Fundamentals Section

TYPES OF TRANSISTOR

Introduction

The transistor consists of a single crystal of germanium or silicon, the properties of which are modified to produce a PNP or NPN sandwich. Many different techniques are now used in producing this result. Partly, these have arisen because of the search for techniques by which transistors could be produced cheaply, and to satisfactorily close tolerances, in spite of the extremely small sizes involved and the almost fantastically high standards of chemical purity which are necessary. The other factor is that many of the properties required for various transistor applications are conflicting.

A high power transistor must withstand relatively large currents and voltages. These cause it to heat up so that it must also withstand high temperatures and be able to quickly dissipate heat to the surroundings. This requires that the collector base junction shall have a large area, but this increases C_{CB} so that the high frequency response suffers. Again, variation of the collector-base voltage varies the width of the collector-base depletion layer. Unless the base thickness is sufficient, the depletion layer can, when V_{CB} is high, expand right across it so that the base virtually disappears. This effect is called "punch through". There are various ways of preventing this, the most obvious being to increase the base thickness, but this effects the transit time and again, the high frequency response suffers. For good high frequency response, carriers must cross the base as quickly as possible and internal capacitances must be minimised.

Although many production techniques have been tried only a few appear to be well established. The main types of transistors produced by these techniques are now described.

The Alloy Junction Transistor

At present the majority of transistors are of this type. In the example of Fig. 1, a germanium PNP transistor is shown, indium pellets having been alloyed to either side of a small disc of N type germanium which forms the base. In manufacture these pellets are held in position by a jig while the temperature is raised so that they melt and dissolve some of the germanium with which they are in contact. When cooled this part of the germanium recrystallizes with an indium impurity so that it becomes P type material. The thin region of N type germanium which remains at the centre of the sandwich forms the base which is usually one or two thousandths of an inch thick. The area of the collector junction is usually made largest to enable it to efficiently collect current carriers and to dissipate heat easily. The transistor is finally sealed into a light and damp proof capsule.

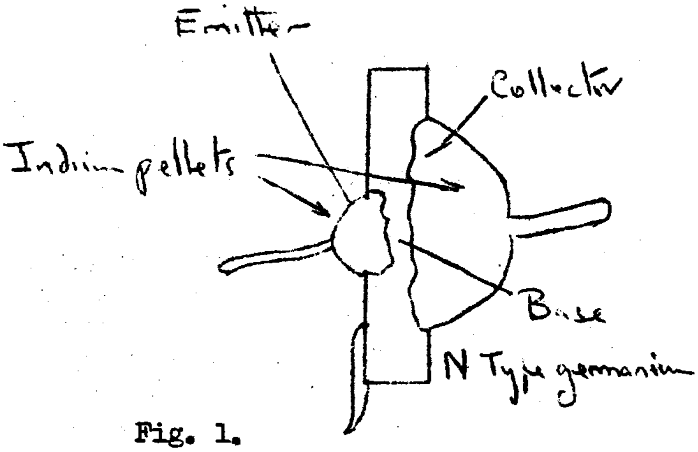


Fig. 1.

Although the majority of alloy junction transistors are of the PNP germanium type, NPN germanium and PNP and NPN silicon transistors are also available. Because of the relatively thick base region their frequency response is limited, the maximum obtainable f_α being about 15Mc/s for germanium and 5Mc/s for silicon transistors.

Transistors Produced by Diffusion Processes

The process of diffusion is used in various ways to produce several types of transistor each having a distinctive name. Impurity elements will diffuse into, and form a part of, a crystal structure, if the temperature is high enough. Two methods can be used to do this. A thin slice of the crystal may be put into an oven in which the impurity is present in gaseous form, or, the impurity may be coated onto the crystal surface before it is placed in the oven.

Although diffusion is a slow process it has the great advantage that its results can be very accurately controlled and extremely thin base regions obtained so that excellent high frequency performance results.

Fig. 2 illustrates the steps in the production of a Diffused Alloy Transistor. Fig. 2(a) shows a single slice of P or N type material which eventually forms the collector. Fig. 2(b) shows one surface of this modified to form the base region. Fig. 2(c) shows the emitter region which is formed, in this case, by the alloy process.

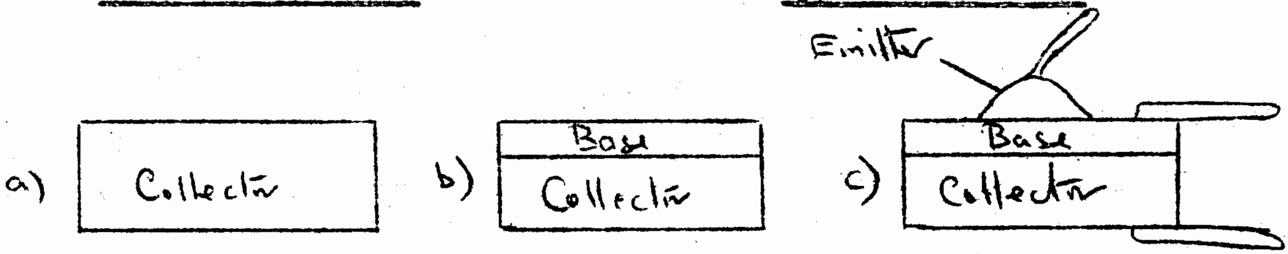
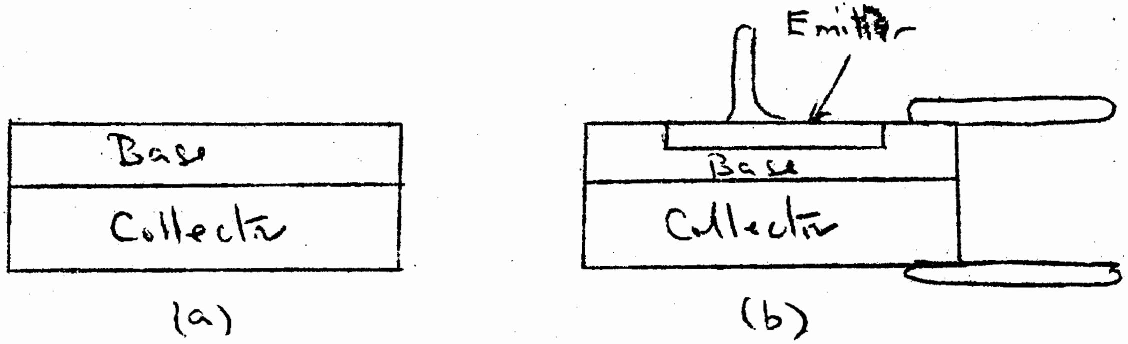


Fig. 2.

An alternative arrangement is the Double Diffused Transistor. This is illustrated in Fig. 3. Fig. 3(a) shows the stage at which the collector and base have been completed as in Fig. 2(b). Fig. 3(b) shows the finished transistor the emitter in this case also having been produced by diffusion.



(a) (b)

Fig. 3.

The Drift Field Transistor

Diffusion tends to produce a greater concentration of impurity elements on one side of the base than the other. In the drift transistor this effect is emphasised and used to produce an accelerating field across the base so that carriers move from emitter to collector more rapidly than by diffusion alone.

By this means transistors with an f_1 of about 1000Mc/s can be made. A power transistor of this type with $f_1 \doteq 120$Mc/s and capable of dissipating 4•5W is available although most of these transistors are intended for low power applications. At present all available transistors of this type are PNP germanium. Micro-alloy and micro-alloy Diffused (MADT) transistors are other examples of this general type. Both of these have extremely thin base regions, that of the latter also having a drift field.

The Mesa Transistor

This is also produced by the diffusion process, the purpose of its construction being partly to minimise C_{CB} and partly to facilitate the mass production of a cheap high frequency transistor. The production technique used permits as many as 1000 transistor elements to be simultaneously produced on a single slice of germanium or silicon, these being separated when most of the manufacturing process is complete. Fig. 4(a) shows such a slice, consisting of P type germanium or silicon, with one side modified by diffusion to be N type. In Fig. 4(b) alternate strips of this surface have been further modified by diffusion to form the emitter and base connections respectively. Fig. 4(c) shows the result of dividing the slice to form many single transistors. Finally surplus material is removed by etching to reduce the value of C_{CB}. This operation results in the mesa-like shape, shown in Fig. 4(d), from which this type of transistor gets its name.

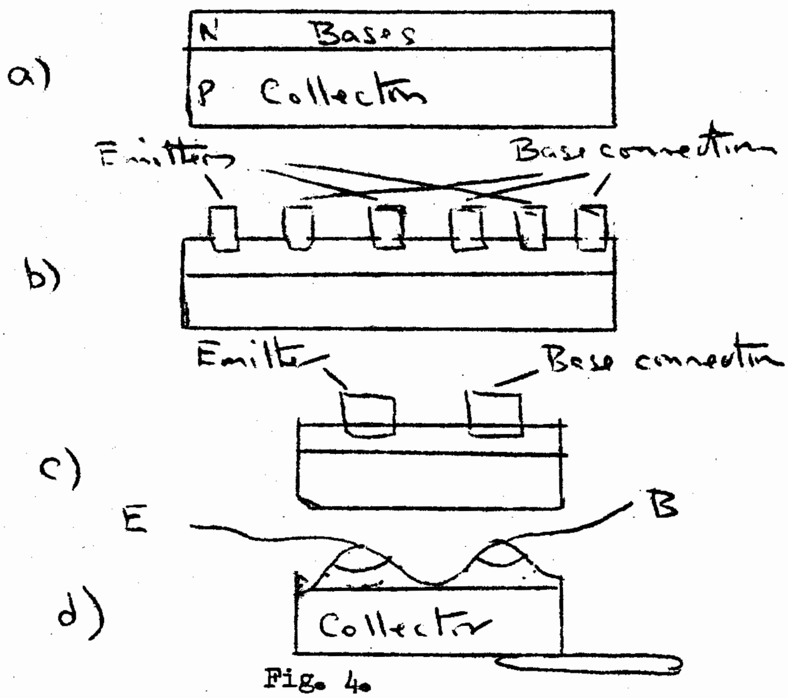


Fig. 4.

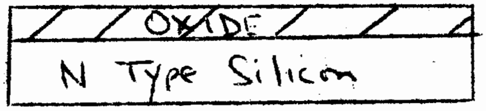
For a germanium mesa transistor f_1 can be 1000Mc/s, the corresponding value for silicon types being about 100Mc/s. This method of construction permits a very bulky collector which can be in direct contact with the metal casing which acts as a heat sink. Thus the maximum power dissipation of these transistors can be fairly high. Transistors with f_1 = 15Mc/s and maximum power dissipation of 2 Watts are easily available.

Planar Transistors

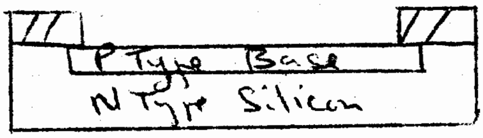
In this type of transistor an oxide coating is applied to the surface of a single slice of a suitably doped crystal which forms the collector region. Parts of this coating are then removed by etching to form a window through which a base region is produced by diffusion. An emitter region is then added using the same technique. This process has many advantages. Other processes usually cause surface effects which can adversely influence many aspects of transistor performance (for example noise, current gain and reverse current). The oxide film largely prevents this and leads to greatly improved reliability. The technique also lends itself well to mass production, it being possible to make many transistors on a single crystal slice, separating them afterwards.

Fig. 5 illustrates the production of a transistor by this method.

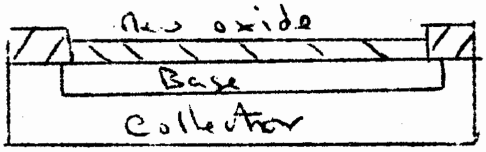
(a) N type collector region with oxide layer.



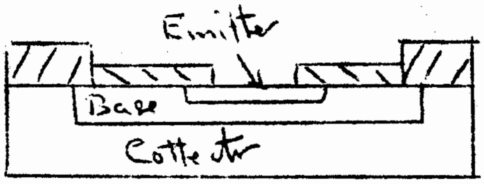
(b) Portion of oxide layer removed by etching. P type base region produced by diffusion.



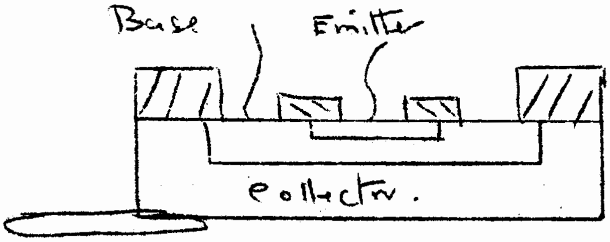
(c) Base region covered by new oxide layer.



(d) Portion of new oxide layer removed by etching. N type emitter region obtained by diffusion.



(e) Portion of base layer revealed by etching and suitable connection made to all three transistor regions.



Transistors operating up to about 1000Mc/s are possible using this technique which also permits quite large values of maximum power dissipation. Both germanium and silicon transistors of this type are possible but silicon types predominate.

Transistors Produced by the Epitaxial Process

These form the third and most promising group at present. The word epitaxy comes from two greek words which mean "arrange upon". This is an excellent clue to the technology employed in this case. A silicon crystal slice which has been very heavily doped so that its resistance is low, is placed in an oven in an atmosphere containing vapourised silicon and the required dopart. Atoms of these materials fall onto the crystal slice and slowly build up a suitably doped extension of it. A part of this layer forms the true collector, the original crystal, which is called the substrat, merely providing mechanical strength and a very low resistance connection to it. Diffusion techniques may then be used to introduce base and emitter regions into the epitaxial layer so that subsequent stages are similar to those described for the planar transistor. Fig. 6 is a diagram of a typical silicon planar epitaxial transistor. These are usually NPN at present. The process can be very accurately controlled and is capable of producing transistors with f_T of 1000Mc/s, or more, and with maximum power dissipations of tens of watts. The technique is well suited to mass production and may be expected to lead to great reductions in the cost of high frequency and high power transistors.

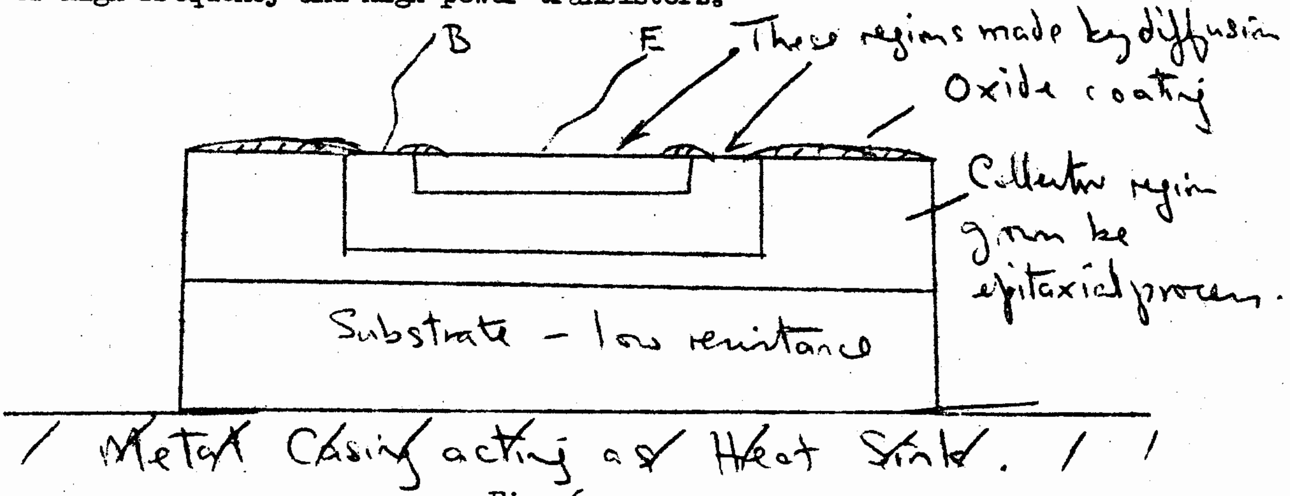


Fig. 6.

Other combinations of the techniques which have been mentioned are also used. For example epitaxy can be used to produce all three regions in one transistor. A drift field base can be produced in any transistor of the diffusion or epitaxial types. It is probable that in a few years time one or two techniques will prove superior to the others but at present the search for satisfactory production methods continues.

Comparison of the Properties of Germanium and Silicon Transistors

It must be realised that the wide range of production techniques which is available makes a direct comparison between silicon and germanium transistors very difficult. It is however, possible to deduce certain features from the known properties of these materials and in general these are a reasonable guide, though in any particular case the production technique used may alter conditions to a very large extent.

The two major differences between germanium and silicon are as follows: (a) It is more difficult to remove an electron from a silicon than a germanium atom. (b) In the germanium crystal holes and electrons move much more easily (are more mobile) than with silicon. Differences between transistors of these materials are usually due to these features.

The difficulty of removing an electron from the silicon atom means that temperature has less effect upon the parameters of a silicon than a germanium transistor. I_{CBO} is small, the device can be operated at a higher temperature and consequently its maximum power dissipation limit is higher than with germanium. Silicon devices also have higher breakdown voltages (typically up to 120 volts with silicon against 40 volts with germanium) but, which is usually a disadvatnage, a higher bottoming voltage.

The greater mobility of current carriers in germanium than silicon means that inherently the frequency response of the germanium transistor is better. This is not always obvious at present since many of the newer techniques of transistor manufacture (planar-epitaxy for example) have so far only been applied to the manufacture of silicon transistors. This has resulted in silicon transistors even being in some cases, superior in this respect of those of germanium. The question of mobility also effects transistor current gain, germanium generally giving higher values of h_{fe} than silicon.

Since the performance of silicon transistors is now adequate for most purposes the modern trend is to use these in preference to germanium since the problem of temperature stabilisation is so much easier. It would seem that the dost of silicon tramsistors will be generally comparable to germanium within a very few years.

JHB/RAS
3.11.64.