \left(\alpha_{04} \cdot \tau_{T4} \cdot \tau_{E4} \cdot \frac{R_{B4}'}{r_{M4}} + \tau_{E4} \cdot \tau_{\beta 4} \cdot \left(\frac{r_{e4}}{r_{M4}} \right) \right)$$

The two terms in parentheses combine to become $\tau_{E4} \cdot \tau_{\beta 4} \cdot (r_{E4}/r_{M4})$ and a becomes

$$a = \tau_{\beta 4} \cdot \left(\alpha_{04} \cdot \tau_{T4} \cdot \frac{R_0}{r_{M4}} + \tau_{E4} \cdot \frac{r_{E4}}{r_{M4}} \right)$$

The products of a are the sums of b and $D(s)$ can be factored into two real poles with time constants $\tau_{\beta 4}$ and

$$\left(\alpha_{04} \cdot \tau_{T4} \cdot \frac{R_0}{r_{M4}} + \tau_{E4} \cdot \frac{r_{E4}}{r_{M4}} \right)$$

The first pole, $(s \cdot \tau_{\beta 4} + 1)$ cancels the zero, leaving as derived, $A_{v2bi}(s)$ for $\tau_{B4} = \tau_{\beta 4}$. Under the additional constraint that $\tau_{E4} = \alpha_{04} \cdot \tau_{T4}$,

$$a = (\alpha_{04} \cdot \tau_{T4}) \cdot \tau_{\beta 4}; \; b = (\alpha_{04} \cdot \tau_{T4}) + \tau_{\beta 4}$$

The pole of time constant $\alpha_0 \cdot \tau_T$ cancels the pole in $\alpha(s)$ and the pole of τ_β is cancelled by the zero, leaving a quasistatic $A_{v2b}(s)$. In the circuit diagram, setting $\tau_{E4} = \alpha_0 \cdot \tau_T$ forms a compensated RC divider between Z_π and Z_i. In addition, if the emitter-referred base circuit is all-pass ($\tau_B = \tau_\beta$), then the resulting circuit is essentially resistive, or quasistatic in response.

In the case of an unbypassed base (C_{B4} removed again), a larger R_0 (such as a current source), a relatively small R_B and a relatively large C_{cs3}, then the conditions for spurious oscillation can exist and can be predicted using the circuit diagram. A parallel resonance formed by $\tau_T \cdot R_{B4}'$ and $C_{b4} + C_{cs3}$ was previously calculated to have a resonant impedance of

$$Z_n = \sqrt{\tau_T \cdot R_{B4}' / (C_{b4} + C_{cs3})} = 592\ \Omega$$

and is damped by a resistance of about $R_0 \| R_{B4}' = 257\ \Omega$, small enough to slightly overdamp ($\zeta = 1.15$). Had R_0 been a current source instead,

the resonance would be underdamped and Q4 oscillation a possibility. The additional requirement for a hf-gyrated oscillation is that the resonant frequency be within the hf region, and it is;

$$\tau_n = \sqrt{\tau_T \cdot R_{B4}' \cdot (C_{b4} + C_{cs3})} = 46.73 \text{ ns} \Rightarrow 3.41 \text{ MHz}$$

This approximate resonance analysis considers $Z_\pi \approx 0\,\Omega$. The omission from z_{E4} of $Z_\pi/(\beta(s) + 1)$ is valid whenever $r_e << r_M$. In the general gain equation, the calculated values of the last terms in *a* and *b* of the pole-pair are small and could be omitted without much error. Those terms become zero (or near-zero in *b*) for zero r_e. For the hf model, not only is $Z_\pi = 0\,\Omega$, R_B/β_0 and $\beta_0 \cdot C_B$ are also set to zero (shorted) so that, in effect, $\alpha_0 = 1$.

When all is said and done for the CB sub-stage of the cascode, we can expect a 2 % dipole rise in magnitude between 28.2 and 29.54 kHz and a bandwidth influenced by the uncancelled dominant pole of A_{v2b} and the added pole of C_{c4};

$$1/\sqrt{1/(7.34 \text{ MHz})^2 + 1/(5.39 \text{ MHz})^2} = 4.34 \text{ MHz}$$

The assumption in use of the OCTC bandwidth formula is that the poles are real or almost so, having a low pole angle ($\zeta \approx 1$).

The simpler OCTC bandwidth neglects the *a* coefficients and calculates f_{bw} from the OCTCs of *b*, which are dominated by τ_B. The τ_B pole is cancelled by the zero, leaving τ_{c4}, τ_{e4}, and τ_{E4};

$$\tau_{bw2} = \sqrt{(123.3 \text{ ns})^2 + (90.85 \text{ ns})^2 + (0.66 \text{ ns})^2} = 153.2 \text{ ns}$$
$$\Rightarrow 1.04 \text{ MHz}$$

The OCTC bandwidth is over 4 times lower than the quadratic estimate and the measured bandwidth.

Dynamic Design: Second Stage CE

The emitter peaking of Q3 can be designed to compensate for τ_{I2}, $\alpha_3(s)$, or both. To compensate for $\alpha(s)$ (so that the emitter impedance appears capacitive at the base) requires an emitter capacitance having a value of $C_{E3} = \tau_{TNPN}/R_{E3} = 0.26$ pF. This is small enough that the parasitic capacitance of a through-hole, quarter-watt R_{E3} can provide it and no discrete C_E component is needed. The time constant it

compensates is $\alpha_{NPN} \cdot \tau_{TNPN} \approx 567$ ps. C_E only compensates for $\alpha(s)$ and leaves τ_{L2} uncompensated.

The lf transfer function is

$$A_{v2a} = -\alpha_{NPN} \cdot \frac{R_0}{r_{M3}} \cdot \frac{s \cdot (R_{E3} + R) \cdot C + 1}{(s \cdot [r_{E3} \parallel R_{E3} + R] \cdot C + 1) \cdot (s \cdot \tau_{L2} + 1)}$$

where

$$r_{E3} = r_{e3} + R_{L1}'/(\beta_{NPN} + 1) = 13\ \Omega + 2551\ \Omega/391 = 19.5\ \Omega$$

and

$$r_{M3} = r_{E3} + R_{E3} = 19.5\,\Omega + 2150\,\Omega = 2170\,\Omega$$

To compensate for $\tau_{L2} = 225$ ns with a series RC shunting R_{E3} (and let the base impedance be complex), set $\tau_z = \tau_{L2}$;

$$(R_{E3} + R) \cdot C = \tau_{L2} = 225 \text{ ns}$$

The remaining pole is as widely separated from the compensated pole as possible. To achieve this, R must not be made too large yet it must be large enough to dominate $r_{E3} \| R_{E3} = 19.3\ \Omega$. Hence,

$$r_{E3} \parallel R_{E3} << R << R_{E3}$$

R is maximally separated from both at the geometric mean;

$$R = \sqrt{(r_{E3} \parallel R_{E3}) \cdot R_{E3}} = \sqrt{(19.3\,\Omega) \cdot (2.15\,\text{k}\Omega)} = 204\,\Omega \rightarrow 200\,\Omega$$

Given the value of R, C follows immediately;

$$C = \frac{225 \text{ ns}}{R_{E3} + R} = \frac{225 \text{ ns}}{2150\,\Omega + 200\,\Omega} = 95.7 \text{ pF}$$

To adjust the response, C is adjusted using a variable capacitor with a range of 5 to 20 pF, centered at 12.5 pF. A 5 % value of 82 pF is added in parallel with it for a range of 82 to 102 pF.

The remaining pole is at

$$(r_{E3} \parallel R_{E3} + R) \cdot C = (19.3\,\Omega + 200\,\Omega) \cdot (95.7 \text{ pF}) = 21 \text{ ns} \Rightarrow 7.58 \text{ MHz}$$

This pole is above the specified bandwidth and is close to the dominant Q4 CB pole of 7.34 MHz. If this were a feedback amplifier, close poles like this are an alert for loop instability.

The third and best emitter peaking design compensates for both $\alpha(s)$ and τ_L while also specifying ζ for the resulting pole-pair. For MFA amplifier response, what must the pole angles be for each

stage? This is a nontrivial question, answered with engineering detail (and tables of pole angles per stage) in the book *Wideband Amplifiers* by Peter Starič and Erik Margan (Springer, p. 4.35). As cascaded MFA stages are increased, the frequency response remains flat and the bandwidth remains constant. This is desirable for frequency response when amplifying steady-state sine-waves. However, step overshoot and risetime increase with increasing stages - rather undesirable for time-domain response. Two stages, each with MFA response, will have about 4 % overshoot, and this is acceptable for many applications, including amplifier 3.

If stages have repeated pole-pairs, we found before that the bandwidth decreases with *n*. If the poles are instead *staggered*, bandwidth can be maintained and rolloff slope steepens. DeMoivre's formula from trigonometry supplies the *n*th complex root of a number and is the conceptual basis for staggering the *n* poles. For MFA amplifier response, the poles of each stage are regularly spaced around a LHP semicircle in the *s*-plane. For a MFA pole-pair, $n = 2$ and $\phi = 45°$. For a pole-pair and single real pole, the angles are 0°, ±60°. For two pole-pairs, $n = 4$ and $\phi = 22.5°$ and 67.5°. For the design of amplifier 3, we will use suboptimal repeated poles and make both stages MFA, with pole-pairs having $\phi = 45°$ and $\zeta = \sqrt{2}/2$.

Applying the emitter-peaking formulas to Q3 with $\zeta = \sqrt{2}/2$ and $R_{L1}' = R_{B3} = 2551\ \Omega$,

$$\tau_R = R \cdot C = 2 \cdot \zeta \cdot \sqrt{\alpha_{03} \cdot \tau_{T3} \cdot \frac{R_{B3}}{r_{M3}} \cdot \tau_L} - \left(\alpha_{03} \cdot \tau_{T3} \cdot \frac{R_{B3}}{r_{M3}} + \tau_L \cdot \frac{r_{E3}}{r_{M3}} \right)$$

$$= \sqrt{2} \cdot \sqrt{(567\ \text{ps}) \cdot \frac{2551\ \Omega}{2170\ \Omega} \cdot (225\ \text{ns})} - \left((567\ \text{ps}) \cdot \frac{2551\ \Omega}{2170\ \Omega} + (225\ \text{ns}) \cdot \frac{19.5\ \Omega}{2170\ \Omega} \right)$$

$$= 17.31\ \text{ns} - (666\ \text{ps} + 2.02\ \text{ns}) = 14.63\ \text{ns}$$

$$\Rightarrow 10.87\ \text{MHz}$$

$$C_E = \frac{\alpha_0 \cdot \tau_T \cdot \tau_L}{R_E \cdot \tau_R} = \frac{(567\ \text{ps}) \cdot (225\ \text{ns})}{(2150\,\Omega) \cdot (14.63\ \text{ns})} = 4.06\ \text{pF} \rightarrow 3.9\ \text{pF}$$

$$C = \frac{\alpha_0 \cdot \tau_T + \tau_L - \tau_R}{R_E} - C_E = \frac{567\ \text{ps} + 225\ \text{ns} - 14.63\ \text{ns}}{2150\,\Omega} - 4.06\ \text{pF} = 94.1\ \text{pF}$$

$$R = \frac{\tau_R}{C} = \frac{14.63\ \text{ns}}{94.1\ \text{pF}} = 156\ \Omega \rightarrow 160\ \Omega$$

C is 82 pF in parallel with a 5 pF to 20 pF variable capacitor. After poles and zeros cancel, emitter peaking leaves one pole at τ_R.

The C_{c3} OCTC results in one more stage-2 pole:

$$\tau_{c3} = \tau_{cc3} + \tau_{cb3} = C_{c3}\cdot[(R_{L1}' + r_{b3}') + (r_{E4}\|R_0)\cdot(1 + K_{v3})]$$

$$= (0.5\text{ pF})\cdot[2551\ \Omega + 46.4\ \Omega\cdot(1)] = 1.3\text{ ns} \Rightarrow 123\text{ MHz}$$

where the base-to-collector voltage gain is

$$K_{v3} \approx \alpha_{NPN}\cdot\frac{r_{E4}\,\|\,R_0}{r_{M3}} = (0.997)\frac{46.4\,\Omega}{2170\,\Omega} = 0.021 \approx 0$$

The Q3 collector impedance is low - about $r_{E4}\|R_0$ - and the Miller effect of the CE stage is negligible.

The effects of C_{e3} have been accounted for in the emitter-peaking design. Of itself,

$$\tau_{e3} = (\alpha_{NPN}\cdot\tau_{TNPN})\cdot\frac{2601\,\Omega + 2150\,\Omega}{2170\,\Omega} = 1.24\text{ ns} \Rightarrow 128\text{ MHz}$$

Both of the Q3 poles are over 1.5 decades above the design bandwidth and have negligible effect on amplifier dynamic response.

Dynamic Design: First Stage

The stage-1 gain C_{c1} pole has time constant

$$\tau_{c1} = [R_{L1}' + R_b\cdot(1 + K_v)]\cdot C_{c1}$$
$$= [2551\,\Omega + (75\,\Omega)\cdot(5)]\cdot(0.5\text{ pF}) = 1.46\text{ ns} \Rightarrow 108.8\text{ MHz}$$

This pole is over a decade above the second-stage bandwidth and has little effect on amplifier bandwidth. It is high in frequency because R_{B1} is small.

The pole of C_{e1} has a time constant of

$$\tau_{e1} = (\alpha_{NPN}\cdot\tau_{TNPN})\cdot\frac{R_{B1} + R_E}{r_{M1}} = (567\text{ ps})\cdot\frac{75\,\Omega + 625\,\Omega}{638.2\,\Omega} = 621\text{ ps}$$

$$\Rightarrow 256\text{ MHz}$$

This pole also has little influence relative to the second-stage pole.

C_{e1} also entangles us in hf considerations with the addition of C_{E1}. For $\alpha(s)$ compensation with a capacitive base impedance,

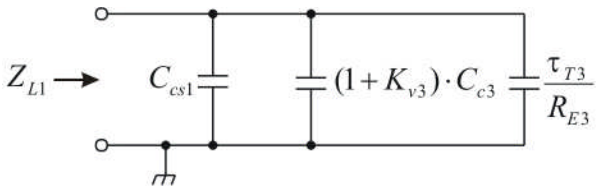
$$C_{E1} = \tau_{TNPN}/R_E = (475\text{ ps})/(625\ \Omega) = 0.76\text{ pF}$$

R_{E1} and R_{E2} each have about 1/3 pF to ground. An additional 0.1 pF is required but is small enough that we will omit it. Q1 is practically the ideal (fast, quasistatic) BJT; its poles are high in frequency.

Now that stage-1 resistor values have been determined, we turn attention to the question of shunt inductive peaking by Q2, based on R_{B2}'. For the 3 NPN BJTs, all biased at 2 mA, the hf region is

$$718\text{ kHz} = f_{\beta NPN} < f_{hf} < f_{TNPN} = 280\text{ MHz}$$

The Q1 collector node capacitance, $Z_{L1} = 1/s{\cdot}C_{L1}$, is shown below.



Because the BJTs are integrated, the collector-substrate capacitance of the NPNs is included: $C_{cs} = 2$ pF at 5 V and $V_{CS1} \approx 5$ V. Q3 is compensated by emitter peaking to have an input impedance that is capacitive;

$$C_{be3} = \tau_{T3}/R_{E3} \approx 568\text{ ps}/2.15\text{ k}\Omega = 0.26\text{ pF}$$

The capacitive loading of the Q3 base is

$$C_{b3} = (1 + K_{v3}){\cdot}C_{c3} + \tau_{T3}/R_{E3}$$

We have already determined that $K_{v3} \approx 0$. At the Q1 collector,

$$C_{L1} = C_{cs1} + C_{b3} = C_{cs1} + [(1 + K_{v3}){\cdot}C_{c3} + \tau_{T3}/R_{E3}]$$

$$= 2\text{ pF} + [(1){\cdot}(0.5\text{ pF}) + 0.26\text{ pF}] = 2.8\text{ pF}$$

In parallel with C_{L1} through the Q2 branch is the gyrated inductance, $L_e = \tau_T{\cdot}R_{B2}$'. The base of Q2 also has C_{c2} in parallel with R_{B2}'. C_{c2} refers (by $\beta_{hf} + 1$) to the emitter as C_{c2} shunting a gyrated resistance of $\tau_T/C_{c2} = 1136\ \Omega$. C_{c2} is on the top side of R_{L1} and in parallel with L_e. This modifies (and complicates) the shunt-peaking circuit in that it resonates undamped with L_e yet its value, though small (0.5 pF), contributes to resonance determination and can be combined with C_{L1}. Adding C_{c2}, the total

$$C_{L1} = 3.3\text{ pF}$$

The uncompensated Q1 collector pole thus has a time constant of

$$\tau_{L1} = R_{L1}'C_{L1} = (2551\,\Omega)\cdot(3.3\,\text{pF}) = 8.42\ \text{ns} \ \Rightarrow 18.9\ \text{MHz}$$

In view of the second-stage poles of 10.87 MHz and 4.34 MHz, extension of the first-stage bandwidth might seem pointless. The second-stage poles combine for a bandwidth of

$$f_{bw2} = \frac{1}{\sqrt{1/(10.87\ \text{MHz})^2 + 1/(4.34\,\text{MHz})^2}} = 4.03\ \text{MHz}$$

An additional pole at 18.9 MHz decreases amplifier bandwidth by about 2.2 % to 3.94 MHz. In a modified design, the second stage bandwidth might be extended (with shunt peaking, a CC output stage, or both) and first-stage peaking would then be of greater value.

A different use for C_{E1} is for emitter peaking in compensation of the Q1 collector pole. To compensate τ_{L1} in the emitter, set

$$C_{E1} = \frac{\tau_{L1}}{R_E} = \frac{7.156\ \text{ns}}{625\,\Omega} = 11.5\,\text{pF} \rightarrow 12\ \text{pF}$$

C_{E1} is not needed to satisfy the bandwidth specification, and cost and space are reduced by omitting it. If bandwidth for a different specification for this amplifier were insufficient, it is an option.

MFA response is used to find the value of inductance for Q2:

$$L_e = \left(\frac{R_{L1}'}{2\cdot\zeta}\right)^2 \cdot C_{L1} = \left(\frac{2551\,\Omega}{2\cdot\sqrt{2}/2}\right)^2 \cdot (3.3\,\text{pF}) = 10.74\,\mu\text{H}$$

Then Q2 R_{B2}' has to be 10.74 μH/(568 ps) = 18.90 kΩ, a resistance that would add 48.4 Ω to R_{L1}' and cause it (and stage-1 gain) to be slightly β-dependent. Some margin (toward flat response) on pole angle is effected by rounding R_{B2}' downward to correspond to R_{B2}' = 18776 Ω. Then L_e = 10.64 μH and the pole angle is 44.7°. The base divider resistors are R_1 = 20.0 kΩ, 1 % and R_2 = 294 kΩ, 1 %.

Given L_e, the resonant frequency is

$$f_n = 1/2\cdot\pi\cdot\sqrt{L_e\cdot C_{L1}} = 26.86\,\text{MHz}$$

with Z_n = 1796 Ω. Bandwidth of the collector pole (τ_{L1}) is extended to

$$1.27\cdot f_n = 1.27\cdot(26.86\ \text{MHz}) = 34.1\ \text{MHz}$$

or by $1.8\cdot f_p$ of the uncompensated pole; 1.8·(18.8 MHz) = 33.8 MHz. This is the dominant first-stage pole. The bandwidth of the first stage is thus

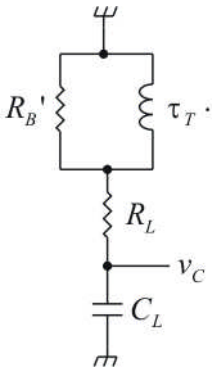
$$f_{bw1} = 1/\sqrt{(1/108.8\,\text{MHz})^2 + (1/256\,\text{MHz})^2 + (1/33.8\,\text{MHz})^2} = 32.1\,\text{MHz}$$

For inductive peaking of a stage to extend amplifier bandwidth, the extended bandwidth must be close to the bandwidth of the other stages. If it is too much higher, it has negligible effect; the lower-bandwidth stages dominate. In this case, first-stage bandwidth extension is truly marginal in benefit though useful as an amplifier design template. With a discrete BJT design, or at least with discrete PNP BJTs, the CB-stage speed can be vastly increased and first-stage peaking becomes significant.

The equivalent circuit for shunt peaking with L_e is as shown below. The collector node impedance is

$$Z_{L1} = [(s \cdot \tau_{T2} \cdot R_{B2}') \parallel R_{B2}' + R_{L1}] \parallel (1/s \cdot C_{L1})$$

where $R_{B2}' = R_{B2} + r_b'$. Thus, $R_{B2}' = 18726\ \Omega + 50\ \Omega = 18776\ \Omega$.



After some algebra, this becomes

$$Z_L = R_L \cdot \frac{s \cdot \tau_T \cdot \left(\frac{R_B' + R_L}{R_L}\right) + 1}{s^2 \cdot \tau_T \cdot (R_B' + R_L) \cdot C_L + s \cdot (\tau_T + R_L \cdot C_L) + 1}$$

This impedance has the form of shunt peaking, with a zero and a quadratic pole-pair. R_L is determined by choice of quasistatic gain. We have design control over the inductor value of $\tau_T \cdot R_B'$ by adjusting R_B. The damping,

$$\zeta = \frac{b}{2 \cdot \sqrt{a}} = \frac{1}{2} \cdot \frac{\tau_T + R_L \cdot C_L}{\sqrt{\tau_T \cdot (R_B' + R_L) \cdot C_L}}$$

ζ can be set to the design value and solved for R_B;

$$R_B = \frac{(\tau_T + R_L \cdot C_L)^2}{(2 \cdot \zeta)^2 \cdot \tau_T \cdot C_L} - R_L - r_b'$$

Plugging in the values with $\zeta = \sqrt{2}/2$,

$$R_{B2} = \frac{[568\text{ ps} + (2490\,\Omega) \cdot (3.3\text{ pF})]^2}{2 \cdot (568\text{ ps}) \cdot (3.3\text{ pF})} - 2490\,\Omega - 50\,\Omega = 18.05\text{ k}\Omega$$

and $L_e = \tau_T \cdot R_B' = 10.25\ \mu\text{H}$. R_B' adds damping by shunting L_e. Then L_e must be increased for desired pole angle and this requires an increase in R_B'. One additional refinement step in adjusting R_{B2} is thus possible, though the existing design will be left with the value of R_{B2} that damps slightly more. We really do not need the extra bandwidth.

The amplifier bandwidth from the stage bandwidths is

$$f_{bw} = 1/\sqrt{(1/f_{bw1})^2 + (1/f_{bw2})^2} = 1/\sqrt{(1/32.1\text{ MHz})^2 + (1/4.03\text{ MHz})^2}$$
$$= 4.0\text{ MHz}$$

The 3.5 MHz bandwidth design goal has been achieved with a margin of 13 %. The looser tolerance on dynamics calculations (which are dependent on rather elusive parasitic capacitances) suggests that bandwidth margins be greater than those for static or quasistatic parameters. A 20 % margin is not ill-advised for dynamics design of circuits with two stages of gain.

To make bandwidth measurements of the prototype circuit, C_{E2} was first adjusted for a flat step response on the 'scope, with no overshoot. To achieve overshoot in the adjustment range of 5 - 18 pF, 33 pF had to be added to C. Then both risetime and bandwidth measurements were made.

Amplifier risetime with a ±50 mV step input having a risetime of about 11.5 ns was 85 ns. For a dominant-pole amplifier, this corresponds to a bandwidth of

$$f_{bw} = \frac{0.35}{t_r} = \frac{0.35}{85\text{ ns}} = 4.1\text{ MHz}$$

When adjusted to take into account the input risetime,

$$t_r = \sqrt{t_{ro}^2 - t_{ri}^2} = 84.2\text{ ns} \Rightarrow 4.16\text{ MHz}$$

In the frequency domain, slewing was observed to appear at 0.85 MHz on a TDS220 (100 MHz) 'scope and at 1.0 MHz on a

TDS360 (200 MHz) 'scope, in good agreement with the f_{BW} calculation. Because of the large-signal limitation of this amplifier, bandwidth was measured for several sine-wave amplitudes, as given in the following table.

v_o, V pk	f_{bw}, MHz	
	TDS360	TDS220
0.4	4.0	4.0
2.0	3.7	3.6
4.0	3.0	2.9
8.0	1.9	1.85

As expected, the 100 MHz TDS220 has somewhat lower amplitude readings because its own bandwidth slightly reduces them. For small-signal gain, the bandwidth measurements agree well with the calculated values and diverge from them as amplitude is increased.

The important lesson in this is that the assumption underlying circuit-dynamics modeling is that of incremental (small-signal) models. As waveforms increase in amplitude, these models begin to fail so that at 4 V of amplitude, the calculated incremental bandwidth of amplifier 3 is in error by –25%. And at the rated full-scale output of 8 V amplitude, the bandwidth measures only half the incremental value. As large-signal effects become significant, the assumptions underlying both incremental circuit theory and measurement methods begin to lose validity.

In retrospect, the choice of gain assignments for the two cascaded stages resulted in a much faster first stage. To increase overall bandwidth, the first-stage gain could be increased somewhat and the second-stage gain decreased so that the overall gain remains the same while the stage bandwidths become closer in value. The two stages are not well-matched in bandwidth for dynamic design optimization, though if the output pole were decreased substantially by adding a CC buffer stage after it, then the amplifier would be much closer to optimum (and much faster). As a design template, amplifier 3 is included for applications not requiring the static precision of a feedback amplifier.

Intrinsic Noise

What other factors can affect circuit behavior? There are two physical processes that are known to create electrical noise. *Thermal*

noise (or Johnson noise) arises from the random paths taken by charge in resistance. The larger the resistance, the more thermal noise. The number to remember for it is

$$\text{Thermal noise: } 129\ \text{pV}/\sqrt{\text{Hz}\cdot\Omega}\ \text{, } 300\ \text{K}$$

For amplifier 3, the largest resistance in the waveform path is R_{L2} = 15 kΩ. With a noise bandwidth of about 6.3 MHz, then the rms thermal noise generated by that resistor is

$$(129\,\text{pV}/\sqrt{\text{Hz}\cdot\Omega})\cdot\sqrt{(6.3\,\text{MHz})\cdot(15\,\text{k}\Omega)} = 39.6\,\mu\text{V}$$

The second kind of self-generated (or *intrinsic*) circuit noise, *shot noise*, is generated in circuits from the random fluctuation of charge distribution in electric current. The number to remember for it is

$$556\ \text{pA}/\sqrt{\text{Hz}\cdot\text{A}}\ \text{, } 300\ \text{K}$$

The NPN BJTs conducting 2 mA over a 6.3 MHz noise bandwidth generate an rms value of

$$(556\ \text{pA}/\sqrt{\text{Hz}\cdot\text{A}})\cdot\sqrt{(6.3\,\text{MHz})\cdot(2\ \text{mA})} = 62.3\ \text{nA}$$

This current develops across the largest resistor in the waveform path, R_{L1}', a voltage of about (62.3 nA)·(2.55 kΩ) = 159 μV. Adding rms contributions of thermal and shot noise from the two largest sources in the circuit, the total is about

$$\sqrt{(39.6\ \mu\text{V})^2 + (159\ \mu\text{V})^2} = 164\,\mu\text{V rms}$$

This sets a *noise floor* below which input amplitudes of waveforms become dominated by noise. If 164 μV is a least amplitude and the full-scale value is 8 V, then the useful dynamic range of the amplifier is no more than

$$\frac{8\ \text{V}}{164\ \mu\text{V}} = 48780 = 4.69\ \text{decades} = 93.8\ \text{dB}$$

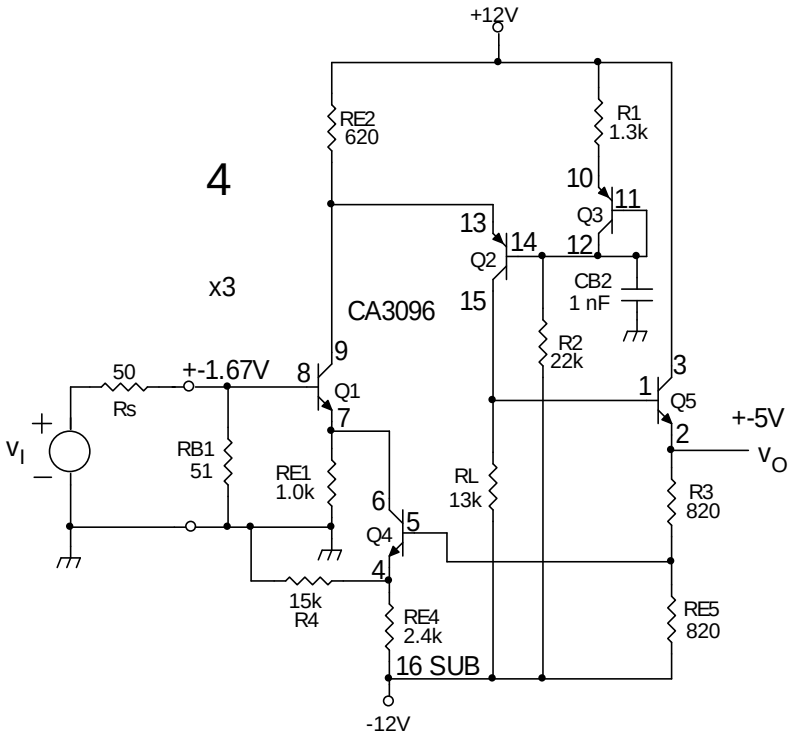
(The dB scaling for voltage or current ratios is $20\cdot\log_{10}(\text{ratio})$.)

The *noise bandwidth* used in the above formulas for single-pole rolloff is not f_{bw} as we have been calculating it, but must also include the noise above f_{bw}, though it is attenuated by the frequency rolloff. The total noise for single-pole rolloff is that of an effective frequency bandwidth of $(\pi/2)\cdot f_{bw} \approx 1.57\cdot f_{bw}$. For amplifier 3 noise calculations,

the noise bandwidth is 4 MHz·(1.57) or about 6.3 MHz. For *narrowband* amplifiers (those without static gain and with a zero at the origin), the bandwidth is $(\pi/2)\cdot\Delta f$, where Δf is the amplified band with single-pole rolloff at both its high and low frequencies.

Noise voltage and current values are too small to be significant in many applications with large waveforms, yet they need to be considered for amplifiers with wide dynamic ranges. They are reduced by reducing resistance or current, or else temperature with artificial cooling. The given noise numbers are for 300 K.

Inverting Cascode Feedback Amplifier



Complementary cascode stages are often used in fast amplifiers. The biasing allows for 0 V in and out. For more precision than amplifier 3, amplifier 4 uses feedback. In feedback amplifier 4, Q1 is the input CE stage of the cascode and Q2 is the CB second stage. They share current from R_{E2}. Q5 is a buffer, the CC output third stage. In the biasing divider R_1, R_2, Q3 statically compensates for the Q2 *b-e*

junction, to stabilize the voltage across R_{E2} and keep the current through it constant by keeping its voltage drop constant. It is the same CB biasing scheme as in amplifier 3.

The amplifier 4 design specifications (or "specs" for short) are

- $A_v = 3$, ±3 %
- $f_{bw} \geq 15$ MHz, MFA response, $C_o = 15$ pF
- 0 V out, ±5 V output range, $r_o \leq 10\,\Omega$
- 0 V in, ±1.67 V input range, $r_i = 50\,\Omega$
- ±12 V power supplies

The feedback path (H) of this amplifier is an active path, a CE stage (Q4) with gain. We have been accustomed to H as a voltage divider with a gain of less than one. For the closed-loop gain to exceed one, the feedback gain must be less than one. Open-loop stages with low gain can be made to be quite stable and fast.

The feedback stage, Q4, drives the emitter of Q1 with a waveform in phase with v_i (plus loop delay). As v_i increases, Q1 current increases, diverting R_{E2} current from Q2. With less current in Q2, the Q5 base voltage decreases as does v_o. Q4 input voltage from v_o decreases, and Q4 inverts it, causing the voltage at the Q1 emitter to increase and Q1 current to decrease, opposing the effect of v_i. Feedback is thus negative and the loop can be stable.

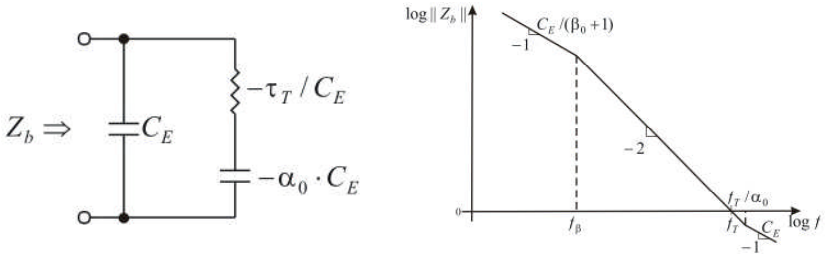
The cascode stage is not much different from amplifier 3 so that if a ×3.75 increase in bandwidth (over $f_{bw} = 4$ MHz of amplifier 3) is to be achieved, then feedback is needed. With one stage of voltage gain, if we try to make $G_0 \approx -10$, then for a closed-loop voltage gain of –3,

$$A_{v0} = -3 = \frac{G_0}{1 + G_0 \cdot H_0} = \frac{-10}{1 + (-10) \cdot H_0} \Rightarrow H_0 = -0.233$$

The loop gain is $G_0 \cdot H_0 = 2.33$ and the feedback factor is 3.33. The bandwidth can thus be increased by about 3.33 times that of the open-loop bandwidth of G, and must be $f_{bw}(G) \geq 15$ MHz/3.33 = 4.5 MHz. The cascode stage of amplifier 3 has a gain of 5 and a 4 MHz bandwidth. Doubling its gain (assuming a dominant pole) halves the bandwidth to 2 MHz. With feedback, the closed-loop bandwidth is estimated as (2 MHz)·(3.33) = 6.67 MHz.

However, the Q5 output stage of amplifier 4 buffers the cascode output node from C_o and increases its bandwidth. The reduction of $C_{E5} = C_o$ can be estimated based on hf BJT theory. The total-frequency equivalent circuit and reactance plot of C_E at the base are

shown below (from the "Transistor Dynamics" chapter). In the hf region, the asymptotic approximation of the reactance plot has a slope of –2 and a value of $\tau_T \cdot C_E$. For $f \geq f_T/\alpha_0$, the base capacitance from the emitter is C_E. In the lf region, it is $C_E/(\beta_0 + 1)$. Starting from f_T/α_0 and decreasing in frequency, the effective value of C_E referred to the base decreases linearly from C_E to $C_E/(\beta_0 + 1)$ at f_β.



For $f_T/\alpha_0 \approx 300$ MHz and $C_E = 15$ pF, at 15 MHz the base-referred emitter capacitance is

$$\frac{C_E}{\left(\frac{300\,\text{MHz}}{15\,\text{MHz}}\right)} = \frac{15\,\text{pF}}{20} = 0.75\,\text{pF}$$

The CC output stage is effective by a ×20 factor in buffering the Q2 collector node from C_o.

A second dynamic design choice that affects the static design is the BJT currents. From the CA3096 data of amplifier 3, the PNP BJT current is chosen to be 1 mA and Q1 NPN current also at 1 mA. For PNPs, 1 mA is the maximum current before β_0 and f_T decrease appreciably. The maximum under β_0 and f_T constraints is desired to maximize slew rate and f_{BW} at the cascode output node. The Q1 current has been halved from amplifier 3 design to decrease power dissipation (which also decreases thermal distortion) and yet maintain a high $f_T = 280$ MHz ($\tau_T = 568$ ps). At 1 mA, NPN $V_{BE} = 0.69$ V and typical $\beta_0 = 390$; PNP $V_{BE} = 0.66$ V, and $\beta_0 = 47$. To drive C_o at its specified value, I_{C5} must be greater. From the slew-rate formula, solving for I,

$$I = 2 \cdot \pi \cdot f_{BW} \cdot V_{pk} \cdot C = (2 \cdot \pi) \cdot (15\ \text{MHz}) \cdot (5\ \text{V}) \cdot (15\ \text{pF}) = 7.1\ \text{mA}$$

When I_{E5} is set to be somewhat greater than 7.1 mA, β_0 and f_T are comparable to what they are at 2 mA. Both peak at around 4 mA.

Static and Quasistatic Design

Operating range is an early consideration affecting the static design. (From amplifier 3, we found that it also affects the dynamic performance; range and bandwidth are conflicting design criteria, though at this early stage in the design, we do not have quantitative knowledge of the tradeoff.) Given output range and gain, then the voltage range of the Q1 base is ±5 V/3 = ±1.67 V. This requires that the Q1 emitter, Q4 collector node be able to linearly operate as low as –2.36 V. The base of Q4 must be at less than this voltage. C_{c4} will increase appreciably at the high end of the output range for low V_{BC}. The Q5 emitter resistance, $R_3 + R_{E5}$, has an upper limit set by the slew-rate condition. By choosing

$$R_3 = R_{E5} = 820\ \Omega$$

then the range of $V_{B4} \in$ –6 V ± 2.5 V with a maximum of –3.5 V. This leaves enough margin for V_{BC4}, both to prevent saturation of Q4 and to keep C_{c4} down in value. This choice of R_3, R_{E5} results in a quiescent Q5 emitter current of 12 V/1640 kΩ = 7.32 mA > 7.1 mA. At this current, $V_{BE5} = 0.744$ V.

The Q5 emitter resistors form a divider in the feedback path. The divider attenuation to the base of Q4 is

$$H_1 = \frac{v_{B4}}{v_O} = \frac{820\ \Omega}{1640\ \Omega} = 0.5$$

assuming, and including in H_1, the loading of r_{B4}, that it is negligible ($r_{B4} >> r_{Ho} = 410\ \Omega$). For $A_v = –3$, as previously derived, $H_0 = –0.233$. The gain of Q4 must then be

$$A_{v4} = \frac{H_0}{H_1} = \frac{-0.2333}{0.5} = -0.466$$

$I_{E1} \approx 1$ mA is given and thus $V_{BE1} \approx 0.69$ V. We can freely choose one of the element values in the error circuit; choose

$$R_{E1} = 1.0\ \text{k}\Omega$$

and we will see where it leads.

Two circuit equations of the error-summing part of the circuit involving the Q1 emitter and Q4, will help us determine the other parameter values. First, KCL is applied at the Q1 emitter:

$$\frac{V_{BE1}}{R_{E1}} + I_{E1} = \alpha_{04} \cdot I_{E4} \text{ where } I_{E4} = \frac{V_{RE4}}{R_{E4}}$$

From this equation,

$$I_{E4} = \frac{I_{E1} + V_{BE1}/R_{E1}}{\alpha_{04}} = \frac{1\text{mA} + (0.69\text{ V}/1.00\text{ k}\Omega)}{0.9974} = 1.69\text{ mA}$$

Then V_{BE4} = 0.69 V + (60 mV)·log(1.69 mA/1 ma) = 0.704 V.

The Q4 collector current source in parallel with R_{E1} forms a Norton circuit and can be converted to a static Thevenin equivalent with a resistor of R_{E1} = 1.0 kΩ in series with a source voltage of

$$(-1.69\text{ mA})\cdot(1.0\text{ k}\Omega) = -1.69\text{ V}$$

R_E of Q4 consists of R_4 and R_{E4}. They form a Thevenin equivalent source that makes it possible to satisfy both the static and gain requirements of the Q4 stage. The Q4 emitter port resistance is

$$r_{E4} = r_{e4} + R_{B4}/(\beta_0 + 1) = \frac{26\text{ mV}}{1.69\text{ mA}} + \frac{410\ \Omega}{391} = 15.4\,\Omega + 1\Omega = 16.4\,\Omega$$

Returning to static calculations, the V_{B4} range is

$$V_{B4} \in [-3.5\text{ V}, -6\text{ V}, -8.5\text{ V}]$$

$$V_{E4} \in [-4.204\text{ V}, -6.704\text{ V}, -9.204\text{ V}]$$

The voltage across R_{E4} has the range,

$$V_{RE4} \in [7.796\text{ V}, 5.296\text{ V}, 2.796\text{ V}]$$

The values correspond to [–fs, midrange (0 V), +fs] input values.

The second of the two circuit equations of the error-summing part of the circuit involving the Q1 emitter and Q4 is the Q4 voltage gain,

$$A_{v4} = -\alpha_{04} \cdot \frac{R_{E1}}{R_{E4} \parallel R_4 + r_{E4}}$$

At the Q4 collector node, R_{E1} is the load resistor for Q4, and

$$r_{E4} = r_{e4} + (R_3 \| R_{E5})/(\beta_{NPN} + 1) = 15.4\ \Omega + 1\ \Omega = 16.4\ \Omega$$

Substituting into A_{v4} and solving,

$$R_{E4}' = R_{E4} \parallel R_4 = \frac{\alpha_{04} \cdot R_{E1}}{-A_{v4}} - r_{E4} = \frac{(0.9974)\cdot(1.0\,\text{k}\Omega)}{0.4667} - 16.4\,\Omega = 2121\,\Omega$$

The Thevenin equivalent voltage with I_{E4} = 1.69 mA is

$$V_{EE4}' = I_{E4}\cdot R_{E4}' - V_{E4} = (1.69\text{ mA})\cdot(2121\ \Omega) - (-6.704\text{ V}) = 10.23\text{ V}$$

Solving for the resistor values,

$$\frac{R_4}{R_4 + R_{E4}} = \frac{R_{E4}'}{R_{E4}} = \frac{10.23\text{ V}}{12\text{ V}} = 0.8573$$

$$\Rightarrow R_{E4} = \frac{2121\,\Omega}{0.8573} = 2474\,\Omega \rightarrow 2.4\text{ k}\Omega$$

$$R_4 = 14.866\text{ k}\Omega \rightarrow 15\text{ k}\Omega$$

From these 5 % resistor values the more accurate Thevenin values are calculated:

$$R_{E4}' = (2.4\text{ k}\Omega)\|(15\text{ k}\Omega) = 2069\ \Omega$$

$$V_{EE4}' = -10.345\text{ V}$$

Refining Q4 parameters further,

$$V_{RE4} \in [6.14\text{ V}, 3.64\text{ V}, 1.14\text{ V}]$$

The range of Q4 emitter current is then

$$I_{E4} \in [2.97\text{ mA}, 1.76\text{ mA}, 0.551\text{ mA}]$$

At the –fs end of the range, f_T of the NPN BJT decreases to about 120 MHz and β_0 to 200. The dynamic behavior often changes significantly over the dynamic range.

The fed-back Thevenin open-circuit voltage, v_{Ho}, of $v_{C4} = v_{E1}$ is the output of Q4 and the *H* block, and is the input to the error voltage loop which is completed by r_{E1} in series with v_i. The choice of error voltage is

$$v_E = v_i - v_{Ho} = v_i - i_{c4}\cdot R_{E1} = v_i - \left(\frac{i_{c4}}{v_o}\cdot R_{E1}\right)\cdot v_o$$

The refined value of quasistatic Q4 gain is

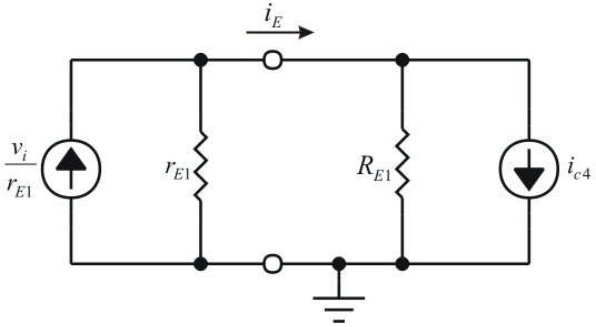
$$A_{v4} = \frac{v_{Ho}}{v_o} = \frac{v_{Ho}}{i_{c4}}\cdot\frac{i_{c4}}{v_o} = R_{E1}\cdot\frac{-\alpha_{NPN}}{r_{e4} + R_{E4} + (R_3 \parallel R_{E5})/(\beta_{NPN}+1)} \approx -0.478$$

$$H_0 = H_1\cdot A_{v4} = (0.5)\cdot(-0.478) = -0.239$$

From basic feedback theory, we know that error and fed-back quantities are often subject to our choice. An alternative choice of

feedback-loop error leaves the Q4 collector current source in parallel with R_{E1} as a Norton circuit and the input source is nortonized as shown below. The resulting error current, $i_E = i_{e1}$;

$$i_E = \frac{v_i}{r_{E1}} \cdot \left(\frac{r_{E1}}{r_{E1} + R_{E1}} \right) - i_{c4} \cdot \left(\frac{R_{E1}}{r_{E1} + R_{E1}} \right)$$



The incremental resistance looking into the emitter of Q1 is

$$r_{E1} = r_{e1} + R_{B1}'/(\beta_{NPN} + 1) = 24.5\,\Omega$$

Choosing i_E for x_E results in T_i - the coefficient of v_i in i_E - that the choice of v_E for error avoids.

From the refined values for the gain parameters, the design goal for the open-loop gain of the forward path is

$$G_0 = \frac{A_v}{1 - A_v \cdot H_0} = \frac{-3}{1 - (-3) \cdot (-0.2391)} = -10.61$$

To calculate G_0, we return to Q1. The Q1 current is now

$$I_{E1} = \frac{V_{BE1}}{R_{E1}} - \alpha_{04} \cdot I_{E4} = \frac{V_{BE1}}{1.0\,\text{k}\Omega} - 1.76\ \text{mA}$$

This nonlinear equation iterates to

$$V_{BE1} = 0.692\ \text{V}\ ;\ I_{E1} = 1.07\ \text{mA}$$

R_{B1}' = 75 Ω (with r_b' = 50 Ω) is referred to the emitter (as 0.192 Ω) and will drop about 0.20 mV. Adding this to $-V_{BE1} = -0.692$, then $V_{E1} = -0.692$ V. Before calculation of G_0 can be completed, the CB static circuit needs attention.

The complementary cascode is more complicated than a cascode in that R_{E2} supplies current for both transistors. It functions as a static

current source when the voltage across it is kept constant by the Q2 base biasing divider. Q3 is a matching PNP to Q2. To design the base bias circuit, we choose as a target $I_{E2} = I_{E3} \approx 1$ mA. Then $V_{BE3} \approx 0.66$ V. Set also as a given target $V_{B2} \approx 10$ V. From this,

$$R_1 = \frac{(12\text{ V} - 10\text{ V}) - (0.66\text{ V})}{1\text{ mA}} = 1.34\text{ k}\Omega \rightarrow 1.3\text{ k}\Omega$$

For $I_{E2} = 1$ mA, its base current is $I_{B2} = 20.8$ μA and adds to I_{E3} to flow through R_2. Its value is thus

$$R_2 = \frac{10\text{ V} + 12\text{ V}}{1.028\text{ mA}} = 21.4\text{ k}\Omega \rightarrow 22\text{ k}\Omega$$

The external thevenized Q2 base resistance is

$$R_{B2} = (R_1 + r_{E3}) \| R_2 = (1.3\text{ k}\Omega + 27\ \Omega) \| (22\text{ k}\Omega) = 1251\ \Omega$$

Adding $r_{b2}' = 50\ \Omega$,

$$R_{B2}' = 1301\ \Omega$$

Then recalculating the bias circuit with the 5 % resistor values,

$$V_{B2} = (1251\ \Omega)\cdot(20.8\ \mu\text{A}) + (24\text{ V} - 0.66\text{ V})\cdot\frac{22\text{ k}\Omega}{23.3\text{ k}\Omega} - 12\text{ V}$$
$$= 10.06\text{ V}$$

$$I_{E3} = \frac{12\text{ V} - 10.06\text{ V} - 0.66\text{ V}}{1.3\text{ k}\Omega} = \frac{1.280\text{ V}}{1.3\text{ k}\Omega} = 0.985\text{ mA}$$

The current through R_{E2}, shared by Q1 and Q2, is

$$I_{RE2} = \frac{1.28\text{ V}}{620\ \Omega} = 2.065\text{ mA} \Rightarrow I_{E2} = I_{RE2} - I_{C1} = 0.997\text{ mA}$$

I_{E2} and I_{E3} are mismatched by about 1.2 %, an amount that will not be considered significant. Calculating more static values,

$$I_{C2} = 0.976\text{ mA}\ ;\ V_{RL2} = (0.976\text{ mA})\cdot(13\text{ k}\Omega) = 12.694\text{ V}$$

The voltage drop across R_L is 12.694 V, leaving 0.694 V at the base of Q5. Its current is

$$I_{E5} = (12\text{ V})/[2\cdot(820\ \Omega)] = 7.3\text{ mA} \Rightarrow V_{BE5} = 0.744\text{ V}$$

The output thus has a –50 mV offset, an amount within V_{BE} variation among CA3096 parts. The prototype measured value was –27 mV.

The quasistatic gain of the cascode stage includes the loss of some of i_{c1} to R_{E2} which forms a current divider with the emitter branch of Q2. The gain with open-circuit output (disconnected from the Q5 base) at R_L is

$$G_1 = -\alpha_{NPN} \cdot \alpha_{PNP} \cdot \left(\frac{R_{E2}}{r_{E2} + R_{E2}} \right) \cdot \frac{R_L}{r_{M1}}$$
$$\approx -(0.997)(0.979) \cdot (0.921) \cdot (12.68) \approx -11.41$$

where

$$r_{E2} = r_{e2} + R_{B2}'/(\beta_{PNP} + 1) = 26.1\,\Omega + 1301\,\Omega / 48 = 53.2\,\Omega$$

$$r_{M2} = r_{E2} + R_{E2} = 673.2\ \Omega$$

$$r_{M1} = r_{E1} + R_{E1} = 24.5\,\Omega + 1000\,\Omega = 1025\,\Omega$$

The CC stage at the output has a voltage gain of

$$G_2 = \frac{R_{E5}'}{R_{E5}' + r_{e5} + R_L/(\beta_{NPN} + 1)} = \frac{1640\,\Omega}{1640\,\Omega + 36.9\,\Omega} = 0.978$$

where $R_{E5}' = R_3 + R_{E5} \| r_{B4} \approx R_3 + R_{E5} = 820\ \Omega + 820\ \Omega = 1640\ \Omega$. The loading of Q4 can be ignored because

$$r_{B4} = (\beta_{NPN} + 1) \cdot (r_{e4} + R_{E4}) = (391) \cdot (3615\ \Omega) = 1.414\ \text{M}\Omega$$

and $r_{B4} >> r_{Ho} = R_{E5} \| R_3$.

Combining cascaded stage gains,

$$G_0 = G_1 \cdot G_2 = (-11.41) \cdot (0.978) = -11.16$$

The quasistatic loop gain is $G_0 \cdot H_0 = 2.67$. This is certainly not an op-amp. The contribution of the Q4 stage results in a feedback factor of

$$1 + G_0 \cdot H_0 = 1 + (-11.\ 41) \cdot (-0.239) = 3.667$$

This small amount of feedback is not insignificant; r_i is 3.667 times greater and r_o is reduced similarly. Static quantities are also stabilized. A small but significant feedback factor enhances the performance of what is otherwise an open-loop amplifier.

The quasistatic closed-loop gain is now

$$A_{v0} = \frac{G_0}{1 + G_0 \cdot H_0} = \frac{-11.16}{3.667} = -3.04$$

The measured values of the amplifier 4 prototype had static voltages within a few mV of the calculated values. The gain for a full-scale (±5 V peak-to-peak) output was –2.97, in error by 2.4 % and within measurement error tolerance.

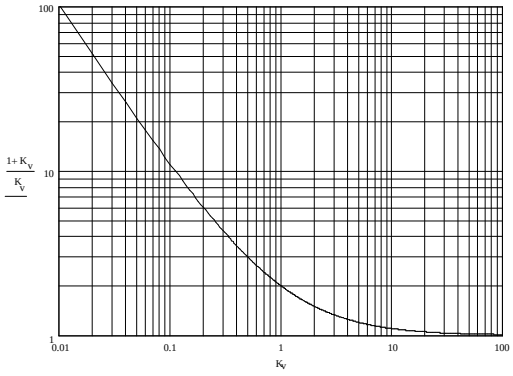
Dynamic Design

The dynamic analysis begins with Q1. The CA3096 specifications show the 1 mA, 5 V value of f_T to be about 280 MHz, or $\tau_{T1} = 568$ ps. Then the Q1 hf range is from $f_\beta = 718$ kHz to $f_T = 280$ MHz. No large reactances exist in the circuit to cause poles in the lf region (below f_β); all the dynamic activity is in the hf region or above.

The cascode generally has high bandwidth because the CE base-collector voltage gain, K_v, is reduced to reduce input Miller capacitance $(1 + K_v)\cdot C_c$. The CE stage of a typical cascode amplifier has $K_v << 1$ and an *inverse Miller effect* occurs. The output Miller transform gives Miller capacitance at the output as

$$\left(1+\frac{1}{K_v}\right)\cdot C_c$$

As K_v decreases, it increases. As $K_v \to 0$, Miller capacitance at the collector becomes infinite, a *b-c* short. The C_c Miller-effect multiplier is plotted below.



Actually, the opposite behavior occurs and bandwidth is maximized by $K_v = 0$. Circuit intuition tells us that something is wrong because

circuit speed is increased by decreasing R_L which decreases τ_L and K_v. The base-to-collector voltage gain of the CE is

$$K_v \approx \frac{R_L}{r_M}$$

This inverse Miller-effect paradox is resolved upon closer inspection of the Miller-effect collector time constant,

$$\tau_{cM} = R_L \cdot \left[\left(1 + \frac{1}{K_v} \right) \cdot C_c \right]$$

Substituting K_v into τ_{cM},

$$\tau_{cM} \approx R_L \cdot \left(1 + \frac{r_M}{R_L} \right) \cdot C_c = (R_L + r_M) \cdot C_c$$

As $R_L \to 0\ \Omega$, $\tau_{cM} \to r_M \cdot C_c$ and contributes a pole factor at

$$(s \cdot r_M \cdot C_c + 1)$$

This pole factor combines with the RHP zero factor contributed by the base-to-collector passive path, G_p, through C_c to form an all-pass filter for the case of $R_E = 0\ \Omega$. Then $r_M = r_m$ of the BJT and an all-pass filter results:

$$\frac{(-s \cdot r_m \cdot C_c + 1)}{(s \cdot r_m \cdot C_c + 1)}$$

A real pole and a real zero of the same magnitude but of positive sign combine to have no effect on the transfer function magnitude though they contribute linear phase delay which does not alter the shape of the waveform. Each pole and zero contributes delay of the same amount which has a total of $-\pi/2$ ($-90°$) at $-1/R_m \cdot C_c$. For $r_M >> r_m$, this collector-side Miller time constant is still small;

$$\tau_{cM1} \approx (R_{E2} \| r_{E2} + r_{M1}) \cdot C_{c1} \approx R_{E1} \cdot C_{c1} = 0.54 \text{ ns} \Rightarrow 296 \text{ MHz}$$

and is above f_T. It is significant that the Miller output time constant does not appear in the basic circuit derivations that result in the exact OCTC time-constant components of τ_c and τ_L. While there is a passive path, G_p, through C_c to the collector (with a RHP zero at τ_{cM}), there is no output Miller capacitance of C_c to include in τ_L or τ_c. The Miller output capacitance, given our formulation of BJT dynamic analysis,

disappears in that the effects of C_c poles (but not the RHP zero) are included in the OCTC of C_c.

In contrast to the output Miller capacitance, the input Miller effect appears in τ_c as τ_{cb}. The lf Q1 collector node resistance, $R_{E2}\|r_{E2} \approx 49\ \Omega$, and the Miller effect of C_{c1} results in an effective capacitance at the Q1 base of

$$(1 + K_{v1})\cdot C_{c1} = (1 + 49\ \Omega/1025\ \Omega)\cdot(0.5\ \text{pF}) = 0.52\ \text{pF}$$

With such a low $K_{v1} = 0.048$, the Miller effect is barely evident at the input. The resulting pole at the base is at 4 GHz - well beyond f_T - and contributes essentially nothing to the dynamic behavior of the circuit.

The CB stage (Q2), unlike amplifier 3 (where C_B was made large), is compensated at the base so that $\tau_B = \tau_\beta$:

$$C_{B2} = \frac{\tau_{\beta PNP}}{R_{B2}'} = \frac{(47)\cdot(24.5\ \text{ns})}{1301\Omega} = 885\ \text{pF} \rightarrow 1\ \text{nF}$$

Although 910 pF is a standard 5 % value, it is not often used. In contrast, decade values are the most commonly used. This is not a sufficient reason to *not* use them, though the 13 % error is well within the range of tolerance for dynamic performance. The somewhat larger value of C_{B2} will ensure that the emitter impedance rolls off within the high end of the lf region to its hf-region value. This is better than to have the rolloff occur in the hf region, where z_E will have increased somewhat (as an inductance) first. The decade value is chosen here to avoid inductance gyration from part tolerance error, and it also reduces parts inventory, thus increasing product profit.

C_{bs2} adds a negligible 2.5 pF to the Q2 base that is dominated by C_{B2} which causes Q2 emitter impedance to appear resistive. The hf approximation is quasistatic. Into the CB emitter, we see

$$z_{E2}(\text{hf}) = r_{e2} + \frac{R_{B2}'}{\beta_0 + 1} = 26.1\Omega + \frac{1301\Omega}{48} = 26.1\Omega + 27.1\Omega = 53.2\ \Omega$$

At $I_{E2} \approx 1$ mA, $\tau_{T2} = 24.5$ ns. The inductive gyration of R_{B2}' is $\tau_{T2}\cdot R_{B2}' = 31.9\ \mu\text{H}$. The external Q2 resistance at the emitter node (looking out of the emitter at the external circuit) is R_{E2} and it is shunted by collector-to-substrate capacitance $C_{E2} = C_{cs1}$.

From the gain formulas derived for amplifier 3, the CB gain with $\tau_B = \tau_\beta$ for the input loop is

$$A_{v2i}(s)\big|_{\tau_{B2}=\tau_{\beta 2}} = \alpha_{PNP}\cdot\frac{R_L'}{r_{M2}}\cdot\frac{1}{s\cdot\left[\alpha_{PNP}\cdot\tau_{TPNP}\cdot\left(\frac{R_{E2}}{r_{M2}}\right)+(r_{E2}\parallel R_{E2})\cdot C_{E2}\right]+1}$$

It has a single pole at a time constant of

$$\tau_{i2} = \alpha_{PNP}\cdot\tau_{TPNP}\cdot\left(\frac{R_{E2}}{r_{M2}}\right)+(r_{E2}\parallel R_{E2})\cdot C_{cs1}$$

$$= (0.979)\cdot(24.5\text{ ns})\cdot(0.921) + (49\ \Omega)\cdot(2\text{ pF}) = 22.1\text{ ns} + 98\text{ ps}$$

$$= 22.2\text{ ns} \Rightarrow 7.17\text{ MHz}$$

The complete gain expression includes the effects of C_{c2}, and for amplifier 3, the additional poles and zeros were found by applying the EET to result in a template formula for the CB gain:

$$A_{v2}(s) = A_{v2i}(s)\cdot\frac{-s^2\cdot[\tau_\beta\cdot(R_{B2}'/\beta_0)\cdot C_{c2}]+s\cdot[\tau_{B2}-(R_{B2}'/\beta_{PNP})\cdot C_{c2}]+1}{s^2\cdot[\tau_{B2}\cdot R_L'\cdot C_{c2}]+s\cdot[\tau_{B2}+(R_L'+R_{B2}'\cdot(1+K_{v2}))\cdot C_{c2}]+1}$$

In the pole coefficients, the time constants of interest are

$$\tau_{B2} = R_{B2}\cdot C_{B2} = (1251\ \Omega)\cdot(1\text{ nF}) = 1.25\ \mu\text{s} \Rightarrow 127\text{ kHz}$$

C_{c2} = 2 pF, and with a Q5 base load shunting R_L of about 656 kΩ,

$$R_L' = (13\text{ k}\Omega)\|(656\text{ k}\Omega) = 12.75\text{ k}\Omega$$

which leads to

$$K_{v2} = (0.979)\cdot(12.75\text{ k}\Omega/673.2\ \Omega) = 18.54$$

The time-constant components of C_{c2} are

$$\tau_{cb2} = (1+K_{v2})\cdot[R_{B2}'\cdot C_{c2}] = (1+18.54)\cdot[(1301\,\Omega)\cdot(2\text{ pF})] = 50.85\text{ ns}$$

$$\tau_{cc2} = R_L'\cdot C_{c2} = (12.75\text{ k}\Omega)\cdot(2\text{ pF}) = 25.5\text{ ns}$$

$$\tau_{c2} = \tau_{cb2} + \tau_{cc2} = 76.35\text{ ns} \Rightarrow 2.08\text{ MHz}$$

Both τ_{B2} and τ_{cc2} are in *a* but τ_{cb2} is only in *b*. τ_{B2} dominates both coefficients, which are

$$a = \tau_{B2}\cdot\tau_{cc2} = (178.5\text{ ns})^2,\ b = \tau_{B2} + \tau_{c2} = 1.326\ \mu\text{s}$$

for which f_n = 891.5 kHz and ζ = 3.715 with values of

$$p_{1,2} = -122.3\text{ kHz}, -6.50\text{ MHz}$$

The lower pole and the zero at f_B = 127 kHz cancel, leaving a small dipole dip in the response and a dominant C_{c2} pole at 6.50 MHz corresponding to

$$\tau_{o2} = 24.48 \text{ ns}$$

At the Q2 collector, the capacitance forming τ_L with R_L is

$$C_{L2} = C_{b5} = C_{c5} + \tau_{T5}/R_{E5}' + C_o \cdot [f_{bw2}/f_{T5}]$$

$$= 0.5 \text{ pF} + (500 \text{ ps})/(1640\ \Omega) = 0.8 \text{ pF} + (15 \text{ pF}) \cdot [f_{bw2}/280 \text{ MHz}]$$

The value of the emitter-referred C_o (the last term in C_{b5}) depends on the frequency at which the Q2 collector time constant is affected and is iterative. A first guess is based on zero C_o and is

$$\tau_L = R_L' \cdot C_L = (12.75 \text{ k}\Omega) \cdot (0.8 \text{ pF}) = 10.2 \text{ ns} \Rightarrow 15.6 \text{ MHz}$$

Then letting f_{bw2} = 15 MHz, 0.80 pF is added to C_L which increases to

$$C_L = 1.6 \text{ pF} \Rightarrow \tau_L = 20.40 \text{ ns} \Rightarrow 7.80 \text{ MHz}$$

Iterating to convergence with the new bandwidth,

$$C_L = 1.3 \text{ pF} \Rightarrow \tau_L = 16.8 \text{ ns} \Rightarrow 9.5 \text{ MHz}$$

Applying the OCTC bandwidth formula, the bandwidth of the Q2 sub-stage is

$$\tau_{bw2} \approx \sqrt{\tau_{i2}^2 + \tau_{o2}^2 + \tau_L^2} = \sqrt{(22.2 \text{ ns})^2 + (24.48 \text{ ns})^2 + (16.8 \text{ ns})^2} = 37.07 \text{ ns}$$
$$\Rightarrow 4.3 \text{ MHz}$$

The Q1 time constants are

$$\tau_{e1} = \alpha_{NPN} \cdot \tau_{TNPN} \cdot \frac{R_{B1}' + R_{E1}}{r_{M1}} = (567 \text{ ps}) \cdot \frac{1075\ \Omega}{1025\ \Omega} = 594 \text{ ps} \Rightarrow 268 \text{ MHz}$$

$$\tau_{c1} = [(1 + K_{v1}) \cdot R_{B1}' + r_{E2} \| R_{E2}] \cdot C_{c1} \approx [R_{B1}' + r_{E2} \| R_{E2}] \cdot C_{c1}$$
$$= [(75\ \Omega + 49\ \Omega] \cdot (0.5 \text{ pF}) = (124\ \Omega) \cdot (0.5 \text{ pF})$$
$$= 62 \text{ ps} \Rightarrow 2.58 \text{ GHz}$$

Both of these poles are far above those of Q2 and can be omitted from bandwidth consideration. That leaves

$$f_{bw}(G) = f_{bw2} = 4.3 \text{ MHz}$$

The *gain-bandwidth product* of G is

$$(11.16) \cdot (4.3 \text{ MHz}) = 48 \text{ MHz}$$

Because the rolloff is not single-pole, the unity-gain frequency, $f_T(G)$, is less than the gain-bandwidth product.

In the feedback path, C_{c4} adds a pole to Q4 (and H) at

$$\tau_{c4} = [(1 + K_{v4})\cdot R_{B4}' + r_{E1}||R_{E1}]\cdot C_{c1}$$

$$= [(1.478)\cdot(469\ \Omega) + (24.5\ \Omega)||(1.0\ \text{k}\Omega)]\cdot(0.5\ \text{pF}) = 0.359\ \text{ns}$$

$$\Rightarrow 444\ \text{MHz}$$

where $K_{v4} = 0.478$. This pole is also far above loop bandwidth, as is $\tau_{e4} \approx \tau_{TNPN} \approx 530$ ps ($f_{T4} \approx 300$ MHz at 1.72 mA). $C_{b4} = \tau_T/R_{E4}'$ forms a pole with R_{B4}' around 1.5 GHz. H has no poles or zeros within a range of significant influence on the loop dynamics, and

$$H(s) \approx H_0$$

With such a low R_{B4}', C_{c4} is nearly grounded at the base terminal, and is nearly decoupled from the Q4 input loop. It parasitically adds emitter peaking to Q1 of 0.5 ns $\Rightarrow$ 318 MHz. This is too high to compensate for the Q2 dominant pole but it does come very close to providing the correct $\alpha(s)$ compensation for Q1 so that its input appears capacitive at 0.53 pF.

The feedback path is fast, though active. H is not bandwidth-limiting; the loop-gain bandwidth is that of G:

$$f_{bw}(G\cdot H) = f_{bw}(G) = 4.3\ \text{MHz}$$

The bandwidth of the loop gain is thus determined by $f_{bw}(G)$. From the section, "Feedback Increases Bandwidth", the closed-loop bandwidth for single-pole rolloff of loop gain is increased by the feedback factor;

$$f_{bw}(\text{cl}) = f_{bw}(G)\cdot(1 + G_0\cdot H) = (4.3\ \text{MHz})\cdot(3.667) = 15.8\ \text{MHz}$$

This is a significant improvement in speed over amplifier 3.

In evaluating the dynamics calculations, it is evident that $f_{bw}(\text{cl})$ barely meets the specified bandwidth and also is not in agreement with the bandwidth measurement of the amplifier 4 prototype. For sine-wave amplitudes of 0.5 V or less, it was only $f_{bw}(\text{cl}) = 6.4$ MHz. (When the CA3096 was replaced with discrete PN3904 and PN3906 BJTs, bandwidth increased to 9.4 MHz.)

What has been missed? The assumption in calculating $f_{bw}(\text{cl})$ from feedback is single-pole (or dominant-pole) loop-gain response. However, poles were found at 6.5, 7.2, and 9.5 MHz. These poles are within a decade of each other and the rolloff is not single-pole.

Frequency-response measurements on the bench confirmed this; once the rolloff began, it proceeded quickly with increasing frequency. Consequently, the bandwidth must be lower than the single-pole f_{bw}(cl) estimate, and it is.

A less obvious assumption is that the higher-degree terms in the pole polynomial are negligible. With closely-spaced poles, this is not a valid assumption, and the OCTC root-sum-of-squares bandwidth formula loses accuracy.

Feedback introduces another factor affecting bandwidth. The effect of multiple poles in $G{\cdot}H$ can be found by beginning with the single-pole formula with $H = H_0$:

$$A_v = \frac{\dfrac{G_0}{s \cdot \tau_G + 1}}{1 + \dfrac{G_0}{s \cdot \tau_G + 1} \cdot H_0} = \frac{G_0}{1 + G_0 \cdot H_0} \cdot \frac{1}{s \cdot \left(\dfrac{\tau_G}{1 + G_0 \cdot H_0} \right) + 1}$$

The open-loop bandwidth, $\omega_{bw} = 1/\tau_G$, is increased by the feedback factor, $1 + G_0{\cdot}H_0$. For an n-pole polynomial of $G(s)$ with real poles,

$$A_v = \frac{\dfrac{G_0}{s^n \cdot \tau_{G1} \cdots \tau_{Gn} + \cdots + s \cdot (\tau_{G1} + \cdots + \tau_{Gn}) + 1}}{1 + \dfrac{G_0}{s^n \cdot \tau_{G1} \cdots \tau_{Gn} + \cdots + s \cdot (\tau_{G1} + \cdots + \tau_{Gn}) + 1} \cdot H_0}$$

This becomes

$$A_v = \frac{G_0}{1 + G_0 \cdot H_0} \cdot \frac{1}{s^n \cdot \left(\dfrac{\tau_{G1} \cdots \tau_{Gn}}{1 + G_0 \cdot H_0} \right) + \cdots + s \cdot \left(\dfrac{\tau_{G1} + \cdots + \tau_{Gn}}{1 + G_0 \cdot H_0} \right) + 1}$$

For n equal real poles, this can be expressed as

$$A_v = \frac{G_0}{1 + G_0 \cdot H_0} \cdot \frac{1}{s^n \cdot \left(\dfrac{\tau_G}{\sqrt[n]{1 + G_0 \cdot H_0}} \right)^n + \cdots + s \cdot \left(\dfrac{n \cdot \tau_G}{1 + G_0 \cdot H_0} \right) + 1}$$

The n-degree-coefficient time constant decreases by only the nth root of the feedback factor while the linear-term time constant decreases by the feedback factor. As n increases,

$$\lim_{n\to\infty} \sqrt[n]{1 + G_0 \cdot H_0} = 1$$

and the *n*-degree coefficient becomes larger relative to the linear coefficient. The higher-degree terms thus influence the bandwidth more with increasing *n.* If only the linear term is used to calculate bandwidth (as is the case for the OCTC formula), then the higher terms are ignored and bandwidth is calculated as f_{bw}(cl) = 15.8 MHz, a value that is too high.

For loop gain with a single-pole, the bandwidth improvement is the feedback factor, but with three-pole rolloff, it is approximately

$$f_{bw}(A_v) = \sqrt[3]{1 + G_0 \cdot H_0} \cdot f_{bw}(G{\cdot}H) = \sqrt[3]{3.667} \cdot (4.3\ \text{MHz}) = 6.63\ \text{MHz}$$

The measured value of f_{bw} = 6.4 MHz is close to this value. The high-frequency slope of $GH(f)$ is –3 on a reactance plot because of the three poles. The approximation assumes repeated poles, which is a worst-case (low-bandwidth) approximation.

With multiple poles and feedback, analytical methods for calculating dynamic response reach a point of diminishing returns, and computer simulation is required for more accuracy. However, feedback amplifiers are generally unstable with multiple closely-spaced poles. This amplifier is stable (and not even underdamped) only because the loop gain is so low. This situation, of low loop gain and multiple poles, is a worst-case example for bandwidth approximation. Much of the time, a dominant pole or pole-pair followed by single-pole rolloff (–1 slope) around f_T makes the bandwidth easier to calculate more accurately.

Fast BJT Arrays

Cascode stages are an important building block of amplifiers, especially fast amplifiers. The static and quasistatic design is not difficult and can predict performance of actual amplifiers to less than a percent or two. Dynamics is otherwise, even with only two BJTs in the stage. For refined dynamics performance, circuit simulation is advisable at this point because the major design considerations have been thought through and dynamics calculations have provided expectations on performance so that the subsequent simulation is not a blind effort.

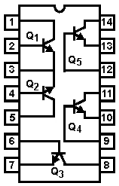
With relatively slow (CA3096) BJTs, low-cost test equipment can be used to measure bandwidth and risetime. The resulting amplifiers

have only a few megahertz of bandwidth but “fast” is relative; the major considerations in dynamics design are applicable, whatever the speed, and if faster BJTs are used instead, the bandwidth scales upward significantly, as does the cost of faster test equipment.

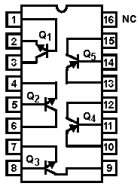
The Intersil HFA3000 series have multi-GHz f_Ts for both NPN and PNP BJTs (as shown below, taken from Intersil part data). The HFA-series PNP BJTs are not of lateral construction but are *dielectrically isolated* by an insulator (such as silicon dioxide) between transistors. They have f_T values comparable to the NPNs: 8 GHz NPNs and 5.5 GHz PNPs. β_0 for NPNs is typically 130 and is 60 for PNPs. The r_o values are lower than the CA3096 BJTs; the NPN Early voltage is 50 V and is only 20 V for PNPs. A tradeoff for higher speed is lower breakdown voltage: 8 V open base to 12 V open emitter for NPN and 10 V to 15 V for PNP. A choice of supply voltages of ±5 V is more compatible with these ratings than ±12 V used with the CA3096.

Pinouts

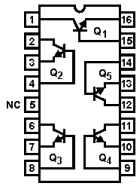
HFA3046
TOP VIEW



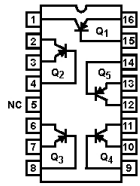
HFA3096
TOP VIEW



HFA3127
TOP VIEW



HFA3128
TOP VIEW



HFA3046, HFA3096, HFA3127, HFA3128

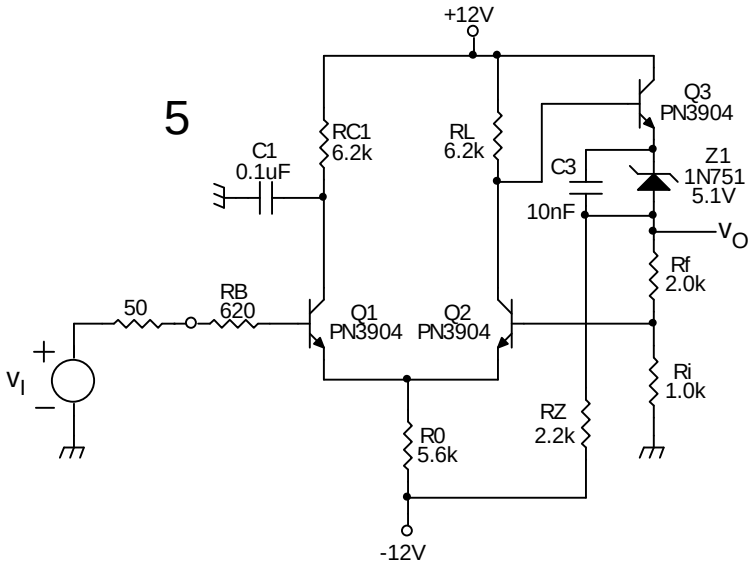
Electrical Specifications $T_A = 25°C$ **(Continued)**

PARAMETER	TEST CONDITIONS	DIE			SOIC, QFN			UNITS
		MIN	TYP	MAX	MIN	TYP	MAX	
DIFFERENTIAL PAIR MATCHING CHARACTERISTICS FOR THE HFA3046								
Input Offset Voltage	I_C = 10mA, V_{CE} = 5V	-	1.5	5.0	-	1.5	5.0	mV
Input Offset Current	I_C = 10mA, V_{CE} = 5V	-	5	25	-	5	25	µA
Input Offset Voltage TC	I_C = 10mA, V_{CE} = 5V	-	0.5	-	-	0.5	-	µV/°C

Differential and Cascade Amplifiers

The 2-BJT differential amplifier, or "diff-amp" or *diff-pair* for short, is often used as the input stage of amplifiers because of its high PSR and its ability to amplify differentially with low offset voltage, temperature drift, and distortion. This makes it a good choice not only for op-amp input stages but also for feedback amplifiers generally. In this chapter, amplifiers with diff-pair input stages and cascade CE stages are designed.

3-NPN Amplifier Dynamic Response

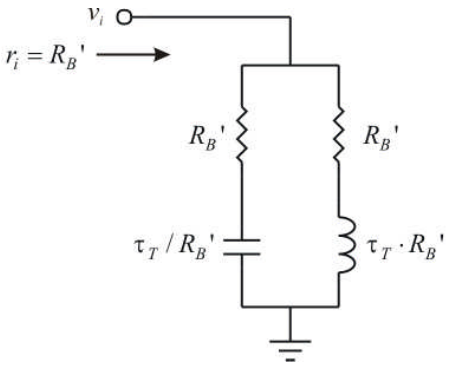


We resume where the section, "3-NPN Feedback Amplifier", at the end of the "Feedback Amplifiers" chapter left off, and call it amplifier number 5, repeated here. The static and quasistatic design was performed, and those results will be used in the dynamic design.

The three discrete PN3904 NPN BJTs have a typical $\beta_0 = 150$ and are rated at $f_T = 300$ MHz at 10 mA and at a reduced 225 MHz at 1 mA ($\tau_T = 707$ ps). $C_c = 2$ pF at 5 V.

The Q1 collector bypassing requires some method for determining the value of the bypass capacitor. C_1 reduces the voltage gain at the Q1 collector to near-zero and v_{C1} is nearly a static voltage. C_1 forms with $R_{C1} = 6.2$ kΩ a time constant that for 100 nF is $\tau = 620$ μs with a pole frequency of 257 Hz, a low frequency at which $C_{c1} \approx 2$ pF has a reactance of 155 MΩ and can be considered an open circuit. Thus, C_1 need only be much larger than C_{c1} to bypass i_{c1} to ground. Consequently, Q1 has no input Miller effect.

In the diff-pair input stage, there is no external R_E, thereby alerting us to hf effects; the pole of C_π is significant for Q1 and Q2. R_0 is large enough that it does not significantly shunt the emitter circuit and can be ignored. The hf equivalent circuit is shown below.



At the emitter of Q2, $R_f||R_i + r_{b2}' = R_{B2}'$ is gyrated +90° in the hf region to result in an emitter-referred circuit of hf inductance $\tau_T \cdot R_{B2}'$ in parallel with R_{B2}'. This impedance is gyrated –90° at the Q1 base so that R_{B2}' becomes R_{B2}' in series with capacitance τ_T/R_{B2}'. This is in parallel with the gyrated inductor, a series RL of R_{B2}' and $\tau_T \cdot R_{B2}'$. The circuit is symmetrical and this is the same impedance as at the Q2 base, so that $R_B' = R_{B2}' = R_{B1}' = 720$ Ω.

The values of L and C meet the requirements for constant resistive impedance:

$$Z_n = \sqrt{\frac{L}{C}} = \sqrt{\frac{\tau_T \cdot R_B'}{\tau_T / R_B'}} = R_B'$$

At the Q1 base, R_{B1}' is in parallel with r_{B1}(hf) = Z_n = R_{B2}' ≈ R_B' for a total hf resistance at the base node (to ground) of about R_B'/2. C_{c1} forms an RC integrator at the base of Q1 and connects to the base where the two R_B' branches join. The resulting base resistance forms a base-node time constant of (R_B'/2)·C_c. The Q1 collector is a dynamic ground (because of C_1), making (R_B'/2)·C_{c1} the OCTC of C_{c1}. From 2N3904 data, typical C_c = 2 pF at 5 V. The Q1 base time constant with C_{c1} = 2 pF is

$$\tau_{c1} = \tau_{b1} = [(r_b' + R_B)/2]\cdot C_{c1} = (360\ \Omega)\cdot(2\ \text{pF}) = 0.72\ \text{ns} \Rightarrow 221\ \text{MHz}$$

As for C_{e1}, Q1 external emitter resistance is

$$R_{E1} = r_{e2} + R_{B2}'/(\beta_{NPN} + 1) = 30\ \Omega$$

The open-circuit *b*-*e* resistance is

$$R_{be1} = r_{\pi 1} \parallel \frac{R_{B1}' + R_{E1}}{1 + \frac{R_{E1}}{r_{m1}}} = 319\ \Omega$$

The OCTC of C_{e1} is thus

$$\tau_{e1} = R_{be1}\cdot C_{e1} = R_{be1}\cdot\frac{\alpha_0\cdot\tau_{T1}}{r_{e1}} = (319\ \Omega)\cdot\frac{527\ \text{ps}}{26\ \Omega} = 6.47\ \text{ns}$$
$$\Rightarrow 24.6\ \text{MHz}$$

Turning to Q2, from base to collector the lf gain is K_{v2}. The Q3 base loads R_L by about 193 kΩ so that the combined collector resistance is

$$R_L' = R_L \| r_{B3} = 6.0\ \text{k}\Omega$$

What complicates Q2 dynamics is that along the *G* path, it functions as a CB stage, yet its base is driven from the feedback divider of the *H* path as a CE stage. Which is it? Because of base resistance, the base is not grounded and varies with input voltage. The CB path of *G* includes the pole of C_{c2}. A similar situation occurred in the amplifier 3 CB sub-stage, where τ_B and τ_c interacted to form a quadratic pole. However, τ_B was large relative to τ_c; in this amplifier, C_B is negligible as is τ_B. What is left is a single real OCTC of τ_{c2}.

Applying the general BJT stage open-circuit *b*-*c* resistance template to the C_{c2} port,

$$R_{bc2} = R_L + R_{b2}\cdot(1 + K_{v2})$$

where

$$r_{B2} = (\beta_{02} + 1) \cdot (r_{e2} + r_{e1} + R_{B1}'/(\beta_{01} + 1))$$

$$r_M = r_{e1} + R_{B1}'/(\beta_{01} + 1) + r_{e2} + R_{B2}'/(\beta_{02} + 1) = 61\ \Omega$$

The Q2 base resistance (including r_b' in each of the R_B') is

$$R_{b2} = R_{B2}' \| r_{B2} = [R_{B2}' \| (\beta_{02} + 1) \cdot (r_{e2} + r_{e1} + R_{B1}'/(\beta_{01} + 1))]$$

$$= (667\ \Omega + 50\ \Omega) \| (151) \cdot [52\ \Omega + (670\ \Omega + 50\ \Omega)]$$

$$R_{b2} = (717\ \Omega) \| (117\ \text{k}\Omega) = 713\ \Omega$$

The Q2 voltage gain is

$$K_{v2} = \alpha_{02} \cdot \frac{R_L'}{r_M} = (0.993) \cdot \frac{6.0\ \text{k}\Omega}{61\ \Omega} = 98$$

Then R_{bc2} = 76.6 kΩ and the C_{c2} OCTC components are

$$\tau_{cb2} = (1 + 98) \cdot (713\ \Omega) \cdot (2\ \text{pF}) = (70.6\ \text{k}\Omega) \cdot (2\ \text{pF}) = 141\ \text{ns}$$

$$\tau_{cc2} = 6.0\ \text{k}\Omega \cdot (2\ \text{pF}) = 12\ \text{ns}$$

$$\tau_{c2} = \tau_{cb2} + \tau_{cc2} = 153\ \text{ns} \Rightarrow 1.04\ \text{MHz}$$

Most of this slow time constant is caused by the Miller effect in τ_{cb2}.

To calculate τ_L, Q3 contributes

$$C_{b3} = C_{c3} + \tau_{T3}/R_{E3} + C_{E3}/(\beta_3(s) + 1)$$

where

$$R_{E3} = r_z + (R_f + R_i) \| R_z = 5\ \Omega + 3.0\ \text{k}\Omega \| 2.2\ \text{k}\Omega = 1274\ \Omega$$

Q3 conducts about 5.5 mA which yields $\tau_{T3} \approx 606$ ps. Emitter $\alpha(s)$ compensation of Q3 is achieved by the parasitic capacitance across the emitter resistors. A ¼ W axial-lead resistor has a parasitic capacitance of about ⅓ pF. Two in parallel (R_f and R_z) result in about 0.67 pF that form a time constant with the emitter external resistance of about 850 ps, close to τ_T. Then

$$C_{b3} = C_{c3} + \tau_{T3}/R_{E3} = 2\ \text{pF} + 0.48\ \text{pF} \approx 2.5\ \text{pF}$$

The parasitic $C_E \approx 0.67$ pF and output load capacitance from a 'scope probe of C_o = 15 pF appear at the base with these values only at f_{T3}. Below f_T, they decrease by $1/(\beta(s) + 1)$ so that at f_β = 1.5 MHz, the

base-referred C_E is 15.67 pF/151 = 0.104 pF and can be ignored. At 25 MHz, it is

$$\frac{C_E}{\beta_0 + 1} \cdot \frac{f}{f_\beta} = (0.104\,\text{pF}) \cdot \frac{25\,\text{MHz}}{1.49\,\text{MHz}} = 1.74\,\text{pF}$$

At 25 MHz, base-referred C_E is an appreciable fraction of C_{b3}. If the bandwidth calculates to be anywhere near 25 MHz, then C_{b3} must be iterated. As a first guess, the Q1 collector time constant will include the effect of C_{E3} around 25 MHz and the referred C_E is 4.25 pF;

$$\tau_L = R_L' \cdot C_{b3} = (6.0\,\text{k}\Omega) \cdot (4.25\,\text{pF}) = 25.5\,\text{ns} \Rightarrow 6.24\ \text{MHz}$$

Finally, the combined time constants of the OCTC bandwidth is

$$\tau_{bw} = \sqrt{\tau_{e1}^2 + \tau_{c2}^2 + \tau_L^2} = \sqrt{(6.47\ \text{ns})^2 + (153\,\text{ns})^2 + (25.5\,\text{ns})^2} = 155\,\text{ns}$$

$$\Rightarrow 1.025\ \text{MHz}$$

The G path has a dominant pole caused by τ_{c2}.

The effect of τ_{c2} was included in G and removes it from being included in H as a C_{b2} loading impedance on its resistive divider. The effect of $r_{Ho} = R_f \| R_i$ of H was included in τ_{c2} as base resistance, leaving H an unloaded open-circuit output from the divider. Thus H is purely quasistatic, having no reactances. This leaves the amplifier as a simple, dominant-pole feedback amplifier with an open-loop bandwidth (from G) of f_{bw} = 1.025 MHz. The quasistatic gains have been calculated as $G_0 \approx 98$ and $H_0 = 1/3$. Then $(G{\cdot}H)_0 \approx 32.7$. Because of the dominant pole at the open-loop f_{bw}, the amplifier closed-loop response is

$$A_v = \frac{\dfrac{G_0}{s \cdot \tau_{bw} + 1}}{1 + \dfrac{G_0}{s \cdot \tau_{bw} + 1} \cdot H_0} = \frac{G_0}{1 + G_0 \cdot H_0} \cdot \frac{1}{s \cdot \left(\dfrac{\tau_{bw}}{1 + G_0 \cdot H_0} \right) + 1}$$

The closed-loop quasistatic gain is A_{v0} = 2.91, and the pole has moved out to 34.5 MHz. This is greater than 25 MHz and the base-referred C_{E3} is consequently somewhat greater, though it will have only a minor, non-dominant effect. A prototype circuit was measured with a 2 V output amplitude as having a bandwidth of 38 MHz with a lf gain of 2.9, and peaking of M_m = 1.20 at f_m = 19.4 MHz.

The risetime was also measured as 8.6 ns with a pulse generator input having a square-wave risetime of about 4.8 ns. Using the more accurate OCTC formula for risetime and solving for the amplifier t_r,

$$t_r = 1.1 \cdot \sqrt{t_{ro}^2 - t_{ri}^2} = 1.1 \cdot \sqrt{(8.6\,\text{ns})^2 - (4.8\,\text{ns})^2} = (1.1) \cdot (7.13\,\text{ns}) = 7.85\,\text{ns}$$

The response was somewhat underdamped, though if it were approximated by a real, dominant pole, then the risetime-bandwidth formula would apply;

$$f_{bw} \approx \frac{0.35}{t_r} = \frac{0.35}{7.85\,\text{ns}} = 44.6\,\text{MHz}$$

This indirect measurement of bandwidth is high by 17 %. Possible causes are that the response is not that of a single pole and that linear (OCTC) bandwidth approximation is too inaccurate.

From a feedback control standpoint, the peaking in the response at f_m indicates a closed-loop complex pole-pair. The pole angle can be found from the closed-loop damping;

$$\zeta_{cl} = \sqrt{\frac{1}{2} \cdot \left(1 - \sqrt{1 - \frac{1}{M_{mcl}^2}}\right)} = \sqrt{\frac{1}{2} \cdot \left(1 - \sqrt{1 - \frac{1}{(1.20)^2}}\right)} = 0.473 \Rightarrow \phi = 61.8°$$

Then having ζ_{cl},

$$f_{ncl} = \frac{f_{mcl}}{\sqrt{1 - 2 \cdot \zeta_{cl}^2}} = \frac{19.4\,\text{MHz}}{1.666} = 26.1\,\text{MHz}$$

The closed-loop bandwidth is

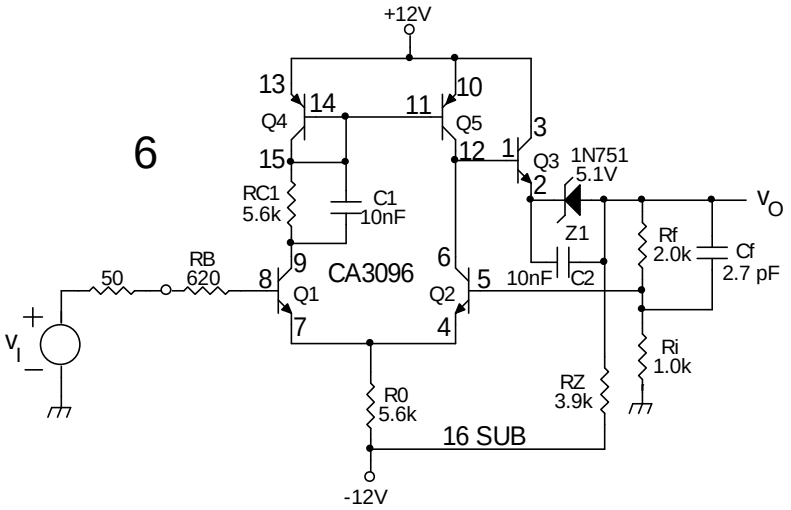
$$f_{bw}(\text{cl}) = (1.30) \cdot f_n = 34\ \text{MHz}$$

and compares favorably to the measured bandwidth of 38 MHz. As $G_0 \cdot H_0$ decreases, so does the pole angle and the peaking in the frequency response. However, decreased quasistatic loop gain also reduces quasistatic accuracy. A better scheme is to introduce frequency compensation in the feedback divider, $H(s)$, as in amplifier 6 of the next section.

BJT-Array Feedback Amplifier

From the section, “Noninverting BJT-Array Amplifier” of the chapter “Matched-Transistor Circuits”, the circuit diagram of the

BJT-array amplifier 6 is repeated below, to be analyzed in more detail. It is similar to the 3-NPN feedback amplifier but has increased gain by using i_{C1}. A current mirror consisting of PNP BJTs includes i_{C1} in the waveform path through G. The mirror increases loop gain and quasistatic performance over amplifier 5, taking advantage of matched PNP BJTs in the CA3096 array. The CA3096 NPN BJTs have an f_T of 335 MHz at 4 mA and are comparable to the PN3904, with f_T = 300 MHz at 10 mA. The CA3096 PNPs have β = 47 and are slow, rated at f_T = 6.8 MHz at 0.1 mA.



The array not only matches the *b-e* junctions of the current mirror (Q4, Q5) but also matches Q1, Q2 *b-e* voltages at equal currents. This minimizes the static *input offset-voltage error*, V_{IOS}, of the amplifier. $V_{IOS} = v_I$ when v_O = 0 V. If the input were shorted (v_I = 0 V), then the output voltage v_O is the output offset-voltage error, V_{OOS}, and is the noninverting closed-loop gain of the amplifier loop times V_{IOS}.

Instead of wasting i_{c1}, it is added algebraically to i_{c2} by the current mirror, resulting in nearly twice the gain. The SUB pin of the array is the connection to the common monolithic substrate and is connected to the lowest voltage node of any of the array BJT terminals so that p-n junctions formed with the substrate are reverse-biased.

Working through the circuit in more detail than in the "Matched-Transistor Circuits" chapter, the static design makes use of CA3096 data; NPN V_{BE}(1 mA) = 0.69 V and PNP V_{BE}(1 mA) = 0.66 V. The emitter bias current of the diff-pair stage is I_0 = 2.02 mA. For

$\alpha_{NPN} = 0.997$, Q1 and Q2 collector currents are $I_{C1} = I_{C2} = 1.01$ mA and

$$V_{C1} = 12\text{ V} - (0.66\text{ V} + (1.01\text{ mA})\cdot(5.6\text{ k}\Omega)) = 5.70\text{ V}$$

From current-mirror theory, we found that the current gain of the 2-BJT (Widlar-type) mirror is $\beta_{PNP}/(\beta_{PNP} + 2) = 47/49 \approx 0.96$. Then the incremental gain of current to the base node of Q3 is $1.96\cdot i_{c2}$. The node has two collector current sources driving the Q3 base input resistance, r_{b3};

$$r_{b3} = (\beta_{NPN} + 1)\cdot(r_{e3} + r_z + R_Z \,\|\, (R_f + R_i)) = 668.3\text{ k}\Omega$$

The collector resistance, r_{b3}, is so large that BJT r_o becomes a factor. For Q5, $I_{E5} \approx 1$ mA and

$$r_{o5} = \frac{65\text{ V} + 10\text{ V}}{1\text{ mA}} = 75\text{ k}\Omega$$

where the PNP $V_A = 65$ V and $V_{BC5} \approx 10$ V. The Q2 collector is in parallel and has

$$r_{o2} = \frac{75\text{ V} + 5\text{ V}}{1\text{ mA}} = 80\text{ k}\Omega$$

The r_{o5} emitter end is dynamically grounded at +12 V, but r_{o2} is a positive feedback path (with gain << 1) in Q2 which has the effect of "bootstrapping" r_{o2} so that at the collector, r_{c2} appears somewhat larger than r_{o2}. The value of r_{o2} is a worst-case (minimum r_o) value.

The NPN $V_A = 75$ V and $V_{CB2} \approx 5$ V. Then $r_{o2} \| r_{o5} = 38.71\text{ k}\Omega$ and

$$R_{b3} = r_{b3} \| r_{o5} \| r_{o4} = 36.6\text{ k}\Omega$$

Using the same choice of feedback quantities as before, the loading of the feedback-divider resistors is included in G. For the r_{b3} calculation, the base of Q2 is opened to open the H output port or else the Q2 base loading would be in H.

The value of r_{e3} depends on I_{E3}. For 0 V input and negligible feedback error (or high loop gain, $G\cdot H$), then $v_O = 0$ V, $I_{E3} = 12\text{ V}/3.9\text{ k}\Omega = 3.08$ mA, and $r_{e3} = 26\text{ mV}/3.08\text{ mA} = 8.5\ \Omega$. Set $r_z \approx 5\ \Omega$ and the above value of r_{b3} results.

The first-stage quasistatic voltage gain of G is based on diff-pair symmetry ($r_{e1} = r_{e2}$, $R_{B1} = R_{B2}$) and is

$$G_1 = \alpha_{NPN} \cdot (1 + A_i) \cdot \frac{R_{b3}}{2 \cdot [R_{B1}/(\beta_{NPN} + 1) + r_{e1}]}$$

$$\approx (0.997) \cdot (1.96) \cdot \frac{36.6\ \text{k}\Omega}{55.7\ \Omega} = 1285$$

The second-stage gain is in the form of a CC voltage divider;

$$G_2 = \frac{R_{E3}}{r_{e3} + R_{E3}} = \frac{1.70\ \text{k}\Omega}{13.5\ \Omega + 1.70\ \text{k}\Omega} = 0.992$$

where $R_{E3} = r_z + R_Z \parallel (R_f + R_i)$

$$= 5\ \Omega + (3.9\ \text{k}\Omega) \| (2.0\ \text{k}\Omega + 1.0\ \text{k}\Omega) = 1701\ \Omega$$

The total quasistatic gain, G_0, is then

$$G_0 = G_1 \cdot G_2 = 1275$$

The loop gain $G_0 \cdot H_0 = 425$ and from it and the feedback formula, the closed-loop $A_v = 2.993$. Although BJT r_o significantly reduces G_1 (by 18.3 times), it is still large enough to make the amplifier somewhat of an op-amp. The error caused by finite loop gain is much less than resistor tolerances. Though G is reduced by output loading, it is high enough that A_v will be insensitive to all but a heavily-loaded output. The Q3 CC stage buffers or isolates the effect on amplifier gain from the output load. This is important not only for gain precision but also for feedback stability, that the loop gain be fixed in value.

In both amplifiers 5 and 6, R_0 is assumed open when calculating incremental resistance. R_0 reduces the gain slightly, by the current-divider ratio $R_0/(R_0 + r_{E2})$, where $r_{E2} \approx 27.8\ \Omega$. For both amplifiers, it reduces G by a factor of 0.995 and consequently has a very small effect on loop gain.

The static voltage at the Q2 collector is maintained by feedback to keep the input error close to zero and hence the output voltage close to zero. By feedback, $V_{C2} \approx 5.1\ \text{V} + 0.7\ \text{V} = 5.8\ \text{V}$. The current-mirror bases (of Q4, Q5) are at about 11.34 V and this gives enough voltage range for an output range of at least ±5 V. Below 0 V output, R_Z is the passive path for output current, though Q3 continues to conduct and the incremental output resistance is kept low by feedback. The amplifier open-loop output resistance, $r_o(G)$, is high because r_{B3} is high so that

$$r_o(G) = R_{b3} \| R_{E3} = (36.6\ \text{k}\Omega) \| (1701\ \Omega) = 1.625\ \text{k}\Omega$$

The loop gain reduces $r_o(G)$ by $(1 + G_0 \cdot H_0)$ to 1.625 kΩ/426 = 3.8 Ω.

Dual-Path Dynamics

In the dynamics, a new situation arises. The *G* path has two parallel paths with different dynamic responses. The path through Q1 and the current mirror includes the current-mirror response, but the path through Q2 is faster. The two *G* paths add at the Q3 base node: the low-speed path, $G_{1l}(s)$, and $G_{1h}(s)$, the high-speed path;

$$G_1(s) = G_{1l}(s) + G_{1h}(s)$$

The quasistatic gains of the paths are separated as

$$G_{1l} = \alpha_{NPN} \cdot A_i \cdot \frac{R_{b3}}{2 \cdot [R_{B1}/(\beta_{NPN} + 1) + r_{e1}]} = (0.997) \cdot (0.96) \cdot \frac{36.6\ \text{k}\Omega}{55.7\ \Omega} = 629.2$$

$$G_{1h} = \alpha_{NPN} \cdot \frac{R_{b3}}{2 \cdot [R_{B1}/(\beta_{NPN} + 1) + r_{e1}]} = (0.997) \cdot \frac{36.6\ \text{k}\Omega}{55.7\ \Omega} = 655.4$$

Including G_2 for complete forward-path gains,

$$G_{0l} = 624.4\ ;\ G_{0h} = 650.4 \Rightarrow G_0 = G_{0l} + G_{0h} = 1275$$

Using the quasistatic data, we can approximate the OCTC poles for both paths. The input circuit of the first stage is similar to that of amplifier 5. For Q1, $K_v << 1$ and dynamic $R_L \approx 0\ \Omega$; there is negligible input Miller effect and $C_{cb1} \approx C_{c1}$. R_{C1} is bypassed by C_1 to reduce Q1 Miller effect.

As with amplifier 5, the emitter circuit of Q1 refers $R_{B2}' = r_{b2}' + R_f \| R_i$ to the Q1 base in the hf region as $R_{B2}' = 717\ \Omega$ and the Q1 base resistance is the parallel combination. Include $r_{b1}' = 50\ \Omega$ with R_{B1} as R_{B1}' and

$$R_{b1} = R_{B1}' \| R_{B2}' = (720\ \Omega) \| (717\ \Omega) = 359\ \Omega$$

The Q1 collector has $R_{L1} \approx r_{e4} = 26\ \Omega$ and

$$K_{v1} \approx (26\ \Omega)/(55.7\ \Omega) = 0.47$$

The Miller multiplier is thus 1.47 and

$$\tau_{c1} \approx \tau_{bc1} = (359\ \Omega) \cdot (1.47) \cdot (0.5\ \text{pF}) = 0.264\ \text{ns} \Rightarrow 603\ \text{MHz}$$

For the slow path, the current mirror has $R_{E4} = R_{E5} = 0\ \Omega$ and a pole is formed by Z_π in (dynamic) parallel with C_{bs} of both Q4 and

Q5. Although the C_{bs} substrate terminal connects to the –12 V supply while C_e connects to +12 V, the voltage sources are dynamically shorted to ground through the low impedance of the supply bypass capacitors and the somewhat higher impedance of practical sources.

The PNP f_T at 1 mA is 6.5 MHz. This places the Z_π pole at

$$\text{PNP } f_\beta = f_T/\beta_0 \approx 6.5 \text{ MHz}/47 = 138 \text{ kHz}$$

C_π in Q4 and Q5 are not independent capacitors because they are in parallel and can be treated as a single capacitance. Shunting them is $2{\cdot}C_{bs}$ = 5 pF. At 1 mA, r_π of Q4, Q5 is

$$r_\pi = (48){\cdot}(26\ \Omega) = 1248\ \Omega$$

The value of

$$C_\pi = \tau_\beta / r_\pi = 1/2{\cdot}\pi \cdot f_\beta \cdot r_\pi = 922 \text{ pF}$$

and verifies that the current-mirror pole is at

$$\tau_{Ai} = (r_\pi/2){\cdot}[2{\cdot}(C_\pi + C_{bs})] = 1.154\ \mu\text{s} \Rightarrow 138 \text{ kHz}$$

The *b-s* capacitance has little effect; C_π dominates.

Proceeding to the OCTC of C_{c5} = 2 pF, this capacitance in the slow path is part of the Q5 gain and is in the current mirror. The fast path does not include it in C_{L2} because it has its own OCTC as part of Q5. (This assumes no capacitive interaction that forms pole-pairs.) The open-circuit resistance across C_{c5} is

$$R_{bc5} =$$

$$(r_\pi/2){\cdot}(1 + K_{v5}) + R_{b3} = (624\ \Omega){\cdot}\left(1 + \alpha_{PNP} \cdot \frac{R_{b3}}{r_{e5}}\right) + 33.6\ \text{k}\Omega = 824\ \text{k}\Omega$$

$$\tau_{c5} = R_{bc5}{\cdot}C_{c5} = 1.65\ \mu\text{s} \Rightarrow 97 \text{ kHz}$$

At least 5 capacitances can be identified as attached to the node of the Q2, Q5 collectors and they all interact in an exact circuit solution. To simplify this mess, we can make an isolation assumption that allows us to apply dynamic superposition to the analysis. We can find the OCTCs of Q5 and Q2 for their respective paths as though they were isolated stages with a common C_L.

The Q5 collector node has a high resistance (despite r_o of Q2 and Q5) resulting in a dominant fast-path pole frequency for any capacitance attached to it. Collector capacitance for Q5 and Q2 is

$$C_L = C_{c3} + C_{cs2} \approx 0.5 \text{ pF} + 1.5 \text{ pF} = 2.0 \text{ pF}$$

Then the shared collector pole has time constant

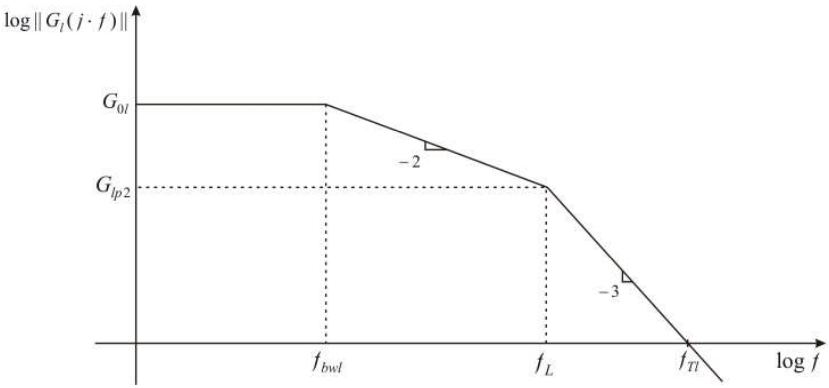
$$\tau_L = R_{b3} \cdot C_L = 67.2 \text{ ns} \Rightarrow 2.37 \text{ MHz}$$

The poles of Q3 are high in frequency because $R_{E3} >> r_{e3}$ and can be disregarded. Output loading - and in particular, $C_o = 15$ pF - is considered in its effect on C_L, but is $C_L/(\beta_{NPN} + 1) = 38$ fF below f_β of Q3, somewhat under 1 MHz.

Applying the OCTC root-sum-of-squares bandwidth formula (with frequencies instead of time constants) for G_{1l},

$$f_{bwl} \approx 1/\sqrt{1/(97\,\text{kHz})^2 + 1/(138\,\text{kHz})^2 + 1/(2.37\,\text{MHz})^2} = 79.3\,\text{kHz}$$

The frequency-response magnitude for the slow path has three poles, one shared with the fast path (f_L). The lowest two poles are only 0.15 decade apart, and the magnitude can be roughly approximated as shown in the plot below, where $f_{bwl} \approx 116$ kHz (the geometric mean of the two lowest poles); $f_L = 2.37$ MHz and $G_{lp2} = 624.4/20.48 = 30.48$. Then $f_{Tl} = f_L \cdot G_{lp2}^{1/3} = 7.4$ MHz.



Three poles in a feedback loop alert us to the possibility of feedback instability of the slow path. However, the fast path mitigates, as we shall see, by adding compensating zeros.

We now consider the fast path, G_h. It shares not only G_2 and the Q1 input loop with G_l but also the C_L pole. It differs in that it has its own C_{c2} pole and no current-mirror poles. The C_{c2} OCTC resistance is

$$R_{bc2} = R_{b3} + R_{B2}' \cdot (1 + K_{v2}) = (33.6\,\text{k}\Omega) + (717\,\Omega) \cdot (656) = 504.2\,\text{k}\Omega$$

where $K_{v2} \approx G_{1h} = 655.4$. The C_{c2} OCTC is

$$\tau_{c2} = R_{bc2} \cdot C_{c2} = 252 \text{ ns} \Rightarrow 631 \text{ kHz}$$

The OCTC of C_L was calculated as

$$\tau_L = R_{b3} \cdot C_L = 67.2 \text{ ns} \Rightarrow 2.37 \text{ MHz}$$

and when the poles are combined for a bandwidth estimate of G_h,

$$\tau_{bwh} = \sqrt{\tau_L^2 + \tau_{c2}^2} = \sqrt{(67.2\,\text{ns})^2 + (252\,\text{ns})^2} = 261\,\text{ns} \Rightarrow 610\,\text{kHz}$$

This is almost a decade faster than f_{bwl}. The gain rolls off from $G_{0h} = 650.4$ at f_{c2} with a –1 slope to f_L, where the gain is 173.2. The slope becomes –2 and crosses a gain of one at $f_{Th} = 31.2$ MHz.

The slow and fast paths combine by adding, and this adds zeros to the response. The transfer functions, as derived, are

$$G_l = G_{0l} \cdot \frac{1}{(s \cdot \tau_{Ai} + 1) \cdot (s \cdot \tau_{c5} + 1) \cdot (s \cdot \tau_L + 1)}$$

$$G_h = G_{0h} \cdot \frac{1}{(s \cdot \tau_{c2} + 1) \cdot (s \cdot \tau_L + 1)}$$

Then

$$G(s) = G_l + G_h = \frac{G_{0l} \cdot (s \cdot \tau_{c2} + 1) + G_{0h} \cdot (s \cdot \tau_{c5} + 1) \cdot (s \cdot \tau_{Ai} + 1)}{(s \cdot \tau_{Ai} + 1) \cdot (s \cdot \tau_{c5} + 1) \cdot (s \cdot \tau_{c2} + 1) \cdot (s \cdot \tau_L + 1)}$$

This normalizes to

$$G(s) = G_0 \cdot \frac{s^2 \cdot \left(\tau_{c5} \cdot \tau_{Ai} \cdot \frac{G_{0h}}{G_0} \right) + s \cdot \left(\tau_{c2} \cdot \frac{G_{0l}}{G_0} + (\tau_{Ai} + \tau_{c5}) \cdot \frac{G_{0h}}{G_0} \right) + 1}{(s \cdot \tau_{Ai} + 1) \cdot (s \cdot \tau_{c5} + 1) \cdot (s \cdot \tau_{c2} + 1) \cdot (s \cdot \tau_L + 1)}, \; G_0 = G_{0l} + G_{0h}$$

Substituting design values, the quadratic zero polynomial is

$$N_G(s) = s^2 \cdot (985.6\,\text{ns})^2 + s \cdot (1.554\,\mu\text{s}) + 1$$

This has $\zeta = 0.7883$ corresponding to a pole angle of $\phi = 38°$. The complex zero-pair has $f_{nz} = 161.5$ kHz. For $\phi < 45°$, no peaking in the time or frequency response occurs. The zero-pair f_{nz} is somewhat higher than the two real poles of the slow path. Its real component is at $\zeta \cdot f_{nz} = 127$ kHz, near the geometric-mean frequency of the two slow poles of 116 kHz. The zero-pair compensates for them, so that at the poles of the fast path (including the shared pole of f_L), the slow-path dynamics have essentially been removed from the G path.

We turn now to the feedback path of H. The loop gain has an additional pole and zero because of the RC phase-lead circuit, R_f, C_f, R_i, which has a transfer function of

$$-H = \frac{R_i}{R_f + R_i} \cdot \frac{s \cdot R_f \cdot C_f + 1}{s \cdot (R_f \parallel R_i) \cdot C_f + 1}$$

Then

$$H_0 = \frac{R_i}{R_f + R_i} = \frac{1}{3}$$

For C_f = 2.7 pF, the H divider z = 29.5 MHz and p = 88.4 MHz. Its value can be adjusted for optimal response. This additional real zero in the loop lessens the descent of the phase toward its loop-unstable value of –180° while the magnitude continues to decrease.

The loop-gain can be simplified by cancellation of the G zero-pair with the two slow-path poles, leaving the fast-path poles in a feedback loop with a loop gain of

$$G_0 \cdot H_0 = (1275) \cdot (1/3) = 425$$

and a feedback factor of $1 + G_0 \cdot H_0 = 426$. Ignoring H for now, the loop is left with two poles. With feedback they come together at a frequency from the quadratic feedback formula,

$$f_{ncl} = f_n \cdot \sqrt{1 + G_0 \cdot H_0}$$

where f_n = 1.27 MHz and f_{ncl} = 26.2 MHz. The real loop poles, taken as a pair, have p_1/p_2 = 2.57 MHz/361 kHz = 4.073. Then (from "Pole Separation" in "Amplifier Design"),

$$\zeta = \frac{1}{2} \cdot \sqrt{\frac{((p_1 / p_2) + 1)^2}{p_1 / p_2}} = 1.257 \Rightarrow \zeta_{cl} = \frac{\zeta}{\sqrt{1 + G_0 \cdot H_0}} = 0.0609$$

with a pole angle of 86.5°. Loop gain reduces ζ by about 20.64 times. This highly underdamped loop is compensated by the zero of H that is placed near f_{ncl}. The lower pole is terminated by the zero and the higher pole, f_L, moves upward to determine bandwidth;

$$f_{bw}(\text{cl}) \approx f_L \cdot \sqrt{1 + G_0 \cdot H_0} = (2.57\ \text{MHz}) \cdot (20.64) = 53\ \text{MHz}$$

On the bench with the amplifier 6 prototype circuit, the static voltages readily agree with the design values. Amplifier 6 measures

with an input sine-wave at 50 kHz and amplitude of 200 mV an output amplitude of 590 mV to result in a quasistatic gain of 2.95. (The Tektronix SG503 leveled sine-wave generator used for the measurement has a 50 kHz setting for the "quasistatic" calibration of the amplitude.)

Without C_f and an output sine-wave of 0.5 V amplitude (measured as 1 V pk-pk), the bandwidth is f_{bw}(cl) = 43 MHz but with peaking in the frequency response at f_{mcl} = 20 MHz or M_{mcl} = 2.14. This is serious enough to require loop frequency compensation, and C_f is installed. From "Frequency Response Characterization" in the "Amplifier Design" chapter,

$$\zeta_{cl} = \sqrt{\frac{1}{2}\cdot\left(1-\sqrt{1-\frac{1}{M_{mcl}^2}}\right)} = 0.241 \Rightarrow 76.1°$$

Also from the same section,

$$f_{mcl} = f_{ncl}\cdot\sqrt{1-2\cdot\zeta_{cl}^2}$$

$$\Rightarrow\ f_{ncl} = f_{mcl} / \sqrt{1-2\cdot\zeta_{cl}^2} = 20\ \text{MHz}/0.940 = 21.3\ \text{MHz}$$

Instead of 2.7 pF as shown on the diagram, 3 1 pF capacitors were placed in parallel for C_f = 3 pF in the prototype. The frequency response magnitude dipped slightly at 11 MHz to 0.96 of its low-frequency value, still well above the bandwidth magnitude. The magnitude then peaked slightly at f_{mcl} = 32.7 MHz to M_{mcl} = 1.01. Thereafter, the magnitude decreased to a bandwidth of f_{bw} = 57 MHz. This exceeds the predicted value of 53 MHz by 7.5 % and is caused by peaking. Using the above formula for $\zeta(M_m)$ and substituting M_{mcl} = 1.01, then ζ_{cl} = 0.656 and the quadratic bandwidth formula gives a bandwidth extension of 1.072, resulting in a bandwidth of 56.8 MHz. What was not anticipated in the design was the dip, then rise, in the frequency response. This requires more accurate determination of the open-loop transfer function, and in particular, identification of the quadratic poles that cause peaking (and dips).

For C_f = 10 pF, the dip grew enough to cause bandwidth to be at 9.3 MHz. The bottom of the dip at 16 MHz was 0.62 times the low-frequency magnitude, with a second "bandwidth" crossing, f_{bw2} at 44 MHz. Increasing C_f to 18 pF reduced the bandwidth further and made the dip larger. By making the zero in H lower than f_{c2}, the lower pole migrates toward it while f_L increases. The response is better

optimized by placing the compensating zero between the two lowest poles of *G*. Then the lower pole migrates upward in frequency to the zero while the higher pole moves even higher.

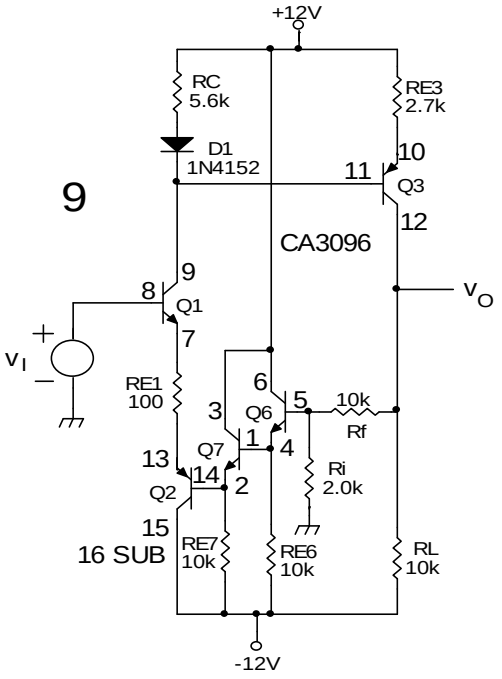
We have seen for a real pole and for a pole-pair how the pole frequencies shift with loop gain. In general (from control theory), the poles of the closed-loop response begin for small loop gain at the open-loop poles. If there is no loop gain there is no feedback, and the amplifier is an open-loop amplifier with open-loop poles. As the loop gain increases, the closed-loop amplifier poles move from the locations of the open-loop poles toward the LHP open-loop zeros and terminate on them. Poles are always real (on the real axis) to the left of an odd number of poles plus zeros in the left half-plane. Closed-loop response cannot have a pole at the origin (because $A(0)$ is always zero or finite, whether G is zero, finite, or infinite), and RHP zeros can be disregarded in determining how the poles migrate in the LHP. The full development of this topic is found in control theory under *root-locus* analysis - the study of pole migration in feedback loops with $G_0 \cdot H_0$ (often called K) as a parameter.

Complementary Differential Amplifier

Amplifier 9 is also based on the versatile CA3096 BJT array. A *complementary diff-pair* stage is followed by a simple current mirror comprised of diode, D1, which roughly matches the *b-e* junction of Q3, a complementary CE stage and current-mirror output. The feedback is different, however, with 2 BJTs configured as a Darlington stage that comprise an active feedback path.

The input stage is a *complementary* BJT diff-pair, Q1 and Q2. No emitter current source is required (though it is a design option) and, being in the same branch, the two BJTs conduct the same current. Unlike unipolar-BJT diff-pairs, however, the V_{BE} voltages add instead of subtract as does temperature drift. The BJTs are in the input loop of the first stage. The base of Q2 is driven by the Darlington stage (Q6, Q7) which is driven by the passive feedback divider. To eliminate bias voltage across the divider, the V_{BE} drops of Q6, Q7 are configured to cancel the V_{BE} of Q1, Q2. This is a suboptimal bias compensation scheme because Q6 and Q7 are both NPN whereas Q2 is PNP and the junctions are not matched. This mismatch is used in setting the first stage current.

The *G* path has two stages, both CE. The second is PNP and is complementary to the NPN first stage. As a feedback amplifier, we again have the choice of feedback and error quantities at the two ends of the loop. For v_f, the choice is easy: $v_f = v_o$. The feedback divider is loaded at its output by the high input resistance of Q6, Q7.



The feedback error quantity, v_B, is the open-circuit output voltage of the Darlington, $v_{E7} = v_{Ho}$. Error summing occurs by KVL around the Q1, Q2 input loop. The error quantity is chosen to be

$$v_E = v_i - v_{Ho}$$

The Darlington gain is in *H* and is very close to one. The input loop of *G* is the Darlington output (as a Thevenin source) in series with the base of Q2. The transresistance of Q1, Q2 across which v_E develops the emitter current of Q1 in the *G* input stage is

$$r_{M1} = r_{e1} + R_{E1} + r_{e2} + R_{B2}'/(\beta_{PNP} + 1)$$

where the incremental output resistance of the Darlington is

$$R_{B2} = R_{E7} \parallel \{r_{e7} + [(r_{b7}' + R_{E6} \parallel (r_{e6} + [r_{b6}' + (R_f \parallel R_i)]/(\beta_{NPN6} + 1))/(\beta_{NPN7} + 1)]\}$$

This unwieldy expression collapses for $R_E >> r_e$ to

$$R_{B2} \approx r_{e7} \Rightarrow R_{B2}' \approx r_{b2}' + r_{e7}$$

The incremental feedback path is found by circuit inspection;

$$H = \frac{v_{Ho}}{v_o} = \left(\frac{R_i}{R_f + R_i} \right) \cdot \left(\frac{R_{E6}}{R_{E6} + r_{e6} + [r_{b6}' + (R_f \parallel R_i)]/(\beta_{NPN} + 1)} \right)$$
$$\cdot \left(\frac{R_{E7}}{R_{E7} + r_{e7} + \{r_{b7}' + R_{E6} \parallel [r_{e6} + [r_{b6}' + (R_f \parallel R_i)]/(\beta_{NPN} + 1)]\}/(\beta_{NPN} + 1)} \right)$$

This evaluates to

$$H = \frac{1}{6} \cdot [(0.9970) \cdot (0.9974)] = \frac{1}{6} \cdot (0.9944) = \frac{1}{6} \cdot \frac{1}{1.006} = \frac{1}{6.034} = 0.1657$$

The β transform reduces the Q6 base resistance of 1717 Ω referred to the emitter of Q6 to

$$r_{E6} = r_{e6} + [r_{b6}' + (R_f \parallel R_i)]/(\beta_{NPN} + 1) = 30.39\ \Omega$$

The emitter-referred base resistance of Q7 is

$$r_{E7} = r_{e7} + \{r_{b7}' + R_{E6} \parallel [r_{e6} + [r_{b6}' + (R_f \parallel R_i)]/(\beta_{NPN} + 1)]\}/(\beta_{NPN} + 1)$$
$$= 26\ \Omega + 0.205\ \Omega = 26.21\ \Omega$$

This leaves Q2 with

$$R_{B2}' = r_{b2}' + r_{E7} = 76.21\ \Omega \Rightarrow R_{B2}'/(\beta_{PNP} + 1) = 1.6\ \Omega$$

Referred to the Q2 emitter from the base circuit, it is 1.6 Ω. Much of the emitter-referred resistance comes from r_{b2}'. The Darlington gain is nearly one and it must be stable for stable closed-loop gain. The Darlington input resistance is very high and does not load the feedback divider appreciably;

$$r_{B6} \approx r_{b6}' + (\beta_{NPN} + 1) \cdot [r_{e6} + r_{b7}' + (\beta_{NPN} + 1) \cdot (r_{e7} + R_{E7})] \approx 1.53\,\mathrm{G}\Omega$$

To more accurately determine the r_e (and r_d) values in the amplifier, static current values are needed and we turn to static analysis. The amplifier is specified to have static input and output voltages of 0 V. The feedback divider is at a static 0 V and v_{B6} will nearly be too except for the small voltage drop of its base current across $R_f \| R_i$. V_{B2} will be about 2 NPN *b-e* junction drops down from ground (with $V_I = 0$ V). As a first iteration, we can approximate $V_{B2} \approx -1.4$ V. Then $I_{E7} \approx 10.6$ V/10 kΩ = 1.06 mA, for which $V_{BE7} = 0.692$ V. Similarly, $I_{E6} \approx 11.3$ V/10 kΩ = 1.13 mA. Then $V_{BE6} = 0.693$ V. Q6 base current is 2.9 μA which drops 4.8 mV across $R_f \| R_i = 1667$ Ω. This voltage subtracts from V_{B2}, making it slightly

more negative. Having values for the emitter currents, we can determine that $r_{e7} \approx r_{e6} \approx 26\,\Omega$.

Returning to the V_{BE} matching of Q6, Q7 with Q1, Q2, and working backwards from the output, for a desired static 0 V out, Q3 current is I_{C3} = 1.2 mA. In the current mirror, if the mismatch of V_{BE3} with D1 is small compared to the emitter resistor voltages, then the mirror incremental current gain is

$$A_i = \frac{i_o}{i_i} = \frac{A}{1 + A/\beta_0}\ ,\ A = \alpha_{03} \cdot \left(\frac{R_C + r_d}{R_{E3} + r_{e3}} \right)$$

For large β_0, this reduces to $A_i = A$;

$$A_i \approx \frac{5.6\ \text{k}\Omega + 44\ \Omega}{2.7\ \text{k}\Omega + 22\ \Omega} = 2.07$$

The diode incremental resistance of 22 Ω was iterated from diode current found as I_{E3} = 1.23 mA divided by A_i. When the more exact formula is used, A_i = 1.986 ≈ 2. From this, Q1 collector current must be close to I_{C1} = (1.23 mA)/2 = 0.615 mA so that $I_{E1} \approx 0.617$ mA. The V_{BE} values are thus

V_{BE1} = 0.69 V + (60 mV/dec)·log(0.615 mA/1 mA) = 0.677 V

V_{BE2} = 0.66 V + (60 mV/dec)·log(0.615 mA/1 mA) = 0.647 V

V_{BE6} = 0.69 V + (60 mV/dec)·log(1.13 mA/1 mA) = 0.693 V

V_{BE7} = 0.69 V + (60 mV/dec)·log(1.06 mA/1 mA) = 0.692 V

Then the voltage across the emitter resistor in the input loop of G is

$$\begin{aligned} V_{RE1} &= -V_{BE1} - V_{BE2} + V_{BE7} + V_{BE6} + I_{B6} \cdot (R_f \parallel R_i) \\ &= -0.677\ \text{V} - 0.647\ \text{V} + 0.692\ \text{V} + 0.693\ \text{V} + 5\ \text{mV} = 66\ \text{mV} \end{aligned}$$

The correct voltage is (100 Ω + 1.6 Ω)·(0.617 mA) = 62.7 mV. The 3 mV surplus will be amplified by the noninverting closed-loop gain and the output will be slightly offset. To correct it (assuming these calculations have been accurate enough), either we will need to decrease R_{E1} to 91 Ω, or the Q1, Q2 current by increasing R_L, or decrease the Darlington currents, or increase the mirror gain, any of which change V_{RE1}. Depending on the output resistance this amplifier is intended to drive, the 1.2 mA output-stage current might best be left unreduced; the open-loop output resistance of this (voltage) amplifier is high;

$$R_L' = R_L \| (R_f + R_i) = 5455\ \Omega$$

If we reduce it, the amplifier r_o is reduced proportionally but the output current range is also reduced. If the voltage surplus across R_{E1} of 3 mV is split evenly between Q6 and Q7, then each must have a reduction in V_{BE} of 1.5 mV. The current factor required to produce this is

$$10^{1.5\ \text{mV}/(60\ \text{mV/dec})} = 1.06$$

This results in a decrease in emitter currents to $I_{E6} = 1.067$ mA and $I_{E7} = 1.001$ mA. The 5 % resistor tolerances and even greater $V_{BE}(I_C)$ tolerances suggest that the static design is close enough to optimum. We will find that static current-mirror error dominates.

The static voltage at the D1 anode is

$$12\ \text{V} - (0.615\ \text{mA}) \cdot (5.6\ \text{k}\Omega) = 8.56\ \text{V}$$

A 1N4152 diode drops 0.63 V at 1 mA. At I_{C1},

$$V_{\text{D1}} = 0.63\ \text{V} + (60\ \text{mV/dec}) \cdot \log(0.615\ \text{mA}/1\ \text{mA}) = 0.617\ \text{V}$$

Then $V_{C1} = V_{B3} = 7.94$ V and $V_{E3} = 8.60$ V; $I_{E3} = 1.26$ mA and $V_O = 0.338$ V, a little high.

Proceeding with the quasistatic design, the recalculated

$$r_{e1} = r_{e2} = 26\ \text{mV}/0.617\ \text{mA} = 42.1\ \Omega$$

$$r_{E1} = 42.4\ \Omega\ ;\ r_{E2} = 43.7\ \Omega \Rightarrow r_{M1} = 186\ \Omega$$

and the quasistatic value of G is

$$G_0 = \alpha_{NPN} \cdot \frac{A_i \cdot R_L}{r_{M1}} \approx (0.997) \cdot \frac{2 \cdot (5455\ \Omega)}{186\ \Omega} = 58$$

For the feedback loop, the quasistatic loop gain is $G_0 \cdot H_0 = 9.613$ - not exactly an op-amp. The closed-loop gain is

$$A_v = \frac{G}{1 + G \cdot H} = \frac{58}{1 + 9.613} = 5.465$$

The 3 mV input offset appears at the output as about a 16 mV offset. Input offset error also occurs in unipolar diff-pair input stages and some op-amps provide pins for adjusting (or *trimming*) the error to zero (thereby *nulling* it).

The gain of this amplifier is easily increased at the expense of stage-gain stability. The second stage A_i can be greatly increased by

setting R_{E3} to zero and removing D1. The static design would depend on feedback to correct gain variation in A_i. The tradeoff for higher loop gain is an increased dependence on feedback to compensate for decreased gain stability in G. Fast amplifiers depend mostly on the open-loop gain of G for bandwidth and make corrections to the static values and thermals through feedback. The delay of feedback loops is excessive for the fastest amplifiers.

As usual, the dynamic design begins by identifying the dominant poles and zeros. For the G path, the OCTCs are

$$\tau_{e1} = \alpha_{NPN} \cdot \tau_{T1} \cdot \frac{R_{B1}' + R_{E1} + r_{E2}}{r_{M1}} = (1.3\,\text{ns}) \cdot \frac{243.7\,\Omega}{186\,\Omega} = 1.70\,\text{ns}$$

$$\Rightarrow 93.4\ \text{MHz}$$

$$\tau_{c1} = [5409\ \Omega + (100\ \Omega)\cdot(1 + 28.5)]\cdot C_{c1} = (8356\ \Omega)\cdot(0.5\ \text{pF}) = 4.18\ \text{ns}$$

$$\Rightarrow 38.1\ \text{MHz}\ ;\ K_{v1} = (0.979) \cdot \frac{5409\,\Omega}{186\,\Omega} = 28.5$$

$$\tau_{L1} = (5409\ \Omega)\cdot(2\ \text{pF} + 2.5\ \text{pF}) = 24.34\ \text{ns} \Rightarrow 6.54\ \text{MHz}$$

$$\tau_{e3} = \alpha_{PNP} \cdot \tau_{T3} \cdot \frac{R_{B3}' + R_{E3}}{r_{M3}} = (24\,\text{ns}) \cdot \frac{5642\,\Omega + 2700\,\Omega}{2840\,\Omega} = 70.5\,\text{ns}$$

$$\Rightarrow 2.26\ \text{MHz}$$

$$\tau_{c3} = [5455\ \Omega + (5642\ \Omega)\cdot(1 + K_{v3})]\cdot C_{c3} = (2207\ \Omega)\cdot(2\ \text{pF}) = 4.41\ \text{ns}$$

$$\Rightarrow 36.1\ \text{MHz}\ ;\ K_{v3} = \alpha_{PNP} \cdot \frac{5642\,\Omega}{2840\,\Omega} = 1.95$$

$$\tau_{L3} = (5455\ \Omega)\cdot(15\ \text{pF}) = 81.8\ \text{ns} \Rightarrow 1.95\ \text{MHz}$$

The H divider r_{Ho} forms an output OCTC with C_{c6} at

$$\tau_{Ho} = (1717\ \Omega)\cdot(0.5\ \text{pF}) = 0.86\ \text{ns} \Rightarrow 185\ \text{MHz}$$

With large R_E values, the Darlington H stage is fast, with OCTCs of $\tau_c = r_b' \cdot C_c$ and $\tau_e \approx \alpha_0 \cdot \tau_T$. In the hf region, the Darlington stage gyrates the parallel R, C of z_{Ho} by +180° into a parallel –R, L at the Q2 base. At the Q2 emitter, Z_{b2} is gyrated to a –R, C. A quick calculation gives a value for –R of

$$-\tau_T / C_{c6} = -(568\ \text{ps})/(0.5\ \text{pF}) = -1136\ \Omega$$

and is much less than the positive resistance in the input loop. This is an indication of possible instability. However, the R_E of Q6 and Q7 are large and have parasitic capacitance of about 1/3 pF. Then $\tau_E = (10\ \text{k}\Omega)\cdot(0.33\ \text{pF}) = 3.3$ ns, well below τ_T for α compensation with additional damping of the next CC stage. The excess parasitic C_E and C_c at the base of the next stage results in a gyrated C becoming a dominant parallel R at its emitter.

For Q2, which is in the *G* block as part of the input stage,

$$\tau_{e2} = (24\ \text{ns})\cdot(1.038) = 25\ \text{ns} \Rightarrow 6.4\ \text{MHz}$$

Slow PNP BJTs are the dominant speed constraint of the CA3096.

With the OCTCs identified, we omit the high-frequency OCTC frequencies that are a decade or more above the lowest to leave those affecting dynamic response. *H* has none, and *G* has four, at

$$f_{L1} = 6.54\ \text{MHz}\ ;\ f_{e3} = 2.26\ \text{MHz}\ ;\ f_{L3} = 1.95\ \text{MHz}\ ;\ f_{e2} = 6.4\ \text{MHz}$$

Four closely-spaced poles in a feedback loop are likely to cause instability.

The prototype of amplifier 9 did not oscillate when powered on (after forgotten bypass capacitors were added!). As expected from the static mirror error, the output was offset, by 0.14 V. The mirror input node (pins 9, 11) was at 7.78 V (–2 % of the calculated 7.94 V) and the difference between the V_{D1} and V_{BE3} was about 20 mV, with the Q3 emitter higher than the diode anode, at 8.51 V (8.60 V predicted). V_{B6} measured 13 mV and there was about 60 mV across R_{E1}.

The quasistatic gain was measured at 50 kHz with 400 mV pk-pk input. The output was 2.12 V pk-pk for a gain of 5.3, 3 % below the calculated value.

The frequency response was highly underdamped, with $M_{mcl} = 3.0$ at $f_{mcl} = 3.02$ MHz. The gain rolled off to bandwidth at f_{bw}(cl) = 5.55 MHz. $\zeta_{cl} = 0.169 \Rightarrow$ pole angle, $\phi = 80.3°$;

$$f_{ncl} = f_{mcl} / \sqrt{1 - 2\cdot\zeta^2} = 3.11\ \text{MHz}$$

No attempt was made during the design to derive the closed-loop quadratic pole parameters. (The four real poles, with loop gain, leave the real axis and result in two complex pole-pairs at different f_n, requiring control theory to analyze.) Had the prototype circuit oscillated, the above data would have been difficult to obtain. With it, we have the design option of refining the dynamic response from measured data. This is not unlike what is often done with circuit

simulations. It demonstrates a variation in design style. Previous designs proceeded with a maximum amount of derivation before turning to the bench. This is a generally preferred procedure, for maximum forethought about a design often reduces the time required to find its subtle nuances, no less effect its modifications (even in simulation). Engineers vary in how much theoretical work they perform before building and testing a prototype or simulating it.

To combine previous theoretical insights into amplifier 9 with experimental refinement, we know that some of the multiple poles in G must be cancelled with zeros in the form of frequency compensation circuits. Two possibilities are Q3 emitter peaking and adding C_f around R_f in the feedback divider. Both add zeros and high-frequency poles to the loop gain. The zeros need to be added near - preferably between - the poles. The number of poles involved in the loop makes this a control-intensive problem, but we can address it experimentally instead.

From measurement, we know f_{ncl} and ζ_{cl}. If ζ_{cl} were about 1, as desired, then real closed-loop poles would be around f_{ncl}. The closed-loop parameters can be unraveled to recover the loop-gain parameters by compensating the dominant resonance. Invoking

$$\sqrt{1+G_0 \cdot H_0} \approx 3.26$$

the open-loop pole-pair parameters are

$$f_n = \frac{f_{ncl}}{\sqrt{1+G_0 \cdot H_0}} = \frac{3.11\,\text{MHz}}{3.26} = 955\ \text{kHz}\ ;$$

$$\zeta = \zeta_{cl} \cdot \sqrt{1+G_0 \cdot H_0} = (0.169) \cdot (3.26) = 0.551 \Rightarrow 56.6°$$

The open loop is underdamped. Two or more of the identified poles combine to form underdamped pole-pairs.

Proceeding on the basis that circuit measurement has led us to a dominant pole-pair having the above parameters, a zero is placed at f_n by adding C_{E3} to form a time constant of 167 ns corresponding to 955 kHz;

$$C_{E3} = 167\ \text{ns}/2700\ \Omega = 62\ \text{pF}$$

The resulting response is somewhat less highly underdamped; more zeros are needed.

Variable (or *trim*) capacitors were placed across R_f and R_{E3}, and adjusted for flat response out to maximum bandwidth. Manual

adjustment of the two parameters is iterative (though not difficult) and resulted in the following measured values of capacitances:

$$C_{E3} = 3.3 \text{ pF} \; ; \; C_f = 6.8 \text{ pF}$$

These values place zeros at

$$z_G = 18 \text{ MHz} \; ; \; z_f = 2.34 \text{ MHz}$$

(The resulting pole in the Q3 emitter is over a gigahertz.) The response now has 1 to 3 % magnitude variation out to a bandwidth of f_{bw}(cl) = 5.5 MHz.

The 18 MHz zero begins to have an effect on phase as low as 1.8 MHz, and the lower zero similarly reaches down a decade to 234 kHz. The phase-lead effect of these zeros extends far above bandwidth. The excessive peaking was trimmed by z_G while z_f flattened the response in the lower frequency range up to about 2 MHz.

Although the above testing could have been simulated, it is more cumbersome to try to adjust parameters and observe effects on a computer simulator. In real time on the bench, these adjustments and the observation of their effects took about a minute or two. The circuit construction took somewhat more time than entering the circuit diagram and configuring a simulator, but not much. For circuits of comparable complexity to those in this book, bench prototyping still has its advantages.

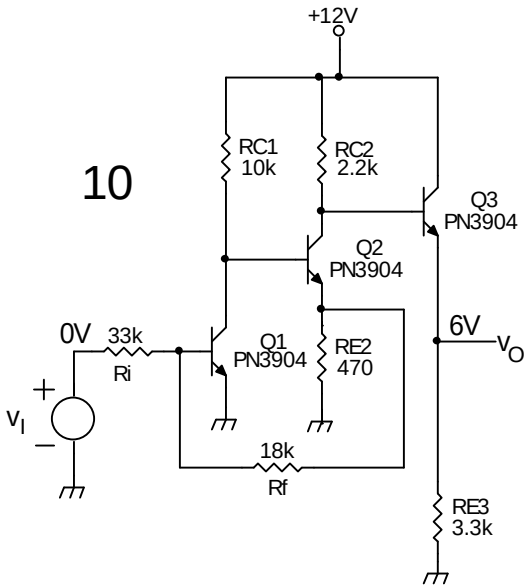
Does amplifier 9 have any particular merits? It is being left in a somewhat unrefined state. The more thorough set of design considerations applied to the cascode amplifiers - especially amplifier 3 - provides a template for further design-oriented analysis of this amplifier. Some of its disadvantages are: its input stage adds rather than subtracts undesirable V_{BE} effects (of Q1 and Q2), its current mirror V_{BE} tracking offsets the output by over 100 mV, and its dynamic response is complicated. It does have at least a ±5 V output range, better PSR than some previous amplifiers, and fairly stable loop gain. It can be improved by reducing the current-mirror resistor values to reduce time constants and also increase positive range; by adding some thermal balancing resistors (with bypass capacitors); and by adjusting for better V_{BE} compensation at the input. As in previous CA3096-based amplifiers, f_{bw} is limited by Q3 and its limited f_T.

The complementary diff-pair input stage seems like an inferior circuit that should not be used in good design. This is certainly true of its thermal drift characteristics. However, the complementary diff-pair

input stage inherently extends input range when the Q2 base voltage moves along with the input voltage to accommodate it. With another two PNP BJTs, it would have been possible to replace D1 and Q7 with them to improve *b-e* matching. We might conclude that this amplifier is better used in applications with relatively large input and output waveform requirements, undemanding static stability specifications, and requiring moderate bandwidth.

Single-Supply Feedback Amplifier

The single-supply feedback amplifier 10, shown below, illustrates the use of T_o in the general feedback scheme. It has a unipolar (positive) output voltage range and a bipolar input-voltage range, though it uses only one supply voltage.



As a single-supply amplifier, the output static voltage is chosen for wide voltage range, midway between the supply voltage of 12 V and ground. Working through the static design, given V_O = 6 V, then I_{E3} = 6 V/3.3 kΩ = 1.82 mA. (Of course, this is analysis of the given design. We would have chosen I_{E3} and calculated R_{E3}.) For a PN3904,

$$V_{BE3} \approx (0.65\text{ V})\cdot(60\text{ mV/dec})\cdot\log(1.82) \approx 0.67\text{ V}$$

I_{C2} = (12 V – 6.67 V)/(2.2 kΩ) = 2.42 mA and $I_{E2} = I_{C2}/\alpha_2$ = 2.44 mA. At the Q2 emitter, V_{E2} = 1.15 V. V_{BE2} is greater than 0.65 V (at 1 mA) by (60 mV/dec)·log(2.44 mA/1 mA) = 23 mV and V_{C1} = 1.82 V. Then I_{C1} = 1.02 mA, $r_{e1} \approx 26\ \Omega$ and $V_{B1} \approx 0.65$ V. Given this circuit framework, the BJT currents and voltages can be adjusted in the above equations. A tradeoff is voltage gain against range. This design does not leave much low-end range for the Q1 collector voltage, and the gain is higher as a consequence of a larger R_{C1}.

The biasing of Q1 is affected by its base current and feedback divider, R_f, R_i. $I_{B1} = I_{C1}/\beta_0 \approx 1.02$ mA/150 ≈ 6.8 μA. The static current flowing into the v_i source (at 0V) from the Q1 base node is 0.65 V/R_i = 19.7 μA. Current through R_f must then be 26.5 μA and

$$R_f = (1.15\text{ V} - 0.65\text{ V})/26.5\ \mu\text{A} = 18.87\ \text{k}\Omega \text{ or } 18\ \text{k}\Omega \pm 5\ \%$$

The static design forces the quasistatic voltage gain to the Q2 emitter to be about (18 kΩ/33 kΩ) ≈ 0.55 if the loop gain is high. Nominal bias current with the chosen R_f value is 27.8 μA. (Note: *nominal* does not mean the same as *typical*. *Nominal* is what is true "as a rule" whereas *typical* is what represents the entire category. Nominal values are what are desired whereas typical values are what generally can be expected.)

Both cascade CE stages (Q1, Q2) have NPN BJTs and thus must share output voltage range. The first-stage waveforms are not as large in amplitude as stage 2, and most of the range optimally goes to Q2. The Q3 input range is $V_{BE1} + V_{BE3}$ to 12 V. The Q2 gain is about

$$-(0.993)\cdot(2.2\ \text{k}\Omega/(10.6\ \Omega + 470\ \Omega \| 18\ \text{k}\Omega)) = -4.66$$

The low end of the Q2 collector range is at V_{CB2} = 0 V or when $v_{C2} = v_{B2}$:

$$v_{C2} = V_{CC} - K_{v2} \cdot v_{B2} = v_{B2}$$

Solving, $v_{B2} = V_{CC}/(K_{v2} + 1)$ = 2.12 V. The Q2 output-voltage range is maximum (as derived in "Design Refinement of a CE Amplifier" in the chapter, "Transistor Dynamics") when

$$V_{C2} = \left(\frac{K_{v2} + 1}{2 \cdot K_{v2} + 1} \right) \cdot (V_{CC} - V_{B2}) + V_{B2}$$

for nonzero base bias voltage V_{B2}. The optimal V_{C2} = 7.42 V. To meet the static design requirement of 6 V out, the range is reduced

somewhat. Working from this constraint, $V_{C2} = 6.67$ V and the v_{C2} negative range is

$$V_{C2} - V_{C1} = 6.67 \text{ V} - 1.82 \text{ V} = 4.85 \text{ V}$$

The maximum symmetrical range is ±4.85 V. The minimum voltage of the Q1 collector range is limited by $V_{BE2} = 0.67$ V. Its symmetrical range is $\pm(V_{C1} - V_{BE2}) = \pm 1.15$ V. The maximum range of the Q2 collector waveform allowed by the Q1 range is

$$(4.66)\cdot(\pm 1.15 \text{ V}) = \pm 5.36 \text{ V} > \pm 4.85 \text{ V}$$

The range of Q1 will not constrain the range of Q2.

The feedback loop error quantity is that of an inverting op-amp, chosen to be added by superposition of v_i and v_f across opposite ends of the feedback divider. The G input is the error voltage that drives the base of Q1,

$$v_E = T_i \cdot v_i - H \cdot v_o$$

This voltage is the divider output at the Q1 base node, unloaded by the Q1 base. The effect of base loading is included in G in that the gain of Q1 includes the Thevenin equivalent divider output resistance, $r_{Ho} = R_f \| R_i$, as the emitter-referred base resistance in r_{M1}.

The G input is preceded by the thevenized source with attenuation by the feedback-divider resistors;

$$T_i = \frac{R_f}{R_f + R_i} = 0.353$$

In reverse, the feedback-divider has a Thevenin equivalent source voltage of $v_f = v_{e2}$ attenuated by the feedback divider,

$$H_0 = -R_i/(R_f + R_i) = -0.647$$

The Q1 external base resistance is the Thevenin output resistance of the feedback divider. It is included in the G_{01} calculation, allowing H_0 to be calculated as unloaded (open-circuited at the Q1 base). (See "Two-Port Loading" in the chapter, "Feedback Amplifiers".)

The quasistatic gain of the first stage, with input loading on $R_f \| R_i$ and the output loading of the Q2 stage, is

$$G_{01} = -\alpha_1 \cdot \frac{R_{C1} \| (\beta_2 + 1) \cdot (r_{e2} + R_{E2} \| R_f)}{r_{e1} + (R_i \| R_f)/(\beta_1 + 1)} = -(0.993) \cdot \frac{8.76 \text{ k}\Omega}{103 \Omega} = -84.4$$

Choose the feedback quantity to be $v_f = v_{e2}$; then

$$T_o = \frac{v_o}{v_{e2}} = T_{o1} \cdot T_{o2}$$

The second stage of the feedback loop has the quasistatic gain

$$G_{02} = \frac{v_f}{v_{c1}} = \frac{v_{e2}}{v_{c1}} = \frac{R_{E2} \parallel R_f}{R_{E2} \parallel R_f + r_{e2}} = \frac{458\,\Omega}{469\,\Omega} = 0.977$$

In the denominator, r_{M2} does not include emitter-referred base resistance of R_{C1} because interstage loading was included in G_{01}. The output of the loop is v_f, though it is not the output of the amplifier. The forward-path gain is

$$G_0 = G_{01} \cdot G_{02} = (-84.4) \cdot (0.977) = -82.5$$

The quasistatic loop gain is

$$G_0 \cdot H_0 = (-82.5) \cdot (-0.647) = 53.38$$

The closed-loop quasistatic voltage gain is thus

$$\frac{v_{e2}}{v_i} = T_i \cdot \frac{G_0}{1 + G_0 \cdot H_0} = (0.353) \cdot \frac{-82.5}{1 + 53.38} = (0.353) \cdot (-1.517) = -0.536$$

With infinite loop gain, the result would instead have been –0.546.

Q2 is both a CC second stage of G and a CB first stage of T_o. The input quantity of T_o is the output quantity of G, and is the feedback quantity, v_{e2}. Thus, the first stage gain of T_o, or T_{o1}, is from the Q2 emitter to its collector. The Q2 gain from v_{e2} to v_{c2} (with output loading from Q3) is

$$T_{o1} = \alpha_2 \cdot \frac{R_{C2} \parallel (\beta_3 + 1) \cdot (r_{e3} + R_{E3})}{R_{E2} \parallel R_f} = (0.993) \cdot \frac{2.190\,\text{k}\Omega}{458\,\Omega} = 4.75$$

The CC second-stage gain is

$$T_{o2} = 3.3\,\text{k}\Omega / (3.3\,\text{k}\Omega + 26\,\text{mV}/1.82\,\text{mA}) = 0.996$$

Then

$$T_o = T_{o1} \cdot T_{o2} = (4.75) \cdot (0.996) = 4.73$$

The total amplifier gain is

$$A_v = \frac{v_o}{v_i} = T_i \cdot \frac{G}{1 + G \cdot H} \cdot T_o = (-0.536) \cdot (4.73) = -2.54$$

From the static analysis, the biasing of this amplifier is very β-dependent on Q1. If the base current is not quite correct, a small change in base current will cause an amplified change in V_{E2}, correcting the bias offset. To minimize I_{B1} sensitivity, R_f and R_i should not be made large. With a loop gain of only 53.4, the offset error can be as much as 2 %. R_{C1} can be increased for added gain at the expense of speed.

The addition of an R_{C3} in the Q3 collector can reduce Q3 power dissipation and thermals. The base resistance of Q3 (R_{C2}) damps parasitic inductance in the Q3 collector supply wire or circuit-board trace that can resonate with C_{c3} if not adequately bypassed. A 10 cm conductor has about 10 nH/cm or 100 nH. It resonates in series with C_{c3} = 4 pF at 252 MHz ($< f_T$) and takes at least 316 Ω to critically damp. A bypass capacitor at the Q3 collector to ground reduces Z_n and f_n to where oscillation of Q3 from stray collector inductance does not occur.

A prototype circuit was built from junk parts including two with a silver (10 %) tolerance band. (The resistance of these parts was later measured for a more accurate calculation of circuit behavior.) Nowadays, 5 % carbon film resistors typically measure to better than 1 % tolerance, allowing for more accurate circuit design.

The static voltages were roughly as expected; V_O was not 6 V but was low: 5.25 V. V_{C2} measured 5.89 V (instead of 6.67 V), V_{E2} was 1.18 V (1.15 V calculated), V_{C1} measured 1.866 V instead of 1.82 V, and V_{BE1} was 0.668 V instead of 0.65 V. Unless matched for V_{BE}, BJTs of the same part number from the factory can vary in V_{BE} by as much as 50 mV at the same current - all the more reason to not depend on unmatched V_{BE} values in static design!

The quasistatic gain at 1 kHz with ±0.2 V input produced ±0.615 V out, for a gain of –3.08. This is considerably more gain than calculated, and resistor values were measured: R_f = 19.1 kΩ; R_i = 32.3 kΩ; R_{C2} = 2.41 kΩ; and R_{E2} = 464 Ω. When these values are substituted into the design equations, the resulting G_0 = –82.5. Also, H_0 = –0.630; $G_0 \cdot H_0$ = 52; T_i = 0.370; T_o = 5.18 and the loop itself, $G_0/(1 + G_0 \cdot H_0)$ = –1.594. Then from input to loop output, v_{e2}, the gain calculates to be –0.59. The measured gain at v_{e2} was –0.588. To the output, the recalculated gain is –3.06. This compares favorably to the measured –3.08. Do not always believe resistor markings.

The output had a slightly asymmetrical range of Δ7.5 V with the low end of the range saturating first. Q1 did not saturate before Q2. No slewing was observed at the output. These characteristics were not

addressed in this design, though in product-design projects, they should be verified.

Single-Supply Amplifier Dynamics

Turning to dynamics, the base-collector voltage gains, K_v, are needed for the Miller effect in the time-constant calculations. PN3904 β_0 is typically 150 over the currents of Q1 - Q3. For the first stage, with R_{E1} = 0 Ω, the voltage gain is high and with such a high gain in one CE stage, it is likely that Q1 r_o has some influence on the collector resistance. It will be included. PN3904 Early voltage, $V_A \approx 100$ V. At 1 mA of current and about 2 V = V_{CB1}, then

$$r_{o1} = \frac{100\,\text{V} + 2\,\text{V}}{1\,\text{mA}} = 102\ \text{k}\Omega$$

$R_{B1}' = r_b' + R_i||R_f$ = 50 Ω + 11.65 kΩ = 11.70 kΩ

$R_{b1} = R_{B1}'||(\beta_0 + 1)\cdot r_{e1}$ = (11.70 kΩ)||(151)·(26 Ω) = 2939 Ω

$r_{M1} = R_{B1}'/(\beta_0 + 1) + r_{e1}$ = (11.70 kΩ)/(151) + 26 Ω = 103.5 Ω

$R_{E2}' = R_{E2}||R_f$ = (470 Ω)||(18 kΩ) = 458 Ω

$R_{B2}' = R_{C1}||r_{o1} + r_b'$ = (10 kΩ)||(102 kΩ) + 50 Ω = 9157 Ω

$R_{C1}' = R_{b2} = R_{B2}'||(\beta_0 + 1)\cdot(r_{e2} + R_{E2}')$
= (9157 Ω)||(151)·(10.7 Ω + 458 Ω) = (9157 Ω)||(70.77 kΩ) = 8108 Ω

$$K_{v1} = \alpha_0 \cdot \frac{R_{C1}'}{r_{e1}} = (0.9934) \cdot \frac{8108\,\Omega}{26\,\Omega} = 309.8 \approx 310$$

$R_{C2}' = R_{C2}||(\beta_0 + 1)\cdot(r_{e3} + R_{E3})$ = (2.2 kΩ)||(151)·(3314 Ω) = 2190 Ω

$$K_{v2} = \alpha_0 \cdot \frac{R_{C2}'}{r_{e2} + R_{E2}'} = (0.9934) \cdot \frac{2190\,\Omega}{468.7\,\Omega} = 4.64$$

$r_{M2} = R_{C1}'/(\beta_0 + 1) + r_{e2} + R_{E2}'$ = 53.7 Ω + 10.7 Ω + 458 Ω = 522 Ω

$r_{M3} = R_{C2}/(\beta_0 + 1) + r_{e3} + R_{E3}$ = 3328 Ω

The list of 2 OCTCs per BJT and their frequencies is as follows:

$\tau_{e1} = \alpha_0 \cdot \tau_T \cdot (R_{B1}'/r_{M1})$ = (0.993)·(707 ps)·(11.70 kΩ/103.5 Ω) = 79.4 ns
$\Rightarrow$ 2.00 MHz

$\tau_{c1} = (R_{C1}' + R_{b1}\cdot(1 + K_{v1}))\cdot C_{c1}$

$= [8108\ \Omega + (2939\ \Omega)\cdot(1 + 310)]\cdot(2.5\ \text{pF}) = 2.305\ \mu\text{s} \Rightarrow 69.0\ \text{kHz}$

$\tau_{e2} = \alpha_0\cdot\tau_T\cdot((R_{B2}' + R_{E2}')/r_{M2})$

$= (0.993)\cdot(665\ \text{ps})\cdot[(9157\ \Omega + 458\ \Omega)/522\ \Omega] = 12.17\ \text{ns}$
$\Rightarrow 13.1\ \text{MHz}$

$\tau_{c2} = (R_{C2}' + R_{b2}\cdot(1 + K_{v2}))\cdot C_{c2}$
$= (2190\ \Omega + (8108\ \Omega)\cdot(1 + 4.64))\cdot(2\ \text{pF}) = 95.84\ \text{ns} \Rightarrow 1.66\ \text{MHz}$

$\tau_{L2} = R_{C2}'\cdot C_{c3} = (2190\ \Omega)\cdot(2\ \text{pF}) = 4.38\ \text{ns} \Rightarrow 36.3\ \text{MHz}$

$\tau_{e3} = \alpha_0\cdot\tau_T\cdot((R_{C2}' + R_{E3})/r_{M3})$
$= (0.993)\cdot(665\ \text{ps})\cdot((2190\ \Omega + 3300\ \Omega)/3328\ \Omega) = 1.09\ \text{ns}$
$\Rightarrow 146\ \text{MHz}$

Quasistatic partitioning of circuits into blocks of the feedback diagram can be challenging for circuits with spaghetti-like interconnections. It can be even more difficult to separate dynamic effects into the right blocks. The OCTCs of G include τ_{e1}, τ_{c1}, and τ_{e2}. Those of $T_o(s)$ include τ_{L2} and τ_{e3}. What of τ_{c2}? Is it in G or T_o? It will only be in one or the other because it is only one time constant. C_{c2} appears as Q2 base capacitance that forms a time constant with R_{C1}'. It also involves R_{C2}'. If it is within the loop, loop gain will cause its pole to migrate. If in T_o, it is open-loop and is fixed in the amplifier transfer function at its calculated value.

We know from earlier circuit analysis that C_c does not affect the collector-node time constant but it does have a Miller effect at the base. The base of Q2 is in the feedback loop and C_c will affect the loop behavior. Thus it is in the loop and is in G.

H has no poles or zeros. T_i also has no poles because τ_{e1} is in G. Both T_i and H consist of Thevenized feedback-divider ratios and are only resistive.

The output block, T_o, has two real, open-loop (non-migrating) poles. Its OCTC bandwidth is

$$\tau_{bw} = \sqrt{(4.38\ \text{ns})^2 + (1.09\ \text{ns})^2} = 4.51\ \text{ns} \Rightarrow 35.3\ \text{MHz}$$

The feedback-loop OCTCs are all in G and are at frequencies of $f_{c1} = 69$ kHz, $f_{c2} = 1.66$ MHz, $f_{e1} = 2.00$ MHz, and $f_{e2} = 13.1$ MHz. The open loop has all real poles and the OCTC bandwidth of $G\cdot H$ is

$$\tau_{bw} = \sqrt{(79.4\ \text{ns})^2 + (2.305\ \mu\text{s})^2 + (12.17\ \text{ns})^2 + (95.84\ \text{ns})^2} =$$
$$2.308\ \mu\text{s} \Rightarrow 69\ \text{kHz}$$

f_{c1} dominates the loop dynamics. It is 1.38 decades below the next lowest pole (f_{c2}). If the loop is approximated by a dominant single pole at f_{bw}, then

$$f_{bw}(\text{cl}) \approx f_{bw} \cdot (1 + G_0 \cdot H_0) = 3.68 \text{ MHz}$$

Using instead the more accurate measured resistor values, $f_{bw}(\text{cl}) = 3.64$ MHz - hardly much difference. When $f_{bw}(T_o)$ is included, the amplifier bandwidth is still 3.64 MHz.

The number of circuit poles equals the number of OCTCs, but the OCTC frequencies are not pole frequencies if their reactances interact. The BJT C_e and C_c capacitances interact and for more accurate design cannot be regarded as independent; their OCTCs cannot be regarded as poles. The poles of Q1 (treating them as though they were isolated from Q2) are found by finding τ_{n1} and ζ_1;

$$\tau_{n1} = \sqrt{\tau_{e1} \cdot \tau_{cc1}} = 40.12 \text{ ns}$$

$$\zeta_1 = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_e}{\tau_{cc}}} + \sqrt{\frac{\tau_{cc}}{\tau_e}} + \frac{\tau_{cb}}{\tau_n} \right) = 29.72$$

The pole frequencies, in Hertz, are

$$\frac{1}{2 \cdot \pi} \cdot \frac{\zeta_1}{\tau_{n1}} \cdot \left(1 \pm \frac{\sqrt{\zeta_1^2 - 1}}{\zeta_1} \right) = 66.76 \text{ kHz}, 235.8 \text{ MHz}$$

The poles are widely separated by 3.55 decades. Invoking the same general equations for Q2 (as though it were isolated from Q1 and Q3),

$$\tau_{n2} = 7.30 \text{ ns} \; ; \; \zeta_2 = 7.397 \Rightarrow 1.48 \text{ MHz}, 321 \text{ MHz}$$

A wide separation of poles is typical in single-stage CE amplifiers. Consequently, the dominant pole can be used in bandwidth calculation. Using the lower-frequency poles of the two stages,

$$f_{bw} \approx 1/\sqrt{(1/(66.76 \text{ kHz})^2 + 1/(1.48 \text{ MHz})^2} = 66.7 \text{ kHz}$$

The result is slightly lower than the OCTC bandwidth. Then the closed-loop single-pole bandwidth is

$$f_{bw}(\text{cl}) \approx f_{bw} \cdot (1 + G_0 \cdot H_0) = (66.76 \text{ kHz}) \cdot (54.4) = 3.63 \text{ MHz}$$

The closed-loop bandwidth calculations are in close agreement.

The measured closed-loop frequency response of the amplifier 10 prototype was flat to

$$\text{measured } f_{bw} = 2.85 \text{ MHz}$$

This is 22 % lower than the last calculated value. The bandwidth to the loop output at v_{e2} measured 2.95 MHz. The risetime measured 114 ns and the step response was without overshoot, corresponding to a single-pole-equivalent bandwidth of 0.35/(114 ns) = 3.07 MHz at the output. The 'scope probe has little effect on any of the critical nodes in the circuit because the circuit capacitances are large relative to 15 pF of probe capacitance. At emitters, the node resistance is too small to matter.

With increasing loop gain, the four real poles come together in pairs on the *s*-plane, then leave the real axis and the closed-loop response becomes underdamped. The lower-frequency pair will dominate the response as they arc toward the right-half plane. The higher pair arcs toward the left, to even higher frequencies, and does not have a significant influence on the response.

Feedback Damping

We have yet to predict the shape of the frequency or step response, as expressed by the damping. As a first attempt, we assume that the two dominant-stage poles, at 66.76 kHz and 1.48 MHz, are real and independent poles of the feedback loop. Thus, they are regarded as an open-loop pole-pair of the loop-gain pole polynomial,

$$D(s) = s^2 \cdot (\tau_1 \cdot \tau_2) + s \cdot (\tau_1 + \tau_2) + 1$$

$D(s)$ factors exactly into the two poles because they are independent. The general expression for $D(s)$ in quadratic parameters τ_n and ζ is

$$D(s) = s^2 \cdot \tau_n^2 + s \cdot (2 \cdot \zeta \cdot \tau_n) + 1$$

Next, express τ_n and ζ in OCTCs τ_1, τ_2, as

$$\tau_n = \sqrt{\tau_1 \cdot \tau_2} \;\; ; \;\; \zeta = \frac{\tau_1 + \tau_2}{2 \cdot \tau_n} = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_1}{\tau_2}} + \sqrt{\frac{\tau_2}{\tau_1}} \right)$$

Let s be in Hz and apply the conversion of $\tau = 1/2 \cdot \pi \cdot f$:

$$f_n = \sqrt{f_1 \cdot f_2} \;\; ; \;\; \zeta = \frac{1}{2} \cdot \left(\sqrt{\frac{f_2}{f_1}} + \sqrt{\frac{f_1}{f_2}} \right)$$

These are the open-loop parameters.

For the feedback loop, the closed-loop pole polynomial from the quadratic feedback formula is

$$s^2 \cdot \left(\frac{\tau_n}{\sqrt{1 + G_0 \cdot H_0}} \right)^2 + s \cdot \left(2 \cdot \frac{\zeta}{\sqrt{1 + G_0 \cdot H_0}} \cdot \frac{\tau_n}{\sqrt{1 + G_0 \cdot H_0}} \right) + 1$$

The closed-loop pole parameters are

$$\tau_{ncl} = \frac{\tau_n}{\sqrt{1 + G_0 \cdot H_0}} \Rightarrow f_{ncl} = f_n \cdot \sqrt{1 + G_0 \cdot H_0}$$

$$\zeta_{cl} = \frac{\zeta}{\sqrt{1 + G_0 \cdot H_0}} = \frac{\left(\sqrt{\frac{f_2}{f_1}} + \sqrt{\frac{f_1}{f_2}} \right)}{2 \cdot \sqrt{1 + G_0 \cdot H_0}}$$

The pole magnitude from the s^2 coefficient is

$$f_{ncl} = \sqrt{f_{c1} \cdot f_{c2}} \cdot \sqrt{1 + G_0 \cdot H_0} = (314.3\,\text{kHz}) \cdot \sqrt{54.4} = 2.32\,\text{MHz}$$

Given that $f_{c2}/f_{c1} = 22.17$, the damping is

$$\zeta_{cl} = \frac{\left(\sqrt{\frac{f_2}{f_1}} + \sqrt{\frac{f_1}{f_2}} \right)}{2 \cdot \sqrt{1 + G_0 \cdot H_0}} = 0.334 \Rightarrow 70.5°$$

The closed-loop bandwidth from the quadratic bandwidth formula (at the end of the "Circuit Dynamics" chapter) follows as

$$f_{bw}(\text{cl}) = f_{ncl} \cdot \sqrt{1 - 2 \cdot \zeta_{cl}^2 + \sqrt{4 \cdot \zeta_{cl}^4 - 4 \cdot \zeta_{cl}^2 + 2}} = (1.43) \cdot f_{ncl} = 3.32\,\text{MHz}$$

This is high by about 16 %.

An equivalent derivation of damping applies the pole separation equation, for which the pole separation ratio,

$$p_2/p_1 = 1.48\,\text{MHz}/66.76\,\text{kHz} = 22.17$$

$$\zeta_{cl} = \frac{\zeta}{\sqrt{1 + G_0 \cdot H_0}} = \frac{\frac{1}{2} \cdot \sqrt{\frac{[(p_2/p_1) + 1]^2}{p_2/p_1}}}{\sqrt{1 + G_0 \cdot H_0}} = \frac{2.460}{7.376} = 0.334$$

This equation gives the same value because it also assumes that p_1, p_2 are independent open-loop poles. The low value of calculated damping has a high pole angle that should result in an oscillatory step response and peaking in the magnitude of the frequency response. Yet both time and frequency responses are flat and show no peaking, indicating a major disagreement between the above theoretical results and what is actually observed of amplifier 10. We can conclude that the amplifier loop is not approximated well by assuming that the lowest two open-loop poles are independent. The discrepancy lies in the assumption that the cascade stages of Q1 and Q2 are isolated.

From "Cascade CE Stages" in "Amplifier Design", we found that cascade stages interact to produce real poles, and that the quadratic coefficient, a, is less for interacting poles, making ζ larger than for independent poles. We can use the cascade-CE solution from that section to derive the pole locations in G. Hence, we need one more time constant, the SCTC, $\tau_{c2;1}$, which is τ_{c2} with C_{c1} shorted. With it, we can calculate the interaction and OCTC pole separation factors,

$$k = \frac{\tau_{c2}}{\tau_{c2;1}} \text{ , } k \geq 1 \text{ ; } \xi = \sqrt{\frac{\tau_{c1}}{\tau_{c2}}}$$

The damping is

$$\zeta(\xi, k) = \frac{\sqrt{k}}{2} \cdot \left(\xi + \frac{1}{\xi} \right)$$

The derivation of $R_{c2;1}$ in $\tau_{c2;1} = R_{c2;1} \cdot C_{c2}$ from circuit elements is also given (modified for amplifier 10) as

$$R_{c2;1} \approx R_{C2}' + [R_{B2}' + (R_{C1}' \| r_{M1} \| R_{B1}')] \cdot \left(1 + \frac{R_{L2}'}{r_{M2}} \right)$$

Then substituting amplifier 10 values,

$R_{c2;1} \approx 2190\ \Omega +$

$$[9157\ \Omega + (8108\ \Omega \| 103.5\ \Omega \| 11.70\ \text{k}\Omega)] \cdot (1 + 2190\ \Omega / 522\ \Omega)$$

$$= 2190\ \Omega + [9157\ \Omega + (101.3\ \Omega)] \cdot (5.195) = 50.29\ \text{k}\Omega$$

From this,

$$\tau_{c2;1} = R_{c2;1} \cdot C_{c2} = (50.29\ \text{k}\Omega) \cdot (2\ \text{pF}) = 100.6\ \text{ns}$$

The cascade CE derivation assumes fast ($\tau_T = 0$ s) BJTs having a single pole at τ_c. We included τ_e in calculation of the dominant single

poles of the stages and will use the stage dominant-pole values instead. They calculate as time constants to be

$\tau_{c1} = 1/2\pi\cdot(66.76\text{ kHz}) = 2.384\ \mu\text{s}$; $\tau_{c2} = 1/2\cdot\pi\cdot(1.48\text{ MHz}) = 107.5\text{ ns}$

Then the interaction and separation factors are

$$k = \frac{\tau_{c2}}{\tau_{c2;1}} = 1.069$$

$$\xi = \sqrt{p_2 / p_1} = \sqrt{1.48\,\text{MHz}/66.76\,\text{kHz}} = \sqrt{22.17} = 4.708$$

Finally, the damping is

$$\zeta_{cl}(\xi, k) = \frac{\sqrt{k}}{2}\cdot\left(\xi + \frac{1}{\xi}\right) = 2.544$$

This is an overdamped response with (interacting) poles separated by a factor of 23.84. This result can be taken as only approximate because the OCTCs were replaced by the stage dominant-pole TCs instead. Using the OCTCs, $\tau_{c1} = 2.305\ \mu\text{s}$ and $\tau_{c2} = 95.84\text{ ns}$, and ignoring τ_{e1}, τ_{e2}, then $k = 0.953$, $\xi = 4.904$, and $\zeta_{cl} = 2.43$ - nearly the same.

The poles of the two interacting stages are the roots of

$$s^2\cdot(\tau_{c1}\cdot\tau_{c2;1}) + s\cdot(\tau_{c1} + \tau_{c2}) + 1 = s^2\cdot(481.5\text{ ns})^2 + s\cdot(2.401\ \mu\text{s}) + 1$$

They are the poles of G (and $G{\cdot}H$): $p_{G1,2} = 69.2\text{ kHz}, 1.58\text{ MHz}$. The pole separation ratio is 22.81, not much different than for non-interacting stages. The reason that stage interaction does not affect stage poles much is because the SCTC, $\tau_{c2;1} \approx \tau_{c2}$, and that is because when C_{c1} is shorted, the base resistance R_{B1}' is so high that the collector resistance does not change by much.

The closed-loop damping can alternatively be calculated from the approximate quadratic bandwidth formula, using the calculated closed-loop $f_{bw}(\text{cl}) = 3.63\text{ MHz}$;

$$\frac{f_{bw}(\text{cl})}{f_{ncl}} = \frac{1}{\sqrt{2}\cdot\sqrt{2\cdot\zeta_{cl}^2 - 1}} \Rightarrow \zeta_{cl} = \sqrt{\frac{1}{4}\cdot\left(\frac{f_{ncl}}{f_{bw}(\text{cl})}\right)^2 + \frac{1}{2}} = 0.776 \Rightarrow 39.1°$$

Using the measured bandwidth instead,

$$\zeta_{cl} = 0.816 \Rightarrow 35.3°$$

The pole angle is not over 45° for either measured or calculated ζ_{cl} and the response is unpeaked. In contrast to the pole-independent derivation of damping, the approximate quadratic bandwidth formula makes no assumption about pole dependency, and for it, $D(s)$ does not necessarily factor into OCTC poles. The damping (and pole angle) range for amplifier 10 is in the optimal range of 20° to 40° for a wideband amplifier, with accurate step response for this result.

A pole-pair with $f_{ncl} = 2.32$ MHz and $\zeta_{cl} = 0.8$ is a complex pole-pair at -1.856 MHz $\pm$ 1.39 MHz. The bandwidth, using the approximate quadratic formula, is 3.1 MHz, 8.7 % above the measured frequency-response value and in close agreement with the calculated bandwidth from the measured risetime.

The lesson to be learned is that the OCTCs (or their corresponding frequencies), while needed to compute the OCTC bandwidth, are not necessarily poles of the circuit unless they are noninteractive and thus independent. (The Cochrun-Grabel method uses the OCTCs to find the pole polynomial coefficients.) We have seen that in BJT stages, the BJT capacitances are not separable into their own pole factors and consequently the circuit poles are the roots of higher-degree polynomials. For earlier circuits, we have been content to characterize dynamic response only in OCTC bandwidth, and where resonances (pole-pairs from second-degree polynomials) occur, we have isolated the effects by grouping reactances into pairs to effect quadratic solutions for poles. (There is a procedure for solving cubic polynomials, but it is complicated enough to motivate numerical computation instead!) The pairing is somewhat informal and left to our own judgement about which pairs of reactances interact most.

Inductive peaking isolates the compensating collector L and C_L. In the cascode amplifier 3 CB output response, C_c and C_L are paired for a quadratic response. Its CB input loop groups C_e and C_B, and then combines it with C_L of the CE stage, all the while keeping the math linear or quadratic. Emitter peaking adds capacitances in the CE input loop and the impedances are second-degree. High-frequency gyration also adds reactances and resonances that we have managed to analyze with quadratic s-domain equations. For CE stages, we have calculated OCTC bandwidth, combining C_e and C_c OCTCs (and SCTCs) to find the poles.

The general BJT stage has real poles though they are not the OCTC frequencies of C_e and C_c unless widely separated. When τ_e and τ_c are not far apart in value (such as τ_{e1} and τ_{c1}), though their OCTCs

are comparable, their interaction results in widely separated poles that allow a dominant-pole approximation, as for Q1 and Q2. The wide separation is caused by the large size of τ_{cb} in *b*. The argument for a single CE pole at $\tau_b = R_b \cdot C_b$ is based on dominant-pole approximation. By keeping R_B small, the poles are not very widely separated and the single-pole (C_b) approximation fails. Circuits with only real poles have an overdamped or critical response with zero pole angle. The frequency response is flat out to bandwidth and we have characterized such circuits by their bandwidth alone.

Feedback, however, adds more complexity in that all real poles in a loop change (or "migrate") with changing quasistatic loop gain. Real poles of cascade BJT stages can become complex in a feedback loop so that dynamic response characterization must include not only bandwidth but also pole angle or damping and pole-pair magnitude, f_n, the undamped resonant frequency of the pole-pair. With amplifier 10 we have arrived at the need for fuller characterization.

Returning briefly to the four-pole loop, had we derived $D(s)$ for four interacting poles (using the Cochrun-Grabel method), the closed-loop pole polynomial s^4 term would have a coefficient, a_4, which is the product of 4 time constants (only one being an OCTC) divided by the feedback factor. The closed-loop pole polynomial coefficients are all (except $a_0 = 1$) divided by the feedback factor. When the *n*th-degree coefficient is put into this form,

$$s^n \cdot \left(\frac{\tau^n}{1 + G_0 \cdot H_0} \right) = s^n \cdot \left(\frac{\tau}{\sqrt[n]{1 + G_0 \cdot H_0}} \right)^n$$

As *n* increases, the *n*th root of the feedback factor approaches one. Thus the coefficients of increasingly higher degree decrease less with loop gain and their coefficients remain approximately constant. They affect pole migration less than the lower-degree coefficients. When the s^4 coefficient for amplifier 10 is written as a time constant to the fourth power, the corresponding frequency is 3.58 MHz. This is a fourth-degree "f_{ncl}" and is somewhat higher than bandwidth.

Another possibility for error in more accurately predicting dynamic response is that four poles might not represent the circuit. We know that for each CE stage there is a high-frequency RHP zero, but it is typically far above the frequency range of influence. One possibility for a zero is the parasitic C_f of R_f, a large enough value that 1/3 pF of stray capacitance across a ¼ W axial resistor results in a

zero at (18 kΩ)·(0.33 pF) = 6 ns, or 26.5 MHz. This zero will affect migration of the higher poles somewhat but leave the dominant poles relatively unaffected. The circuit model seems to be correct.

The feedback loop of amplifier 10, with four real poles, has demonstrated the importance of pole interaction in circuit dynamics. We have been able to identify the OCTCs and from them the dominant poles and the amplifier dynamic response. In taking a design-oriented path, we have insight into how the circuit elements affect the dynamics and are ready for circuit simulation.

An intermediate design step before simulation is to numerically plot the transfer function, given the known poles, in a math program with complex-number math capability such as MathCAD or Matlab. Not only is execution faster than calculator button fingering, the math equations retain the symbols of circuit elements that effect behavior. If the analysis is comprehensive enough, it might even replace the need for simulation.

A MathCAD program that generates (Bode) frequency-response magnitude and phase plots is given on the next page. The plots of $G{\cdot}H$ for the independent-stages assumption of amplifier 10 are shown because they illustrate that even two dominant, separated poles in a feedback loop can result in insufficient stability and an excessive pole angle. The plots reveal that the unity-gain (0 dB) frequency for the magnitude is at about $f_T = 2.094$ MHz. At the closed-loop gain magnitude of 0.59, or –4.58 dB, the frequency is very close to the measured bandwidth of 2.85 MHz. At f_T, the phase is –144°, leaving a phase margin of stability, *PM*, of about $PM = 180° - 144° = 36°$. Invoking the approximation for quadratic closed-loop response,

$$\zeta_{cl} \approx \frac{PM}{100°} = 0.36 \Rightarrow 69° \; ; \; M_{pcl} \approx (75° - PM) \cdot (1\,\%/°) = 39\,\%$$

The damping agrees with the prior result for independent stage poles. Removal of the highest two poles did little to affect *PM*. For closed-loop quadratic response, the exact

$$PM = 90° - \tan^{-1}\left\{ \frac{\sqrt{2}}{2} \cdot \sqrt{\sqrt{\frac{1}{4 \cdot \zeta_{cl}^{\,4}} + 1} - 1} \right\}$$

For $\zeta_{cl} \approx 0.8$, then $PM \approx 70°$, corresponding to a phase at f_T of –110°.

Amplifier 10 - Single-Supply Amplifier

s in units of Hz

$$K := 54.4$$

$$GH(s, K) := K \cdot \frac{1}{\left(\frac{s}{66.76 \cdot 10^3} + 1\right) \cdot \left(\frac{s}{1.48 \cdot 10^6} + 1\right) \cdot \left(\frac{s}{325.8 \cdot 10^6} + 1\right) \cdot \left(\frac{s}{321 \cdot 10^6} + 1\right)}$$

Bode Analysis:

$$MAG_{GH}(f, K) := 20 \cdot \log\left(\left|GH(j \cdot f, K)\right|\right)$$

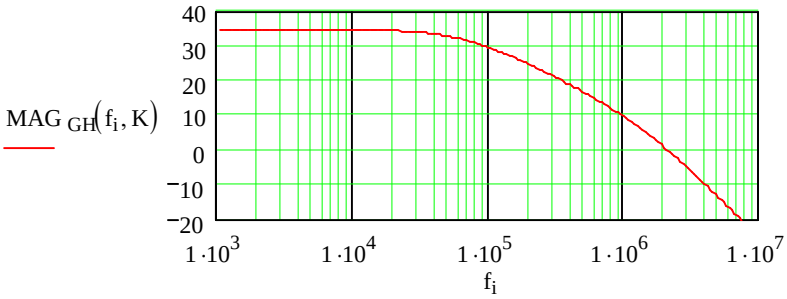
$$phase(T) := \left(if\left(\arg(T) > 0, \arg(T) - 2 \cdot \pi, \arg(T)\right)\right) \cdot \left(\frac{180}{\pi}\right)$$

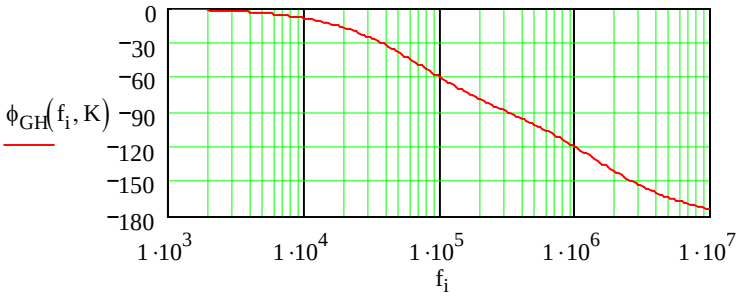
$$\phi_{GH}(f, K) := phase(GH(j \cdot f, K))$$

Number of points: $N := 200$ $i := 0 .. N - 1$

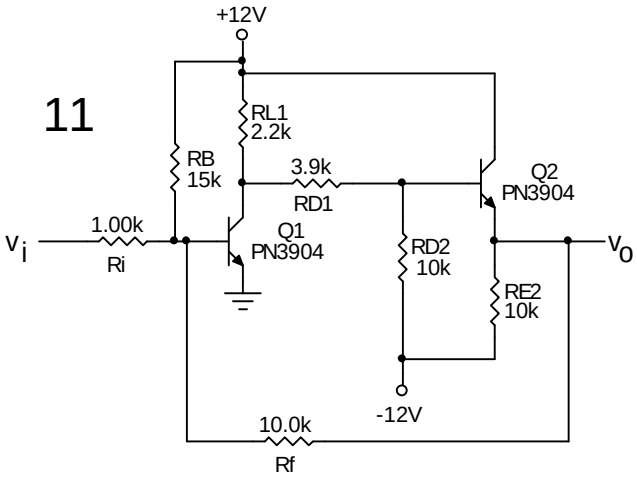
Start: $f_L := 10^3$ End: $f_U := 10^7$ step size: $r := \frac{1}{N} \cdot \log\left(\frac{f_L}{f_U}\right)$

frequency range: $f_i := f_U \cdot 10^{i \cdot r}$





Inverting Feedback Amplifier Dynamics



The earlier appearance of what is now numbered amplifier 11 in the "Feedback Amplifiers" chapter as a lab experiment leaves the more substantial analysis (and design modification, if necessary) to be completed. With only two BJTs, this circuit seems simple enough and offers a break from pushing so many calculator buttons. By circuit inspection (and with the dynamics experience from amplifier 10), we can see that OCTCs at f_{e1} and maybe f_{c1} will be dominant; Q2, as a CC stage, with a large R_E is expected to be fast. Its collector pole has no Miller multiplier yet it forms a time constant with R_{b2}.

Having built a prototype of this amplifier in chapter 2, where A_v and r_o were measured, the complete design will be worked through, though with less emphasis on the easier aspects of static and quasistatic design. The static input and output voltages are 0 V and nominal voltage gain is –10. For the feedback loop quantities, $v_f = v_o$. Following prior analysis of amplifier 10, let the error voltage be a superposition of voltages of the feedback divider. Following the feedback-circuit convention (as in amplifier 10), the effects of input and output loading of G (by H, T_i, T_o) is kept within G.

The numbers pour forth as follows, beginning with static calculation of the interstage divider resistances. To design the divider for maximum range, set $V_{C1} = V_{CC}/2 + V_{BE2} \approx 6.5$ V. This almost maximizes the range while minimizing thermal distortion. Then working back from Q2,

$$I_{E2} = 12\ \text{V}/10\ \text{k}\Omega = 1.2\ \text{mA}\ ;\ r_{e2} = 26\ \text{mV}/1.2\ \text{mA} = 22\ \Omega$$

$$V_{BE2} = 0.65\ \text{V} + (60\ \text{mV})\cdot\log(1.2) = 0.655\ \text{V}\ ;\ V_{B2} = 0.655\ \text{V}$$

$$I_{DIV} = (6.5\ \text{V} - 0.655\ \text{V})/R_{D1} = 1.50\ \text{mA}$$

$$I_{C1} = (12\ \text{V} - 6.5\ \text{V})/(2.2\ \text{k}\Omega) - I_{DIV} = 1.00\ \text{mA}\ ;\ I_{E1} = 1.04\ \text{mA}$$

$$V_{BE1} = 0.656\ \text{V}\ ;\ I_{B1} = I_{C1}/\beta_0 = 6.67\ \mu\text{A}$$

With 0 V at the other end of both R_f and R_i, current flows out from the Q1 base node through them of an amount that, along with base current, is supplied through R_B;

$$R_B = \frac{12\ \text{V} - 0.656\ \text{V}}{\dfrac{0.656\ \text{V}}{(1.0\,\text{k}\Omega\,\|\,10.0\,\text{k}\Omega)} + 6.67\ \mu\text{A}} = \frac{11.34\ \text{V}}{728.3\ \mu\text{A}} = 15.57\ \text{k}\Omega\ \rightarrow 15\ \text{k}\Omega$$

The value of R_B directly affects I_{B1} and the input circuit is β-sensitive. Variation in β_0 will cause a voltage offset at the output.

For a gain of –10, R_f/R_i is chosen to be 10, with R_f made large enough to not appreciably load the output; 10.0 kΩ is chosen. Then R_i is 1.00 kΩ. The unloaded (by $r_{Gi} = [(\beta_0 + 1)\cdot(r_{e1})]\|R_B$) divider feedback gain is

$$H_0 = -\frac{R_i}{R_i + R_f} = -\frac{1}{11} = -0.0909$$

The open-loop gain, G_0, is calculated to include the feedback-divider loading at the input;

$$G_0 = G_{01}\cdot G_{div}\cdot G_{02}$$

The interstage divider gain (or attenuation) is that of the unloaded divider at the input but loaded at the output. Therefore, loading must be included in G_{01} but not G_{02} gains. Including Q1 r_{o1} = 107 kΩ,

$$R_{L1}' = R_{L1}\|r_{o1}\|[R_{D1} + R_{D2}\|(\beta_0 + 1)\cdot(r_{e2} + R_{E2})]$$

$$= 2.2\ \text{k}\Omega\|107\ \text{k}\Omega\ \|13.83\ \text{k}\Omega = 1865\ \Omega$$

$$R_{B1}' = R_B\|R_i\|R_f = 857.1\ \Omega$$

$$r_{M1} = R_{B1}'/(\beta_0 + 1) + r_{e1} = 31.7\ \Omega$$

$$G_{01} = \frac{v_{c1}}{v_E} = -\alpha_0\cdot\frac{R_{L1}'}{r_{M1}} = -\frac{1853\ \Omega}{31.7\ \Omega} = -58.44$$

$$G_{div} = (R_{D2}||(\beta_0 + 1)\cdot(r_{e2} + R_{E2}))/(R_{D1} + R_{D2}||(\beta_0 + 1)\cdot(r_{e2} + R_{E2}))$$
$$= 9934\ \Omega/(3.9\ \text{k}\Omega + 9934\ \Omega) = 0.718$$

$$r_{M2} = [(R_{D1}||R_{D2}) + r_{b2}']/(\beta_0 + 1) + r_{e2} + R_{E2} = 40.58\ \Omega + 10\ \text{k}\Omega$$
$$= 10.04\ \text{k}\Omega$$

$$R_{B2}' = (R_{D1}||R_{D2}) + r_{b2}' = 2856\ \Omega$$

$$G_{02} = \frac{v_o}{v_{b2}} = \frac{R_{E2}}{r_{e2} + R_{E2}} = 0.998$$

Combining the gain stages,

$$G_0 = (-58.44)\cdot(0.718)\cdot(0.998) = -41.88 \approx -42$$

$$G_0 \cdot H_0 = 3.81$$

With such low loop gain, this is hardly an op-amp, though amplifier performance is improved by the feedback factor,

$$1 + G_0 \cdot H_0 = 4.81$$

Preceding the amplifier is the feedback-divider input attenuation,

$$T_i = \frac{R_f}{R_i + R_f} = \frac{10}{11} = 0.9091$$

The closed-loop gain is thus

$$A_{v0} = T_i \cdot \frac{G_0}{1 + G_0 \cdot H_0} = (0.9091)\cdot\frac{-42}{4.81} = -7.94$$

As for dynamics, the usual OCTCs are calculated using the OCTC single-stage templates:

$$\tau_e = \alpha_0 \cdot \tau_T \cdot \frac{R_B' + R_E}{r_M}\ ;\ \tau_c = (R_L + R_b \cdot (1 + K_v)) \cdot C_c,$$

$$K_v = \alpha_0 \cdot \frac{R_L}{r_e + R_E}$$

$$\alpha_0 \cdot \tau_T = (0.9934)\cdot(707\ \text{ps}) = 702.3\ \text{ps}\ ,\ 1\ \text{mA}$$

$\tau_{e1} = \alpha_0 \cdot \tau_T \cdot (R_{B1}'/r_{M1}) = (702.3\ \text{ps})\cdot(27.04) = 19.0\ \text{ns} \Rightarrow 8.38\ \text{MHz}$

$\tau_{c1} = (R_{L1}' + R_{B1}'\cdot(1 + \alpha_0 \cdot R_{L1}'/r_{e1}))\cdot C_{c1}$
$= (1865\ \Omega + (857.1\ \Omega)\cdot(71.7))\cdot(2\ \text{pF}) = (63.35\ \text{k}\Omega)\cdot(2\ \text{pF}) = 127\ \text{ns}$
$\Rightarrow 1.26\ \text{MHz}$

$\tau_{cc1} = R_{L1}' \cdot C_{c1} = 3.73$ ns ; $\tau_{cb1} = 123$ ns

$\tau_{e2} = (702.3 \text{ ps}) \cdot (R_{B2}' + 10 \text{ k}\Omega)/(10.04 \text{ k}\Omega) = 0.90 \text{ ns} \Rightarrow 177 \text{ MHz}$

$\tau_{c2} = R_{B2}' \cdot C_{c2} = (2856\ \Omega) \cdot (2 \text{ pF}) = 5.71 \text{ ns} \Rightarrow 27.87 \text{ MHz}$

Amplifier 11, like amplifier 10, has no poles or influential zeros in the feedback path, *H*. The parasitic C_f across R_f adds a zero at about 48 MHz. The four real OCTCs in the *G* path are grouped into BJT pairs and the poles calculated. The poles of the Q1 CE stage are found by finding τ_{n1} and ζ_1;

$$\tau_{n1} = \sqrt{\tau_{e1} \cdot \tau_{cc1}} = 8.42 \text{ ns} \Rightarrow 18.9 \text{ MHz}$$

$$\zeta_1 = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_{e1}}{\tau_{cc1}}} + \sqrt{\frac{\tau_{cc1}}{\tau_{e1}}} + \frac{\tau_{cb1}}{\tau_{n1}} \right) = 8.654$$

The pole frequencies, in Hertz, are

$$\frac{1}{2 \cdot \pi} \cdot \frac{\zeta_1}{\tau_{n1}} \cdot \left(1 \pm \frac{\sqrt{\zeta_1^2 - 1}}{\zeta_1} \right) = 1.10 \text{ MHz},\ 326 \text{ MHz}$$

The CE stage has one dominant pole at 1.10 MHz $\Rightarrow$ 145 ns.

The second stage is a CC with a high $f_{e2} = 177$ MHz, but the pole formed by C_{c2} with the Q2 base resistance, at $f_{c2} = 27.87$ MHz, is more significant. Q2 has no τ_{cc2} (no collector resistance) and $\tau_{c2} = \tau_{cb2}$. With $\tau_{cc2} = 0$ s, we cannot very well use the CE formula for pole calculation, though it does tell us (with $\zeta \to \infty$ and $\tau_n = 0$ s) that the two poles are infinitely separated and one must therefore be at infinity (as is the pole magnitude) and the other at zero Hz. In effect, the pole-pair collapses into two independent poles; we retain the dominant one which is at f_{c2}.

The loop therefore has two poles, one per stage. The stages are cascaded and the reactances can interact to form poles of different frequencies. The interstage divider provides some isolation. If C_{c1} were shorted, and $\tau_{c2;1}$ found (a SCTC), the difference between short- and open-circuit resistance is 4433 Ω shorted and 6100 Ω open, a difference of about 38 % - not much for time-constant separation. Thus, in the spirit of approximation (and calculator relief) we now venture (after amplifier 10) to consider the two poles in *G* to be independent.

Without requiring pole independence, we can find the OCTC bandwidth of the loop;

$$\tau_{bw} \approx \sqrt{(145\text{ ns})^2 + (5.71\text{ ns})^2} = 145.1\text{ ns} \Rightarrow f_{bw} = 1.1\text{ MHz}$$

The Q1 pole dominates. The closed-loop bandwidth is

$$f_{bw}(\text{cl}) = f_{bw}\cdot(1 + G_0\cdot H_0) = (1.10\text{ MHz})\cdot(4.81) = 5.3\text{ MHz}$$

The pole separation for the two lowest pole frequencies is 25.3 which corresponds to an open-loop $\zeta = 2.62$. Then

$$\zeta_{cl} = \zeta / \sqrt{1 + G_0\cdot H_0} = 2.62/2.193 = 1.2$$

which is slightly overdamped. No peaking was observed in the prototype unit. Based on this ζ_{cl}, the exact closed-loop bandwidth is

$$f_{bw}(\text{cl}) = f_{ncl}\cdot\sqrt{1 - 2\cdot\zeta_{cl}^2 + \sqrt{4\cdot\zeta_{cl}^4 - 4\cdot\zeta_{cl}^2 + 2}}$$
$$= (12.14\text{ MHz})\cdot(0.5025) = 6.1\text{ MHz}$$

This is 15 % higher than the OCTC bandwidth.

On the bench, the prototype provides the following measurements or quantities derived from them. The static values are

$$V_{BE1} = 0.672\text{ V}\ ;\ V_{C1} = 4.98\text{ V}\ ;\ V_{B2} = 0.19\text{ V}\ ;\ V_O = -0.514\text{ V}$$

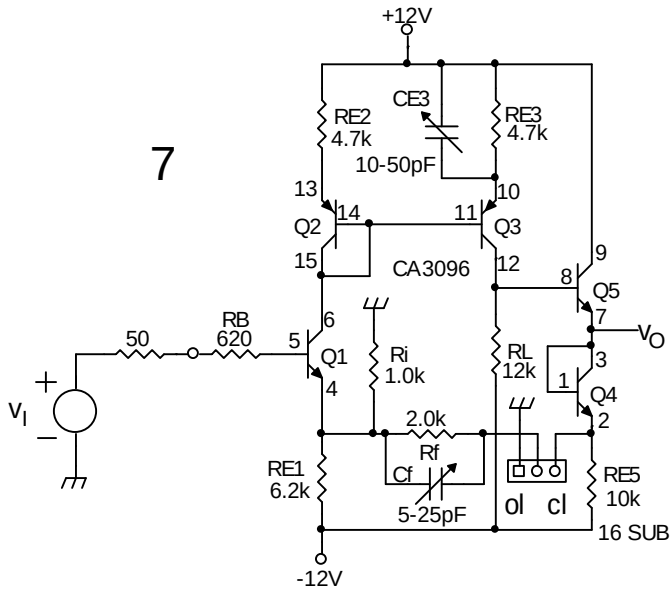
The linear range of the output is symmetrical and is ±3.0 V.

Applying ±0.1 V, the output voltage measures ±0.86 V for a gain magnitude of 8.6, about 8 % higher than the predicted 7.94 value. Old 5 % resistors were used and no attempt will be made to measure them and refine the calculated gain; the accuracy of the quasistatic gain method is not in doubt.

The bandwidth measures 5.5 MHz. This agrees well with f_{bw}(cl) of 5.3 MHz and is low relative to the ζ_{cl} bandwidth-based calculation. The latter should be expected to have a wider tolerance because ζ_{cl} is hard to calculate accurately with so much undetermined about pole interaction. At this point, the critical design effects on the dynamics have been identified and the circuit is ready for either a simulator or more bench experimentation.

Current- and Shunt-Feedback Amplifiers

Emitter-Feedback Amplifier



Amplifier 7 illustrates the concept of the *current-feedback amplifier* (CFA). Before introducing the concept, we will transition into it by analyzing amplifier 7 as we have others, by choosing a voltage error quantity.

Static Design

The *G* path consists of the Q1 CE input stage followed by a 2-BJT current mirror (Q2, Q3) and an output CC stage (Q5). The static design for 0 V out with $v_I = 0$ V requires that the Q1 *b-e* voltage drop be matched (by Q4) so that no static voltage is applied to the feedback divider, R_f, R_i. Q4 matches Q1 with equal currents. Their junctions

drop the same voltage, leaving 0 V across R_f and $-V_{BE}$ across R_i. With a static 0 V across R_f, it conducts zero bias current. R_{E1} and R_i can be combined into a Thevenin equivalent circuit. The design equation for R_{E1} will be derived.

First, for minimization of thermal distortion, we want $V_{C1} \approx 6$ V. Then $V_{E2} \approx 6.69$ V and I_{E2} = (12 V – 6.69 V)/4.7 kΩ = 1.13 mA. The emitter current of Q3 will be about the same and $I_{B3} \approx I_{B1}$. I_{B3} adds to I_{C1} = 1.127 mA so that the loss of I_{B1} is compensated, and $I_{E1} \approx I_{E2}$. The thevenization of $-V_{EE}$, R_{E1}, and R_i result in

$$V_{EE}' = (12\ \text{V}) \cdot \left(\frac{1.0\ \text{k}\Omega}{1.0\ \text{k}\Omega + R_{E1}} \right)\ ;\ R_{EE}' = 1.0\ \text{k}\Omega \parallel R_{E1}$$

The Q1 external base resistance is R_B plus the source resistance of 50 Ω, or R_B = 670 Ω. R_B' includes r_b' = 50 Ω; R_B' = 720 Ω. Moving R_B' to the emitter with the β transform and applying KVL around the Q1 input loop,

$$(V_{EE}' - V_{BE1}) = I_{E1} \cdot \left(R_i \parallel R_{E1} + \frac{R_B'}{\beta_{NPN} + 1} \right)$$

Solve for R_{E1}. After some algebraic manipulation, we have the design formula;

$$R_{E1} = \frac{V_{EE}' - (V_{BE1} + I_{E1} \cdot R_B'/(\beta_{NPN} + 1))}{\left(\frac{V_{BE1} + I_{E1} \cdot R_B'/(\beta_{NPN} + 1)}{R_i} \right) + I_{E1}}$$

The polarity of I_{E1} is chosen as positive coming out of the emitter (and not according to port convention), to minimize the number of negative signs in the algebra. Substituting values after calculating that

$$V_{BE1} + I_{E1} \cdot R_B'/(\beta_{NPN} + 1) = 0.69\ \text{V} + (1.13\ \text{mA} \cdot \frac{720\ \Omega}{391}) = 0.692\ \text{V}$$

the value is

$$R_{E1} = \frac{12\ \text{V} - (0.692\ \text{V})}{\frac{(0.692\ \text{V})}{1.0\ \text{k}\Omega} + 1.13\ \text{mA}} = 6.206\ \text{k}\Omega$$

The nearest 5 % resistor value is 6.2 kΩ (a close fit) and R_{EE}' = 861 Ω.

We assumed V_{BE} at 1 mA, not at I_{E1} = 1.13 mA. The assumed V_{BE1} = 0.69 V is correct at 1 mA. Silicon p-n junction voltage changes 60 mV/dec of current. A ×1.13 difference changes V_{BE} by (60 mV/dec)·log$_{10}$(1.13) = 3.2 mV. Then the corrected V_{BE1} is about 0.693 V and the assumption of 0.69 V introduces little error. The logarithmic change in voltage with current makes V_{BE} rather insensitive to changes in current and a static circuit solution converges quickly. By substituting 0.693 V for 0.69 V in the equation for I_{E1}, the new current value of 1.127 mA is slightly smaller and within the 5 % tolerance of R_{E1}.

The mirror current gain is

$$A_i = \frac{\beta_{PNP}}{\beta_{PNP} + 2} \cdot \frac{R_{E2}}{R_{E3}} = \frac{47}{49} \cdot 1 = 0.96$$

Then I_{C3} = (0.96)·I_{C1} = (0.96)·(1.127 mA) = 1.082 mA. For 0 V out, then V_{C3} = 0.69 V (assuming again that $I_{E5} \approx$ 1 mA) and we can calculate the value of

$$R_L = \frac{12.69\ \text{V}}{1.082\ \text{mA}} = 11.73\ \text{k}\Omega$$

The closest 5 % value is 12 kΩ. Then the nominal voltage at the Q5 base is 0.984 V and V_O = 294 mV, slightly higher than zero. (For more precision, all the resistors should be ±1 % tolerance instead.)

To match the V_{BE} of Q1 and Q4, they must conduct the same current. Then I_{E5} is set by design to be 1.13 mA. Applying Ohm's Law again,

$$R_{E5} = \frac{12\ \text{V} + 0.294\ \text{V} - 0.69\ \text{V}}{1.13\ \text{mA}} = 10.27\ \text{k}\Omega$$

The closest 5 % value is 10 kΩ. The static design of the amplifier is completed.

Quasistatic Design

This voltage amplifier has three stages: the CE input, the current mirror, and the CC output. Before analyzing circuitry, choose the feedback quantities, x_f and x_E. The feedback quantity is easy to choose: $x_f = v_o$. The error quantity can be a voltage or a current.

Choose $x_E = v_E$. We can (as for previous voltage-feedback amplifiers) thevenize the feedback divider;

$$v_E = v_i - v_B = v_i - \left(\frac{R_i \parallel R_{E1}}{R_f' + R_i \parallel R_{E1}} \right) \cdot v_o$$

Q4 r_{e4} and R_{E5} present a Thevenin resistance of $23\,\Omega$ in series with R_f and will be considered a part of it as

$$R_f' = R_f + r_{e4} \| R_{E5} = 2023\ \Omega$$

The voltage divider formed by r_{e4} and R_{E5} has an attenuation of

$$H_1 = \frac{R_{E5}}{R_{E5} + r_{e4}} = 0.998$$

The feedback divider includes R_{E1} and R_i must be adjusted in value accordingly;

$$H_2 = \frac{R_{E1} \parallel R_i}{R_{E1} \parallel R_i + R_f'} = 0.2986$$

Then v_B is the open-circuit voltage of H_1 in series with R_f to the Q1 emitter node, through H_2, and

$$H = H_1 \cdot H_2 = 0.298 = 1/3.356$$

With the given choice of error voltage, we can solve for G beginning with the first stage, G_1;

$$G_1 = \frac{i_{C1}}{v_E} = \alpha_{NPN} \cdot \frac{1}{(R_{E1} \parallel R_i \parallel R_f') + r_{e1} + R_B'/(\beta_{NPN} + 1)} = \frac{1}{631\,\Omega}$$

The first stage is partitioned to be a transconductance amplifier so that the current amplifier (the mirror) following it is current-driven at its input. If R_L is included as the load resistor of Q3, then the mirror stage is a transresistance amplifier with a gain of

$$G_2 = \frac{v_{C3}}{i_{C1}} = A_i \cdot R_L = (0.96) \cdot (12\ \text{k}\Omega) = 11.52\ \text{k}\Omega$$

The Q3 collector node of G_2 is unloaded by Q5 and is the open-circuit (Thevenin) voltage. To include the loading, the CC stage includes R_L in its base resistance;

$$G_3 = \frac{v_o}{v_{C3}} = \frac{r_{e4} + R_{E5} \parallel (R_f + R_i \parallel R_{E1})}{[r_{e4} + R_{E5} \parallel (R_f + R_i \parallel R_{E1})] + r_{e5} + R_L/(\beta_{NPN} + 1)}$$

Substituting values,

$$G_3 = \frac{2.24\ \text{k}\Omega}{2.24\ \text{k}\Omega + 53.7\ \Omega} \approx 0.977$$

Then combining the stage gains,

$$G_0 = G_1 \cdot G_2 \cdot G_3 = 17.84$$

A value of 18.8 was measured in the prototype circuit with an input sine-wave amplitude of 0.1 V and by setting the jumper to the open-loop (ol) position. This value is higher than calculated by about 5 %, within the tolerance of the resistors and measurements. The amplifier has a rather low forward-path gain and a loop gain of

$$G_0 \cdot H_0 = 17.84/3.356 = 5.32$$

Performance depends on good open-loop characteristics and a × 6.32 feedback improvement. The calculated closed-loop gain is

$$A_{v0c} = \frac{G_0}{1 + G_0 \cdot H_0} = \frac{17.84}{6.32} = 2.82$$

The measured value with 0.5 V pk sine-wave input is $A_{v0c} = 2.80$.

Dynamic Design

We now enter the dynamic analysis of amplifier 7 with some trepidation. As E. James Angelo Jr., the author of a prominent circuits textbook, *Electronics: BJTs, FETs, and Microcircuits* (McGraw-Hill, 1969) has written (page 449): "… the analysis [of interacting stages in cascade] is quite complicated, and a complete solution is not all that practical." Not only are two stages in cascade in amplifier 7, it also has a feedback loop. Nevertheless, we will endeavor to see how close we can calculate the bandwidth.

Apply the following OCTC formulas to the three amplifier stages:

$$\tau_e = (\alpha_0 \cdot \tau_T) \cdot \frac{R_B' + R_E}{r_e + R_E + R_B'/(\beta_0 + 1)} = (\alpha_0 \cdot \tau_T) \cdot \frac{R_B' + R_E}{r_M}$$

$$\tau_c = [R_L + R_b \cdot (1 + K_v)] \cdot C_c,\ K_v = \alpha_0 \cdot R_L/(r_e + R_E)$$

$$\tau_L = R_L \cdot C_L$$

The resistance and gain values for the first stage are

$$R_{L1} = (r_{e2} + R_{E2}) \,\|\, [(\beta_{PNP} + 1) \cdot (r_{e3} + R_{E3}) + r_{b2}' + r_{b3}'] = 4631\,\Omega$$

$$R_{E1}' = R_{E1} \,\|\, R_i \,\|\, R_f' = 604\,\Omega$$

$$R_{b1} = R_{B1}' \,\|\, (\beta_{NPN} + 1) \cdot (r_{e1} + R_{E1}') = 718\,\Omega$$

$$K_{v1} = (0.997) \cdot \frac{4631\,\Omega}{718\,\Omega} = 6.43$$

$$r_{M1} = r_{e1} + R_{E1}' + R_{B1}'/(\beta_{NPN} + 1) = 632\,\Omega$$

Then the three input-stage OCTCs are

$$\tau_{e1} = (\alpha_{NPN} \cdot \tau_{TNPN}) \cdot \frac{R_{B1}' + R_{E1}'}{r_{M1}} = [(0.997) \cdot (568\text{ ps})] \cdot \frac{720\,\Omega + 604\,\Omega}{632\,\Omega} = 1.187\text{ ns} \Rightarrow 134\text{ MHz}$$

$$\tau_{c1} = (4631\,\Omega + [718\,\Omega \cdot (1 + 6.43)]) \cdot (0.5\text{ pF}) = 4.98\text{ ns} \Rightarrow 31.94\text{ MHz}$$

$$\tau_{cc1} = 2.32\text{ ns} \; ; \; \tau_{cb1} = 2.67\text{ ns}$$

$$\tau_{L1} = (4631\,\Omega) \cdot (C_{cs1} + C_{bs2} + C_{bs3}) = (4631\,\Omega) \cdot (7\text{ pF}) = 32.42\text{ ns} \Rightarrow 4.91\text{ MHz}$$

For the second stage,

$$R_{B3}' = (r_{e2} + R_{E2}) + r_{b2}' + r_{b3}' = 4726\,\Omega + 100\,\Omega = 4826\,\Omega$$

$$R_{b3} = R_{B3}' \,\|\, (\beta_{PNP} + 1) \cdot (r_{e3} + R_{E3}) = 4725\,\Omega$$

$$r_{M3} = r_{e3} + R_{E3} + R_{B3}'/(\beta_{PNP} + 1) = 4827\,\Omega$$

$$K_{v3} = (0.979) \cdot \frac{12\text{ k}\Omega}{4726\,\Omega} = 2.49$$

$$\tau_{e3} = (\alpha_{PNP} \cdot \tau_{TPNP}) \cdot \frac{R_{B3}' + R_{E3}}{r_{M3}} = [(0.979) \cdot (24.5\text{ ns})] \cdot \frac{4826\,\Omega + 4700\,\Omega}{4827\,\Omega} = 47.34\text{ ns} \Rightarrow 3.36\text{ MHz}$$

$$\tau_{c3} = [(12\text{ k}\Omega + (4825\,\Omega) \cdot (1 + 2.49)] \cdot (2\text{ pF}) = 57.7\text{ ns} \Rightarrow 2.76\text{ MHz}$$

$$\tau_{cc3} = 24\text{ ns} \; ; \; \tau_{cb3} = 33.7\text{ ns}$$

$$\tau_{L3} = (11.96\,\text{k}\Omega)\cdot[C_{c5} + C_{E5}'] \\ = (11.96\,\text{k}\Omega)\cdot[0.5\,\text{pF} + 1.4\,\text{pF}] = 22.72\,\text{ns} \quad \Rightarrow 7.00\ \text{MHz}$$

The Q5 emitter-node capacitance consists of the collector-substrate capacitance of Q4, C_{bs4}, plus C_o, that of the 'scope probe used to make the measurement. The lower end of the NPN hf range is f_T/β_{NPN} = 857 kHz. Thus f_{L3} is in the hf region of Q5, and emitter impedance at the Q5 base is gyrated. The capacitance is effectively reduced by the β transform at 857 kHz to $C_E/(\beta_{NPN} + 1)$, or about 45 fF (0.045 pF), a negligible capacitance. As frequency increases in the hf region, β decreases and the base-referred C_E effectively increases. (At a given frequency in the hf region, set $Z_b(s) = X_C(s)$ and solve for C. It is the capacitance plot intersecting $Z_b(s)$ on a reactance graph.) C_{c5} will limit f_{L3} to 26.6 MHz, or 31.06·(f_T/β_{NPN}), at which the referred emitter capacitance is

$$C_{E5}' = (31.06)\cdot(45\ \text{fF}) = 1.4\ \text{pF}$$

In using this value, f_{L3} will itself be reduced, leading to a lower value. It is a worst-case maximum and the value to use in design calculations short of iterating C_{E5}'.

The OCTC bandwidth of G is calculated from the above six time constants, assuming that others (such as Q5 $r_{b5}'\cdot C_{c5}$ = (50 Ω)·(0.5 pF), which is 25 ps ⇒ 6.4 GHz) are well above the amplifier f_T. With no additional reactances in the circuit, the poles are real and

$$\tau_{bw}^2 = (\tau_{e1}^2 + \tau_{c1}^2 + \tau_{L1}^2) + (\tau_{e3}^2 + \tau_{c3}^2 + \tau_{L3}^2) = \tau_1^2 + \tau_3^2$$

in which τ_1 = 32.82 ns ⇒ 4.85 MHz and τ_3 = 78.0 ns ⇒ 2.04 MHz, resulting in τ_{bw} = 84.64 ns from which

$$f_{bw} = 1.88\ \text{MHz}$$

The feedback path has no significant poles or zeros, and $H = H_0$. Then the loop bandwidth is that of G.

The G path is reminiscent of the cascade CE stages of amplifiers 9 and 10. Each stage has interacting C_e and C_c that are combined to produce the two real poles, split widely apart and leaving one as a dominant low-frequency pole. The amplifier will have two of these dominant poles, one per stage, which themselves interact. The stage poles are derived from the OCTCs by applying the CE formulas. The poles of Q1 are found by finding τ_{n1} and ζ_1;

$$\tau_{n1} = \sqrt{\tau_{e1} \cdot \tau_{cc1}} = 1.66 \text{ ns} \Rightarrow 96 \text{ MHz}$$

$$\zeta_1 = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_{e1}}{\tau_{cc1}}} + \sqrt{\frac{\tau_{cc1}}{\tau_{e1}}} + \frac{\tau_{cb1}}{\tau_{n1}} \right) = 1.86$$

The first-stage pole frequencies, in Hertz, are

$$\frac{1}{2 \cdot \pi} \cdot \frac{\zeta_1}{\tau_{n1}} \cdot \left(1 \pm \frac{\sqrt{\zeta_1^2 - 1}}{\zeta_1} \right) = 27.96 \text{ MHz}, 329 \text{ MHz}$$

For the second stage,

$$\tau_{n3} = \sqrt{\tau_{e3} \cdot \tau_{cc3}} = 33.71 \text{ ns} \Rightarrow 4.72 \text{ MHz}$$

$$\zeta_3 = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_{e3}}{\tau_{cc3}}} + \sqrt{\frac{\tau_{cc3}}{\tau_{e3}}} + \frac{\tau_{cb3}}{\tau_{n3}} \right) = 1.558$$

The second-stage pole frequencies, in Hertz, are

$$\frac{1}{2 \cdot \pi} \cdot \frac{\zeta_3}{\tau_{n3}} \cdot \left(1 \pm \frac{\sqrt{\zeta_3^2 - 1}}{\zeta_3} \right) = 1.715 \text{ MHz}, 13.0 \text{ MHz}$$

The second stage is slower than the first, so slow that the higher pole of stage 2 is lower than the stage-1 dominant pole. This is expected; the CA3096 PNP BJTs are much slower (lower f_T) than the NPN BJTs. Furthermore, the collector OCTC at f_{L1} is only 4.91 MHz, comparable to the OCTC frequencies of Q3, and must be included in any attempt to find interstage pole locations or to construct ζ. f_{L3} at 7.00 MHz might also be included.

The second-stage pole (assuming no interaction with the first stage) of 1.715 MHz is dominant in that it has a nearest OCTC that is ×2.86 higher, at 4.91 MHz. The second pole reduces the bandwidth by 5.5 % and results in a bandwidth of less than the dominant pole. We can thus expect the open-loop bandwidth to be less than 1.7 MHz. The OCTC bandwidth, calculated from the six time constants including the two BJT pole-pairs is 1.56 MHz.

From our previous derivations, not only have we involved the above formulas relating τ_e and τ_c to stage poles, but there is also a cascade CE relationship. All of the six capacitances interact and how they should be combined to avoid a four- to six-degree polynomial

requires some higher-level insight. For amplifier 11, we encountered the same problem, of both intra- and inter-stage capacitance interactions. We eluded the interstage interaction on account of the isolation afforded by the intervening divider. In amplifier 7, the stages are more coupled. What can we do?

One place to turn is to try to apply the fast-BJT (no C_e) cascaded CE template to the two stages. Besides the OCTCs, we need an SCTC. It can be found without much difficulty from the circuit. Short C_{c1} and the SCTC resistance is

$$R_{c3;1} = R_{L2} + R_{b3;1} \cdot \left(1 + \alpha_{03} \cdot \frac{R_{L3}}{r_{M3}} \right) \text{ , } C_{c1} \text{ shorted}$$

where

$$R_{b3;1} = (\beta_{PNP} + 1) \cdot r_{M3} \,\|\, [R_{B3}' + (R_{L1} \,\|\, r_{M1} \,\|\, R_{B1})] \text{ , } C_{c1} \text{ shorted}$$

Substituting parts values,

$$R_{b3;1} = (391) \cdot (4827\ \Omega) \| [4826\ \Omega + (4631\ \Omega \| 632\ \Omega \| 720\ \Omega)] = 5126\ \Omega$$

$$R_{c3;1} = 12\ \text{k}\Omega + (5126\ \Omega) \cdot (1 + 2.49) = 29.89\ \text{k}\Omega$$

Then the SCTC of C_{c3} is

$$\tau_{c3;1} = 59.78\ \text{ns} \Rightarrow 2.66\ \text{MHz}$$

The OCTC is

$$\tau_{c1} = R_{c1} \cdot C_1 = 4.98\ \text{ns}$$

The quadratic coefficient,

$$a = \tau_{c1} \tau_{c3;1} = (17.25\ \text{ns})^2 \Rightarrow \tau_n = 17.25\ \text{ns} \Rightarrow 9.22\ \text{MHz}$$

and the linear coefficient is

$$b = \tau_{c1} + \tau_{c3} = 4.98\ \text{ns} + 57.7\ \text{ns} = 62.68\ \text{ns} \Rightarrow 2.54\ \text{MHz}$$

From the quadratic coefficients, $\zeta = 1.816$, $\zeta/\tau_n = 16.76$ MHz, and the poles are

$$p_{1,2} = 2.77\ \text{MHz}, 30.74\ \text{MHz}$$

The OCTC bandwidth of these two poles is $f_{bw} = 2.76$ MHz, 77 % higher than the dominant-pole BJT estimation. This is expected because the PNP BJT Q3 is not fast and τ_{e3} has been omitted.

Another method for determining circuit poles from OCTCs is to apply the Cochrun-Grabel method. The three slowest OCTCs are chosen. They are, from slowest to fastest,

$$\tau_{c3} = 57.7 \text{ ns} \; ; \; \tau_{e3} = 47.34 \text{ ns} \; ; \; \tau_{L1} = 32.42 \text{ ns}$$

If they are stacked in that order in a Rosenstark table, then the following SCTCs are needed:

$$R_{b3}' = (4826\ \Omega)||(12\ \text{k}\Omega) = 3442\ \Omega$$

$$\tau_{e3;c3} = \alpha_{PNP} \cdot \tau_{TPNP} \cdot \frac{R_{b3}'+R_{E3}}{r_{M3}} = (24 \text{ ns}) \cdot \frac{3442\,\Omega + 4700\,\Omega}{4827\,\Omega} = 40.46 \text{ ns}$$

$$\tau_{L1;e3} = (4631\ \Omega||4700\ \Omega)\cdot(7 \text{ pF}) = (2333\ \Omega)\cdot(7 \text{ pF}) = 16.33 \text{ ns}$$

$$\tau_{L1;c3,e3} = (4631\ \Omega||12\ \text{k}\Omega)\cdot(7 \text{ pF}) = (1953\,\Omega)\cdot(7 \text{ pF}) = 13.67 \text{ ns}$$

The resulting Rosenstark table is constructed as shown.

$\tau_{c3} = 57.7$ ns	$\tau_{e3;c3} = 40.46$ ns	$\tau_{L1;c3,e3} = 13.67$ ns
	$\tau_{L1;c3} = 23.4$ ns	
$\tau_{e3} = 47.34$ ns	$\tau_{L1;e3} = 16.33$ ns	
$\tau_{L1} = 32.42$ ns		

137.5 ns	$(66.77 \text{ ns})^2$	
$\Rightarrow 1.157$ MHz	$\Rightarrow 2.384$ MHz	$(31.72 \text{ ns})^3$

The sums at the bottom of the columns are calculated by the Cochrun-Grabel method. Each entry in the middle column is multiplied by the value in the same row of the first column. Then these products are added. The first (leftmost) column is the sum of the OCTCs.

By truncating the table at the second column, the resulting polynomial is

$$s^2 \cdot (66.77 \text{ ns})^2 + s \cdot (137.5 \text{ ns}) + 1$$

The damping is $\zeta = 1.030$, close to one; the poles will be nearly equal. The polynomial roots are the poles;

$$-p_{1,2} = 2.45 \text{ MHz} \cdot (1 \pm 0.05845) = 2.311 \text{ MHz}, 2.598 \text{ MHz}$$

The OCTC bandwidth based on these poles is 1.73 MHz.

If the table is shrunk to include only the first two OCTC entries, τ_{c3} and τ_{e3}, then $a = (48.32 \text{ ns})^2$, $b = 105$ ns, $\zeta = 1.087$, and

$$-p_{1,2} = 3.58 \text{ MHz} \cdot (1 \pm 0.3918) = 2.18 \text{ MHz}, 4.98 \text{ MHz}$$

and the OCTC $f_{bw} = 2.0$ MHz. The poles are split slightly wider in the more approximate two-OCTC table. In both cases, the bandwidth is not hugely different from that of the dominant-pole BJT bandwidth of 1.56 MHz, or even that of the interacting-cascade, of 2.76 MHz.

Last but not least, stage interaction for quasistatic analysis took into account the resistive loading of the second stage on the first, either in calculating the first- or the second-stage gain. The resistive loading can be extended to impedance because the analysis is based on basic circuit laws that apply to both. The input impedance of the second stage, applied to the output of the first, can be used as a method of isolating the stages for analysis. Then (disregarding C_{E3}) the Q3 input resistance is $(1 + 47) \cdot (4700\ \Omega) = 226\ \text{k}\Omega$, large enough to disregard. The capacitance is

$$C_{b3} \approx (1 + K_{v3}) \cdot C_{c3} + \tau_{PNP}/R_E$$

$$= (3.49) \cdot (2 \text{ pF}) + 24.5 \text{ ns}/4700\ \Omega \approx 12.2 \text{ pF}$$

This adds to C_L and modifies

$$\tau_{L1} = (4631\,\Omega) \cdot (C_{cs1} + C_{bs2} + C_{bs3} + C_{b3}) = (4631\ \Omega) \cdot (19.2 \text{ pF}) = 88.9 \text{ ns}$$
$$\Rightarrow 1.79 \text{ MHz}$$

Then

$$\tau_1 = \sqrt{\tau_{e1}^2 + \tau_{c1}^2 + \tau_{L1}^2} = 89.03 \text{ ns} \Rightarrow 1.79 \text{ MHz}$$

Combined with $\tau_3 = 78.0$ ns, the resulting $\tau_{bw} = 118.4$ ns from which

$$f_{bw} = 1.35 \text{ MHz}$$

This value - the lowest yet - is still 87 % higher than the measured open-loop bandwidth.

The first stage has input loading by an RC integrator with a pole time constant of

$$\tau_i = R_{B1}' \cdot C_{b1} \approx (720\ \Omega) \cdot [(7.43) \cdot (0.5 \text{ pF}) + 0.94 \text{ pF}] = 3.35 \text{ ns}$$
$$\Rightarrow 47.5 \text{ MHz}$$

and does not significantly affect the bandwidth.

Prototype circuit measurement of the voltage gain with 0.1 V pk input and 1.88 V pk output, shows an open-loop gain of $G = 18.8$ and closed-loop gain of $A_{v0} = 2.93$. These quasistatic values are close enough to the calculated values. The open-loop bandwidth, however, is lower than the dominant-pole BJT value at $f_{bw} = 720$ kHz for 0.5 V to 2 V pk output and with no observed slewing. This is about half of the dominant-pole-BJT estimate of 1.56 MHz. The trim capacitors, C_{E3} and C_f, have been added for closed-loop response compensation, but even if left in, do not alter the above open-loop results by much; the compensated open-loop prototype measured $f_{bwcl} = 865$ kHz for a 0.5 V amplitude sine-wave out.

The prototype does not oscillate when the loop is closed though it is highly underdamped. Without the additional capacitive trims, the closed-loop frequency-response has a highly underdamped resonance of $M_m = 5.17$ with a peak at $f_m = 4.4$ MHz.

From the measured M_m at f_m, the closed-loop damping, ζ_c and resonant frequency, f_{nc} are found as

$$\zeta_c = \frac{\sqrt{2}}{2} \cdot \sqrt{1 - \sqrt{1 - \frac{1}{M_m^2}}} = \cos \phi_c \ , \ M_m \geq 1$$

$$f_{nc} = \frac{f_m}{\sqrt{1 - 2 \cdot \zeta_c^2}}, \ 0 \leq \zeta_c \leq \frac{\sqrt{2}}{2}$$

Using these formulas, $\zeta_c = 0.115 \Rightarrow \phi_c = 83.4°$, and $f_{nc} = 5.65$ MHz. For quadratic poles (pole-pairs) feedback-loop phase margin can be derived (in radians) as

$$PM = \frac{\pi}{2} - \tan^{-1}\left\{ \frac{\sqrt{2}}{2} \cdot \sqrt{\sqrt{\frac{1}{4 \cdot \zeta_c^4} + 1} - 1} \right\}$$

Substituting ζ_c calculated from measurements, $PM = 13.12°$ - barely stable.

To improve the stability of the amplifier, adjustable or *trim* capacitors are subsequently added as shown in the circuit diagram. C_{E3} has a range of 10 to 50 pF to implement the emitter-peaking technique used in "CE Design Refinement". For $C_{E3} = 30$ pF, this places a zero at a time-constant of $R_E \cdot C_E$ that for Q3 is at 141 ns $\Rightarrow$ 1.13 MHz. (The pole accompanying the zero is at a high

frequency and can be ignored.) From open-loop measurement, f_{bw} = 865 kHz, and is increased by the zero.

Within the range of C_{E3}, the response is still too underdamped. The closed-loop frequency peaking decreases to $M_m = 1.85$ at $f_m = 4.7$ MHz, resulting in $\zeta_c = 0.282$, $\phi_c = 73.6°$, and $f_{nc} = 5.12$ MHz. A second zero can be added to the loop by adding trim capacitor C_f with a range of 5 to 25 pF. At midrange (15 pF), C_f adds a zero at 5.3 MHz and when adjusted, it flattened ($M_m = 1$) the closed-loop magnitude response to f_{bwc}. The closed-loop bandwidth for 0.5 V amplitude output measures

$$f_{bwc} = 10.4 \text{ MHz}$$

Then

$$f_{bw} \approx f_{bwc}/(1 + G_0 \cdot H_0) = 10.4 \text{ MHz}/6.32 = 1.65 \text{ MHz}$$

which is in acceptable agreement with the calculated values. As closed-loop gain increases, $1/H$ increases and the loop gain, $G \cdot H$ decreases; the amplifier becomes more stable and less underdamped. It also has a loss of the benefit of feedback with a lower loop gain.

Amplifier 7 pushes the design-oriented dynamics methods to their limits. Various bandwidth calculation methods were applied yet none of them resulted in values of open-loop bandwidth that agreed well with measured bandwidth. Any additional advancement in dynamics for design purposes requires new insights or methods. Dynamic design beyond these methods relies on computer simulation for more accurate values, and also on experimental insights.

Current-Feedback Amplifiers

Amplifier 7 differs from amplifiers 5 and 6 of the previous chapter in that the feedback returns to the emitter of the input BJT, a relatively low-resistance node. The popular engineering name given to an amplifier for which the feedback, x_B, is at a low-impedance node in the error circuit is *current-feedback amplifier* (CFA) and x_B is thus chosen to be a current, i_B.

For a *voltage feedback amplifier* (VFA) having a single pole ($1/\tau_{bw}$) in the forward path and a frequency-independent (no poles or zeros) voltage divider for $H = H_0$, then the closed-loop bandwidth is inversely proportional to the closed-loop gain. We have derived this previously by letting

$$G = G_0 \cdot \frac{1}{s \cdot \tau_{bw} + 1}$$

and substituting it into the feedback formula for closed-loop gain,

$$A_v = \frac{G}{1 + G \cdot H} = \frac{\dfrac{G_0}{s \cdot \tau_{bw} + 1}}{1 + \dfrac{G_0}{s \cdot \tau_{bw} + 1} \cdot H_0} = \left(\frac{G_0}{1 + G_0 \cdot H_0} \right) \cdot \frac{1}{s \cdot \left(\dfrac{\tau_{bw}}{1 + G_0 \cdot H_0} \right) + 1}$$

The closed-loop bandwidth is increased from that of the loop gain by the quasistatic feedback factor, $1 + G_0 \cdot H_0$. The closed-loop gain-bandwidth product is the unity-gain frequency,

$$f_T = 2 \cdot \pi / \tau_T = f_{bw} \cdot (G_0 \cdot H_0) \Rightarrow \tau_T = \tau_{bw} / (G_0 \cdot H_0)$$

The closed-loop pole has a time constant of

$$\frac{\tau_{bw}}{1 + G_0 \cdot H_0} = \left(\frac{G_0 \cdot H_0}{1 + G_0 \cdot H_0} \right) \cdot \tau_T = A_{v0} \cdot H_0 \cdot \tau_T \approx \tau_T \text{ , } G_0 \cdot H_0 >> 1$$

Now consider what happens in amplifier 7 when v_i develops a voltage across the feedback divider, $R_f \| R_i$ (assuming R_{E1} is very large or else included in R_i), resulting in an error current of i_E which is then amplified by $G = Z_m$, a transimpedance. The output voltage is fed back to the input of H which is the R_f, R_i divider. Then T_i appears in the feedback formula (as other than 1) because of the transfer function from v_i to error current, i_E;

$$T_i = \frac{i_E}{v_i} = \frac{1}{R_f \| R_i}$$

The port resistance at the emitter node,

$$r_{E1} = r_{e1} + R_B' / (\beta_{NPN} + 1)$$

is assumed to be small so that $r_{E1} << R_f \| R_i$, and we can set $r_{E1} \approx 0\ \Omega$. Then the feedback path, H, must be a transconductance, i_B / v_o, and with the emitter node a short to ground ($r_{E1} \approx 0\ \Omega$), $i_B = v_o / R_f$ and $H = 1/R_f$. The closed-loop response for a single-pole transimpedance, Z_m, in the forward path is

$$A_v = T_i \cdot \frac{G}{1+G \cdot H} = \left(\frac{1}{R_f \parallel R_i} \right) \cdot \frac{Z_m}{1+Z_m \cdot (1/R_f)}$$

This can be rewritten as

$$A_v = \left(\frac{R_f}{R_f \parallel R_i} \right) \cdot \frac{Z_m}{Z_m + R_f} = \left(\frac{R_f + R_i}{R_i} \right) \cdot \frac{Z_m}{Z_m + R_f} = A_{v0} \cdot \frac{Z_m}{Z_m + R_f}$$

where A_{v0} is recognized as the quasistatic noninverting-op-amp voltage gain. Substitute for the forward path a transimpedance with a single pole at $1/\tau_{bw}$;

$$Z_m = \frac{R_{m0}}{s \cdot \tau_{bw} + 1}$$

Reducing the algebra, the result is the closed-loop voltage gain of a noninverting CFA:

$$A_v = A_{v0} \cdot \frac{\dfrac{R_{m0}}{s \cdot \tau_{bw} + 1}}{\dfrac{R_{m0}}{s \cdot \tau_{bw} + 1} + R_f} = A_{v0} \cdot \left(\frac{R_{m0}}{R_{m0} + R_f} \right) \cdot \frac{1}{s \cdot \tau_{bw} \cdot \left(\dfrac{R_f}{R_{m0} + R_f} \right) + 1}$$

As the quasistatic forward-path transresistance, R_{m0}, becomes large, the feedback-divider ratio to the right of A_{v0} approaches one and the closed-loop gain becomes A_{v0}. More interesting is what happens to bandwidth. As R_{m0} becomes large, the bandwidth increases to infinity! There is no gain-bandwidth product. Ideally, the closed-loop bandwidth is independent of loop gain. In practice, $r_{E1} \neq 0\ \Omega$ and the bandwidth has a limit, but is typically much higher than that of VFAs.

CFA op-amp gain is set by the same feedback-divider formulas for inverting and noninverting configurations as for VFAs. However, dynamic compensation for increased bandwidth of VFA and CFA op-amps is different. The feedback divider that is H for integrated VFAs has the input capacitance of the op-amp, C_i, to slow it down. By placing a compensating C_f across R_f that is chosen so that the time constants of the upper RC and lower RC of the divider are equal,

$$R_f \cdot C_f = R_i \cdot C_i$$

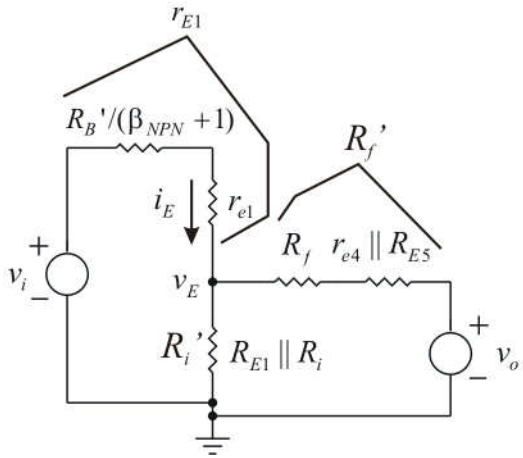
the divider has an *all-pass* (zero cancels pole) transfer function. The CFA R_f is a voltage-to-current converter and R_i has a different

purpose, that of determining i_E as part of the transresistance of the input error circuit. Thus, CFAs are not frequency-compensated by placing a shunt C_f across R_f. R_f is chosen to set the bandwidth (which is finite when $r_E \neq 0\ \Omega$) and R_i is chosen to set the quasistatic gain.

It has been possible to derive the closed-loop gain of amplifier 7 by choosing a voltage error because $r_E \neq 0\ \Omega$. A CFA interpretation of the amplifier can also be given by another choice of x_E, that of i_E.

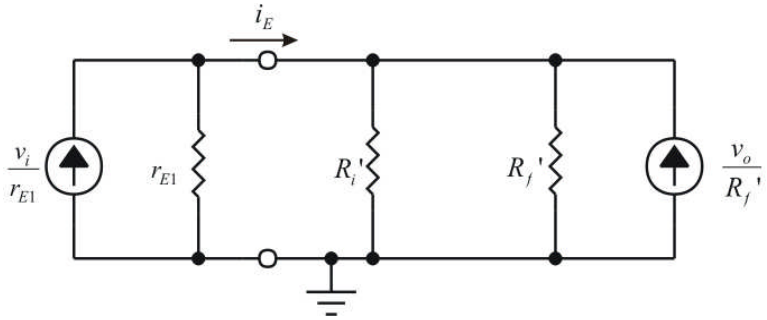
For static analysis, Thevenize $-V_{EE}$, R_{E1}, and R_i. Then $R_i' = R_{E1} \| R_i$ corresponds to R_i in the CFA model developed above and as before,

$$R_f' = R_f + r_{e4} \| R_{E5}$$



For incremental analysis, the input loop of Q1 is equivalent to the above circuit. Instead of using KVL to sum voltages around a loop, error current, i_E, is chosen as the Q1 emitter current.

The error loop can be nortonized by converting the Thevenin sources to Norton sources, as shown.



The error current is

$$i_E = i_{e1} = \frac{v_i}{r_{E1}} \cdot \left(\frac{r_{E1}}{r_{E1} + r_{Ho}} \right) - \frac{v_o}{R_f'} \cdot \left(\frac{r_{Ho}}{r_{E1} + r_{Ho}} \right) = \frac{v_i}{r_{E1} + r_{Ho}} - \frac{v_o}{R_f'} \cdot \left(\frac{r_{Ho}}{r_{E1} + r_{Ho}} \right)$$

where r_o of the *H* block is

$$r_{Ho} = R_i' \| R_f' = (R_{E1} \| R_i) \| (R_f + r_{e4} \| R_{E5})$$

and

$$r_{E1} = r_{e1} + R_B'/(\beta_{NPN} + 1)$$

As each source and series resistance is converted to a Norton equivalent circuit, it can be helpful to put terminals (small circles) at the nodes that remain the same, as in the above diagram. Then i_E does not flow through the nortonized r_{E1} but as shown, as before - from the emitter terminal in the same direction (polarity).

To approximate the ideal CFA, $r_{E1} \approx 0\ \Omega$, and the error current simplifies to

$$i_E = i_{e1} = \frac{v_i}{r_{Ho}} - \frac{v_o}{R_f'} = T_i \cdot v_i - H \cdot v_o$$

From this equation, $T_i = 1/r_{Ho}$ and $H = 1/R_f'$. The forward-path transimpedance, with G_2 as the current-mirror gain to the Q5 base, is

$$G = Z_m = \frac{v_o}{i_E} = \frac{v_o}{i_{e1}} = \alpha_{NPN} \cdot G_2 \cdot G_3$$

The closed-loop gain is then

$$A_v = T_i \cdot \frac{G}{1 + G \cdot H} = \frac{1}{r_{Ho}} \cdot \frac{\alpha_{NPN} \cdot G_2 \cdot G_3}{1 + \alpha_{NPN} \cdot G_2 \cdot G_3 / R_f'}$$

The same circuit has been analyzed in the VFA analysis and A_v must be equivalent. The above A_v can be written in VFA form as

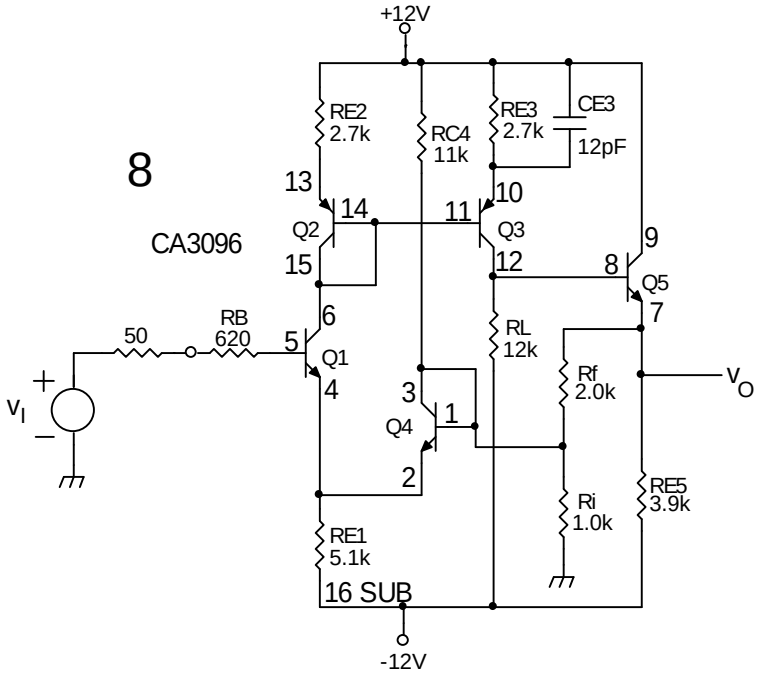
$$A_v = \frac{1}{r_{Ho}} \cdot \frac{\alpha_{NPN} \cdot G_2 \cdot G_3}{1 + \alpha_{NPN} \cdot G_2 \cdot G_3 / R_f'} = \frac{\frac{\alpha_{NPN} \cdot G_2 \cdot G_3}{r_{Ho}}}{1 + \frac{\alpha_{NPN} \cdot G_2 \cdot G_3}{r_{Ho}} \cdot \left(\frac{R_i'}{R_f' + R_i'} \right)}$$

In effect, r_{Ho} has been slid into *G*, making T_i disappear while $G \cdot H$ is factored differently in the denominator for the *H* divider. It is the same circuit structure, only partitioned into different blocks.

Amplifier 7 has a significant r_{E1} and is not close to being an ideal CFA. Op-amp CFAs approach the ideal in that the inverting input has near-zero input impedance, and is usually implemented as the output of a CC stage with small R_B'.

Noninverting 3-NPN, 2-PNP CFA

A variation on the previous CFA is amplifier 8, shown below.



The feedback divider is now taken directly off the output and the bias shift of the Q1 *b*-*e* junction is compensated by BJT Q4, connected as a diode. It is biased by R_{C4} at equal current to Q1, resulting in equal and opposing V_{BE} cancellation. Then 0 V at the bases of Q1, Q4 should result in 0 V output.

Static quantities calculate as follows: $I_0 = 2.22$ mA, $I_{E1} = 1.13$ mA, and $V_{CE1} = 8.27$ V. $I_{E2} = 1.09$ mA, $I_{E3} = 1.08$ mA and $V_{B5} = 0.94$ V and $I_{E5} = 3.15$ mA. Then $V_O \approx 0.22$ V. The static output error is corrected largely by feedback.

The dynamic response for amplifier 8 (as for CFA amplifier 7) has a Miller-effect pole for Q1 because of load resistance $R_{E2} + r_{e2}$,

reduced slightly by the loading of Q3. The current-mirror emitter resistors have been reduced from those of amplifier 7 to reduce Q1 Miller capacitance at the expense of greater thermal distortion. Amplifiers 7 and 8 differ in circuit structure from that of amplifier 6 by making Q4 a diode instead of a diff-amp BJT, as is Q2 in amplifier 6. Consequently, r_{Ho} is much lower.

There can be significant C_c poles (lower than a decade above f_T) for Q1, Q3, and Q5. Because Q5 is a CC, C_{c5} affects the output pole of Q3, τ_{L3}. C_{e5} has a high-frequency pole because its emitter circuit has high resistance, causing $\tau_{T5}/[R_{E5}\|(R_f + R_i')]$ to be a small gyrated capacitance (0.3 pF). It forms a time constant (with R_{b5} = 11.84 kΩ) of about 3.55 ns ⇒ 45MHz.

The second-stage forward-path gain is

$$G_2 = -11.84\ \text{k}\Omega/2.73\ \text{k}\Omega \approx -4.34$$

where R_L is loaded by the Q5 base input resistance of 663 kΩ. The Q5 output stage $G_3 \approx 0.975$. Then $G_0 \approx 20.3$ (much like amplifier 7).

The output time constant is estimated as

$$\tau_{L3} = R_L \cdot (C_{c5} + C_{cs3}) \approx (11.84\ \text{k}\Omega)\cdot(0.5\ \text{pF} + 2\ \text{pF}) = 29.6\ \text{ns} \Rightarrow 5.38\ \text{MHz}$$

where $K_{v3} \approx -G_2$. The time constant of C_{c3} is

$$\tau_{c3} = \tau_{cc3} + \tau_{cb3}$$

where

$$\tau_{cc3} = R_L \cdot C_{c3} = (11.84\ \text{k}\Omega)\cdot(2\ \text{pF}) = 23.68\ \text{ns}$$

$$\tau_{cb3} = R_{b3} \cdot [C_{c3} \cdot (1 + K_{v3})] = (2.73\ \text{k}\Omega)\cdot[(0.5\ \text{pF})\cdot(5.34)] = 7.29\ \text{ns}$$

These are not two separable time constants but are added to give

$$\tau_{c3} = \tau_{cc3} + \tau_{cb3} = 23.68\ \text{ns} + 7.29\ \text{ns} = 30.97\ \text{ns} \Rightarrow 5.14\ \text{MHz}$$

A third time constant of Q3 is that of C_{b3}. At I_{E3} = 1 mA, τ_{T3} = 568 ps and the hf R_E-gyrated capacitance at the base of Q3 is (568 ps)/2.7 kΩ = 0.21 pF. This capacitance is in series with R_E;

$$\tau_{b3} = R_{b3}\cdot(C_{bs2} + C_{bs3}) + [R_{E3} + (r_{e2} + R_{E2} + 2\cdot r_b')]\cdot(0.21\ \text{pF}) = (2.72\ \text{k}\Omega)\cdot[2\cdot(2.5\ \text{pF})] + (5526\ \Omega)\cdot(0.21\ \text{pF}) = 13.6\ \text{ns} + 1.16\ \text{ns} = 14.76\ \text{ns} \Rightarrow 10.78\ \text{MHz}$$

Moving backward through the forward path to Q1, the $b \rightarrow c$ voltage gain,

$$K_{v1} \approx \frac{R_{E2} + r_{e2}}{r_{e1} + r_{Ho}} = \frac{2.73\,\text{k}\Omega}{26\,\Omega + 560\,\Omega} = 4.66$$

where r_{Ho} is approximately the resistance of the four external emitter resistors in parallel (R_{E1}, R_{C4}, R_f, R_i), or

$$r_{Ho} = 560\ \Omega$$

The Q1 collector time constant is

$$\tau_{L1} \approx (2.72\,\text{k}\Omega) \cdot (C_{c1} + C_{cs1}) = 6.8\,\text{ns} \ \Rightarrow 23.4\ \text{MHz}$$

The pole of τ_{L1} is located at the collector node and is within the G path. The CFA feedback interpretation of the circuit places T_i at the input, posing for us the problem of how to handle the time constant of C_{c1} which involves both base and collector nodes. Should it be placed in T_i or G? T_i is the transfer function defining $v_i \rightarrow i_E$, and i_E is affected by the rolloff of v_{b1} because $i_E = v_{b1}/(r_{e1} + r_{Ho})$. Thus T_i is affected by the Q1 time constants that affect the base node: τ_{c1} of C_{c1} and τ_{e1} of C_{e1}. If they are placed in T_i, then G and H have no poles or zeros from them and the CFA form of closed-loop voltage gain is

$$A_v = \frac{1}{z_{Ho}(s)} \cdot \frac{\alpha_{NPN} \cdot G_2 \cdot G_3}{1 + \alpha_{NPN} \cdot G_2 \cdot G_3 / R_f'}$$

T_i includes not only the C_{c1} OCTC but also that of C_{e1};

$$\tau_{c1} = [(R_{E2} + r_{e2}) + R_{B1}'\cdot(1 + K_{v1})] \cdot C_{c1} = 3.4\ \text{ns} \Rightarrow 46.8\ \text{MHz}$$

$$\tau_{cc1} = (R_{E2} + r_{e2}) \cdot C_{c1} = 1.36\ \text{ns}\ ;\ \tau_{cb1} = 2.038\ \text{ns}$$

$$\tau_{e1} = (\alpha_{NPN} \cdot \tau_{T1}) \cdot \frac{R_B' + r_{Ho}}{r_{Ho} + r_{e1} + R_B'/(\beta_{NPN} + 1)} = 1.23\ \text{ns} \Rightarrow 129\ \text{MHz}$$

These OCTCs interact to form poles with quadratic parameters

$$\tau_{ni} = \sqrt{\tau_{e1} \cdot \tau_{cc1}} = 1.295\,\text{ns} \Rightarrow 123\ \text{MHz}$$

$$\zeta_i = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_{e1}}{\tau_{cc1}}} + \sqrt{\frac{\tau_{cc1}}{\tau_{e1}}} + \frac{\tau_{cb1}}{\tau_{ni}} \right) = 1.788$$

The pole frequencies, in Hertz, are

$$\frac{1}{2\cdot\pi}\cdot\frac{\zeta_i}{\tau_{ni}}\cdot\left(1\pm\frac{\sqrt{\zeta_i^2-1}}{\zeta_i}\right) = 37.6 \text{ MHz, } 402 \text{ MHz}$$

The dominant pole is in T_i Any bandwidth limitation in T_i - at the input of the feedback loop - limits CFA bandwidth. Because T_i and the feedback loop are effectively cascaded, their time constants can be combined with the OCTC formula.

The loop has pole time-constants at: τ_{L3} = 53.3 ns (2.99 MHz); τ_{c3} = 30.97 ns (5.14 MHz); τ_{e3} = 7.4 ns (21.5 MHz); τ_{L1} = 6.8 ns (23.4 MHz); and τ_1 = 3.36 ns (47.4 MHz). The loop-gain bandwidth is only roughly approximated by the lowest-frequency pole, at 2.99 MHz because it is only about a quarter decade below the next-lowest pole frequency and two more poles are within a decade. Consequently, single-pole open-loop response approximation will not be accurate enough for this amplifier. All the poles are real and the OCTC formula can be used; open-loop

$$\tau_{bw} \approx \sqrt{\tau_{L3}^2+\tau_{c3}^2+\tau_{e3}^2+\tau_{L1}^2+\tau_1^2} \approx 62.6 \text{ ns} \Rightarrow 2.54 \text{ MHz}$$

The fastest time constant, τ_1, of T_i does not affect τ_{bw} much. The lowest two poles are separated by only 0.24 decade and that of τ_{c3} is 0.62 decade from the next highest pole (of τ_{e3}) which is nearly equal to that of τ_{L1}. With the lowest two poles so close, some loop compensation is needed for stability. A zero can be added to cancel τ_{c3} by adding C_{E3} = (30.97 ns)/(2.7 kΩ) = 11.5 pF → 12 pF across R_{E3}. The remaining pole is at about 160 MHz, a decade beyond all but that of τ_1. The low loop gain aids stability.

H is a passive voltage divider with no poles or zeros in the CFA feedback formula. Hence the loop bandwidth is also the bandwidth of *G*. The quasistatic

$$H = 1/R_f' = 1/(R_f + R_{E5}||r_{s5}) \approx 1/2039\ \Omega$$

and the quasistatic

$$T_i = 1/r_{Ho} = 1/587\ \Omega$$

The quasistatic closed-loop gain is

$$A_{v0} = T_i \cdot \frac{G}{1 + G \cdot H} = \frac{1}{r_{Ho}} \cdot \frac{\alpha_{NPN} \cdot R_{E2} \cdot G_2 \cdot G_3}{1 + \alpha_{NPN} \cdot R_{E2} \cdot G_2 \cdot G_3 / R_f'}$$

$$= \frac{1}{587\ \Omega} \cdot \frac{11.52\ \text{k}\Omega}{1 + (11.52\ \text{k}\Omega)/(2039\ \Omega)} = 2.95$$

or slightly less than the ideal set by $1/H = 3$.

For a dominant pole in G, an approximation of the closed-loop bandwidth can be made from

$$A_v = A_{v0} \cdot \left(\frac{R_{m0}}{R_{m0} + R_f} \right) \cdot \frac{1}{s \cdot \tau_{bw} \cdot \left(\frac{R_f}{R_{m0} + R_f} \right) + 1}$$

The resistance values are $R_{m0} = 11.52\ \text{k}\Omega$ and $R_f = R_f' = 2039\ \Omega$. Using the OCTC $f_{bw} = 2.54$ MHz, then the bandwidth pole is at

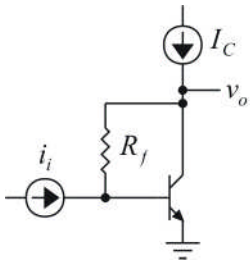
$$f_{bwc} = \frac{f_{bw}}{\left(\frac{R_f}{R_{m0} + R_f} \right)} = \frac{2.54\ \text{MHz}}{6.65} = 16.9\ \text{MHz}$$

and is somewhat higher than the measured $f_{bwc} = 10.4$ MHz. The calculated open-loop bandwidth is also high and comparable to that of the fast-BJT cascade approximation.

Shunt-Feedback Amplifiers

The *shunt-feedback amplifier* (SFA) is named after the *c-b* impedance, Z_f, shunting the BJT across its input and output nodes. In the "Feedback Amplifiers" chapter, the gain formulas were derived along with some laboratory experience with a simple SFA.

In this section, we look mainly at the dynamics, and start with a simplified shunt-feedback stage, as shown.



R_L has been replaced by a current source, as has the input source. The ideal SFA approaches that of an inverting op-amp, with minimal r_M and maximum R_L to maximize gain. The feedback circuit, Z_f, is intended to set the closed-loop gain and increase bandwidth. The base is the input node and is driven by a current source so that none of the feedback current of R_f is diverted from the base.

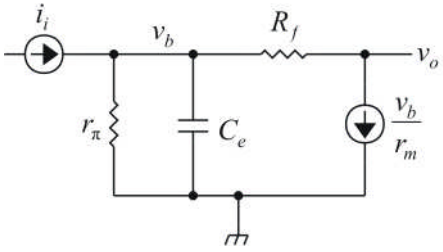
The transresistance of the above stage was previously derived as

$$R_m = \frac{v_o}{i_i} = -\alpha_0 \cdot R_f + r_e$$

and in its simplest form, when $\beta_0 \to \infty$ and $r_e = 0\ \Omega$, is $R_m = -R_f$. This can be derived intuitively from the circuit diagram. The feedback current, v_o/R_f, and i_i add to become the input to the BJT which is, in this simplified view, a transconductance amplifier with infinite R_m. Consequently, it needs only an infinitesimal (0+) input error current to output v_o, and the feedback and input currents must be equal, or

$$\frac{v_o}{R_f} = -i_i \Rightarrow \frac{v_o}{i_i} = -R_f$$

The simplest dynamic SFA includes C_e. The equivalent circuit is shown below.



The shunt r_π, C_e circuit has impedance

$$Z_\pi = r_\pi \,\|\, (1/s \cdot C_e) = \frac{r_\pi}{s \cdot \tau_\beta + 1},\ \tau_\beta = r_\pi \cdot C_e$$

Apply KCL at base and collector:

$$\text{KCL @ } b\text{: } \frac{v_b}{Z_\pi} = i_i + \frac{v_o - v_b}{R_f} \Rightarrow \frac{v_b}{r_m \,\|\, Z_\pi} = i_i$$

$$\text{KCL @ } c\text{: } \frac{v_o - v_b}{R_f} = -\frac{v_b}{r_m} \Rightarrow v_b = \frac{r_m}{-R_f + r_m} \cdot v_o$$

Substitute v_b from the *c* equation into the *b* equation:

$$\frac{r_m}{(-R_f + r_m) \cdot (r_m \parallel Z_\pi)} \cdot v_o = i_i \Rightarrow \frac{v_o}{i_i} = (-R_f + r_m) \cdot \frac{Z_\pi}{r_m + Z_\pi}$$

To simplify the right factor, we will need

$$\frac{r_\pi}{r_\pi + r_m} = \frac{(\beta_0 + 1) \cdot r_e}{(\beta_0 + 1) \cdot r_e + \frac{(\beta_0 + 1)}{\beta_0} \cdot r_e} = \frac{1}{1 + 1/\beta_0} = \alpha_0$$

$$\frac{r_m}{r_\pi + r_m} = 1 - \frac{r_\pi}{r_\pi + r_m} = 1 - \alpha_0 = \frac{1}{\beta_0 + 1}$$

Then

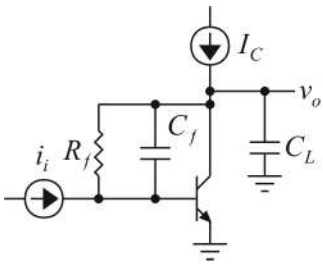
$$\frac{v_o}{i_i} = \alpha_0 \cdot (-R_f + r_m) \cdot \frac{1}{s \cdot \frac{\tau_\beta}{\beta_0 + 1} + 1}$$

Finally,

$$Z_m(s) = (-\alpha_0 \cdot R_f + r_e) \cdot \frac{1}{s \cdot (\alpha_0 \cdot \tau_T) + 1} = \alpha(s) \cdot (-R_f + r_m)$$

The result is that $Z_m(s)$ has a pole at f_T/α_0, showing that fast SFAs are possible despite the lack of R_E to keep τ_e low. However, the two other anticipated parasitic capacitances have yet to be added.

The other capacitances that limit bandwidth are C_c of the *b*-*c* junction and the collector node capacitance, C_L, both shown in the following three-C SFA circuit.



C_f is an external capacitor shunting C_c and is often larger than C_c. In previous circuits, C_c reduced bandwidth; in the SFA, it has a somewhat different effect. To unclutter equations, it is included in C_f.

The effect of C_e is included in the Z_m derivation by using a total-frequency BJT model. Then for $R_f >> r_e$ and high β_0, the hf BJT model can be applied, for which $\beta_0 \to \infty$ and $r_e = 0\ \Omega$.

$$i_c(s) = \beta_{hf} \cdot i_b(s) = \frac{1}{s \cdot \tau_T} \cdot i_b(s)$$

The feedback circuit is

$$Z_f = R_f \parallel (1/s \cdot C_f) = \frac{R_f}{s \cdot R_f \cdot C_f + 1} = \frac{R_f}{s \cdot \tau_f + 1}$$

The base current (like the simple SFA) is

$$i_b = i_i + v_o/Z_f$$

Applying KCL at the collector,

$$\frac{v_o}{Z_f} + \frac{i_b}{s \cdot \tau_T} + \frac{v_o}{1/s \cdot C_L} = 0\,\text{mA}$$

Substituting for i_b,

$$\frac{v_o}{Z_f} + \frac{v_o}{s \cdot \tau_T \cdot Z_f} + \frac{v_o}{1/s \cdot C_L} = -\frac{i_i}{s \cdot \tau_T}$$

The three impedances under v_o (the middle one an inductance) are in parallel. Solving for the transresistance,

$$\text{hf}\ \ Z_m(s) = \frac{v_o}{i_i} = -\frac{Z_f \parallel (s \cdot \tau_T \cdot Z_f) \parallel (1/s \cdot C_L)}{s \cdot \tau_T}$$

Z_f is gyrated +90° by $s \cdot \tau_T$ at the collector, the same as would occur at the emitter for a base Z_f. This impedance forms a resonance with C_L and we can expect a quadratic pole-pair. With some additional algebra, this simplifies to normalized form;

$$\text{hf}\ \ Z_m(s) = -R_f \cdot \frac{1}{s^2 \cdot [\tau_T \cdot R_f \cdot (C_f + C_L)] + s \cdot [\tau_T + R_f \cdot C_f] + 1}$$

The static collector current source in design practice is not uncommonly implemented by a high-value R_L. If the I_C source were replaced by R_L, then a collector time constant, $\tau_L = R_L \cdot C_L$ appears, and when Z_m is derived to include it, the result is

$$\text{hf } Z_m = -R_f \cdot \frac{1}{s^2 \cdot [\tau_T \cdot R_f \cdot (C_f + C_L)] + s \cdot \left[\tau_T \cdot \left(1 + \frac{R_f}{R_L} \right) + R_f \cdot C_f \right] + 1}$$

As R_L decreases from its current-source ideal value, the τ_T term in the *b* coefficient increases, effectively reducing the f_T of the BJT.

What has been analyzed is a general single-stage amplifier that can also be reduced to the textbook CE stage by removing R_f and letting $C_f = C_c$. Then all three capacitances are included in the gain. The RHP zero is absent because of the hf simplification, whereby the base-node impedance is zero and the passive path through C_f to the output is shorted by the active-path (BJT) input. To convert Z_m to a CE, multiply numerator and denominator by $1/R_f$ and let $R_f \to \infty$;

$$Z_m\big|_{R_f \to \infty} = -\frac{1}{s \cdot \left(\frac{\tau_T}{R_L} + C_f \right)} \cdot \frac{1}{s \cdot \left[(\tau_T \parallel R_L \cdot C_f) \cdot \left(1 + \frac{C_L}{C_f} \right) \right] + 1}$$

The transfer function retains the unit of Ω because the left factor is a capacitive reactance. It has a *c*-*b* capacitance of C_f shunting a –90° hf-gyrated R_L of τ_T/R_L. The polarity of gyration views the capacitance from the base. C_f is in parallel, not in series, with it because impedance at the collector node of an SFA is gyrated as though in the emitter when referred to the base. (Keep in mind that ground, or 0 V, is a human convention imposed on the circuit. Through ground, R_L is in series with the emitter.) The pole at the origin has an f_T (or projected f_T) interpretation. It is accompanied by a second pole with a time constant that can be interpreted in the following equivalent form:

$$\left[\frac{\tau_T \parallel (R_L \cdot C_f)}{C_f \parallel C_L} \right] \cdot C_L$$

The capacitance as viewed from the base, $C_f \| C_L$, is the math expression for C_f in series with C_L. The left (bracketed) factor is a collector capacitance gyrated +90° to become a resistance that forms a time constant at the pole frequency with C_L. It is, in other words, a

collector-referred view of the pole, and it is a hf effect - at least in this interpretation. It also demonstrates the way in which C_e, through τ_T, participates in the formation of the pole. The time constant is minimized for $R_L \cdot C_f = \tau_T$.

The hf pole polynomial of the basic 3-C SFA is

$$s^2 \cdot [\tau_T \cdot R_f \cdot (C_f + C_L)] + s \cdot [\tau_T + R_f \cdot C_f] + 1$$

Unlike CE amplifiers in which the linear coefficient, b, has an additional $\tau(b)$ term, τ_{cb}, that makes $\zeta > 1$, the SFA has the opposite; τ_T is common to a and b, or $\tau_T = \tau(a, b)$, and the remaining term in b - that of $\tau_f = \tau(b)$ - also appears in a but along with $R_f \cdot C_L$, and is larger. This has the opposite effect, that of making $\zeta < 1$. With C_L removed, $D(s)$ factors into time constants τ_T and τ_f. With C_L,

$$\zeta = \frac{b}{2 \cdot \sqrt{a}} = \frac{1}{2} \cdot \frac{\tau_T + \tau_f}{\sqrt{\tau_T \cdot (\tau_f + R_f \cdot C_L)}} = \frac{1}{2} \cdot \left(\sqrt{\frac{\tau_T}{\tau_f + R_f \cdot C_L}} + \sqrt{\frac{\tau_f}{\tau_T \cdot (\tau_f + R_f \cdot C_L)}} \right)$$

For $C_L = 0$ pF and $\tau_T = \tau_f$, ζ has its maximum value of 1. As the denominator under the radicals increases, the values of the radicals become smaller and the resulting $\zeta < 1$. SFAs have complex pole-pairs in their response.

For dynamic design, we want to know what the value of C_f should be for a desired ζ. R_f is constrained in design by the given quasistatic transresistance. The equation is solved for C_f and is

$$C_f = \frac{1}{(2 \cdot \zeta)^2 \cdot R_f} - C_L$$

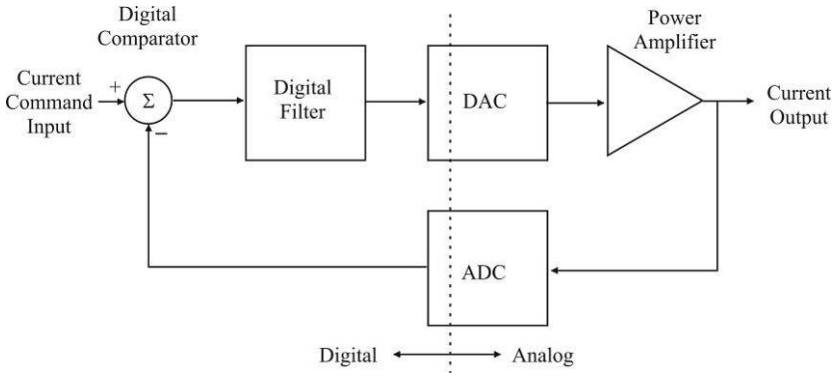
The SFA can be improved for speed by replacing the CE forward path with a cascode amplifier. Another possibility is to replace it with a Darlington “gain cell”. SFAs can also be made differential and are found in the vertical amplifiers of oscilloscopes.

A Wider View of Amplifiers

A/D and D/A Converters in Circuits

One of the confusing aspects of circuit design involves circuits with both digital and analog quantities. Circuits often have both *analog* (continuous) and *digital* (discrete) subsystems and with analog-to-digital (A/D) converters (ADCs) and digital-to-analog (D/A) converters (DACs), waveforms appear in both analog and digital representations.

To appreciate the problems that can arise, consider the general scheme of an analog and digital (or *mixed-signal*) feedback system, shown below.



An analog output device (which could be a motor or a loudspeaker) is within a feedback loop. In the feedback path, an ADC converts the output current to a digital value of *n* bits. This quantity representing the current is then compared to the commanded value, also digital, and the error processed by a digital filter. Its output is D/A-converted to an analog waveform which drives the power amplifier. The digital filter has a digital transfer function with various poles and zeros in its frequency response to compensate the dynamic response of the loop. How can a feedback loop with digital processing in it be analyzed?

Selecting one component, the ADC, we begin our analysis. An ADC has n digital output bit-lines, plus data-transfer control lines, one analog input voltage to be converted, and a full-scale reference input voltage, V_R. A transfer-function view of ADCs is that their digital output ratio equals their analog input ratio;

$$\frac{w_O}{2^n} = \frac{v_I}{V_R}$$

where n is the number of bits of the ADC, and 2^n is the digital range of w_O, the output digital value. This fraction is the same as the (analog) input voltage, v_I, over the reference voltage, V_R. The value of w_O is

$$w_O = \left(\frac{v_I}{V_R}\right) \cdot 2^n$$

This value has no units, though it represents v_I. It is the fraction of 2^n that v_I is of V_R. An ADC is used to acquire v_I in digital form. Yet what we actually have out of the ADC is a unitless fraction.

The ADC digital output, w_O, is not the desired digital voltage value but is a fraction of 2^n, part of a digital ratio. On the digital side of the ADC, we must "undo" the scaling that occurred in the ADC to recover the value of v_I. If the microcomputer (μC) performs the inverse ADC function on w_O the result should be a digital voltage of value equal to v_I;

$$v_I = \left(\frac{V_R}{2^n}\right) \cdot w_O$$

This inverse function is an additional software block following the ADC, though not shown above. It might be labeled "ADC^{-1}". Looked at from this equation, the ADC is an amplifier with a quasistatic gain equal to the scaling ratio $V_R/2^n$. By representing V_R as 2^n, then each least-significant bit (LSB) of digital value - a one-count change in w_O - is equal to

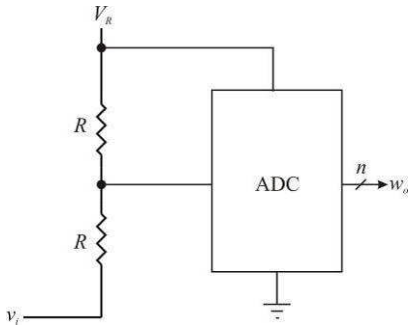
$$V_{LSB} = V_R/2^n = V_R \cdot 2^{-n}$$

For example, an 8-bit ADC ($n = 8$) with a reference voltage of $V_R = 5$ V has a minimum change in voltage, or *voltage resolution*, of

$$V_{LSB} = (5\text{ V})/2^8 = (5\text{ V})/256 \approx 19.53\text{ mV}$$

This is the voltage change that results in a change of one in its digital output. The digital value of the analog input voltage is $(V_R \cdot 2^{-n}) \cdot w_O$, or about (19.53 mV)$\cdot w_O$. For V_R = 5.12 V, V_{LSB} is an even 20 mV.

Suppose the design requires the ADC input voltage to have a range of −5 V at negative full-scale (−fs) and +5 V at +fs with 0 V at zero-scale (zs). On the digital side, digital numbers are usually represented (or *coded*) in *two's complement arithmetic*, though the interpretation of the digital values depends on the corresponding analog values.



The ADC block can include within it not only the ADC core but also a divide-by-2 divider returned to V_R, as shown above. It offsets the analog input voltage, v_I, for a bipolar range. The transfer function is affected by this mid-scale offset and this affects the interpretation of the digital input value so that 0 V is mid-scale of the ADC input, corresponding to $w_O = 2^{n-1}$. The range of v_I is $[-V_R/2, +V_R/2)$.

Three points in the range of the transfer function are given in the table below, with some alternative digital values.

Analog Input, V	Two's Complement Output	Offset Binary Output
−5 V (−fs)	−2048 (800 hex)	0
0 V (zs)	0	2048 (800 hex)
$+V_R - V_{LSB}$ (+fs)	2047 (7FF hex)	4095 (FFF hex)

The quasistatic transfer function of the ADC becomes $(V_R/2^{n-1})$ and the range is $\pm(V_R/2^{n-1})$. The ADC can thus be envisioned as a linear analog to digital amplifier with a gain that is the w_o/v_i transfer function.

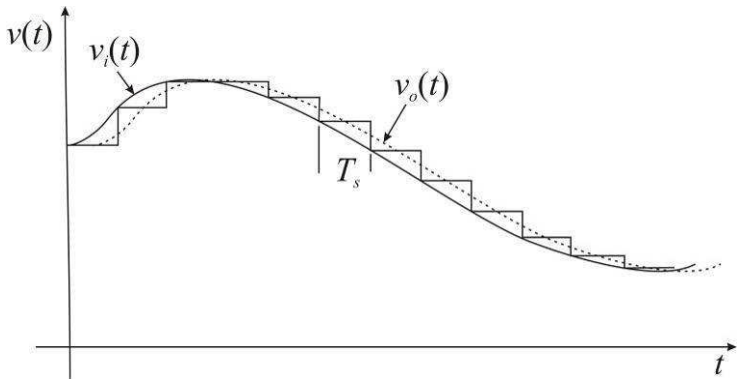
ADCs have a dynamic gain factor involving *s*-domain poles like transistor amplifiers. Dynamic behavior is caused by sampling within ADCs. Their digital output is clocked into a holding register.

Dynamically, ADCs are *sample-and-hold* or *zero-order hold* (*ZOH*) circuits. After the sampling edge of the clock causes storage of the digital value into the ADC register, the new w_o value remains at the output of the ADC and is *held*. The rate of sampling is $f_s = 1/T_s$.

Plotted against time, the digital ADC output waveform,

$$v_o(t) = w_o(t) \cdot \frac{V_R}{2^n}$$

has the stair-step look of a ZOH waveform. The voltage waveform, $v_i(t)$, at the ADC input is shown plotted below with $v_o(t)$. The output voltage is reconstructed as a continuous waveform by shifting the input waveform to the right by $T_s/2$ so that it passes through the centers of the horizontal steps of the sampled waveform. A continuous approximation of $v_o(t)$ is shown as $v_i(t)$ delayed by $T_s/2$.



In the *s*-domain, a waveform can be delayed in time (to happen in the future) by the following *time-shifting* transformation from the time domain to complex-frequency domain:

$$x(t - T_s) \Rightarrow x(s) \cdot e^{-s \cdot T_s} = x(s) \cdot e^{-s/f_s}$$

The subtraction $(t - T_s)$ of T_s from t in x shifts $x(t)$ in time to the right on the time axis by T_s, thereby delaying x by T_s. When transformed to the *s*-domain, a waveform is delayed by T_s by multiplying its *s*-domain transform, $x(s)$, by $e^{-s \cdot T_s}$. Then one sampling period of a unit-step function, $u(t)$ can be isolated from the rest of the waveform in time, or *gated* (or "windowed"), by subtracting from it a delayed version of itself T_s later:

$$u(t) - u(t - T_s)$$

This function is non-zero only during the time interval of one step, 0 s < $t \leq T_s$. This transforms to the *s*-domain as

$$\frac{1}{s} - \frac{e^{-s \cdot T_s}}{s}$$

Multiplying an ADC input waveform by this function (in either the time or complex-frequency domain) gates on and off arbitrary waveforms for one sampling period. The ADC output successively applies gating each sampling period. The gating function for the second clock period is

$$u(t - T_s) - u(t - 2 \cdot T_s)$$

and for the *n*th period, it is

$$u(t - (n-1) \cdot T_s) - u(t - n \cdot T_s)$$

In an ADC, the value of $v_i(t)$ is sampled and then held by integrating the impulse of voltage that was sampled at the instant of the clock edge. Sampling ideally occurs at an instant in time and a waveform that lasts but for an instant is an *impulse*. The *unit impulse* is the derivative of the unit step. At t = 0+ the step has a discontinuity in changing value from zero at t = 0 s to one at 0+ s an instant later. The derivative of the step is infinite and on a plot an impulse is graphed symbolically as an upward arrow;

$$\delta(t) = \begin{cases} \infty, t = 0+ \\ 0, t \neq 0+ \end{cases}$$

The impulse nevertheless has a value that is its "area" and is like an amplitude. The holding of an *impulse* of sampled voltage is accomplished by integrating it over T_s. For the unit impulse, its value in general is

$$\int \delta(t) \cdot dt = 1$$

The ADC sampling function can be described as a sequence (or "train") of unit impulses that are multiplied by $v_i(t)$. Then $v_i(t) \cdot \delta(t)$ is integrated, gated, and output by the ADC as a digital value for each sampling interval. For the *n*th sampling interval,

$$\int_{(n-1)\cdot T_s}^{n\cdot T_s} v_i(t)\cdot\delta(t)\cdot dt = v_o(t)\,,\,(n-1)\cdot T_s < t \le n\cdot T_s$$

The $v_i(t)$ within the integral is continuous whereas $v_o(t)$ is the ZOH stair-stepped or *piecewise-continuous* v_i, the digital (discrete-time) version of the analog (continuous) $v_i(t)$. We now examine this difference mathematically.

The normalized ZOH function in the *s*-domain is

$$H_0(s) = \frac{1}{T_s}\cdot\left(\frac{1}{s} - \frac{e^{-s\cdot T_s}}{s}\right) = \frac{1}{s\cdot T_s}\cdot(1-e^{-s\cdot T_s}) = \frac{1-e^{-s\cdot T_s}}{s\cdot T_s}$$

where $(1-e^{-s\cdot T_s})/s$ is the gated function over T_s and $1/s\cdot T_s$ is integration over any T_s interval as a pole at the origin. The frequency-response magnitude and phase are found by setting $s = j\cdot\omega$;

$$H_0(j\cdot\omega) = \frac{1-e^{-j\cdot\omega\cdot T_s}}{j\cdot\omega\cdot T_s} = \frac{\sin(\omega\cdot T_s/2)}{\omega\cdot T_s/2}\cdot e^{-j\cdot\omega\cdot T_s/2}$$

The rightmost expression is obtained by multiplying the expression by one in the form of $1 = e^{+j\cdot\omega\cdot T_s/2}\cdot e^{-j\cdot\omega\cdot T_s/2}$. Then

$$H_0(j\cdot\omega) = \frac{1-e^{-j\cdot\omega\cdot T_s}}{j\cdot\omega\cdot T_s} = \frac{e^{+j\cdot\omega\cdot T_s/2} - e^{-j\cdot\omega\cdot T_s/2}}{j\cdot\omega\cdot T_s}\cdot e^{-j\cdot\omega\cdot T_s/2}$$

From the "Sine-Waves" section of "Circuit Dynamics", substitute

$$\sin\theta = \frac{e^{j\cdot\theta} - e^{-j\cdot\theta}}{2\cdot j} \Rightarrow e^{j\cdot\theta} - e^{-j\cdot\theta} = 2\cdot j\cdot\sin\theta$$

where $\theta = \omega\cdot T_s/2$. The *j*s cancel and the 2 is moved to the denominator to divide T_s.

The ADC frequency response is the magnitude and phase of $H_0(j\cdot\omega)$. The phase, $\omega\cdot T_s/2$, can be expressed as

$$\omega\cdot T_s/2 = \frac{2\cdot\pi\cdot f}{2\cdot f_s} = \pi\cdot\frac{f}{f_s}$$

The magnitude of $H_0(j\cdot\omega)$ is

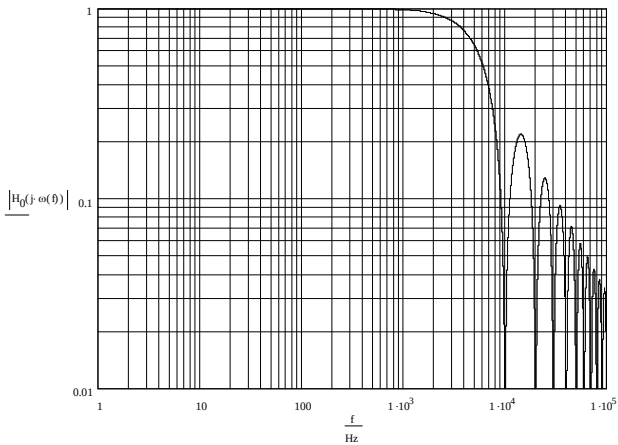
$$\left|H_0(j\cdot\omega)\right| = \left|\frac{\sin(\omega\cdot T_s/2)}{\omega\cdot T_s/2}\right| = \left|\frac{\sin(\pi\cdot f/f_s)}{\pi\cdot f/f_s}\right|$$

and the phase is

$$\angle H_0(j\cdot\omega) = -\frac{\omega\cdot T_s}{2} + \theta = -\pi\cdot\left(\frac{f}{f_s}\right) + \theta \;,\; \theta = \begin{cases} 0, \sin(\omega\cdot T_s/2) > 0 \\ \pi, \sin(\omega\cdot T_s/2) < 0 \end{cases}$$

The phase function is linear and in degrees is $-180°\cdot(f/f_s)$.

The magnitude is plotted below for f_s = 10 kHz. This response looks somewhat like a pole though it does not roll off with an asymptotic slope of –1 but bounces to zero every $n\cdot f_s$ with decreasing amplitude at the peaks of the bounces. The peaks decrease at a –1 slope. $H_0(j\cdot\omega)$ consists of an infinite number of poles. The response appears flat to about $f_s/10$ (1 kHz) and at the *Nyquist frequency* of $f_s/2$ is down to $2/\pi \approx 0.637$.



The Nyquist frequency sets a boundary below which $v_i(t)$ can be recovered from the sampling process. Any equal or higher frequency than $f_s/2$ in the sampled waveform results in a difference (or "beat") frequency with f_s that produces lower-frequency *alias* waveforms that mix with the original $v_i(t)$ and distort it. Aliasing occurs when tuning stringed instruments as the difference frequency that is tuned out (to 0 Hz). It also appears visually as backward-turning wheel spokes on video. The video sampling rate aliases with the frequency at which the spokes are rotating and produces the negative-frequency (backward-turning) aliasing effect.

The ADC response can be compensated in software with an inverse function (not shown, but possibly following the ADC), called a *reconstruction filter*. It can be omitted if the ADC sample rate is so much higher than the loop bandwidth that its amplitude roll-off and phase delay are insignificant. If the highest frequency component of $v_i(t)$ is $f_s/10$ or lower, then the sampled waveform is practically continuous though in digital form, and the ADC is nearly quasistatic in operation.

At the Nyquist frequency of $f_s/2$, the magnitude will have decreased to $(\sin(\pi/2))/(\pi/2) = 2/\pi \approx 0.637$ with a phase delay of $-180°/2 = -90°$. A sampling rate of 10 times f_s results in a normalized amplitude of 0.9836, or about 6 bits in accuracy. Some values are given in the following table.

f/f_s	$\omega \cdot T_s/2 = \pi \cdot (f/f_s)$	$\sin(\omega \cdot T_s/2)/(\omega \cdot T_s/2)$	$\angle H_0(j\omega)$, deg
1.00	$\pi/1$	0	−180
0.50	$\pi/2$	0.637	−90
0.33	$\pi/3$	0.827	−60
0.25	$\pi/4$	0.900	−45
0.20	$\pi/5$	0.936	−36
0.167	$\pi/6$	0.955	−30
0.125	$\pi/8$	0.975	−22.5
0.10	$\pi/10$	0.984	−18

In a feedback control loop, the ADC dynamic response can have a significant effect on loop behavior depending on f/f_s.

DACs perform the inverse function of ADCs and have the same transfer-function characteristics as ADCs: a quasistatic gain of the same form except that the analog ratio is the output with the same dynamic behavior. For the DAC, the magnitude roll-off of $\sin\theta/\theta$ can be compensated digitally by a software function of ZOH^{-1} preceding the DAC if the f/f_s value is known.

The Conceptual Organization of Circuits

This book started by plunging immediately into circuit concepts and has continued in them unabated. We now close in reflection upon what we have been doing and look at the larger picture of design activity.

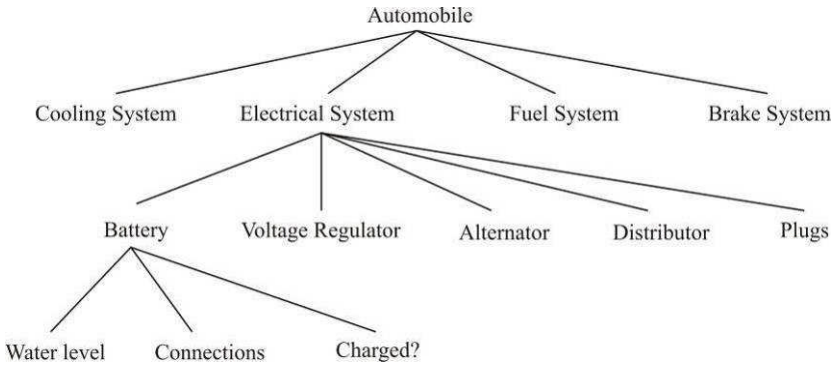
A *system* is anything consisting of a set of interacting elements. Electronic *systems* are often so complicated that we cannot think about all the details at once. Systems are often organized into *hierarchies* having multiple levels of manageable complexity, not unlike human organizations. Electronic systems can be described by a multilevel hierarchy of concepts. At the most concrete level is the physical circuits themselves, represented commonly by a circuit diagram which gives us a *structural* description of the circuit: what the components are and how they are interconnected. From structure, various electrical (and thermal or mechanical) behaviors are deduced through a causal theory of circuits, by analysis. When analyzed, a *behavioral* description of the circuit results, usually as waveforms. At the next more abstract level of description, these behaviors are explained by a *functional* theory that leads to a functional description or *specification* of what its purpose is or what it achieves.

Each of these descriptions in itself may be complicated enough to require a hierarchical organization. For example, a structural description of a DVD player consisting of hundreds of parts is too unwieldy to handle directly. Systems are consequently organized into *subsystems*, graphically described on a *block diagram*. These subsystems consist of *circuits* that, in turn, are composed of circuit *elements*, which are implemented as *components* in actual circuits. (Elements are idealized components.) Structural descriptions are often presented in a way that makes the causal and functional descriptions explicit. Block diagrams not only show which parts are grouped together, but also represent various subsystem functions that help to show the overall function of the system.

As an example of how structure can be organized hierarchically, consider the diagram below of the structure of an automobile. This simple chart illustrates that the system is composed of various subsystems (cooling, electrical, etc.). The electrical subsystem is decomposed into its components (battery, voltage regulator, etc.)

Under battery we have a further decomposition of structure - or do we? The chart has *changed* at this lowest level of decomposition into something out of a repair manual. The language of repair is *functional* language and is concerned about whether the subsystem component (battery, in this case) is behaving within specified constraints. If not, then misbehavior is a clue to structural failure. Therefore, this last level in the chart does not belong there! It is not a

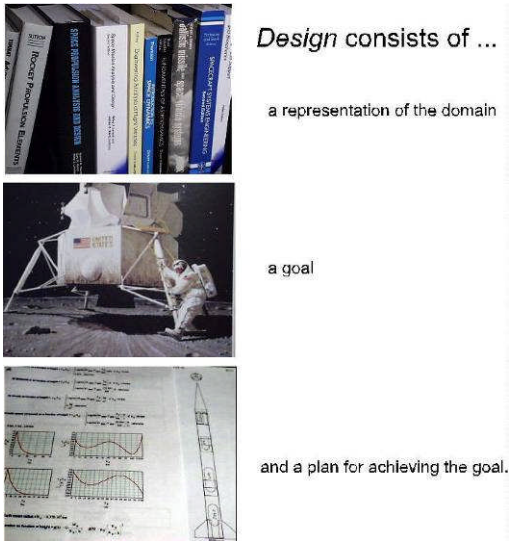
further decomposition of the structure of the automobile but pertains instead to its function, though function is related to structure.



Design begins with a functional description or *specification* of the goals that the designed system is intended to accomplish. It describes function in terms of how the device or system should behave. The designer converts the functional specification into a structural description. Analysis is about how to determine from a given structure its behavior. Design instead goes from function to structure. Function is described in the language of behavior, and behavior relates to structure through the physical laws.

The three elements of *design* can be illustrated as shown below using astronautics for a change; design concepts are not unique to electronics. The functional aspect is expressed as specifications, plans, and goals, or as a problem to be solved. Goals are expressed as constraints on behavior - as *specifications* - and are often achievable by more than one different plan. Choosing the best alternative plan or solution to the problem is a design skill.

The ability to produce plans (which we usually call *designs*) and choose the optimal one is a mark of a good engineer. The plans, designs, or specifications are all expressed in the language and concepts of the domain, such as rocket propulsion and flight principles for astronautics. A manned mission to the moon is the illustrated goal. For electronics, the domain is rather familiar to us. We express specifications in terms of the electrical or thermal quantities which are the measurable behaviors of circuits. The representation of electronic behavior and its causal relationships is the science underlying electronics engineering.



Engineering has *design* as its distinctive subject-matter, though it involves much analysis based on scientific principles, mainly from physics. The reason for this is that analysis makes explicit the constraints on designs. Though the basic circuit laws do not change, alternative circuit analyses of the same circuit are usually possible. Some analyses are not guided by functional interests and blindly spew out large quantities of mathematics which, although ostensibly true, do not provide much simplifying insight into the circuitry. This blind approach also makes it difficult to envision higher-level or more general concepts that apply to the circuit. Regrouping of algebraic expressions can sometimes lead to a simple, insight-laden formulation of circuit behavior that can aid functional thinking about it. Conceptually simplified circuit principles or theorems and meaningful formulas have been a goal of this book.

As an example, consider the simple circuit of a BJT with a shunt external base-emitter resistor, R_{BE}, and an external emitter resistor, R_E, to ground. After some work in deriving the expression for transconductance (collector current, i_c, as output over base voltage, v_b, as input), some rather messy algebra can be ordered into the following form of expression:

$$\frac{i_c}{v_b} = \frac{\beta \cdot i_b}{v_b} = \alpha \cdot \left(\frac{R_{BE}}{R_{BE} + r_e} \right) \cdot \left(\frac{1}{R_E + R_{BE} \parallel r_e} \right)$$

$$\Uparrow \qquad \Uparrow$$

fraction of current in R_E that is i_e — v_b times this is the current in R_E

As the annotated equation shows, with the groupings of factors as given, each factor has a physical interpretation that allows us to remember this otherwise unwieldy formula by remembering the higher-level description of it instead. Even the resistances involved are expressed in a circuit-recognizable (structurally explicit) form. We can see that the first factor in parentheses on the right-hand side is that of a current divider and the second is a transconductance (1/transresistance). From this formulation of the algebra, a simple, memorable description results that lets us reconstruct the otherwise obscure mathematics simply by remembering the functional description. This is the "intuitive circuit theory" that is the quest of every good electronics engineer. It is what R. D. Middlebrook called "low-entropy equations".

A clear understanding of the organization of electronics results from distinguishing between its functional, behavioral, and structural aspects, though it takes a while to think through the difference between the two theories relating structure to behavior (physics) and function to behavior (design). Functional theory is not merely the organization of behavior. It is a separate theory from physics with its own principles or rules about optimizing and devising plans and seeking solutions, and has its own organization of plan details. It is what is distinctively *engineering.* Plans are about how to achieve the desired behavior, and functional theory is expressed in the language of behavior. Similarly, behavior cannot be determined apart from the structure that exhibits it. This is the physics - the science - that sets the constraints on how plans can be devised and specifications satisfied. It is Ohm's Law, KCL, KVL, and other principles of electricity and magnetism. The transresistance method of finding amplifier gain is not required by physics but is instead a useful functional principle for determining a behavioral parameter (gain) that is related to performance, a function of the design.

Both *circuit theory* (structure ⇔ behavior or equivalently, circuits ⇔ waveforms) and *design theory* (function ⇔ behavior or specs ⇔ waveforms) can often be complicated enough to need organization, and this is often hierarchical, where lower levels in the hierarchy present ever more detailed descriptions of the overall theory. The familiar representations of electronics, as circuit diagrams organized into block diagrams, illustrate the three aspects of electronics. The components each have symbols and the circuit diagram shows how they are interconnected. This is structure. On some circuit diagrams, the oscilloscope waveforms are included for selected nodes. This is behavior. And the circuits are grouped into block diagrams for more explicit identification of function - of what they are supposed to accomplish. These three conceptual levels are often not adequately distinguished, nor are physics and design as theories. By keeping them separate in our thinking, design problems and projects are often clarified and simplified.

The Engineering Challenge

One of the general goals of engineering is to make the overall performance of a system better than its components. This at first might seem like an unachievable goal but it is possible by applying various design concepts that have already been presented in one form or another. Engineering presents us with a set of components and some techniques for combining them into something that meets a functional goal. The components and their interconnection are the *structure*. In electronics, these are circuits. The structure has a certain set of behaviors under various conditions, constrained by basic physical laws and allowed modes of operation. The desired behaviors are its *function*. Devices or systems are defined by their *specification* which is a list of functions the device is to perform. Generally, every engineering project begins with a desire to achieve certain physical functions. *Engineering* is the activity of designing a structure that achieves the specified functions.

What is of interest among engineers are the limits of performance of functions given the building blocks, tools, and techniques of the trade. Some typical performance parameters in engineering are reliability, stability, accuracy, precision, and the range of performance-specifying parameters. More specifically, in each area are design challenges. For electronic test and measurement

instruments, speed and precision are driving parameters and they often conflict. Amplifier gain and bandwidth are one such *tradeoff*. It is usually the case that to achieve higher bandwidth, precision must be sacrificed, and the best that can be achieved is a *compromise optimum*.

The *state of the art* or "leading edge" of progress in engineering is the best that can be achieved given the set of components and tools. *Engineering research* attempts to extend functions and their performance. New or better components or tools are sometimes referred to as "enabling technology" while new *techniques* are new circuit ideas, refined into concepts about how to use the given components. Component advances are usually slower than advances in techniques of what to do with components. The discovery or invention of techniques is one of the more gratifying aspects of engineering work because it enables more to be achieved without an improvement in the components. The discovery of translinear circuits by Barrie Gilbert is an example. Another is the invention of the bandgap voltage reference by Robert Widlar. Neither required new components, only a new idea about what to do with them.

The central component in much of electronics is the active device - the transistor. It performs amplification and switching functions. The development of active devices summarizes the history of electronics. First, there were *thermionic valves* (as the British more descriptively call them), or *electron tubes* (as RCA called them) or "vacuum tubes" (as they are less descriptively called in America), followed by transistors. The integrated circuit (IC) is merely an extension of discrete transistor circuits, albeit with some additional possibilities such as matched transistors. Although FETs were conceived before the BJT, historical development was reversed from this order. The unpredictability of engineering development adds some mystery and excitement to being an engineer. We are now, it appears, on the threshold of new kinds of transistors or active devices, based on ballistic or quantum tunneling effects or organic chemistry. Perhaps in the future, we will be including proteins in our list of components or applying genetic or quantum-physics principles.

The behavior of circuits is determined most easily by computer simulation. Simulators *analyze* circuits but do not determine *what* to analyze. They do not *design.* Simulator results apply to particular circuits. The kind of analysis that is useful for design is more general, such as the transresistance method for finding BJT gain. Parameters

in equations can take on a range of possible design values. From the previous chapters of this book, we now have methods for writing down such equations from circuit diagrams. Algebra abstracts from specific numbers to variables and parameters which allow us to consider what happens over a range of values and thus a range of circuits from which we can select the optimum one.

Physical or Behavioral Device Models

Newly-discovered physical principles are often the driving cause for the invention of new kinds of devices such as the transistor. It is one of the highlights in the history of electronics, and illustrates how device models develop in engineering and why they are often a key to technological advancement.

The transistor, like any new breakthrough in technology, was not well understood in the 1950s. The very first transistors were bipolar, not field-effect, devices. They were manufactured using a simple, highly obsolete process that gave them their name of “point-contact” transistors. As the benefits of transistors were quickly realized, great effort was put into their development, and the bipolar junction transistor (BJT) soon replaced it. The quirky point-contact transistor could show some of the negative-resistance effects of tunnel diodes under the right conditions, adding to its mysteriousness and obscuring the phenomena essential to transistor behavior. From an engineering viewpoint, what was needed was to understand the essential principles underlying transistors. In this regard, the development of electronics is no different than any other area of engineering and leads to device modeling. Without a clear understanding of the building blocks, one is left in doubt as to how best to stack them. A device *model* is the basis for this kind of engineering understanding.

The full range of possibilities for a new kind of device is also unknown early in its development. Nobody could foresee in the early 1900s that the electron tube would ultimately pose reliability limits for its use, performing functions - namely, computing - that escaped even the speculations of sci-fi writers of the early electron-valve days. Foresight into the fuller exploitation of technology is limited by a murky conception of its devices. Knowledge of the building blocks we have to work with is a key to the refinement of technology, and in engineering this takes the form of a *model*: a representation of a physical device. Sometimes, models themselves are physical

representations, but of a different scale, as architects often build. Essentially all electronics device modeling, however, is mathematical, graphical, or computational. The goal is to capture the full range of behaviors of a device in the model.

The development of device models parallels the growing discoveries of how to apply the devices. As models improve, they provide deeper insight into the possible uses for devices. Better models also can lead to a better understanding of the context in which devices are used. As transistor models improved, the understanding of the circuits they were used in became simplified, as the simpler models led to simpler circuit concepts. This is elegantly illustrated in the history of BJT model development.

There are two basically different kinds of models. A model that shows us what a device is in itself is an ontological or *physical model*. This kind of model is context-independent. It is valid no matter how the device is used or configured or the circuit in which it is embedded. In the early history of a new kind of device, not much is known about it physically. It appears as a "black box" wherein the contents are largely unknown. Semiconductor physics was not very advanced when the BJT made its appearance. Instead of immediately developing a good physical model, engineers were forced to instead revert to a *behavioral model* for the BJT.

A behavioral model is based on what a device *does*, not what it *is*. Suppose you have a black box with three terminals. This was the BJT to engineers in the early '50s. From earlier work, the concept of two-port networks was available. With three leads, a transistor, like an electron-valve triode, could be modeled as a network with two ports, where one of the BJT terminals is shared in common by the ports. The first chapter began with the port concept as basic to BJT circuits.

Modeling then proceeds by characterizing BJTs by their two-port parameters. The parameters are coefficients within network equations based on a given BJT configuration. Electronics books in the 1950s and 1960s often presented transistor theory using *hybrid-* or *h-parameters*. Instead of the physical parameter, β, there was the equivalent parameter, h_{fe} of the common-emitter (CE) configuration. The difference is that whereas β appears in physical models as a consequence of device structure, h_{fe} appears as a characterization of port behavior under certain conditions depending on BJT configuration. Whereas β is derived from insight into the inner working of the BJT, h_{fe} merely captures BJT terminal behavior for a

given external circuit configuration. The *h*-parameters are configuration-dependent, and hence circuit-dependent, and could be applied equally to electron tubes or transistors. This is not a desirable feature, for it fails to account for the differences in physical structure of the devices. To be optimal, models should be specific to the device and not based on a general network characterization technique. Device models should also be *modular* - self-contained and not dependent upon external conditions.

Of the multiple sets of possible two-port parameters (*h*, *y*, *z*, etc.), the *h* parameters were chosen for BJT modeling because the four *h* parameters best characterized BJT behavior. The four port quantities are each a function of the other quantities, leading to two equations with two terms each. For example, the impedance- or *z*-parameter two-port incremental model equations are

$$v_o = z_m \cdot i_i + z_o \cdot i_o$$
$$v_i = z_i \cdot i_i + z_r \cdot i_o$$

The four parameters (all impedances) result from choosing v_o and v_i to be functions of i_o and i_i. Then each *z* parameter can be found from port-quantity measurements. For instance, transimpedance is

$$z_m = \frac{v_o}{i_i}, i_o = 0$$

The condition that $i_o = 0$, or an open output port, is that i_O be held constant, or $\Delta i_O = 0$. This externally-imposed condition allows z_m to be derived from the first equation. The other parameters are derived similarly. In general, for dependent variable $y(x_1, x_2)$, either x_1 or x_2 is nulled in order to find y with respect to the other independent variable.

The *h* parameters are defined for $i_o(i_i, v_o)$ and $v_i(i_i, v_o)$ by two equations:

$$i_o = h_f \cdot i_i + h_o \cdot v_o$$
$$v_i = h_i \cdot i_i + h_r \cdot v_o$$

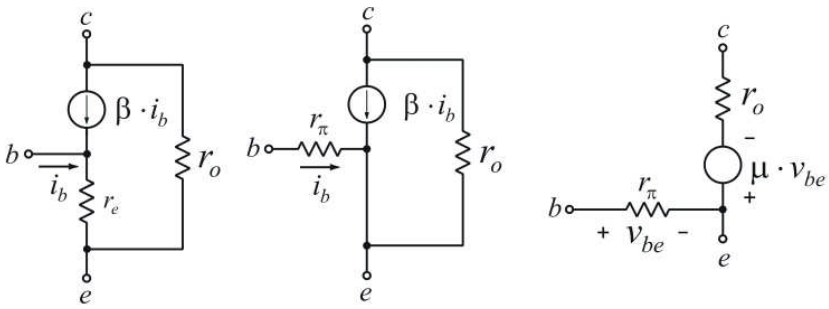
In the CE circuit configuration, BJT quantities $i_o = i_c$ and $i_i = i_b$. Then

$$h_{fe} = \frac{i_c}{i_b}, v_o = 0$$

which in the T or hybrid-π physical models is β. Furthermore for the CE, $h_{oe} = 1/r_o$, $h_{ie} = r_\pi$, and $h_{re} \cong 1/\mu$, where

$$\mu = -\frac{v_{ce}}{v_{be}}, i_c = 0 \Rightarrow \frac{1}{h_{re}} = \frac{v_{ce}}{v_{be}}, i_b = 0$$

The negative sign accounts for base-to-collector voltage inversion, to make μ a positive number. μ and $1/h_{re}$ are not exactly equal (nor are their signs) because the condition, taken from the h_{re} parameter, is not the same as for calculation of μ. For μ, $i_c = 0$ (open-circuit output port), to prevent voltage drop across r_o and allow the dependent collector voltage-source voltage to be applied across the collector (output) terminal. See the equivalent physical models below.



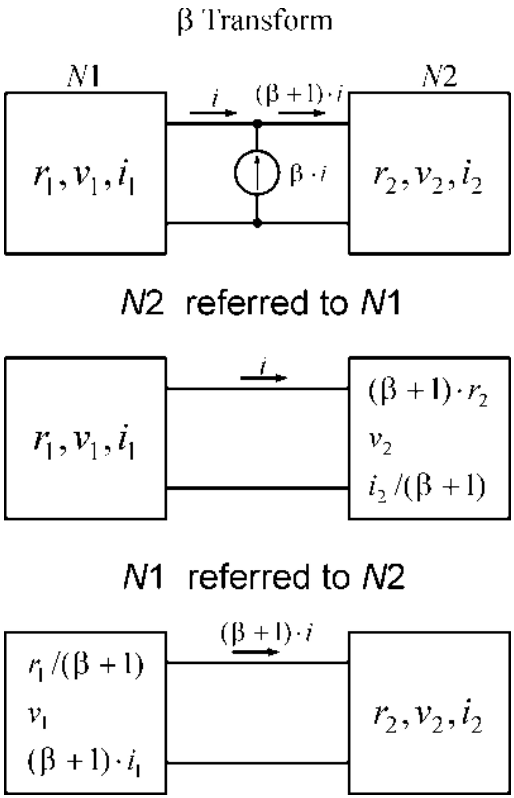
The *h* model and physical model parameters do not correspond exactly and h_{re} is of limited usefulness. Not only are the *h*-parameter models dependent upon BJT configuration ($h_{fb} = \alpha$, not β), they only approximate, if that, the actual physical parameters. Because of this, they tend to obscure rather than reveal basic BJT-circuit insights.

A simple physical model for a BJT is the T model with r_o, as shown. The dependent current source of the T model allows for the application of the β transform. The β transform is half of the *reduction theorem*, a circuit theorem for controlled current sources whereby resistance on the emitter side of the circuit appears to be $\beta + 1$ times larger at the base port;

$$r_B = (\beta + 1) \cdot R_E$$

and from the emitter port,

$$r_E = R_B/(\beta + 1)$$



More generally, the β transform can be diagrammed as shown, where *N1* is the base circuit, and *N2* is the emitter circuit. The dependent current source can be removed and the resistances of one network referred to the other if the resistances are scaled appropriately by $\beta + 1$.

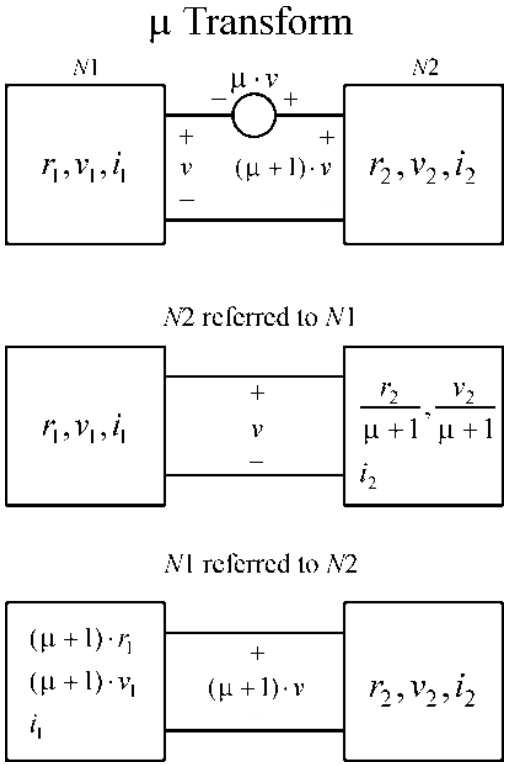
By applying the β transform, we can express the relationship between physical model parameters r_π and r_e as

$$r_\pi = (\beta + 1) \cdot r_e$$

In other words, r_π is simply r_e referred to the base, as shown in the hybrid-π model above. (Note the word "hybrid" in hybrid-π, reflecting historic development of the physical model from the two-port model.)

The Thevenin equivalent hybrid-π model makes use of μ, the dual parameter to that of β. The other half of the reduction theorem is the voltage dual, the *μ transform*, depicted below. This applies to the transistor model (and also to the triode model) in that *N*1 is the base (or gate) circuit and *N*2 the collector (or drain) circuit.

The dependent voltage source can be removed by referring collector resistances to the emitter (middle drawing) or emitter resistances to the collector (bottom). In other words, emitter resistance, R_E, referred to the collector appears as collector resistance,



$$r_C = (\mu + 1) \cdot R_E$$

with $i_b = 0$, or the base open. Also r_o can be referred by the μ transform to the emitter as incremental emitter resistance, r_e. (For electron tubes, the corresponding equation is

$$r_p = (\mu + 1) \cdot r_k$$

where r_p is plate resistance and r_k is the incremental cathode resistance). FET transresistance,

$$r_m = 1/g_m = r_s/\lambda$$

and $r_o = \mu \cdot r_m$, where

$$\lambda = \frac{\mu}{\mu + 1} = \frac{v_s}{v_g}$$

and is the voltage dual of $\alpha = \beta/(\beta + 1)$ of the β transform. (Consequently, for electron tubes, $r_p = \mu \cdot r_m$.) λ is the transmittance from the gate voltage to the source voltage. At the source, the μ transform applies. The FET incremental physical model can be derived simply from the Thevenin-equivalent hybrid-π BJT model by letting $r_\pi \to \infty$. (A fuller explanation can be found in *Designing Amplifier Circuits*.)

What is elegant about the use of parameter r_o over $1/h_{oe}$ is that it results in a simple, physical circuit model for the BJT (or FET) to which the μ transform, a general circuit principle, can be applied. (Electron-tube models by the 1950s were mature and had physical models to which the μ transform was commonly applied.) Instead of trying to think in terms of BJT behavior based on circuit conditions imposed on two-port equations, physical models give us insight into their necessary behavior, represented as circuits. Circuit models are modular and can be placed into any larger circuit, in any

configuration, and be valid. Computer circuit simulator programs make use of physical models at least for this reason.

Another advantage of physical models over behavioral models is that as additional device behaviors, or secondary effects, are accounted for in a more refined model, the circuit model can incorporate them more simply and optimally. For a two-port equivalent model, the number of equations and independent variables would have to grow, and could become unwieldy without the simplifications physical models provide. Neither do two-port models inculcate a causal understanding of circuits because they merely capture behavior. In contrast, physical models, expressed as circuits, capture structure (as interconnected circuit elements), which allows us to infer causes for behavior from our understanding of the circuit elements themselves.

For Further Reading

This book is intended in part to bridge the gap between aspiring engineers who have access to hobbyist or technician literature and that of practicing design engineers. Hopefully, the concepts in it will enable the serious aspiring reader to access much that is in the engineering literature. Recommended as complementary is the *Analog Circuit Design* set of engineering books by SciTech Publishing (www.scitechpub.com), also written by D. Feucht:

Designing Amplifier Circuits
Designing Dynamic Circuit Response
Designing High-Performance Amplifiers
Designing Waveform-Processing Circuits

Basic electronics-engineering circuits courses cover passive circuits and then active circuits. Various textbooks of each kind exist. The list of passive circuits textbooks begins with the author's university textbook when a student, an enduring classic:

Introduction to Circuits, Instruments, and Electronics, James W. Nilsson, Harcourt, Brace & World, Inc., 1968.

A more recent textbook with lighter treatment for technicians should serve for aspiring engineers as background for this book:

Electric Circuit Analysis, C. A. Schuler, R. J. Fowler, Glencoe/Macmillan/McGraw-Hill, 1993.

In the gap between technician and engineering literature is another enduring classic:

Electronic Circuit Analysis: Volume 1 *Passive Networks*, Phillip Cutler, Hardy Hoover, McGraw-Hill, 1960.

The basic passive-circuits math is presented and the book also has some active circuits (including electron tubes for those curious about these once-prominent devices, now seemingly about to return in much smaller form without the power-intensive heaters).

Going back in time, a legacy passive-circuits textbook is

Introductory Circuit Theory, Ernst A. Guilleman, Wiley, 1953.

This book demonstrates the more rigorous treatment of passive circuits as taught at MIT in the early 1950s. It covers some circuits principles not found in more recent passive circuits books such as the reciprocity theorem.

A more recent book that has a mathematical emphasis and covers sampling circuits (sampled-data systems) and transmission lines (distributed-parameter systems) is

Analysis of Linear Systems, David Cheng, Addison-Wesley, 1959.

For active circuits, the best recommendation overall is

Electronics: BJTs, FETs, and Microcircuits, E. James Angelo Jr., McGraw-Hill Book Co., 1969.

Angelo was at the Polytechnic Institute of Brooklyn (New York), or "Brooklyn Tech". This textbook has been used at CalTech and is good at simplifying circuit analysis with higher-level concepts in the style of R. D. Middlebrook of CalTech. Unhappily, it is no exception among most engineering textbooks in not having used the transresistance method for expressing gain, thereby eliminating a confusing mixture of conductances with resistances. Angelo manages (for the most part) to use a hybrid-π or T model for the BJT or FET, and not the less intuitive, configuration-dependent, and obsolete *h*-parameter model. This book comes close to hf BJT theory by hinting at its existence in hf analysis of the CC, but, like all the other circuits textbooks on this list, never develops it.

Phillip Cutler has another book with an industry flair to it:

Semiconductor Circuit Analysis, Phillip Cutler, McGraw-Hill, 1964.

This book is also meant to bridge the gap between technicians and engineers and can be of interest to both. It is an introductory book for semiconductor circuits theory. Cutler was ahead of his time in preferring the T model to the earlier *h*-parameter BJT model, though he does cover two-port models in some detail. He also provides more detailed circuit derivations (as in this book) to guide the reader more carefully through the math. His book was seen around Tektronix in the 1970s and its influence on the Tektronix view of circuits is evident. Implicit in the book is the use of the β transform and the transresistance form of gain expressions. Transformer-coupled and narrow-band, tuned stages are included, and as with most circuits books, also includes capacitively-coupled stages and large emitter bypass capacitances for amplifiers with a zero at the origin (having zero gain at zero Hz).

The convention in many circuits textbooks is to use A for the forward path gain (instead of the conventional G of control theory) and β for H, which becomes confused with BJT β. Cutler runs into this problem and resorts to the *h*-parameter designation of BJT β as h_{FE}.

A book written and used at MIT has good coverage of semiconductor device electronics and was the first to present some of the dynamics formulas in this book. Do not confuse the author Paul E. Gray with Paul R. Gray, who along with Meyer wrote a more recent and well-known circuits textbook on the other coast.

Electronic Principles: Physics, Models, and Circuits, Paul E. Gray, Campbell L. Searle, Wiley, 1967.

This book is carefully written and rigorous but is also an example of "high-entropy" equations - equations put in a form that maximize the difficulty of relating the equation to the circuit. As in most other circuits textbooks, approximating is profuse as an alternative to the arduous derivation of exact (or nearly exact) and more general template formulas.

Gray and Searle go somewhat farther than does Angelo in taking on dynamic circuit analysis before they revert to computer solutions. Both books - and this is typical of circuits textbooks - go as far as deriving a base-node dominant-pole equivalence, with time constant $R_b \cdot C_b$, where C_b includes C_e ($R_E = 0\ \Omega$ in textbook CE amplifiers) in parallel with the Miller-multiplied C_c. To their credit, Gray and Searle

also add τ_{cc} as a third term in C_b, as $(R_L/R_b)\cdot C_c$. They also go farther than most books in addressing the cascade CE interactions.

Also going farther is

Feedback Amplifier Principles, Sol Rosenstark, Macmillan, 1986.

Rosenstark approaches dynamics by using gain derivations to find zeros and the Cochrun-Grabel method for poles, streamlined in tabular form called "Rosenstark tables" in this book. He takes on two- and even three-stage cascade CE amplifiers but truncates the tables for a quadratic approximation or else turns on the computer. Feedback dynamics is emphasized.

Another circuits book from Brooklyn Tech (as it was called) is

Transistor Circuit Analysis, Maurice Joyce, Kenneth Clarke, Addison-Wesley, 1961.

Written in the early 1960s, the book shows the influence of the hybrid-parameter BJT model, though it mixes in the T and hybrid-π models. Static design and thermal feedback are well-developed. It covers somewhat the same ground as Cutler's semiconductor book but with engineering depth. Chapters on power amplifiers and regulated supplies (another form of power amplifiers) lead onward to circuit dynamics and frequency effects on the BJT noise model, a topic not covered much in this book. Bandwidth extension methods (shunt inductive and emitter peaking) appear but little in the way of bandwidth calculation.

A book with advanced mathematics (such as advanced calculus and contour integration of enclosed poles using residues) that covers in detail the design of fast amplifiers, with some emphasis on staggering stages, is written (in excellent English) by Slovenian Peter Starič and Croatian Erik Margan. Starič has worked in the past on wideband (vertical oscilloscope) amplifiers at Tektronix, and Margan has designed particle-detector electronics for the ATLAS experiment of the European Large Hadron Collider of CERN. For those who do not dream much in contour integrals, the book is very much readable after the first chapter or two where the math is explained well. Their book is

Wideband Amplifiers, Peter Starič, Erik Margan, Springer, 2006. (www.springer.com)

One control theory textbook recommendation is the author's favorite for its clarity, rigor, and comprehensiveness:

Introduction to Continuous and Digital Control Systems, Roberto Saucedo, Earl E. Schiring, Macmillan, 1968.

Last but not least for books is another enduring MIT product,

Operational Amplifiers: Theory and Practice, James K. Roberge, Wiley, 1975.

Roberge's book fills in more of the op-amp design picture, including op-amp design at the circuit level. Like this book, the second half is example circuits that illustrate, apply, and extend basic concepts. Op-amps as feedback amplifiers find feedback dynamics (control theory applied to circuits) emphasized, though the book is light on BJT dynamics. Roberge presents circuit diagrams of multiple commercial op-amps with some circuit analysis, mostly static and quasistatic. The versatile op-amp can be used for many more functions than amplification and Roberge presents a number of different waveform-processing circuits.

These books, though dated as listed, are well-written, still available as new and used books, and present clearly many of the concepts underlying transistor amplifiers. Many more recent books are also available including some that can be downloaded free from the Internet. The author's website at www.innovatia.com has some downloadable tutorial webpages.

Finally, a web search on "R. D. Middlebrook" will turn up various sites from which some of his lectures and papers can be downloaded. Middlebrook's forte was the reduction of electronics theory to simplifying clarity as some of his theorems in this book exemplify.

Closure

The transistor amplifiers presented in this book consist of less than a dozen transistors and are structurally simple relative to what is found in many analog integrated circuits or in measurement instruments such as vertical amplifiers of oscilloscopes. Yet even complicated circuits operate from the basic principles that have been presented and applied extensively in the preceding chapters, and their designs decompose into circuits that can be similarly analyzed.

Emphasis on fundamental design and analysis principles has prepared the reader for the challenge of the more complicated cases. The algebra and calculator button pushing that so characterizes design in this book is often replaced by computer circuit simulation, yet an understanding of how these concepts apply to actual circuits is in the back of the minds of (the better) engineers involved in simulation-oriented design.

As the decades of the electronics era have gone by, electronic complexity has increased immensely, and seems to have overwhelmed engineering school curricula and industry practice alike, to the effect that a solid ability in basic circuits skills has suffered. Hopefully, this book will contribute to the renewal of emphasis on the basic principles - both historically well-established and also those that have been neglected - of transistor circuit design.