THE ancestry of many electronic circuits in common use today can be traced back to the invention of the thermionic triode by Lee de Forest in 1907. For example, Eccles and Jordan published the circuit of a "two-state" or bistable electronic switch in 1919, which was based on the triode, and this circuit later became the building brick of modern digital computers. The important thing about de Forest's invention was that it introduced for the first time the active principle of amplification to circuit design.

When transistors began to appear in quantity after 1950, they were initially regarded merely as substitutes for the thermionic triode, and old circuits were adapted to accommodate them. However, increasing knowledge of semiconductor principles soon led to the development of new devices and circuits, which bear little resemblance to those of the triode.

THERMIONIC TRIODE

A basic triode consists of a thin wire filament (cathode), a wire grid, and a metal plate (anode), all contained in a vacuum, see Fig. 3.1. Electrons are thrown off by the vibrating atoms of the heated filament and travel across the vacuum space towards the positively charged anode, thus forming an electric current. There can be no flow in the other direction because the anode does not emit electrons, and there are no other current carriers present in the vacuum.

The function of the grid is to control the electron flow to the anode, and it exerts a large influence on the electrons because it is close to the filament. Thus, a small voltage change on the grid results in a large current change at the anode.

A resistor placed externally, in series with the anode connection, will convert a change of anode current into a change of anode voltage. Thus a small change in grid input voltage results in an amplified change in anode voltage. The valve acts as an amplifier.

TRANSISTOR ACTION

It was explained in Part 2 that a diode is formed by the combination of p and n type semiconductor materials. If a "sandwich" is made with two n type materials on the outside and a p type filling or central layer, this will obviously give two diodes back to back, as shown in Fig. 3.2a. Similarly when the materials are arranged in a sandwich of *pnp*, but the diodes will then be the opposite way round, as in Fig. 3.2b.

TRIODE

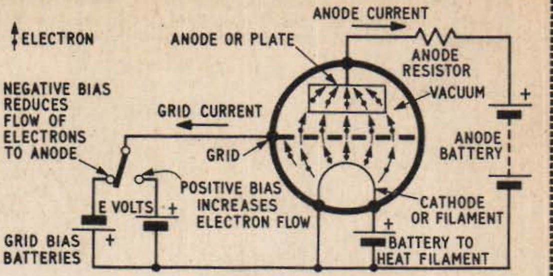


Fig. 3.1a. Working principle of a thermionic triode

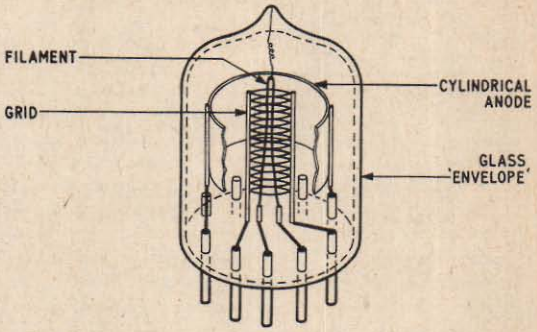


Fig. 3.1b. Physical construction of a triode valve

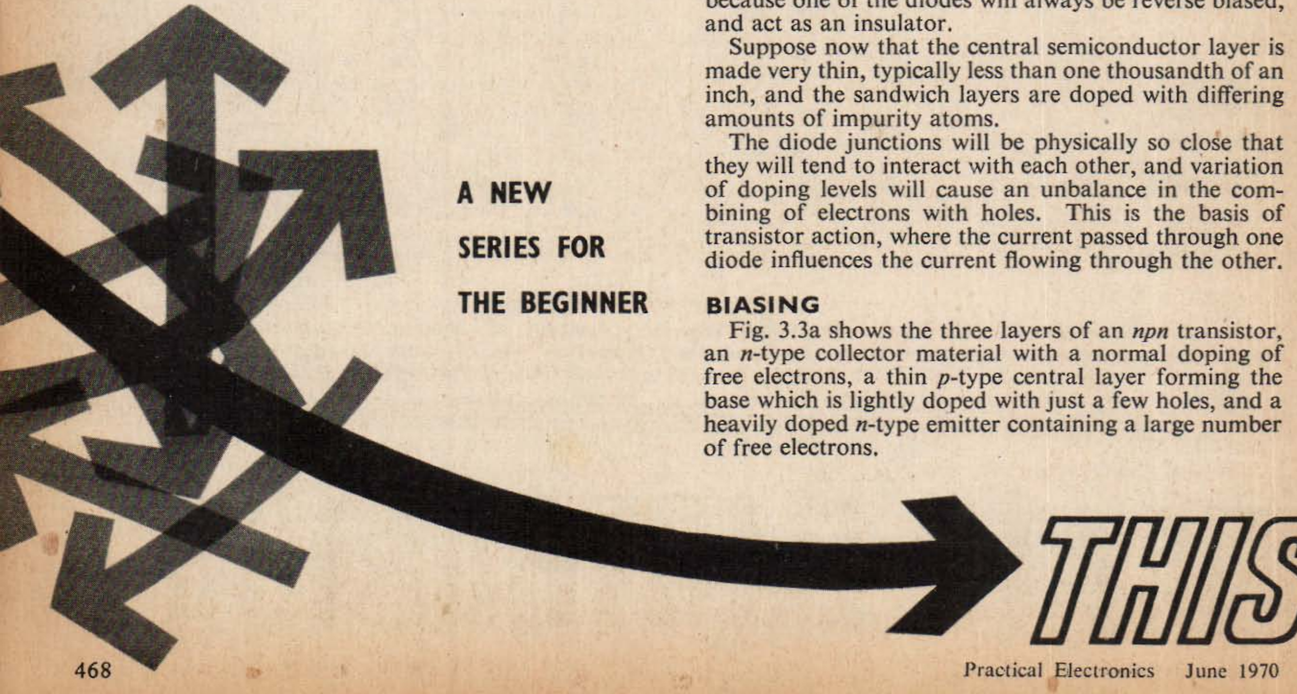
Both devices of Fig. 3.2 are incapable of conducting a significant current between the terminals marked *emitter* and *collector* when the *base* terminal is unconnected because one of the diodes will always be reverse biased, and act as an insulator.

Suppose now that the central semiconductor layer is made very thin, typically less than one thousandth of an inch, and the sandwich layers are doped with differing amounts of impurity atoms.

The diode junctions will be physically so close that they will tend to interact with each other, and variation of doping levels will cause an unbalance in the combining of electrons with holes. This is the basis of transistor action, where the current passed through one diode influences the current flowing through the other.

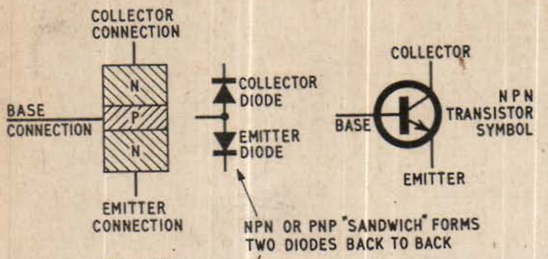
BIASING

Fig. 3.3a shows the three layers of an *npn* transistor, an n-type collector material with a normal doping of free electrons, a thin p-type central layer forming the base which is lightly doped with just a few holes, and a heavily doped n-type emitter containing a large number of free electrons.



TRANSISTOR

Fig. 3.2a. NPN transistor shown in block form (left); theoretical circuit (centre); circuit symbol (right)



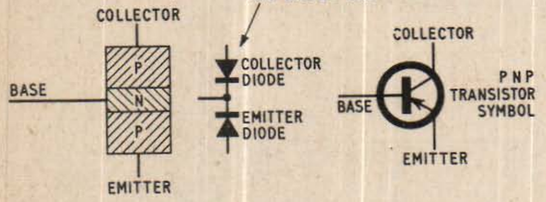


Fig. 3.2b. PNP transistor shown in block form (left); theoretical circuit (centre); circuit symbol (right)

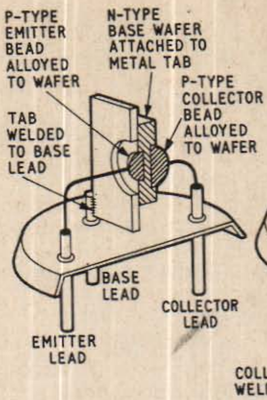


Fig. 3.2c. Cross-section view through a germanium alloy pnp transistor

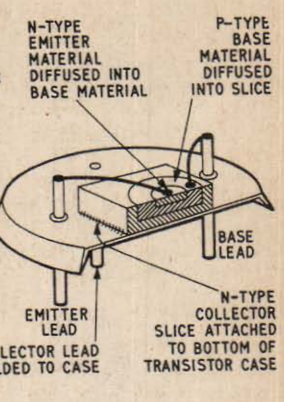


Fig. 3.2d. Cross-section view through a silicon planar npn transistor

Fig. 3.3a. Different impurity doping levels in a transistor sandwich

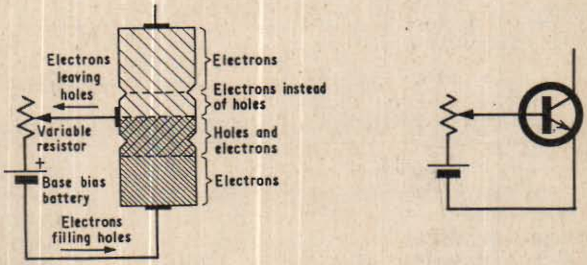


Fig. 3.3c. Free electrons move into collector insulator region when emitter junction is forward biased

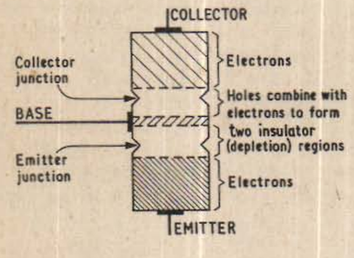


Fig. 3.3b. Two insulator regions formed in an unconnected transistor

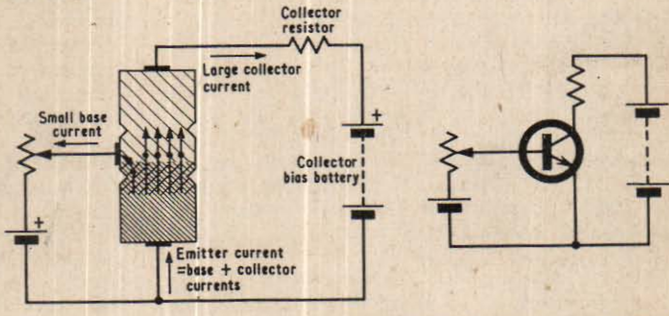
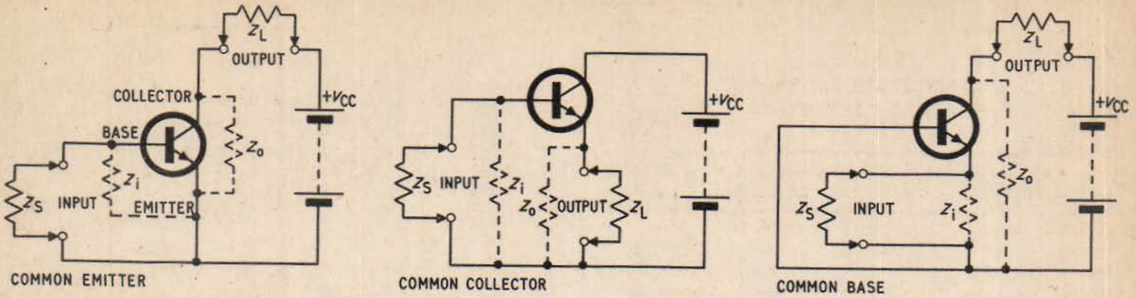


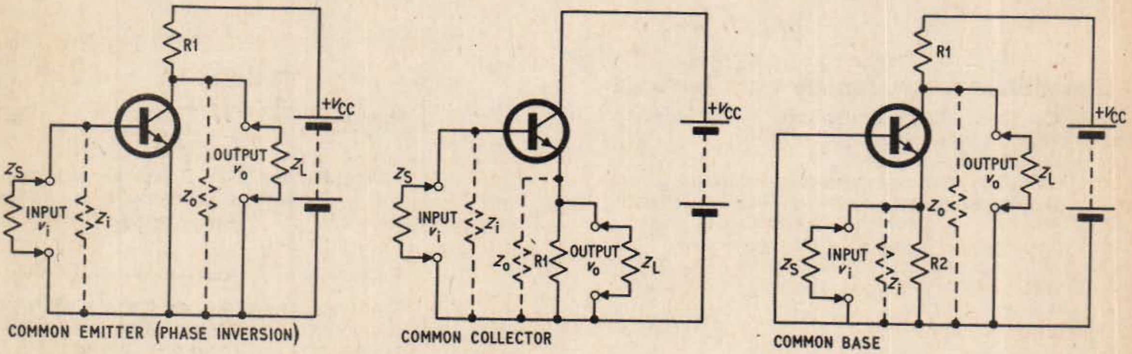
Fig. 3.3d. With base and collector bias, more electrons are carried across the collector junction than are passed through the base bias battery, thus giving amplification of the base current

WAY TO ELECTRONICS 3



COMMON EMITTER — COMMON COLLECTOR — COMMON BASE

Fig. 3.4a. The three transistor configurations connected as current amplifiers



COMMON EMITTER (PHASE INVERSION) — COMMON COLLECTOR — COMMON BASE

Fig. 3.4b. The three transistor configurations connected as voltage amplifiers

When the three layers are merged, free electrons and holes combine to form two insulator or depletion regions; one at the junction between base and collector, the other at the base-emitter junction (Fig. 3.3b).

If a base bias battery is now connected across the base and emitter, as in Fig. 3.3c, with a variable resistor in series to adjust the level of base current, the emitter diode will be forward biased. The emitter insulator region therefore disappears, and electrons will flow by the mechanism of filling and leaving holes.

However, the emitter has many free electrons, while the base material has only a few holes. So, while some electrons from the emitter are kept busy filling holes, others will be swept along by the current to find no holes vacant. These uncommitted electrons tend to repel each other, and quickly diffuse throughout the base material, into the region of the collector junction insulator.

It will be remembered that a diode insulator can only exist as such when there are holes on one side of the junction and free electrons on the other. The presence of free electrons instead of holes in the vicinity of the collector junction tends to "spoil" the diode insulator, and thus converts the junction into a conductor.

AMPLIFICATION

When a battery is coupled to the collector and emitter terminals (Fig. 3.3d) the uncommitted electrons from the emitter proceed to flow across the collector junction, under the influence of a positive charge, thus creating a collector current.

Any increase of base bias current will cause a corresponding increase of collector current, but because the base material is very thin, more electrons tend to find their way to the collector material than are "used up" by the base bias. This is called current gain or amplification. If 50 electrons cross the collector junction for every one taken by the base current, the gain of the transistor will be 50.

A *pnp* transistor functions in much the same way, except that the role of free electrons and holes is exchanged, and base and collector supply polarities are reversed. The arrowheads in the transistor symbols of Fig. 3.2 indicate the direction of "conventional" flow, not electron flow. (See Part 1 for explanations.)

THREE CONFIGURATIONS

A transistor is primarily a current amplifying device, but a current flowing through a resistance will give rise to a voltage drop across that resistance ($V = I \times R$). Therefore, a transistor can be considered as a voltage amplifier when the internal resistances of the device, and the values of external resistors connected to it, are taken into account.

There are three main ways in which a transistor can be employed to amplify small currents or voltages, in circuits termed "common emitter", "common collector", and "common base". Table 3.1 lists the main features of each configuration, and the circuits appear in Fig. 3.4 under the headings current and voltage amplifiers. For the sake of clarity, base biasing has been omitted and will be dealt with later.

IMPEDANCE MATCHING

Although an amplifier is energised by a d.c. supply, it is used to increase the voltage or current from a

Table 3.1. TRANSISTOR AMPLIFIER CHARACTERISTICS

Configuration	Common emitter	Common collector	Common base
Current gain	medium	medium	unity
Voltage gain	high	unity	high
Power gain	high	low	medium
Input impedance Z_i	low	high	very low
Output impedance Z_o	medium	very low	very high

Table 3.2. ABBREVIATIONS USED IN THIS ARTICLE

V_{CC}	Collector bias battery voltage
v_i	Input signal voltage
v_o	Output signal voltage
Z_i	Input impedance of transistor circuit
Z_L	Impedance of load applied to output terminals
Z_o	Output impedance of transistor circuit
Z_S	Impedance of signal source or generator
R_L	D.C. resistance of applied load
R_S	D.C. resistance of signal source
V	Voltage
I	Current
R	Resistance

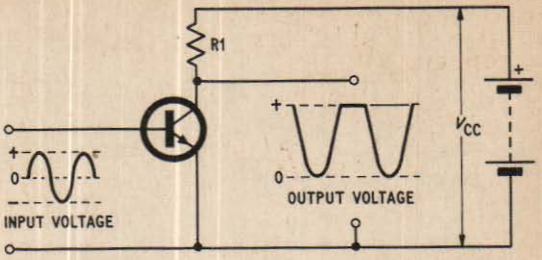


Fig. 3.5a. Common emitter a.c. amplifier without base bias. The transistor only amplifies alternate positive half-cycles and inverts the waveform

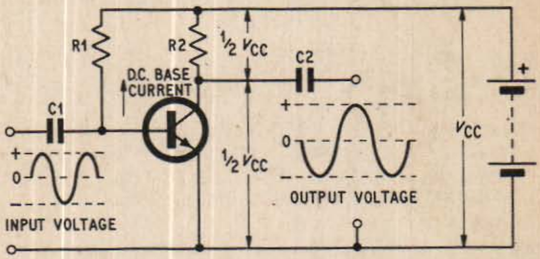


Fig. 3.5b. Common emitter a.c. amplifier with base bias and coupling capacitors. The transistor amplifies the complete sine wave and inverts it

separate a.c. or d.c. source which is connected to its input terminals. Such a source will have a certain known internal resistance.

If the source is d.c., its resistance is of value R_s; if the source is a.c., its resistive effect has to take into account variations according to the inductive and/or capacitive components of the source. In this case, the combined resistive effect is called impedance and is denoted by to the symbol Z_s.

Similarly, the load applied to the output can be a pure resistance and is termed R_L, or in the case of inductive and/or capacitive loads applied to a.c. is termed Z_L.

For simplification, the source and load are considered in Fig. 3.4 as impedances Z_s and Z_L so that they can apply to d.c. or a.c. If a d.c. source is applied, then $Z_s = R_s$ and $Z_L = R_L$.

It would be natural to assume that the input and output impedances of the amplifier itself would be the same as the internal resistances of the transistor, plus the values of any external resistors, but this is not so. The existence of amplification in a transistor circuit has the effect of modifying real resistive values to give an "apparent" value of input and output impedance denoted by dotted lines in the circuits of Fig. 3.4 and marked Z_i and Z_o respectively. If the amplifier is to work efficiently it should be "matched" to the source and load impedance, i.e. Z_s should be nearly the same as Z_i, and Z_L should be approximately equal to Z_o.

With the exception of the common emitter amplifier of Fig. 3.4b, all the circuits in Fig. 3.4 will give an output which increases as the input increases. In the case of the common emitter voltage amplifier, however, the output voltage is at maximum when the input voltage is at minimum, and decreases as the input voltage increases. The term for this is "phase inversion".

UNBIASED BASE

If the circuits in Fig. 3.4 are made up without base biasing, they will be found to amplify only input currents or voltages of single polarity. For example, the common emitter circuit of Fig. 3.4b will accept and amplify positive input voltages, but will ignore negative input voltages.

When the input voltage is zero or negative there will be no collector current, therefore the output voltage will be maximum and equal to almost the full battery voltage, V_{cc}. It follows that the output can only vary between V_{cc} and zero in response to a positive input voltage.

Fig. 3.5a shows what happens when an unbiased common emitter amplifier handles an a.c. signal. Positive half-cycles are amplified and appear at the output upside-down (phase inversion), but negative half-cycles at the input produce no change of output. How then can a complete a.c. cycle be amplified?

BASE BIAS

The following measures are taken to convert the unbiased amplifier into an a.c. amplifier. Firstly, the transistor base is supplied with d.c. bias from the battery positive terminal via R1, see Fig. 3.5b.

This base current is amplified by the transistor to yield a collector current (since the emitter is common to both circuits) which causes about half the total battery voltage to appear across R2 and the other half across

D.C. STABILITY

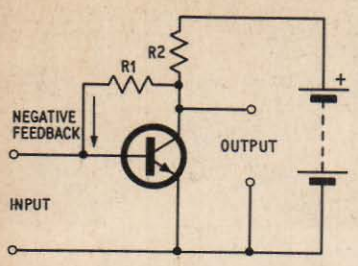


Fig. 3.6a. Base current bias with d.c. negative feedback

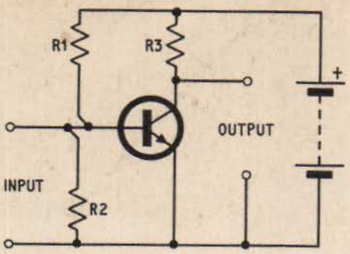


Fig. 3.6b. Base voltage bias provided by a potential divider

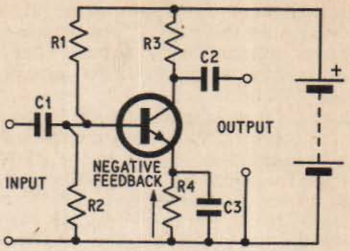


Fig. 3.6c. Base voltage bias with d.c. negative feedback and a.c. decoupling capacitor

the output terminals. So, on receipt of a signal, the output voltage can now either increase or decrease about the mean value of $\frac{1}{2}V_{cc}$.

Having established d.c. bias values, it is important to ensure that they will not be disturbed when an external circuit is connected to the amplifier input or output. A capacitor has the property of preventing a flow of d.c., but will "pass" an a.c. signal.

Capacitors C1 and C2 are therefore placed in series with the input and output terminals, and the amplifier will now respond to a.c. signals, with positive and negative half cycles appearing at the output, as depicted in Fig. 3.5b.

D.C. STABILITY

A single resistor R1 is used to set the d.c. operating conditions of the amplifier in Fig. 3.5b, but this simple method of biasing has two disadvantages. The value of R1 must be altered to suit individual transistors of slightly different current gain. The circuit is also sensitive to changes of temperature. It will be remembered from Part 2 that the resistance of a semiconductor decreases with rising temperature, and tiny changes of base current are, of course, amplified.

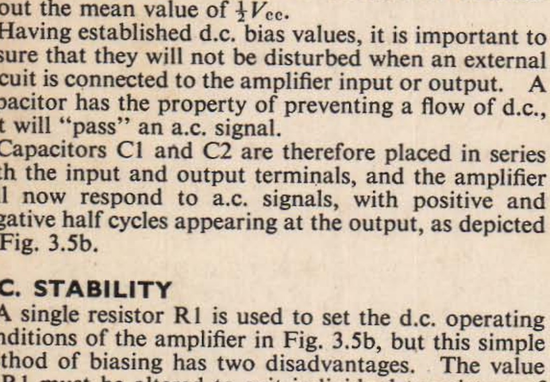


Fig. 3.7. Transistor phase shift oscillator. Frequency of oscillation is approximately 80,000/RC where R is in ohms and C is in microfarads

If R1 is connected to the collector terminal, instead of the positive battery terminal, as shown in Fig. 3.6a, d.c. stability is improved. As ambient temperature increases so does base and collector currents, but the voltage at the collector falls, thus counteracting an increase of base current and nullifying the effects of temperature.

The circuit will now accept transistors of differing gain without the need for adjusting the value of R1. Unfortunately, these improvements are obtained at the expense of amplification. The phase inverted output at the collector is fed back via R1 to the base, and is subtracted from the input; this is called *negative feedback*.

VOLTAGE DIVIDER

A prefered method of biasing is where two resistors, R1 and R2 (Fig. 3.6b) form a voltage divider across the battery, from which the base of the transistor is supplied with a voltage bias. The d.c. operating conditions of the circuit in Fig. 3.6b are moderately stable, but can be much improved if a small amount of amplification is sacrificed in the form of negative feedback.

Instead of taking feedback from the collector, a similar result can be achieved if a low value resistor R4 is inserted in series with the emitter, as in Fig. 3.6c.

To avoid loss of amplification of an a.c. signal (Fig. 3.6c is shown as an a.c. amplifier) R4 can be bypassed by a capacitor C3, without affecting the d.c. stability of the circuit. Thus, R4 limits the d.c. current for stability, while C3 acts as a short for a.c. and infinitely high parallel resistance path to d.c.

TRANSISTOR OSCILLATOR

An amplifier can be made to oscillate by the application of positive feedback. In the circuit in Fig. 3.7, the common emitter amplifier feeds a phase inverted signal to a network of resistors and capacitors. The network has the property of causing a phase inversion only at one particular a.c. frequency. Two successive phase inversions cancel out to leave a non-inverted or in-phase signal, which, when fed back to the amplifier input, reinforces the input signal and causes a build-up of oscillations. The output from the oscillator is sinusoidal, with the same waveform as mains supplies, and is derived from the laws of circular motion.

Next month we shall be looking at more oscillators, and will go on to pulse and switching circuits.