

Transistors As Amplifiers

This discussion will concentrate on bipolar junction transistors (BJTs) in amplifier applications because BJTs are by far the most commonly used amplifying device. Remember though that all amplifying devices operate in a similar way so the same principles that govern the way BJTs amplify govern the use of JFETs, MOSFETs and valves as amplifiers.

A Word About Amplifiers

The purpose of an amplifier is to increase the amplitude of a signal. If one thinks purely in terms either of voltage or of current then it is possible to change the amplitude of a signal by using a transformer. However a transformer offers no possibility of power gain – if a weak signal enters the primary of a transformer it will be at best equally weak when it emerges from the secondary. Imagine voltage is increased by a ratio of five to one. Current will be reduced by a similar ratio and the power of the signal entering the primary will be equal to the sum of the power of the signal leaving the secondary and any power lost in the transformer. The crucial factor about an amplifier is its ability to offer *power gain*. At low frequencies, one is usually more interested in the factor by which the signal (voltage or current) amplitude has been magnified than in the signal power gain which tends to be a more important parameter at higher frequencies (> 50 MHz). Several measures of gain are available:

Voltage Gain is the ratio of the output voltage amplitude and input voltage amplitude. It is used when the parameter of interest is the signal voltage amplitude. It is used at low frequencies (100 MHz or less). An ideal voltage amplifier has infinite input resistance (i.e. it draws zero current from the signal source driving it) and has zero output resistance (i.e. it can supply unlimited current to its load).

Current Gain is the ratio of the output current amplitude and input current amplitude. It is used when the parameter of interest is the signal current amplitude. It is also used at low frequencies. An ideal current amplifier has zero input resistance (i.e. there is no signal voltage at the input) and infinite output resistance (i.e. it can supply unlimited voltage to its load).

Power Gain is the ratio of the output signal power to the input signal power. Power gain is used at high frequencies in “impedance matched” systems where the effects of electromagnetic propagation in the circuit cannot be ignored. In an impedance matched system all output impedances are equal to all impedances at a value known as the “characteristic impedance”. $50\ \Omega$ is a common characteristic impedance in communications and radar applications, television systems use $75\ \Omega$.

Note that in an impedance matches system knowledge of any one of these three gains automatically defines the other two.

There are two other kinds of gain that are of interest in special applications; transconductance and transresistance. Transconductance is the ratio of the output signal current to the input signal voltage and is measured in Amps per Volt (or Siemens but occasionally written as Mhos as well). Transconductance is an important concept for all amplifying devices. Transresistance is the ratio of output voltage to input current and is measured in Volts per Amp or Ohms.

The Mechanism of Amplification

All amplifying devices can be regarded as circuit elements that have their output current controlled by an input voltage. The characteristic that describes this behaviour is known as the transconductance characteristic (or occasionally mutual characteristic) because it relates output current to input voltage. The transconductance characteristics for various devices are shown in Fig. 1. If a signal is regarded as a small change or “perturbation” around some average value (often zero), there are obvious problems with these characteristics from an amplification point of view. For example, a signal with an average value of zero applied to a BJT would cause no change in I_C for all signal voltages below 0.7 V. In other words the signal voltage below 0.7 V would effectively be lost. This is usually not an acceptable state of affairs and consequently the signal is added to a d.c. voltage, known as a bias voltage, to ensure that I_C can respond to the whole of the signal.

The situation is shown in the diagram of Fig. 2. If ΔV_{BE} , the signal, was applied with no bias, i.e. with its average value equal to zero there would be no change of I_C and so $\Delta I_C = 0$. If, on the other hand, a bias voltage, V_{BEB} , is added to the signal, there is a substantial change in I_C as a result of the signal. The same arguments hold for all the other devices although the best choice of bias voltage will be different for each.

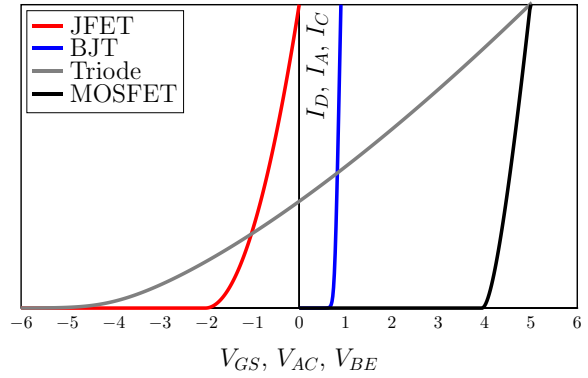


Figure 1: Example transconductance curves for, JFET (red), triode (grey), BJT (blue), MOSFET (black).

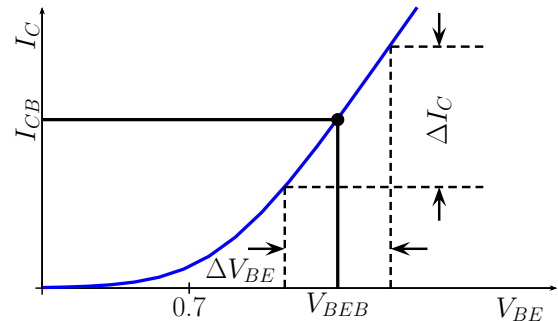


Figure 2: A BJT transconductance curve with quiescent (no signal or d.c.) conditions shown (solid lines) and the extent of signal swing shown (dashed lines)

The relationship between ΔI_C and the signal that caused it, ΔV_{BE} is the “small signal transconductance”, g_m , of the device being used. g_m is the slope of the transconductance characteristic at the bias point (V_{BEB}, I_{CB}) . Since the transconductance characteristic is not a straight line, g_m , varies with V_{BEB} and indeed within ΔV_{BE} if ΔV_{BE} is not sufficiently small. It is usually assumed that ΔV_{BE} is sufficiently small for the transconductance characteristic to be approximated as a straight line over the range of V_{BE} .

In Fig. 3, the changes in collector current, ΔI_C , are converted into an output signal voltage using a resistor, R_L . An input voltage of (1),

$$V_{IN} = V_{BEB} \pm \frac{\Delta V_{BE}}{2} \quad (1)$$

will give a collector current change of (2),

$$I_C = I_{CB} \pm g_m \frac{\Delta V_{BE}}{2} \quad (2)$$

remember,

$$g_m \equiv \frac{\Delta I_C}{\Delta V_{BE}} \quad (3)$$

for a BJT, and this will in turn give rise to a change in collector voltage of,

$$V_O = V_{CC} - I_C R_L \quad (4)$$

$$= V_{CC} - I_{CB} R_L \mp g_m R_L \frac{\Delta V_{BE}}{2} \quad (5)$$

$$\equiv V_{OB} \pm \frac{\Delta V_O}{2} \quad (6)$$

where V_{OB} is the output voltage obtained when the signal is zero,

$$V_{OB} = V_{CC} - I_{CB} R_L \quad (7)$$

and $\frac{\Delta V_O}{2}$ is the component of the output voltage due to the signal perturbation,

$$\frac{\Delta V_O}{2} = -g_m R_L \frac{\Delta V_{BE}}{2} \quad (8)$$

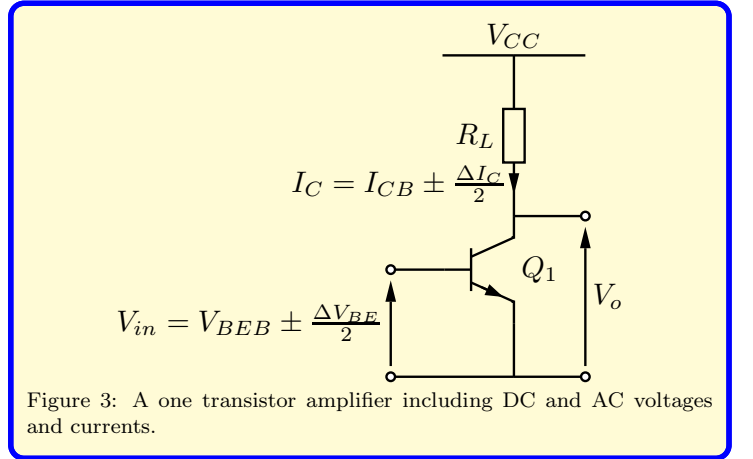
By using the relationship between ΔV_O and ΔV_{BE} it is possible to estimate the voltage gain of the amplifier,

$$\Delta V_O = -g_m R_L \Delta V_{BE} \quad (9)$$

or

$$\frac{\Delta V_O}{\Delta V_{BE}} = -g_m R_L = \text{gain} \quad (10)$$

Note that:



1. The bias conditions V_{BE} , I_C and V_{OB} do not explicitly appear in the expression for gain although it must be remembered that g_m is a function of V_{BE} .
2. The gain is negative. This simply means that an increase in input voltage leads to a decrease in output voltage and vice versa. In signal terms it implies inversion of a 180° phase shift.

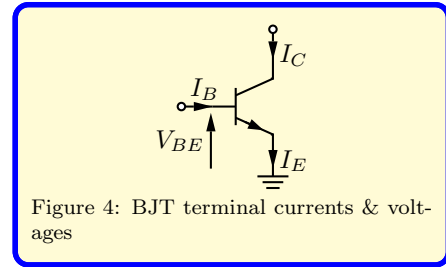
Point 1 above is very important because it suggests that the bias conditions and the signal conditions can be considered separately. Defining a stable set of bias condition is one of the primary objectives of amplifier circuit design.

BJT Biassing

BJTs are the odd ones out in the family of amplifying devices because they need to draw an input current in order to operate. A given collector current I_C will require a base current I_B to support it and the two are related by,

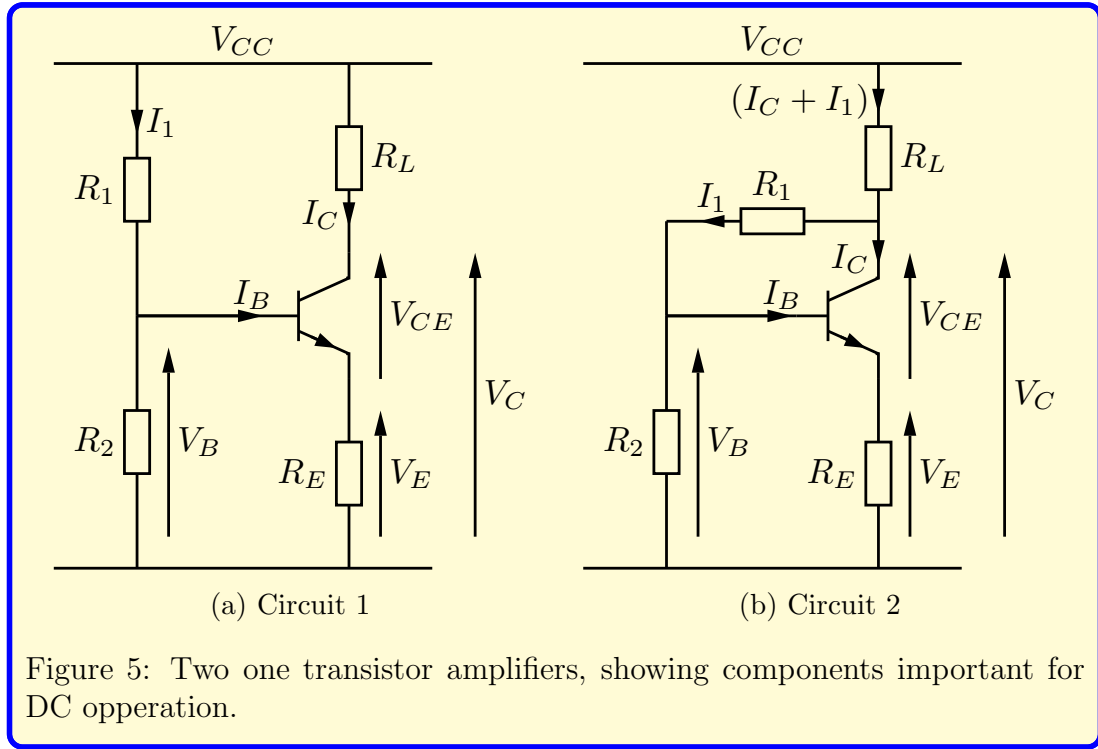
$$\frac{I_C}{I_B} = h_{FE} \quad (11)$$

see Fig. 4. h_{FE} is the large signal static current gain of the BJT. It is approximately independent of I_C but it varies with temperature and there is a large spread of values (typically a factor of five) from device to device of the same type. Control of the bias conditions must not therefore fall to the transistor but should be accomplished by well defined circuit elements such as resistors.



Two types of bias circuit are suitable for single transistor BJT amplifiers (Fig. 5). The objective of both of these bias circuits is to control the collector current, I_C .

In both cases this control is achieved by negative feedback. In circuit 1 the voltage V_B defined by V_{CC} , $R_1 + R_2$, is made up of $V_E + V_{BE}$. If V_E is made large compared to changes expected in V_{BE} (either as a result of temperature changes or device to device variation) the V_E , and hence I_C is substantially constant. In circuit 2 R_E provides negative feedback as in circuit 1 but there is a second source of negative feedback from V_C via R_1 and R_2 . I_C will tend to reduce V_C , hence reducing V_B and counteracting the increase in I_C . Circuit 1 will not operate satisfactorily with $R_E = 0$ because under such a condition, all negative feedback has been removed. Circuit 2 will operate with $R_E = 0$ because there still remains the negative feedback path from V_C via R_1 and R_2 . It is usual in the analysis of both circuit 1 and circuit 2 to assume that I_B is negligible and it is usual in design to make sure that the assumption is valid.



Working Out the Bias Conditions

Circuit 1 – Assume I_B is negligible, $V_{BE} \approx 0.7$ V and $h_{FE} \gg 1$ (i.e. $I_C \approx I_E$).

$$V_B = \frac{V_{CC} R_2}{R_1 + R_2} \text{ by potential division} \quad (12)$$

$$V_B = V_E + 0.7 = V_E + V_{BE} \text{ by Kirchhoff's Voltage Law} \quad (13)$$

$$I_E \approx I_C = \frac{V_E}{R_E} = \frac{V_B - 0.7}{R_E} = \frac{1}{R_E} \left[\frac{V_{CC} R_2}{R_1 + R_2} - 0.7 \right] \quad (14)$$

$$V_C = V_{CC} - I_C R_L \text{ by Kirchhoff's Voltage Law (K.V.L)} \quad (15)$$

Circuit 2 – Assume I_B is negligible, $V_{BE} \approx 0.7$ V and $h_{FE} \gg 1$ (i.e. $I_C \approx I_E$).

$$I_1 R_2 + I_1 R_1 + (I_1 + I_C) R_L = V_{CC} \text{ (K.V.L)} \quad (16)$$

$$\text{or } V_{CC} = I_C R_L + I_1 (R_L + R_1 + R_2) \quad (17)$$

$$I_1 R_2 = V_E + V_{BE} = V_E + 0.7 \text{ (K.V.L)} \quad (18)$$

$$\text{or } I_1 R_2 = I_C R_E + 0.7 \quad (19)$$

either I_1 or I_C may be eliminated from (17) using (19) for example, eliminating I_1 gives,

$$V_{CC} = I_C R_L + \frac{I_C R_E + 0.7}{R_2} (R_L + R_1 + R_2) \quad (20)$$

$$\text{or } I_C = \frac{V_{CC} - \frac{0.7 (R_L + R_1 + R_2)}{R_2}}{R_L + \frac{R_E (R_L + R_1 + R_2)}{R_2}} \quad (21)$$

This result for I_C can be used in (19) to find I_1 . V_C is found using,

$$V_C = V_{CC} - (I_C + I_1) R_L \quad (22)$$

Notes

- It is not the results that are important, but the application of the basic circuit rules that lead to them.
- The only transistor voltage drop that should appear in equations is V_{BE} . V_{CB} and V_{CE} do not and should not appear in equations for amplifiers.
- The assumption “ I_B is negligible” really says that the existence of I_B does not disturb the potential at the transistor base significantly.
- Always check that the solution to equations (17) and (19) in circuit 2 is self consistent.

Design of Bias Circuits

The design process for single transistor amplifiers involves choosing one of the two circuits and deciding on appropriate values of node voltages and transistor collector current and then working out sensible component values. The choice of circuit depends to some extent on the application area. For low frequency applications, either circuit 1 or circuit 2 can be used. For high frequency applications, circuit 2 with $R_E = 0$ tends to be used.

The value of I_B must be considered during the design process to ensure that the design will satisfy the criterion “ I_B is negligible”. The case most likely to violate the criterion is *smallest* h_{FE} . Remember that the manufacturer will specify a maximum and minimum value of h_{FE} for a particular transistor. Remember also that the purpose of the bias circuit is to control I_C . Thus, $I_{B_{max}} = \frac{I_C}{h_{FE_{min}}}$

and $I_{B_{max}}$ is usually taken to be negligible if I_1 , the current at the top of the biasing chain $\geq 10 I_{B_{max}}$.

The value of I_C , V_C , V_E and V_B are a little more complicated to decide on because they will affect the signal properties of the amplifier. A few of the compromises are:

1. The value of collector voltage will affect the output voltage swing available. For example in circuit 1, V_C can lie anywhere between V_{CC} and V_E . To maximise output voltage swing for a symmetrical signal like a sinusoid, V_C should be placed halfway between V_{CC} and V_E . i.e.

$$V_C = \frac{V_{CC} + V_E}{2} \text{ for max symmetrical swing} \quad (23)$$

2. Clearly both V_{CC} and V_E will affect the max symmetrical swing, which is $V_{CC} - V_E$. V_{CC} is usually set by what is available within the rest of the system, V_E can be chosen.
3. Larger V_E gives more precise control of I_C . For a BJT it is unwise to let V_E fall below 1 V in a circuit such as number 1.
4. I_C is chosen by considering the nature of the load, but it also affects the effective input resistance of the transistor¹ and output resistance of the amplifier. In general one would aim for a condition $R_L \ll$ [input resistance of the next stage].
5. R_1 and R_2 should be as large as possible consistent with the maintenance of the appropriate relationship between $I_{B_{max}}$ and I_1 .

To visualise how the supply voltage will be divided up between the various parts of the circuit, it is helpful to draw a chart such as Fig. 6. This makes it clear that increasing V_E reduces the range of voltage that can be occupied by V_C and that the best position for V_C with symmetrical signals is halfway through the available range. Note that in this chart the minimum available value of V_C is V_B whereas in the comments above it is V_E . Most amplifier transistors will work satisfactorily with V_C as low as a few hundred mV above V_E but there are good reasons for saying that ideally V_C should not fall below V_B .

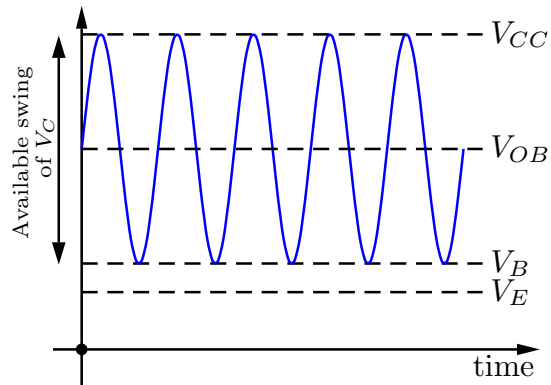


Figure 6: The transistor electrode voltages in the one transistor amplifier circuits.

¹because I_C controls g_m via $g_m = (e I_C)/(k T)$ and $r_{be} = \beta/g_m$ so r_{be} and I_C are linked.

Notes

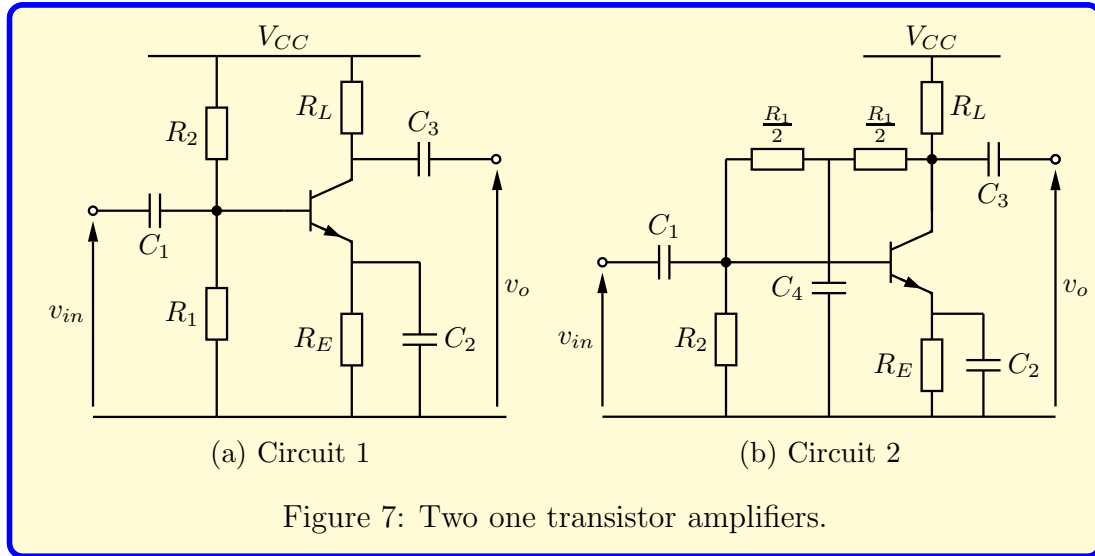
1. The design process is a compromise.
2. No two designers would make identical decisions.
3. Never specify component values more tightly than is necessary.
4. Use “preferred” (E12, E24 etc.) values.

Coupling and Decoupling

Transmitting signals from one place to another in a circuit is called “coupling”. Removing signals from nodes in the circuit is called “decoupling”. Capacitors or transformers can be used for coupling leading to so called “R-C” and “transformer coupled” amplifiers. Amplifiers that are required to amplify d.c. signals, such as strain gauge amplifiers or thermocouple amplifiers, cannot use transformers or capacitors – instead they must be “direct coupled” or “d.c.” coupled. Direct coupled amplifiers use many transistors and will not be considered further at this point. Transformer coupling is attractive at high frequencies or in tuned amplifiers where resonant circuits are used. Capacitor coupling is used at lower frequencies. For example, an audio amplifier will be a combination of d.c. and capacitor coupling; a radio or TV I.F. amplifier will be transformer coupled.

Circuits 1 and 2 are shown in Fig. 7 with coupling and decoupling capacitors included. For the purposes of this discussion, a capacitor may be regarded as an open circuit (infinite impedance) to d.c. and a short circuit (zero impedance) to signals. Note that in Fig. 7 the signal voltage, v_{in} and v_o are in lower case v whereas the bias conditions are in upper case V . In both cases,

- C_1 couples the signal from the signal source to the transistor base without allowing the source to affect the bias conditions or the bias conditions to affect the source.
- C_2 decouples the emitter node of the transistor. In other words C_2 short circuits the emitter node of the transistor to ground as far as signals are concerned. This prevents R_E having the same stabilising effect on the signals as it has on the d.c. conditions. by removing the negative feedback caused by R_E . The circuit voltage gain $\frac{v_o}{v_i}$ is much larger if C_2 is included in the circuit than it would be if R_E was not bypassed by a capacitor.
- C_3 couples the signal from the output (collector node) to the load without allowing disturbance of the bias conditions or the imposition of a d.c. voltage across the load.



In circuit 2,

- C_4 decouples the mid point of R_1 . Since R_1 is also a negative feedback path it will reduce the circuit gain if a.c. as well as d.c. voltages can be transmitted via R_1 to the base. C_4 short circuits the mid point of R_1 to ground as far as a.c. (signal) voltages are concerned hence eliminating any effects of the negative feedback via R_1 on circuit gain.

How the Transistor Interacts with Signals

The transistor is characterised by non-linear characteristic curves (see notes on characteristics). The rest of the amplifier circuit consists of standard circuit elements such as resistors and capacitors so it is convenient to represent the behaviour of the transistor towards the signal in standard circuit terms. A circuit representation of how the transistor behaves towards a signal is called a “small signal model” – it assumes that the signal represents only a small deviation from the bias conditions. All amplifying devices can be represented by a small signal model.

A Small Signal BJT Model

The underlying process of amplification involves the device “transconductance” – i.e. the amplifying device can be considered as a current source whose magnitude is controlled by the input voltage. For small signals it is the slope of the transconductance characteristic at the bias point which is of interest (see Fig. 8). For a BJT,

$$I_C = I_{CO} \left(\exp \left(\frac{e V_{BE}}{k T} \right) - 1 \right) \quad (24)$$

and the slope at the bias point is,

$$\frac{dI_{CB}}{dV_{BEB}} = I_{CO} \frac{e}{k T} \exp \left(\frac{e V_{BE}}{k T} \right) = g_m \quad (25)$$

for a conducting diode,

$$\exp \left(\frac{e V_{BE}}{k T} \right) \gg 1 \quad (26)$$

so

$$I_C = I_{CO} \left(\exp \left(\frac{e V_{BE}}{k T} \right) - 1 \right) \approx I_{CO} \exp \left(\frac{e V_{BE}}{k T} \right) \quad (27)$$

substituting,

$$\therefore \frac{dI_C}{dV_{BE}} = I_{CO} \frac{e}{k T} \exp \left(\frac{e V_{BE}}{k T} \right) = \frac{e I_C}{k T} = g_m \quad (28)$$

$g_m = \frac{e I_C}{k T}$ where I_C is the quiescent or d.c. collector current. It is one of the fundamental BJT relationships and should be remembered. At room temperature, $\frac{e}{k T} \approx 40$. This transconductance consideration leads to the simplest BJT model, shown in Fig. 9. It is a good low frequency model for JFETs, MOSFETs and valves (although these devices and the BJT would probably have a resistor in parallel with the current source to take account of the slope on the output characteristics).

The BJT, however, is unique in having an input resistance that can rarely be ignored. The input resistance is found by working out the slope of the input characteristic, at the operating or quiescent point, in an indirect way,

$$r_{be} = \frac{dV_{BE}}{dI_B} = \frac{dI_C}{dI_B} \cdot \frac{dV_{BE}}{dI_C} \quad (29)$$

$$\frac{dI_C}{dI_B} = \beta = \text{small signal current gain} \quad (30)$$

$$\frac{dV_{BE}}{dI_C} = \frac{1}{g_m} \text{ from (28)} \quad (31)$$

$$\therefore r_{be} = \frac{\beta}{g_m} \quad (32)$$

another vital BJT relationship.

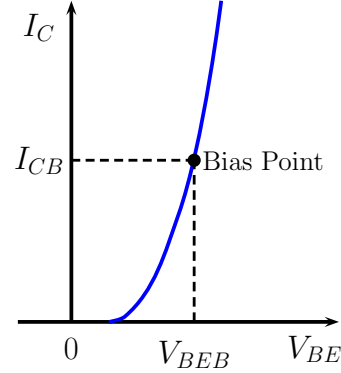


Figure 8: A BJT transfer characteristic showing the bias or quiescent point. The characteristic (blue line) is expressed by (24).

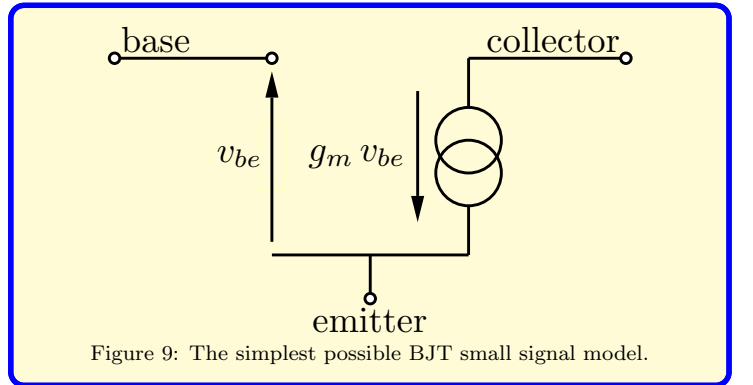
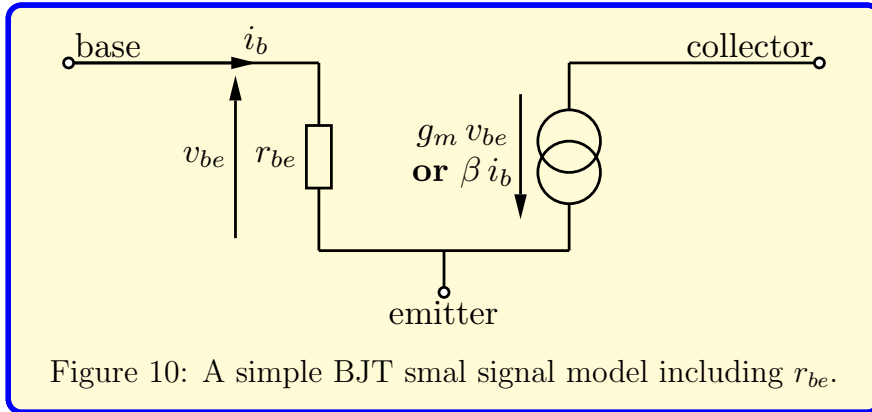


Figure 9: The simplest possible BJT small signal model.



Note that dV_{BE} , dI_B , dI_C are the small changes to the bias conditions and could be represented at small signal quantities v_{be} , i_b and i_c ,

$$r_{be} = \frac{\beta}{g_m} = \frac{dV_{BE}}{dI_B} = \frac{v_{be}}{i_b} \quad (33)$$

$$\text{so } g_m v_{be} = \beta i_b \quad (34)$$

This is an interesting result because it says that the output current generator in the BJT model may be thought of as being controlled by the current through r_{be} or by the voltage across r_{be} . People get very worked up over the question “is a BJT a current or transconductance amplifier?” The answer really is that it doesn’t matter – use whichever is more convenient for any particular problem. The discussion of the BJT in transconductance terms is helpful because the transconductance viewpoint is common to all three terminal amplification devices. No other device can be looked at as a current amplifier. Including r_{be} in the model leads to Fig. 10.

Notes

- Usually $\beta \neq h_{FE}$. β is a small signal parameter and h_{FE} is a large signal parameter.
- β is sometime given as h_{fe} . h_{fe} is derived from a different modelling system and except at high frequencies they can be taken as equal.
- There are other elements one could add to this model to explain details of behaviour. One example is a resistor in parallel with the current generator to model the slope on the output characteristic. The simple model in Fig. 10 consisting of input resistance and output current source is reasonable for a wider range of applications and will be used for the rest of this course.

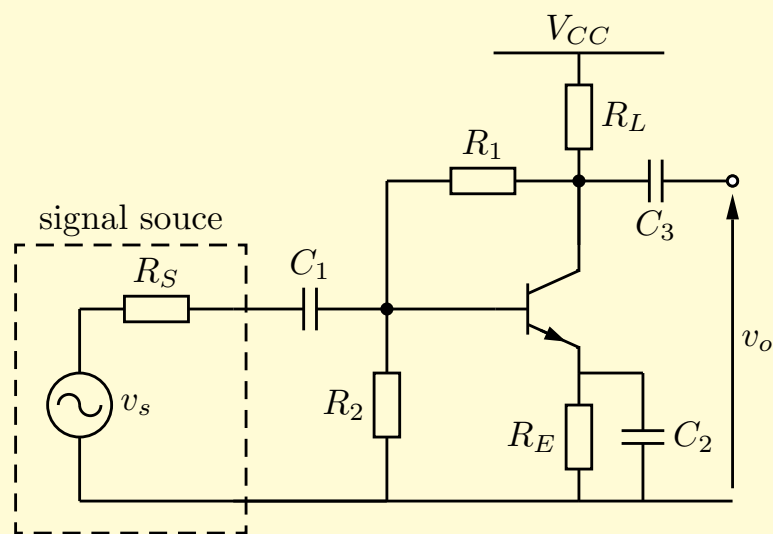


Figure 11: Circuit 1 with AC and DC feedback via R_1 and DC feedback via R_E .

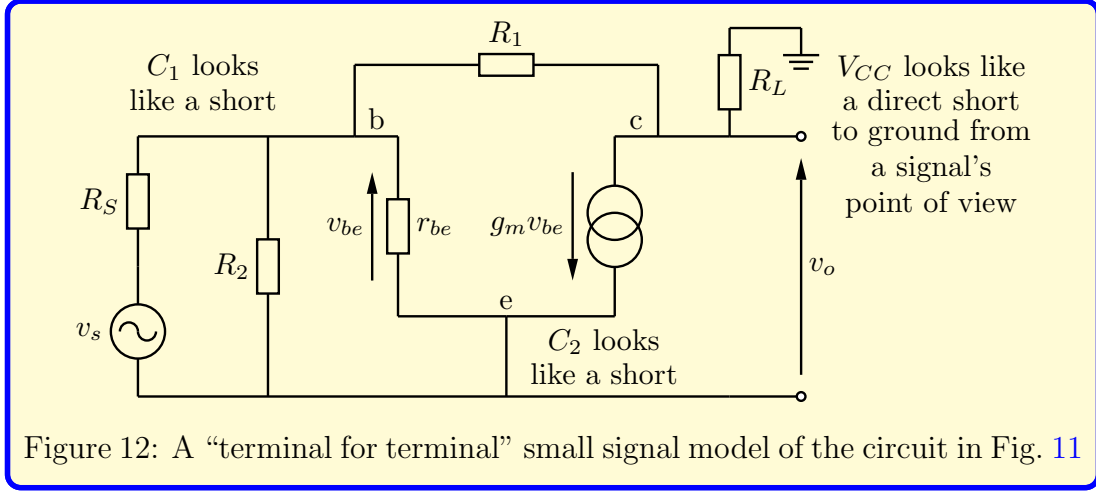
The Small Signal Equivalent Circuit

In principle this is a straightforward task – it is a matter of drawing a circuit which describes what a signal in the circuit would experience, so it is necessary to look at the circuit from the signal's point of view. There are two important consequences of being interested only in the signal's interaction with the circuit.

1. All d.c. voltage sources (such as power supplies) are replaced by their Thévenin equivalent impedance (i.e. 0Ω – a short circuit).
2. All d.c. current sources are replaced by their Thévenin equivalent impedance, i.e. $\infty \Omega$ – an open circuit.

In addition, since for the purposes of this course capacitors are considered as open circuit at d.c. and short circuits to a.c., all capacitors are replaced by short circuits. The transistor is replaced terminal for terminal by its small signal model. Consider circuit 2, without decoupling R_1 , which has the circuit diagram shown in Fig. 11. This circuit has the small signal model shown in Fig. 12. This small signal model can be tidied up to form Fig. 13.

Note that the small signal equivalent circuit will vary according to the circuit it is derived from. Do not attempt to learn the result – attempt instead to acquire the skill of deriving the small signal model for any circuit.



Once the equivalent circuit is obtained, normal circuit analysis methods can be used to evaluate performance. For example, to obtain the overall voltage gain, $\frac{v_o}{v_s}$, one would begin by summing the currents (applying Kirchhoff’s current law) at the output node,

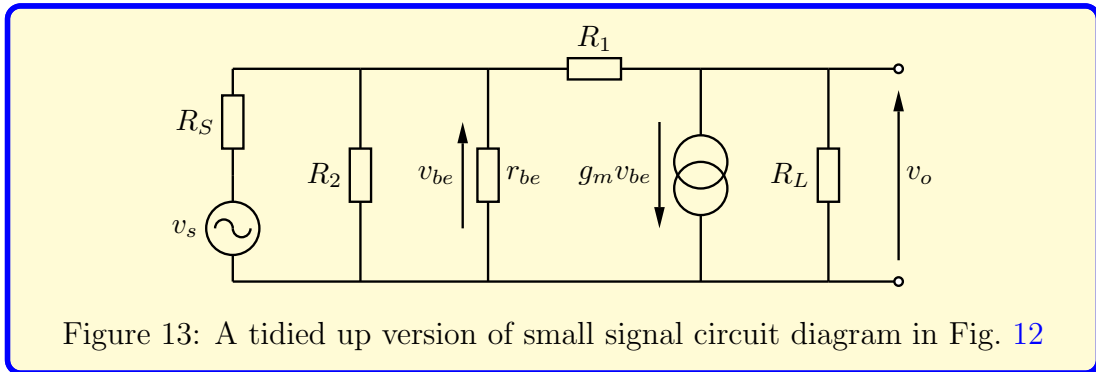
$$\frac{v_o}{R_L} + \frac{(v_o - v_{be})}{R_1} + g_m v_{be} = 0 \quad (35)$$

summing currents (K.C.L) at the input node,

$$\frac{(v_s - v_{be})}{R_S} + \frac{(v_o - v_{be})}{R_1} = \frac{v_{be}}{R_2} + \frac{v_{be}}{r_{be}} \quad (36)$$

Equations (35) and (36) can be transposed to yield respectively,

$$v_{be} = -\frac{v_o (R_1 + R_L)}{g_m R_1 R_L - R_L} \approx -\frac{v_o}{g_m R_1 // R_L} \quad (37)$$



and

$$v_{be} = - \frac{\frac{v_s}{R_S} + \frac{v_o}{R_1}}{\frac{1}{R_2} + \frac{1}{r_{be}} + \frac{1}{R_S} + \frac{1}{R_1}} \quad (38)$$

$$= \frac{v_s (R_2 // r_{be} // R_S // R_1)}{R_S} + \frac{v_o (R_2 // r_{be} // R_S // R_1)}{R_S} \quad (39)$$

Eliminating v_{be} and transposing to obtain the voltage gain, $\frac{v_o}{v_s}$, required gives,

$$\frac{v_o}{v_s} = - \frac{R_1}{R_S} \cdot \frac{1}{1 + \frac{R_1}{g_m R_L (R_2 // r_{be} // R_S)}} \quad (40)$$

and since R_1 is very large the $R_1 / (g_m R_L (R_2 // r_{be} // R_S))$ term will dominate the denominator giving,

$$\frac{v_o}{v_s} = - \frac{R_1}{R_S} \cdot \frac{1}{\frac{R_1}{g_m R_L (R_2 // r_{be} // R_S)}} = -g_m R_L \cdot \frac{R_2 // r_{be}}{R_S + R_2 // r_{be}} \quad (41)$$

This expression consists of a gain term, $g_m R_L$ and an input potential division $(R_2 // r_{be}) / (R_S + R_2 // r_{be})$. Note that the circuit gain is now directly dependent on the transistor parameters g_m and r_{be} ; the negative feedback effects of R_1 have been eliminated. In removing R_1 , the circuit is being changed from a small signal point of view, from circuit 2 in to circuit 1 with the emitter decoupled. The R_1 in circuit 1, which is necessary for correct biasing of the transistor, appears in small signal terms in parallel with R_2 , hence altering the effective value of R_2 but not the form of the result.

Each circuit shape will produce its own result for gain and other performance measures so memorising this result would be unhelpful. The desirable outcome is for the student to practice the skill of deriving small signal circuit diagrams and equations until they can do it for any circuit and then to be able to interpret the results of their analysis.