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 ..... but a transformer offers no possibility of power gain — in other words if a weak signal enters the primary of a transformer it will be at best equally weak when it emerges from the secondary. 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 amplitude has been magnified than in the signal power gain which tends to be a more important parameter in high frequency (50 MHz upwards) applications. There are three gain measures:

<u>Voltage gain</u> — the ratio of output voltage amplitude divided by input voltage amplitude. Used when the parameter of interest is the signal amplitude. Used at low frequencies (100 MHz or less). Ideal voltage amplifier has infinite input resistance (ie it draws zero current from the signal source) and zero output resistance (ie it can supply unlimited current to its load)

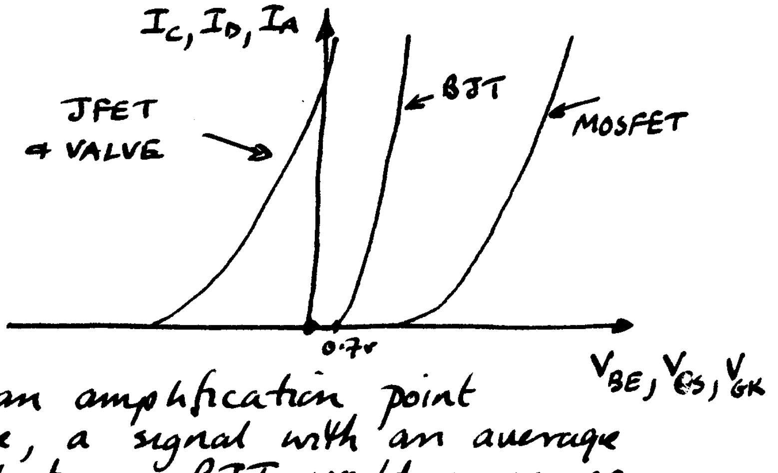
<u>Current gain</u> — the ratio of output current to input current amplitude. Used when the parameter of interest is the signal current amplitude. Used at low frequencies. Ideal current amplifier has zero input resistance (ie there is no signal voltage at the input) and infinite output resistance (ie it can supply unlimited voltage to its load)

Power gain – the ratio of output signal power to input signal power. Used at high frequencies in "impedance matched" systems. In an impedance matched system all output impedances are equal to all input impedances at a value known as the system "characteristic impedance". 50Ω is a common characteristic impedance in communications and radar applications, television systems use 75Ω. Note that in an impedance matched 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 output signal current to input signal voltage and transresistance the ratio of output signal voltage to input signal current. Transconductance is an important concept for all amplifying devices.

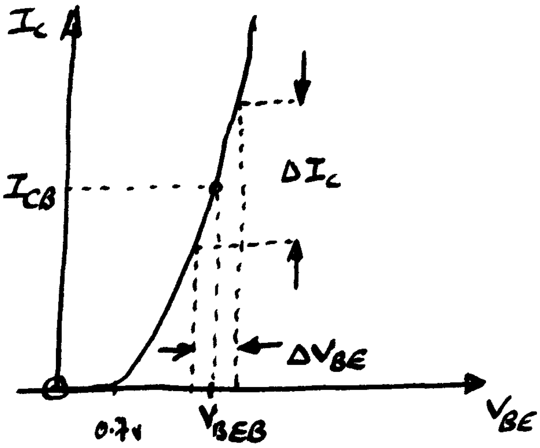
The basic mechanism of amplification

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 because it relates output current to input voltage. The transconductance characteristics for various devices are shown opposite. 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 0.7v –



In other words signal voltages below 0.7v 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 opposite. If ΔV_{BE}, the signal, was applied with no bias, ie 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.

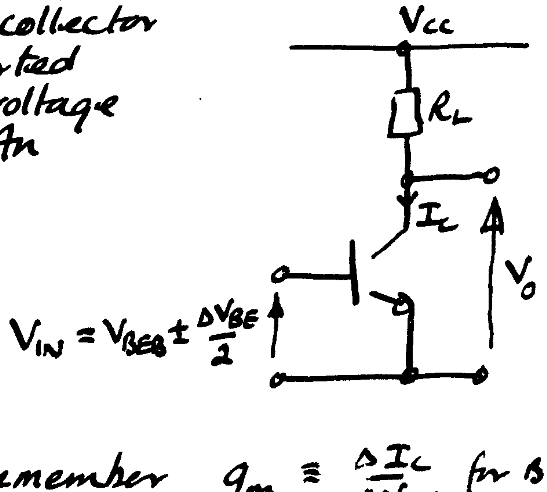


The same argument holds for all the other devices although the best choice of V_{BIAS} will be different for each.

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 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}.

The changes in collector current, ΔI_c, are converted into an output signal voltage using a resistor, R_L. An input voltage of

$$V_{IN} = V_{BEB} \pm \frac{\Delta V_{BE}}{2}$$



will give a collector current change of

$$I_c = I_{CB} \pm g_m \frac{\Delta V_{BE}}{2} \quad \text{(remember } g_m \equiv \frac{\Delta I_c}{\Delta V_{BE}} \text{ for BJT)}.$$

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

$$V_o = V_{cc} - I_c R_L = V_{cc} - I_{CB} R_L \mp g_m R_L \frac{\Delta V_{BE}}{2}$$

$$= V_{OB} \pm \frac{\Delta V_o}{2}$$

where V_{OB} is the output voltage obtained when the signal is zero ($V_{OB} = V_{cc} - I_{CB} R_L$) and $\frac{\Delta V_o}{2}$ is the component of output voltage due to the signal perturbation ($\frac{\Delta V_o}{2} = -g_m R_L \frac{\Delta V_{BE}}{2}$). By using the relationship between ΔV_o and ΔV_{BE} it is possible to estimate the voltage gain of the amplifier:

$$\Delta V_o = -g_m R_L \Delta V_{BE} \quad \text{or} \quad \frac{\Delta V_o}{\Delta V_{BE}} = -g_m R_L = \text{gain}.$$

Note that: (i) The bias conditions V_{BEB}, I_{CB} 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_{BEB}

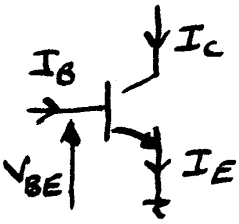
(ii) The gain is negative. This simply means that an increase in input voltage leads to a decrease in output voltage and vice versa. In signal terms it implies inversion or a 180° phase shift.

Point (i) 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 conditions is one of the primary objectives of amplifier circuit design.

BJT biassing

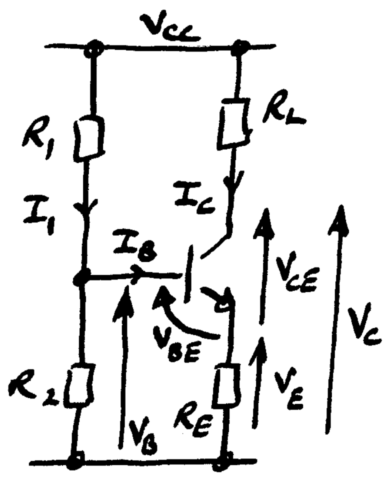
BJTs are the odd ones out in the family of amplifying devices because they need to draw an input current in order to operate. A given collector current I_C will require a base current I_B to support it and the two are related by

$$\frac{I_C}{I_B} = h_{FE}$$

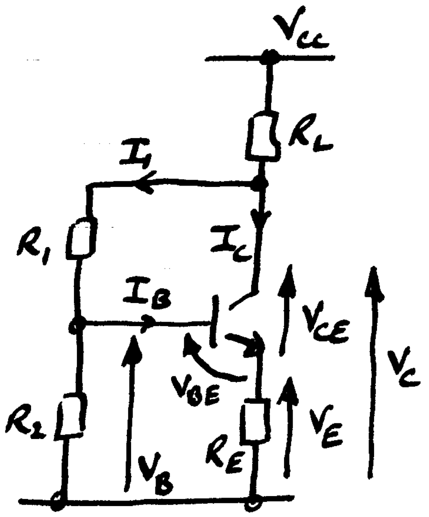


h_{FE} is the large signal static current gain of the BJT. It is approximately independent of I_C but it varies with temperature and there is a large spread of values (typically a factor of 5) from device to device of the same type. Control of the bias conditions must therefore be taken out of the hands of the transistor and placed in the hands of well defined circuit elements such as resistors.

Two types of bias circuit are suitable for single transistor BJT amplifiers....



①



②

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) then 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. Any attempt by the transistor to increase 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 $R_1 + R_2$.

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 bias conditions.

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

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

$$V_B = V_E + 0.7 = V_E + V_{BE} \quad \text{by Kirchhoff's voltage law}$$

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

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

circuit ②. – Assume I_B is negligible, $V_{BE} \approx 0.7v$ and $h_{FE} \gg 1$ (ie $I_C \approx I_E$).

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

$$\text{or } V_{CC} = I_C R_L + I_1 (R_L + R_1 + R_2). \quad \text{—— ①.}$$

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

$$\text{or } I_1 R_2 = I_C R_E + 0.7 \quad \text{—— ②.}$$

either I_1 or I_C may be eliminated from ① using ②..... for example, eliminating I_1 gives....

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

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

this result for I_C can be used in ② to find I_1 and V_C found using:

$$V_C = V_{CC} - (I_C + I_1)R_L \qquad \text{K.V.L.}$$

Notes: – it is not the results here that are important, it is the application of the basic circuit rules that lead to them.

– the only transistor voltage drop that can be used is V_{BE}. V_{CB} and V_{CE} do not and should not appear in your equations.

– the assumption "I_B negligible" really says that the existence of I_B does not disturb the potential at the transistor base to a significant extent

– ALWAYS CHECK THAT YOUR SOLUTION TO THE EQUATIONS ① + ② IN CIRCUIT 2 IS SELF CONSISTENT.

Design of bias circuits.

The design process for single transistor amplifiers involves choosing one of the two circuits, ① or ②, deciding on appropriate values of node voltages and transistor collector current and then working out suitable 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 negligible". The case most likely to violate the criterion is smallest h_{FE}. (remember that a manufacturer will specify a minimum and a maximum value of h_{FE} for a particular transistor and remember also that the purpose of the bias circuit is to control I_C) Thus $I_{BMAX} = I_C/h_{FE\,MIN}$ and I_{BMAX} is usually taken to be negligible if I_1, the current at the top of the bias chain, $\geq 10\,I_{BMAX}$.

The values 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:

- 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} + V_E, ie

$$V_c = \frac{V_{cc} + V_E}{2} \quad \text{for max symmetrical swing.}$$

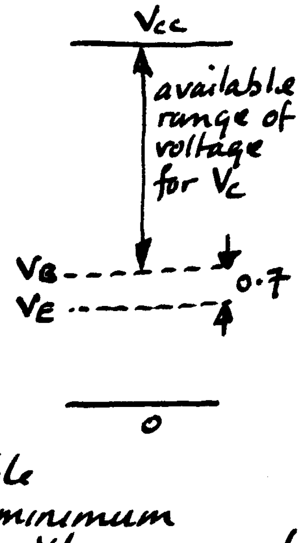
- clearly both V_{cc} and V_E will affect the max symmetrical swing, which is $V_{cc} - V_E$. V_{cc} is usually set by what is available within the rest of the system, V_E can be chosen.

- larger V_E gives more precise control of I_c. For a BJT it is unwise to let V_E fall lower than around 1V in a circuit such as circuit ①.

- I_c is chosen by considering the nature of the load.....but it also affects the effective input resistance of the transistor. In general one would aim for a condition $R_L \ll$ (input resistance of next stage)

- R_1 + R_2 should be as large as possible consistent with maintenance of the appropriate relationship between I_{Bmax} + I_1.

If you have difficulty visualising how the supply voltage will be divided up between the various parts of the circuit, it is sometimes helpful to draw a chart such as the one opposite:- 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.

- Design process is a compromise
  - no two designers would make identical decisions
  - never specify component values more tightly than is necessary
  - use preferred values.

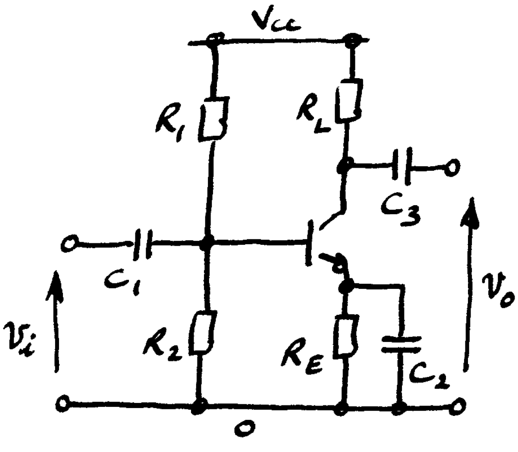
Getting signals in and out

(and removing them from where they are not wanted)

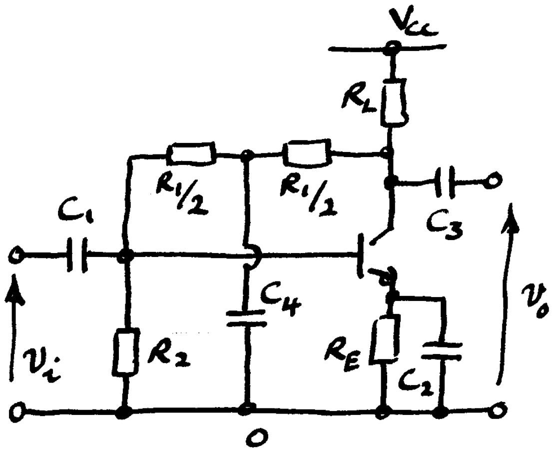
- transmitting signals from one place to another within a circuit is called "coupling"
- removing signals from nodes in the circuit is called "decoupling"

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 + ② are drawn below 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.



circuit ①



circuit ② .

Note that the signal voltages, v_i + v_o, are in lower case v whereas bias conditions are upper case V.

In both cases ....

C_1 couples the signal from the 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 signal as it has on the d.c. conditions by removing the negative feedback caused by R_E. The circuit voltage gain, 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

— transistor is characterised by non-linear characteristic curves (see real device characteristics)

— rest of circuit consists of standard cct elements such as R + C so it would be convenient to represent the behaviour of the transistor towards the signal in similar 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 basic idea of amplification involved the parameter known as transconductance — ie 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

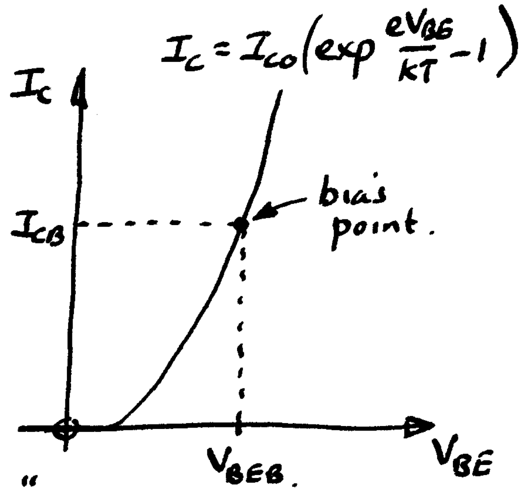
- For a BJT

$$I_c = I_{co}\left(\exp\left(\frac{eV_{BE}}{kT}\right) - 1\right)$$

and the slope at the bias point is

$$\frac{dI_c}{dV_{BE}} = I_{co}\,\frac{e}{kT}\,\exp\left(\frac{eV_{BE}}{kT}\right).$$

= "transconductance" or "mutual conductance", g_m



now for a conducting diode, $\exp\left(\frac{eV_{BE}}{kT}\right) \gg 1$

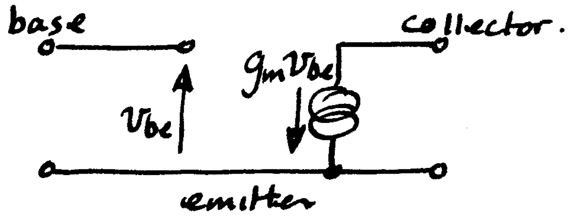
so $I_c = I_{co}\left(\exp\left(\frac{eV_{BE}}{kT}\right) - 1\right) \approx I_{co}\exp\left(\frac{eV_{BE}}{kT}\right)$

$$\therefore \frac{dI_c}{dV_{BE}} = \frac{e}{kT} \cdot I_{co}\exp\left(\frac{eV_{BE}}{kT}\right) = \frac{eI_c}{kT} = g_m.$$

$g_m = \frac{eI_c}{kT}$, where I_c is the collector current bias is one of the fundamental BJT relationships and should be remembered.

At room temperature, $e/kT \approx 40$.

This transconductance consideration leads to the simplest BJT model....



.....which is also a good low frequency model for JFETs, MOSFETs and valves (although these devices, and indeed the BJT, would probably have a resistor in parallel with the current source to take account of the slope on the output characteristic).

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 in an indirect way

$$r_{be} = \frac{dV_{BE}}{dI_B} = \frac{dI_c}{dI_B} \times \frac{dV_{BE}}{dI_c}$$

$\frac{dI_c}{dI_B} = \beta$ = small signal current gain (specified by manufacturers).

$\frac{dV_{BE}}{dI_c} = \frac{1}{g_m}$ from above

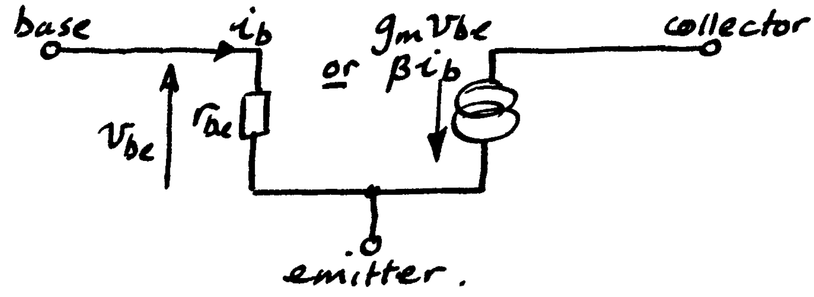
$\therefore \quad r_{be} = \beta/g_m$. – another vital BJT relationship note that dV_{BE}, dI_B, dI_C are the small changes to the bias conditions and could be represented as small signal quantities v_{be}, i_b and i_c

$$r_{be} = \frac{\beta}{g_m} = \frac{dV_{BE}}{dI_B} = \frac{v_{be}}{i_b}$$

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

..... this is an interesting result because it says that the output current generator in the BJT model may be thought of as 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 approach is more convenient. I have talked about a BJT in transconductance terms merely because the ideas can be translated easily into other device applications — no other device can be looked at as a current amplifier.

Including r_{be} in the model leads to .....



Notes - usually $\beta \neq h_{FE}$. β is a small signal parameter and h_{FE} is a large signal parameter.
- β is sometimes 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 above consisting of input resistance and output current source is reasonable for a wide range of applications and will be used for the rest of this

course.

Drawing the small signal equivalent circuit.

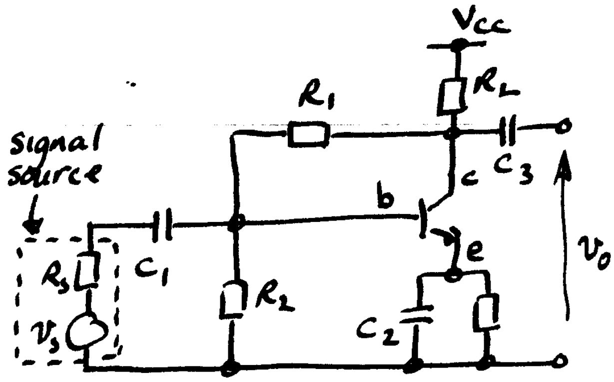
In principle this is a straightforward task – it is a matter of drawing a circuit which describes what happens to the signal alone so it is necessary to look at the circuit from the signal's point of view. There are two important consequences of being interested only in the signal's interaction with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

- All d.c. voltage sources (such as power supplies) are replaced by their Thevenin equivalent impedance, ie 0Ω ..... a short circuit.
- All d.c. current sources are replaced by their thevenin equivalent impedance, ie $\infty\Omega$ – an open cct.

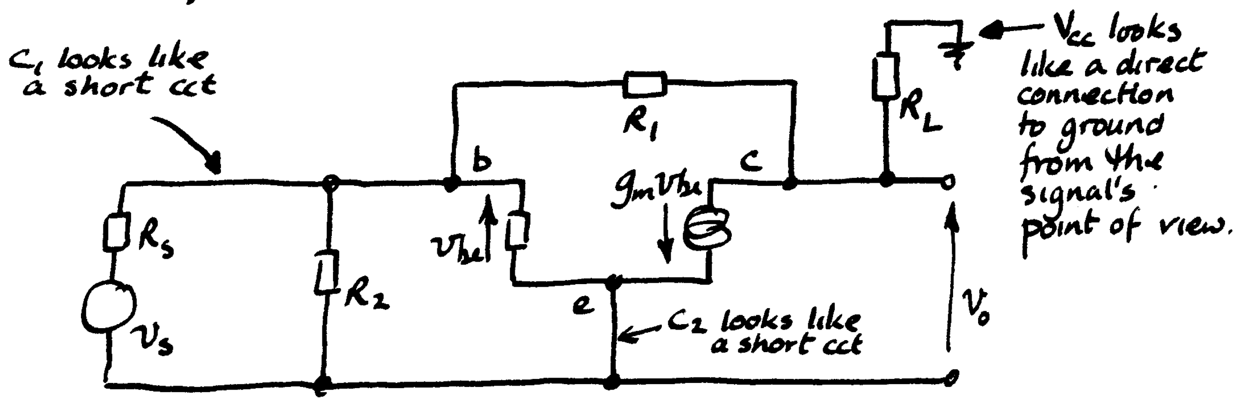
In addition, since for the purposes of this part of the course capacitors are considered as open circuits to d.c and short circuits to a.c., all capacitors are replaced by short circuits

The transistor is replaced terminal for terminal by its small signal model.......

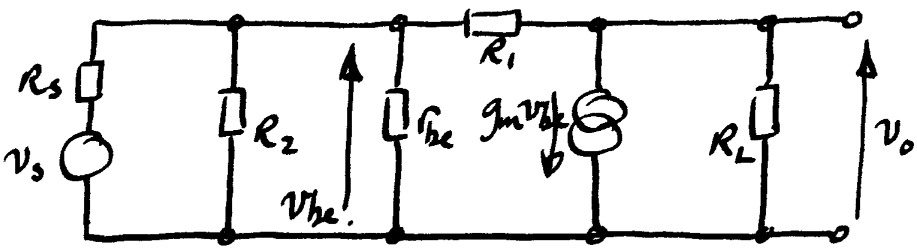
Consider circuit ②...



The S.S. equiv cct is....



or when tidied up a bit ...



<u>NOTE</u> The small signal equivalent circuit will vary according to the circuit it is derived from. Don't try + learn the result – try + understand how it was arrived at so that you can do the same for different circuits.

Once the equivalent circuit is obtained, normal circuit analysis methods can be used to evaluate performance.

<u>eg</u> what is the overall voltage gain, v_o/v_s ?

summing currents at output node...

$$v_o/R_L + (v_o - v_{be})/R_1 + g_m v_{be} = 0.$$

summing currents at input node .....

$$(v_s - v_{be})/R_S + (v_o - v_{be})/R_1 = v_{be}/R_2 + v_{be}/r_{be}.$$

These two equations can be rearranged to give respectively....

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

providing $g_m R_1 \gg 1$ ..... usually true

and $$v_{be} = \frac{v_s/R_S + v_o/R_1}{\frac{1}{R_2} + \frac{1}{r_{be}} + \frac{1}{R_S} + \frac{1}{R_1}}$$

$$= \frac{v_s(R_2 \| r_{be} \| R_S \| R_1)}{R_S} + \frac{v_o(R_2 \| r_{be} \| R_S \| R_1)}{R_1}.$$

eliminating v_{be} and rearranging to obtain the voltage gain, v_o/v_s, required gives:

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

The value of a result like this is not the numerical estimate of gain that it can provide but rather the information it offers about how the gain depends upon circuit and transistor parameters. In this case, if $R_1/(g_m R_1 \| R_L)(R_2 \| r_{be} \| R_S \| R_1) \ll 1$ the gain is controlled by the resistors R_1 and R_S and is largely independent of transistor parameters like g_m + r_{be}.

The feedback caused by R_1 can be eliminated by letting R_1 go to a very high value from a signal point of view. R_1 will then disappear from the parallel combinations to leave:

$$\frac{v_o}{v_s} = -\frac{R_1}{R_S} \frac{1}{1 + \dfrac{R_1}{g_m R_L (R_2 \| r_{be} \| R_S)}}$$

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:

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

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 + 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 to circuit 1 on page 10. 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 is futile. The important skills are the ability to obtain and analyse the small signal equivalent circuit and then to interpret the results of that analysis.