Linearity of the Transistor Amplifier

Effect of base–emitter junction

by W. T. Cocking, F.I.E.E.

The relation between the base and collector currents of a bipolar transistor is substantially linear, whereas the relation between the base-emitter voltage and the collector current is exponential. Because of this, it is often said that the transistor is a current-operated device and should be driven at its base by a high-resistance source. This is, however, a practice which is becoming less and less adopted.

Most people now tend to think in terms of voltage drive and seem rarely to bother about the exponential voltage-current relation. Statements are sometimes made to the effect that this must cause severe waveform distortion, but in practice this does not seem to occur. The reason is that an emitter resistor is nowadays nearly always included for thermal stability and if this is not by-passed at signal frequencies, the resulting negative feedback linearizes the characteristic to a very high degree.

The question arises as to how much feedback is needed. The writer knows of no published information about this, but it is surprising if no one has in fact investigated the matter.

The essentials of a typical earthed-emitter circuit are shown in Fig. 1. The whole of the emitter current flows through the base-emitter junctions and the relation between the base-emitter voltage V_{BE} and the emitter current I_E is given by the basic diode equation

$$I_E = I_S [\exp. (KV_{BE}) - 1] \qquad (1)$$

where I_S is the reverse saturation current. This relation usually holds well over about three decades of current.

The practical difficulty about using equation (1) is that I_S is rarely known. Of course, if I_E is known for a given value of V_{BE}, I_S can be calculated. This can be rather
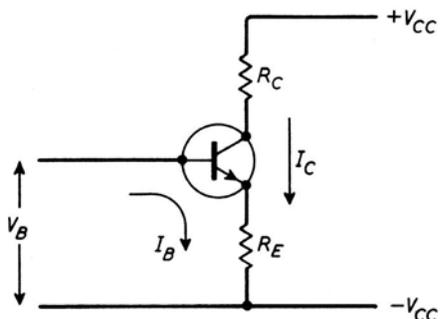
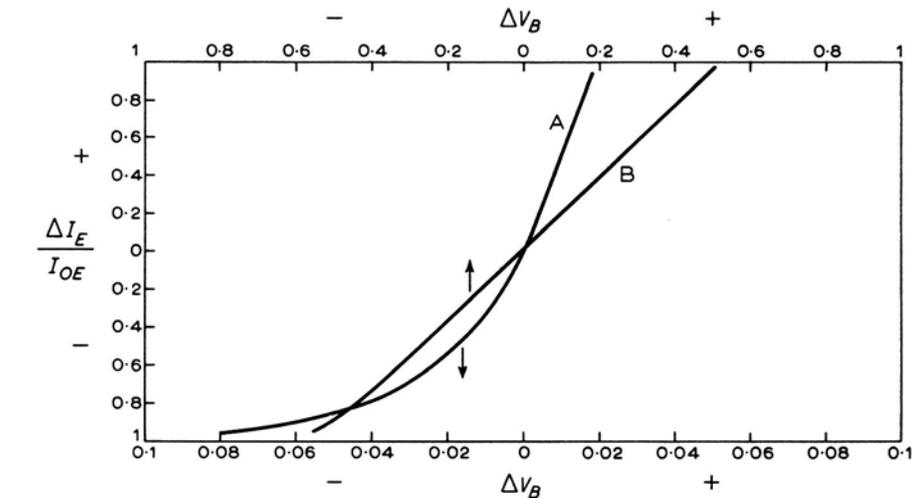*Fig. I. Basic circuit of common-emitter stage.*

Fig. 2. Input/output curves for the stage of Fig. I. Curve A is for $R_E = 0$ and curve B for $I_{OE}R_E = 0.5$ volt.

troublesome because KV_{BE} is large and I_S is very small indeed. It is, however, unnecessary. If we let I_{OE} be the current corresponding to V_{OBE}, then

$$I_E = I_{OE} + \Delta I_E$$
$$= I_S[\exp. K(V_{OBE} + \Delta V_{BE}) - 1]$$
$$= I_S \exp. KV_{OBE} \exp. K\Delta V_{BE}$$
$$= I_{OE} \exp. K\Delta V_{BE}$$

since the unity term is negligibly small. We have therefore,

$$\Delta V_{BE} = \frac{1}{K} \log_e \left(1 + \frac{\Delta I_E}{I_{OE}}\right)$$
$$= 0.06 \log_{10} \left(1 + \frac{\Delta I_E}{I_{OE}}\right) \qquad (2)$$

since $1/K = 0.026V$ at room temperature and the conversion factor from Naperian to common logarithms is 2.3.

Now the circuit equation is

$$V_B = V_{BE} + I_E R_E \qquad (3)$$

$$V_{OB} + \Delta V_B = 0.06 \log_{10} \left(1 + \frac{\Delta I_E}{I_{OE}}\right)$$
$$+ (I_{OE} + \Delta I_E)R_E + V_{OBE}$$

whence $\Delta V_B = 0.06 \log_{10} \left(1 + \frac{\Delta I_E}{I_{OE}}\right)$

$$+ I_{OE}R_E \frac{\Delta I_E}{I_{OE}} \qquad (4)$$

We have thus got rid of the awkward I_S term. The equation relates the change of base voltage from the quiescent value which produces the quiescent emitter current I_{OE} to the fractional change of emitter current $\Delta I_E / I_{OE}$, which may be positive or negative but can never exceed unity. The remaining term $I_{OE}R_E$ is the one which affects linearity. It represents the emitter voltage above $-V_{CC}$ for an unbypassed emitter resistor which carries no current but I_E.

Fig. 2 shows in curve A a plot of ΔV_B against $\Delta I_E / I_{OE}$ when $R_E = O$. This is the basic voltage-current relation for a diode and is highly non-linear. Curve B shows the relation when $I_{OE}R_S = 0.5$ volt. It is an almost perfect straight line. Reading from the curve for $\Delta V_B = \pm 0.42V$, $\Delta I_E / I_{OE}$ is 0.8 and -0.76, which represents second-harmonic distortion of only about 2.5%. Distortion can be further reduced by increasing $I_{OE}R_E$ but at the expense of gain.

It should be noted that the results apply to any transistor in the region where the base-emitter diode follows the fundamental diode equation. This is most small-signal transistors at current up to about 10mA. The peak signal output (current) for low distortion is about 75% of the quiescent current.

Of course, the collector resistance must not be too large in relation to the supply voltage, otherwise bottoming will occur with severe distortion. This is, however, a

quite independent mechanism which need not be considered here.

It is worth noting that an external base resistance R_B also exercises a linearizing effect. It is equivalent to an emitter resistor of value R_B/h_{fe} and, if large enough, gives constant-current drive. It is thus clear that the so-called constant-current drive and linearization by an emitter resistance are fundamentally the same. They both operate by making the actual base-emitter voltage changes very small compared with the signal input voltage applied to the transistor stage as a whole.

When two transistors are in push-pull, much of the basic distortion is balanced out. The true push-pull case is less easily computed than that of a phase-splitter; that is, the case of a push-pull pair with an input to only one transistor. Ignoring the small loss in the coupling resistor R_T, Fig. 3, the two collector signal currents are equal and opposite and so the input for a given current change ΔI_C in Tr_1 is ΔV_{B1} minus the input, $\Delta V_{B2}'$ for the same magnitude of current change in Tr_2. As the latter current change is negative, so is ΔV_{B2}, and the input is really the sum of the magnitudes. Curve A of Fig. 4 shows the resultant for $R_E = 0$ and is much more linear than curve A of Fig. 2. Curve B shows the result when $I_{OE}R_E = 0.1V$ This is not as good as curve B of Fig. 2, but that is for $I_{OE}R_E = 0.5V$. In all cases, for the same value of $I_{OE}R_E$, push-pull operation gives better linearity
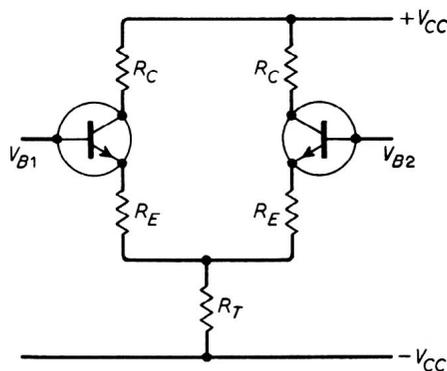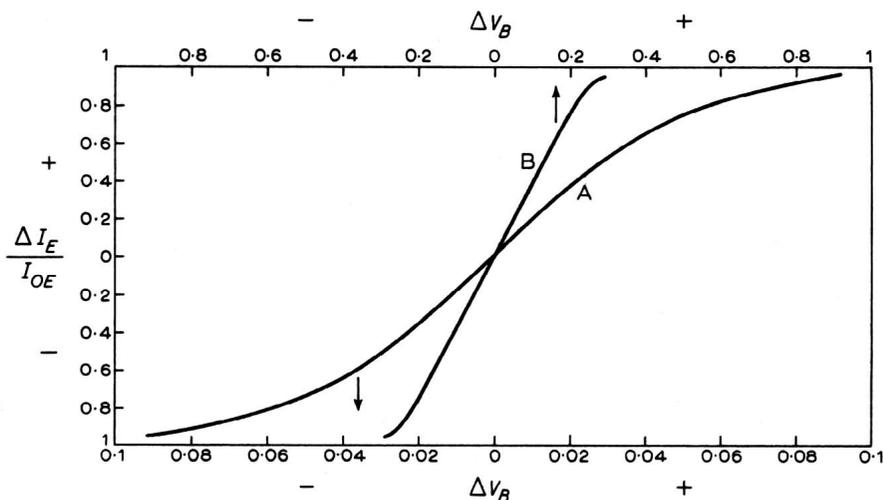

Fig. 3. Basic circuit of a push-pull stage.

Fig. 4. Performance of push-pull stage, curve A with $R_E = 0$ and curve B with $I_{OE}R_E = 0.1V$. The curves are for an input to one base only, the other being held at a fixed potential.

Tunable Phonon Sources

Hypersonic mechanical waves from sources which are 'essentially monochromatic', and can be tuned over the frequency range 150 to 280GHz, have been generated experimentally at Bell Labs in the U.S.A., Powers of up to hundreds of milliwatts are claimed. Hitherto it has been difficult to sustain coherent 'monochromatic' vibrations in a material and the energy has had the nature of wide-band noise, or just heat. One possible use of this new type of mechanical wave source is in phonon spectroscopy at previously inaccessible frequencies (the thermal frequencies). By this method new avenues of study may be opened into the electronic properties of materials, impurities and defects in crystals and the interaction of acoustic vibrations within a solid. Already the method has been used to observe the ground state splitting of vanadium impurity ions in sapphire. (The phonon is the quantum mechanical 'packet of energy' of mechanical waves, analogous to the photon of electromagnetic waves.)

The phonon generating device consists of a heater, in the form of an alloyed metal film of copper and nickel (Constantan), mounted near to a thin superconducting film of tin but separated from it by a thin layer of electrical insulator. The tin film is attached to the material being studied by phonon spectroscopy, and is sealed inside a cryostat (low temperature chamber), which is used to lower the operating temperature to 1.3°K.

The heater is turned on and off to create heat pulses. Each pulse of heat forces phonons into the superconducting film. In a superconductor the loss of resistance results from electrons becoming bound together in pairs. The phonons from the heater break these electron pairs apart, but when the electrons recombine they emit another phonon of a particular energy which is equal to the binding energy of the electron pair. The breaking and recombination of pairs of electrons generates a narrow band of high-frequency phonons.

In order to tune the phonon source, a magnetic field is applied, parallel to the thin film superconductor. The energy gap of the superconducting electrons decreases uniformly as the magnetic field is increased, and because of this the phonon frequencies can be changed by varying the magnetic field. In superconducting tin, for example, phonons can be tuned over the 150 to 280GHz range mentioned above. Bell Labs say they expect to be able to produce phonon frequencies from about 50GHz to 1THz by experimenting with different superconducting materials' and tuning them. They think they may be able to generate phonon powers of more than one watt by this technique.

To detect the phonons the researchers use a second superconductor, operating in a complementary way to the one in the phonon generator. They also make use of a property of an antimony doped germanium crystal: the antimony can be tuned by squeezing the crystal so that its energy gap corresponds to that of the superconduction energy gap. The crystal can be made to absorb phonons of only a particular frequency, and this is detected through a decrease in the amplitude of the observed signal.